Dear Reader!
The publication of this document does not pursue for itself any commercial benefit. But such documents promote the most rapid professional and spiritual growth of readers and are advertising of paper editions of such documents.
You can download (copy) this file for educational purposes only. **PLEASE NOTE** that **ANY COMMERCIAL USE** of this file **IS STRICTLY FORBIDDEN**
You should delete it immediately after your reading is completed. Under international legislation you are responsible for prorer use of this file. All copywrite reserved by owner.

INTERFACE SCIENCE AND TECHNOLOGY

Interface Science in Drinking Water Treatment

THEORY AND APPLICATIONS

Gayle Newcombe
David Dixon
Editors



Interface Science in Drinking Water Treatment

Theory and Applications

INTERFACE SCIENCE AND TECHNOLOGY
Series Editor: ARTHUR HUBBARD

In this series:

Vol. 1: Clay Surfaces: Fundamentals and Applications
Edited by F. Wypych and K.G. Satyanarayana

Vol. 2: Electrokinetics in Microfluidics
By Dongqing Li

Vol. 3: Radiotracer Studies of Interfaces
Edited by G. Horányi

Vol. 4: Emulsions: Structure Stability and Interactions
Edited by D.N. Petsev

Vol. 5: Inhaled Particles
By Chiu-sen Wang

Vol. 6: Heavy Metals in the Environment: Origin, Interaction and Remediation
Edited by H.B. Bradl

Vol. 7: Activated Carbon Surfaces in Environmental Remediation
Edited by T.J. Bandosz

Vol. 8: Tissue Engineering: Fundamentals and Applications
By Y. Ikada

Vol. 9: Particles at Interfaces: Interactions, Deposition, Structure
By Z. Adamczyk

Vol. 10: Interface Science in Drinking Water Treatment: Theory and Applications
Edited by G. Newcombe and D. Dixon

INTERFACE SCIENCE AND TECHNOLOGY – VOLUME 10

Interface Science in Drinking Water Treatment

Theory and Applications

Edited by

Gayle Newcombe

Australian Water Quality Centre
Salisbury, Australia

David Dixon

Department of Chemical and Biomolecular Engineering
The University of Melbourne
Victoria, Australia



ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
84 Theobald's Road, London WC1X 8RR, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2006

Copyright © 2006 Elsevier Ltd. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting *Obtaining permission to use Elsevier material*

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

ISBN-13: 978-0-12-088380-6
ISBN-10: 0-12-088380-5

ISSN: 1573-4285

For information on all Academic Press publications visit our website at books.elsevier.com

Printed and bound in The Netherlands

06 07 08 09 10 10 9 8 7 6 5 4 3 2 1

Working together to grow libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER BOOK AID International Sabre Foundation

Preface

It is difficult to imagine anything more important to the human population than safe drinking water. Lack of clean drinking water is still the major cause of illness and death in young children in developing countries. In more fortunate communities, where water treatment is practiced, the primary aim of water authorities is to provide water that is free from pathogens and toxins. A secondary and very important objective is to provide water that is clear, colourless, and pleasant to taste. These latter objectives, while admirable, are often very difficult to achieve. Most countries now have water quality regulations, or guidelines, which are driving water authorities to produce purer water, with the minimum of contamination from natural or man-made origin. At the same time consumers are demanding that chemicals added during the treatment of drinking water be kept to a minimum. As a consequence, conventional clarification methods are being challenged to comply with the new regulations and restrictions and our understanding of the mechanisms involved is being tested as never before. It is in this environment that the editors felt there was a need for a rigorous review of water treatment practices from a fundamental viewpoint.

Drinking water is produced from sources such as lakes, rivers, reservoirs (surface waters) or is abstracted from below the surface (groundwater). In all cases the water is affected by the external environment. Surface and ground waters contain dissolved, colloidal and large particulate materials, which are composed of minerals or natural organic matter from the terrestrial and aqueous environments. Also present are small organisms such as bacteria, algae, protozoa and diatoms. In many instances microcontaminants from human origin are present. These include industrial and agricultural chemicals, pharmaceuticals and personal care products. It is inevitable that in the next few years our raw water supplies will be further contaminated by the products of the burgeoning nano-technological industries that are generating advanced materials for drug delivery, catalysis, energy storage devices and sensors, amongst others. Whilst it is the ability to engineer these materials at the molecular level which bestows their unique properties and functions, it is this reduced size and vastly enhanced surface area that will also challenge our existing treatment technologies.

The result of all these natural and anthropogenic activities is that our typical raw water is a complex "soup" that often requires a range of treatment processes for the achievement of water quality targets. Once those objectives have been met at the water treatment plant, the water is distributed through a series of pipes of various types and sizes to the consumer tap. Almost every step of the journey from source to consumer tap involves some sort of interaction at a surface or an interface. Therefore an understanding of the processes taking place at interfaces through the series of complex processes is vital for the provision of safe and palatable, drinking water.

As editors we have been extremely fortunate in that we have been able to convince a group of the leading experts in the field of water treatment to spend time reviewing their specific fields of expertise against a background of colloid and surface chemistry.

The first four chapters are dedicated to aspects of coagulation and flocculation; these processes are the workhorse technologies that we rely upon to remove the bulk of the unwanted contaminants. The general aim is to concentrate the particulates and adsorb the solutes on loosely bound aggregates called flocs which are then separated from the product water by processes based on differences in size or density. Despite the fact that we have been using flocs for this purpose for a very long time, there is still much to learn about their production and their properties. Brian Dempsey outlines what we know about their formation and their interactions with other particulates, John Gregory discusses their strength and stability and Bruce Jefferson reviews the way in which we usually measure their size. Brian Bolto describes the types of high molecular weight organic molecules or polyelectrolytes that we often add as a way of changing floc structure and behaviour to assist the overall clarification process. It is almost trivial to add that it is interfacial phenomena which control floc properties and clarification efficiency.

The following two chapters by Gayle Newcombe and Detlef Knappe are devoted to the use of activated carbon for the adsorption of soluble pollutants, natural organic matter or NOM as it is widely known, algal metabolites and synthetic chemicals. The selection and preparation of these adsorbents, their application and their reuse and/or disposal are all intended to maximize the uptake of these more recalcitrant chemicals by adsorption at the carbon-water interface. The current interest in the removal of cosmetic and pharmaceutical residuals can only intensify the need to choose the best possible adsorbent and to provide the optimum adsorption conditions.

The inorganic contaminants of current interest to the water treatment field are manganese and arsenic compounds and for quite different reasons – aesthetics and health. Both can be present in drinking water supplies as any of a number of soluble species of varying oxidation state depending upon the prevailing redox conditions. In each case the challenge is to remove them from solution by the provision of suitable adsorbent surfaces. The presence of manganese in drinking waters is one of the greatest causes of customer complaints because of its visibility, usually in the bathroom and laundry. Barry Chiswell outlines the causes of black bath water and stained clothing and reviews possible solutions. In contrast, arsenic is a purely a health issue, the introduction of new compliance standards in many countries prompting a surge in research projects studying better methods of removal. Martin Jekel and Gary Amy have provided an overview of these technologies.

Jim Edzwald describes the basics of a somewhat more recent treatment technology, dissolved air flotation or DAF, in which all three phases of matter (solid, liquid, gas) are present with a corresponding increase in the number of possible interfacial interactions. In this process, coagulation is still necessary to produce flocs and form aggregates but sedimentation is replaced by flotation and the contact between floc and bubbles is critical to the success of this separation.

It is interesting to trace the history of water quality monitoring from the earliest times when colour (NOM, manganese) and turbidity (particulates) were the sole criteria, to more recent times when taste and odour (algal metabolites) have become important, to today when health is the major determinant. Thus the water industry is now concerned with inorganics such as arsenic, organics such as endocrine disruptors and most recently microbiological contaminants, a good example of which is *Cryptosporidium*. Rob Considine, David Dixon and Calum Drummond have examined the forces of interaction between oocysts and sand media using a microscopic technique known as atomic force microscopy in an attempt to identify the factors that control filtration efficiency.

The latest clarification technology is membrane filtration and Tony Fane reviews the various modes of operation and the advantages and disadvantages of this increasingly popular clarification process. The focus is on surface interactions between the membrane and the contaminants both from a removal viewpoint (under what conditions is it possible to remove solutes, colloids and particles) and a fouling perspective (under what circumstances is the membrane fouled by solutes, colloids and particles).

The end product from every clarification stage is a concentrate composed of the contaminants that have been successfully removed from the water and the additives used to achieve the clarification. This waste or sludge is usually in the form of a slurry and almost regardless of its final destination (to reuse in rare cases and for disposal in most cases) it requires transportation. Dewatering then becomes a critical economic issue and Peter Scales outlines the progress that has been made in characterization and processing by applying a first principles approach, based on rheological fundamentals.

Beyond the boundaries of the water treatment plant lies the reticulation system, the network of pipes that convey the product water to the customers houses. In the final three chapters, Iwona Beech, Simon Parsons and Bruce Jefferson and Jan Vreeberg look at the surfaces present in this part of the system. In the first of these there is a review of the role of microbiology in corrosion of pipe materials; this is followed by a description of scale formation and prevention and finally a study of how the forces of interaction between pipe surfaces and particles (in combination with flow rates) determine the fate of particulates.

The final two chapters are our reference works tackling the science behind particle stability, as outlined by Charles O'Melia, and NOM, as described by Ron Beckett and Jim Ranville and as previewed in the following introductory chapter.

Gayle Newcombe and David Dixon

This Page Intentionally Left Blank

Table of Contents

Preface.....v

Chapter 1: Introduction
Gayle Newcombe and David Dixon.....1

Chapter 2: Coagulant characteristics and reactions
Brian A. Dempsey.....5

Chapter 3: Floc formation and floc structure
John Gregory.....25

Chapter 4: Practical application of fractal dimension
Bruce Jefferson and Peter R Jarvis.....45

Chapter 5: Coagulation and flocculation with organic polyelectrolytes
B. A. Bolto.....63

Chapter 6: Dissolved air flotation in drinking water treatment
James K. Edzwald.....89

Chapter 7: Membrane filtration processes and fouling
A.G. Fane, Wei Xi and Wang Rong.....109

Chapter 8: Removal of natural organic material and algal metabolites using activated carbon
G. Newcombe.....133

Chapter 9: Surface chemistry effects in activated carbon adsorption of industrial pollutants
D.R.U. Knappe.....155

Chapter 10: Manganese removal
B. Chiswell and S.-H. D. Huang.....179

Chapter 11: Arsenic removal during drinking water treatment
M. Jekel and G. L. Amy.....193

Chapter 12: *Cryptosporidium*/sand interactions during filtration
Robert F. Considine, David R. Dixon and Calum J. Drummond.....207

Chapter 13: Dewatering of water treatment plant sludges
Peter Scales.....225

Chapter 14: Biocorrosion in drinking water distribution systems
I.B. Beech and J.A. Sunner.....245

Chapter 15: Adhesion analysis of scaling systems
Simon A Parsons and Bruce Jefferson 257

Chapter 16: Fate of particles in the distribution system
Jan Vreeburg 279

Chapter 17: Natural organic matter
Ronald Beckett and James Ranville 299

Chapter 18: Fundamentals of particle stability
C. R. O'Melia 317

Index **363**

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 1: Introduction

Gayle Newcombe[a] and David Dixon[b]

[a]Cooperative Research Centre for Water Quality and Treatment, PMB 3, Salisbury, S.A. 5108, Australia

[b] Particulate Fluids Processing Centre, Dept of Chemical and Biomolecular Engineering, University of Melbourne, Parkville, Victoria 3010, Australia

1. BACKGROUND

The central tenet of any lecture course devoted to the study of colloids and surfaces is that in any system all the important interactions take place on or near whatever surfaces are present. It is easy to find numerous examples from across the whole spectrum of our activities that demonstrate the wisdom of this approach. From the earliest times to the present day, society has relied on the use of small particles; initially to provide the pigments for primitive paintings and our first paper records and now to supply the tiny vehicles for modern drug delivery systems and the nanoparticles for our latest advanced composite materials.

The technologies used to provide today's society with drinking water of the desired quality are no different, they each rely on a knowledge of what is happening at the various interfaces within the system. The oft-used schematic of the water cycle traditionally begins with clouds in the sky and traces the fate of water molecules as they fall as raindrops into catchment areas, are collected in storage reservoirs, clarified in treatment plants, distributed via pipes, used in industries and homes, discharged into rivers and transported back to the oceans for the cycle to begin again via evaporation. In each stage the key events happen at interfaces, gas-liquid at the start and finish and solid-liquid elsewhere. The extent to which changes occur and the rates at which they occur, are all driven by interfacial phenomena and any improvements to the overall process will be facilitated by improving our understanding of what is happening at the interfaces present. This is true regardless of the treatment process chosen – coagulation, flocculation, adsorption, flotation or filtration.

The focus of this book is to highlight the critical role that colloid and interfacial science plays in the removal of unwanted contaminants from the raw water and the supply of clean water to the customer and this doctrine is unashamedly repeated in each of the following chapters. The two most commonly encountered components of drinking water supplies, particulates and organics or natural organic matter (NOM), are also the most critical for each and every stage of the treatment process. For this reason they are the subject of individual chapters which review the importance of particle stability and the properties and behaviour of NOM in all aspects of water clarification. We believe these two reviews will serve as reference tomes not only for the other chapters of this book but also for researchers and practitioners in the field for years to come. In keeping with the concept and the title of Applications and Theory, we have elected to provide a summary of these reviews in this

introductory chapter and to structure the book with the reference works after the more applied chapters.

2. NATURAL ORGANIC MATTER

NOM is a generic term describing a heterogeneous and unstable mixture of compounds with a vast array of chemical moieties of varying molecular size; it is ubiquitous to virtually all aquatic environments and arises from the slow and random breakdown of organic materials of all sorts and as explained by Ron Beckett and Jim Ranville, it is an extremely complex material, the complete characterization of which still exercises some of our finest scientific minds. It has been the subject of thousands of research projects and hundreds of PhD theses and its detailed structure still defies all the advanced analytical equipment and skills that are available, even today.

We know that it has both hydrophobic and hydrophilic components much like a surfactant molecule, so that it can cause the formation of stable foams. In some circumstances it behaves as if it is stable in aqueous solution whilst in other cases it appears to take every opportunity to escape from the aqueous environment by attaching to whatever surfaces are presented. Our size fractionation efforts have shown that it can be present in solution, or as a colloid or on the surface of larger particles. Our chemical fractionation efforts have shown that it contains a wide variety of chemical groups. We know that it can adsorb on most surfaces via a combination of electrostatic and hydrophobic interactions. There is evidence that it can associate both with hydrophobic pollutants and heavy metals and that it can interact with both oxidation and reduction agents.

In short, NOM can be considered as a surfactant, an adsorbate or an adsorbent, a complexing agent, an oxidant or a reductant, a solute, a colloid or a particulate. It therefore plays a pivotal role in coagulation and flocculation interacting with the additive and also determining the size, structure and strength of the flocs, in separation, controlling both the extent and the rate of the sedimentation, flotation or filtration process and in disinfection, since it can interact directly with the chemical disinfectants often yielding unwanted by-products.

We ought not be surprised then that our efforts to extract it from whatever source, to fractionate it by whatever means (based on its size or its chemistry) and to characterize it by traditional methods, have mainly served to highlight the deficiencies of our approaches. Most of what we do in these pursuits will have altered its properties sufficiently that even if or when the particular sample has been successfully extracted or fractionated or characterized, it won't be exactly the same as the original starting material.

We also know that the nature of NOM varies from place to place and from time to time, we know that NOM is capable of interacting with every other component of the system and we know that it is present at every surface within the treatment plant and reticulation system. In practice the selection and operation of each treatment stage is usually determined by the presence and type of NOM which is why it is the subject of one of the two reference chapters in this book.

3. PARTICLE STABILITY

Particulates can originate from the soils of the catchment areas or can be generated by treatment processes or by corrosion of the pipe materials used in the reticulation systems. In any event they are a primary target of each clarification stage and a primary cause of customer

dissatisfaction with the final product; it is imperative that we understand the factors that influence their stability and ease of removal. In the second of our reference chapters, Charles O'Melia provides a rigorous description of the various types of interparticle forces and of the roles these interactions have in determining particle stability and ultimately in influencing our ability to remove these species in clarification and particularly filtration.

The causes of, and influences on, particle stability are important in drinking water treatment in the context of our efforts to destablize the suspended material. Clearly destabilisation will occur when the attractive forces between particles overcome any repulsive forces. Therefore an understanding of the origin of the forces between particles is vital in the optimisation of separation processes such as coagulation and filtration.

Between any two particles in nature there will be some type of interactive force which may be attractive or repulsive. The force the particles experience will be the sum of a number of types of forces, which individually may also be attractive or repulsive, and will be a function of the distance between the particles. These forces generally act over very short distances which are small even when compared with the radii of the particles themselves. Therefore, particles must come into close proximity to experience this net force, and, under favourable conditions, for destabilisation to occur. The two major forces acting in aqueous systems are electrostatic interactions and Van der Waals forces.

3.1. Electrostatic interactions

Particles in water can have a positive or negative net surface charge caused by charged surface sites, or adsorption onto, or complexation with, the surface of particles present in the aqueous phase; this is called the primary charge of the surface. For example, a clay particle in distilled water may have a positive charge due to surface metal complexes, however, in natural aqueous systems the net charge may be negative due the adsorption of negatively charged NOM. This primary charge is controlled by the pH and the ionic strength of the solution, and can be measured by surface titration. The pH at which the primary charge of the surface is zero is called the point of zero charge, or pH_{pzc}. In all cases of surfaces in water the surface charge is balanced by a cloud of ions (*diffuse layer*) present in solution in the interfacial region adjacent to the particle. The primary charge and the diffuse layer, taken together, form an *electrical double layer*. The electrostatic interaction between two particles occurs when they are close enough that their diffuse layers begin to overlap. For two identical particles in solution (for example two clay particles with adsorbed NOM) this will result in an electrostatic repulsion. The thickness of this diffuse layer is strongly influenced by the ionic strength of the aqueous phase. When the concentration of salt, particularly di- or tri- valent ions, increases in solution the diffuse layer is compressed and the distance at which the electrostatic interactions come into play is reduced. The presence of high levels of salt can also reduce the effect of the repulsive or attractive electrostatic forces. This is called "screening" of the interaction.

3.2. Van der Waals forces

For particles in natural aqueous systems, Van der Waals forces are attractive. They arise from three types of interactions: those between polar molecules (Keesom forces), those between non-polar molecules that can develop a dipole moment in an electric field (Debye forces), and quantum mechanical forces between two atoms which are the result of fluctuations in the charge densities of the electron clouds surrounding the nuclei of the atoms (London or dispersion forces). The latter force is not affected by the ionic strength of the

solution. In contrast the Keesom and Debye forces are electrostatic in nature and can be screened by the free ions in solution.

3.3. Derjaguin-Landau-Verwey-Overbeek (DLVO) theory

In natural aqueous systems the two major interactions that take place between particles are the van der Waals forces and the electrostatic double layer interactions. The theory that is often used to describe the combined interactions is DLVO theory. When two particles and their associated electrical double layers approach each other, their diffuse ion atmospheres overlap and a coulombic force is produced. The classic DLVO theory considers this electrostatic repulsion (or attraction) together with the van der Waals attractive interaction in assessing the interaction forces or the interaction energies between the particles. Essentially, if particles of opposite charge approach one another there will be attractive force due to the opposite charge of the diffuse layers, on closer approach van der Waals forces will come into effect, and destabilisation of the particles will occur. With similar particles in natural aqueous systems the electrostatic interactions will be repulsive, resulting in an energy barrier which must be overcome for close approach of the particles and resultant destabilisation.

The zeta potential measurement, where the speed of particles in an applied electric field is measured, gives the charge at a point close to the surface called the surface of shear. In practice, zeta potential can be monitored by making electrophoretic, streaming current or streaming potential measurements and the details of these techniques are briefly described in our chapter on coagulation.

Particles in drinking water sources producing turbidity will be negatively charged due to the presence of NOM adsorbed on the surface. As with most processes in water treatment this will dominate the charge of the particle, and therefore the stability in water.

Interfaces play a vital role in the supply of drinking water both in its treatment and in its transport. Surfaces provide high accessibility for the reactants and therefore the reactions that are necessary to remove contaminants of all sorts. The low energy barrier for mobility in the plane of the surface not only facilitates some of the complex reaction steps that are needed by encouraging reactant concentration but also enhances reaction rates. The low surface energies allow specific structure orientations to develop and favour certain interactions. These unique environments cannot be manufactured elsewhere in the system and the effectiveness of each of the various treatment technologies described in the following chapters relies on interfacial phenomena. Similarly, further advances in these processes depend on adding to our knowledge of what is happening at the interfaces.

4. CONCLUSION

In the following chapters, the individual treatment steps will be described in detail but readers are asked to keep in mind the impact that the presence of particulates and NOM have on the effectiveness of each method.

It is the editors hope that this book has not only highlighted the role of interfacial science in water treatment but that it will also act as a catalyst for further progress in both the applications and theory of water treatment.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 2: Coagulant characteristics and reactions

Brian A. Dempsey

Department of Civil and Environmental Engineering
The Pennsylvania State University
212 Sackett Building, University Park, PA, 16802, U.S.A.

1. INTRODUCTION AND DEFINITIONS

This chapter focuses on the processes that occur when coagulants are added to water. The following items are considered: composition of coagulants that contain Al or Fe, hydrolysis reactions, formation of reactive polymers that contain Al(III) or Fe(III), precipitation of aluminum or ferric hydroxides or oxides, destabilization of particles, adsorption of natural organic materials (NOM) or other dissolved materials, and the use of zeta potential (ZP) or streaming current (SC) monitors to determine mechanism of reaction and appropriate dosing and pH for coagulation. Other chapters in this book will focus on mixing, frequency of particle collisions, stabilization and destabilization, and other physical aspects of coagulation and flocculation.

Coagulation is the central and most important process in conventional treatment to produce potable water. Chemical coagulants are also used for treatment of wastewaters for removal of phosphorus, enhancement of primary or secondary clarification, or improvements in tertiary treatments such as rapid filtration or membrane filtration. Coagulants are also used for aggregation of hemi-cellulose and cellulose particles into filterable flocs during production of paper.

Coagulation is defined in this chapter as the addition of chemicals and the provision of mixing so that particles and some dissolved contaminants are aggregated into larger particles that can be removed by solids removal processes such as sedimentation, dissolved air flotation, rapid filtration, or membrane filtration. That definition and details of coagulation processes have been reported in several reviews, notably Amirtharaja and O'Melia [1]. Depending on the solids separation process, significantly different coagulation regimes might be required. As a crude overview, sedimentation requires particles that are large and dense enough to settle, dissolved air flotation requires particles that are hydrophobic enough to attach to a non-polar air bubble [2], rapid "filtration" requires chemically destabilized particles that will "stick" to the filter media plus previously attached particles [3], and membrane filtration requires particles large enough to be rejected by the pores of the membrane [4].

Flocculants are usually very high molecular weight (in the millions) straight-chain polymeric organic chemicals, although flocculant behavior can be achieved using lower MW polymers [5]. Flocculants in water treatment often function by "bridging" between several coagulated particles, thus forming larger aggregates that settle faster and that are more resistant to break-up into smaller particles in turbulent zones or during high-shear dewatering processes. Moderate to low molecular weight polymers (e.g., synthetic materials such as poly-

diallyldimethyl ammonium chloride or natural polymers such as chitosan) can assist inorganic coagulants by cumulative charge neutralization or by electrostatic patch destabilization of contaminants [6-8]. Flocculants will be dealt with in a later chapter. However, it is pertinent to note that flocculation is generally unsuccessful unless preceded by effective coagulation [9].

2. COAGULANTS

The most commonly used inorganic coagulants are alum (e.g., $Al_2(SO_4)_3 \cdot 14.4H_2O$) and ferric chloride (e.g., $FeCl_3 \cdot 6H_2O$). The number of water molecules in the empirical formula for alum changes according to source (e.g., $18H_2O$ in analytical grade alum) but the formula weight for commercial alum is usually considered to be 600. Alum is used in approximately 85% of water treatment facilities in the United States [10]. The number of water molecules in the empirical formula for ferric chloride ranges from zero to six. Ferric chloride is used in the majority of wastewater applications.

Fig. 1 illustrates the solubility of amorphous Al(III) hydroxide, $Al(OH)_3(am)$, and provides a stylized representation of the conditions for pH and coagulant dose that are usually required in order to remove contaminants during conventional coagulation, i.e., rapid mix addition of coagulant followed by clarification in some cases and by rapid filtration. This type of graph has been used by Amirtharajah and co-workers [11] and by many others to demonstrate conditions that can be used for conventional coagulation. The U-shaped solid line describes the solubility of $Al(OH)_3(am)$ using hydrolysis constants from Baes and Mesmer [12] and $\log^*K_{sp} = 9.2 = \{Al^{+3}\}\{H^+\}^{-3}$ [13]. All of the shaded areas are zones in which removal of contaminants is possible during conventional treatment, whether by sweep floc (dark shaded area) or by charge neutralisation (CN) (diagonal lines and dotted areas). The white zones above the solubility line represent conditions for which removal of contaminants is usually unsuccessful during conventional treatment, due to excessive positive charge on flocs (labeled the "restabilization zone") or due to insufficient coagulant dose resulting in very negative charge on flocs (right and lower white areas within the solubility diagram).

A diversity of coagulants is commercially available, usually containing Al or Fe or an organic polymer and sometimes a combination of two of these components. Many coagulants with Al(III) or Fe(III) salts are partially neutralized; the degree of pre-neutralization is described by the hydroxy ligand number (n), where n represents the molar ratio of hydroxide to metal in the product (e.g., n = 1.5 for a coagulant with the empirical formula $Al(OH)_{1.5}Cl_{1.5}$), or by percent basicity, i.e. the percentage of the alkalinity required to produce Al(III) or Fe(III) (hydr)oxides that is added during production of the coagulant (i.e., 50% in this example). Based on patent applications and generally accepted definitions, Al coagulants with less than 45% basicity are called basic aluminum chlorides, coagulants with basicity between 45% and 75% are called polyaluminum chlorides (PACl), and coagulants with 82% basicity are called aluminum chlorohydrates (ACH) [14]. Theoretically, ACH consists almost entirely of the thermodynamically stable polymer $Al_{13}(OH)_{32}^{+7}$, also written $AlO_4(Al(OH)_2)_{12}^{+7}$ or abbreviated Al_{13}, and this has been demonstrated using Al-27 NMR.

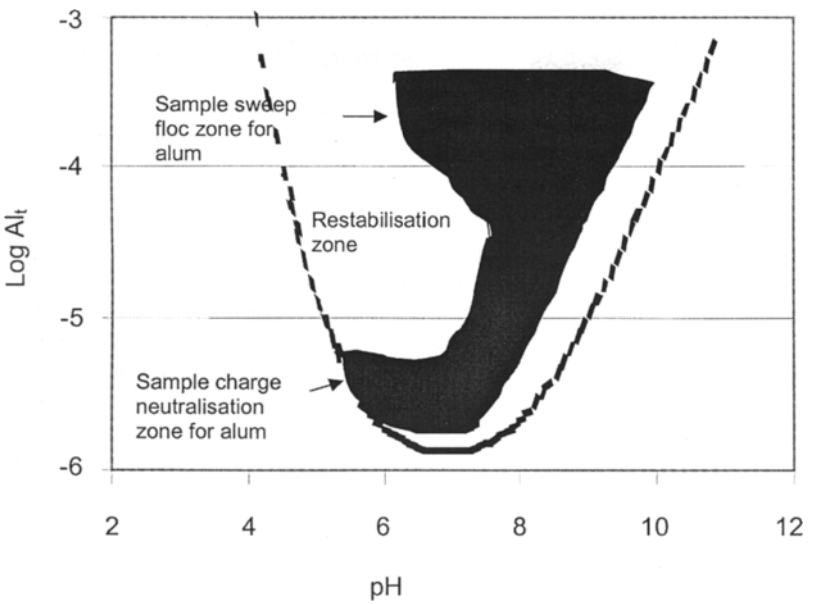


Fig. 1. Alum dose and pH conditions that result in effective removal of contaminants during conventional treatment. The zones are stylized, i.e. the positions will change depending on the concentration and type of contaminants in the raw water

It is relatively easy to produce PACl in the laboratory by addition of various alkalis to a concentrated solution of $AlCl_3$ or $Al_2(SO4)_3$ with some time allowed for dissolution of the precipitates that form [e.g., 15-18], but pre-neutralization and polymerization to produce relatively concentrated commercial PACl or ACH is performed using proprietary procedures that may require pressure reactors, control of temperature, and removal of intermediate products [e.g., 14, 19-22]. Basic aluminum chlorides are rarely marketed. The most common commercial PACl products have either 50% or 70% basicity. ACH products are also widely available. Sulfate is usually added to PACl products, because sulfate results in broadening of the effective zones of coagulation thus allowing destabilization by CN over a broader range of coagulant doses and enlarging the zone for sweep floc coagulation into more acidic pH ranges [18, 23, 24]. Polyferric coagulants are also produced and have been studied, usually with <10% basicity. There is more uncertainty about the chemical speciation in polyferric coagulants due to the much lower solubility of ferric (hydr)oxides compared with aluminum (hydr)oxides and less convincing evidence for the formation of thermodynamically stable hydroxyl polymers of Fe(III) beyond dimers and trimers. This will be discussed later in this chapter.

PACl usually works better than alum for charge-neutralization conditions or when there is inadequate mixing [24-27]. Ferric chloride is sometimes found to be a better coagulant than alum for low pH conditions. Ferric chloride and PACl usually produce flocs that settle better when the water is cold and especially when the temperature is close to freezing. It has been reported that alum works better than PACl for some waters in warm weather. Some municipalities have converted from alum for warm water to PACl for cold water, but the price

differentials between these products have decreased in the last ten years and therefore the economic incentive to make seasonal changes has decreased. Alum is sometimes combined in commercial products with sulfuric acid to produce "acid alum" to decrease the pH even more than would occur with alum alone. This is motivated by the consistent observation that removal of NOM is improved by a decrease in the pH, but application of acid alum must be balanced by the need to maintain low Al residuals. Although these generalizations are often useful, there are numerous contradictions to these rules and it is sometimes difficult to rationalize the different performance of various coagulants.

The solid-line boundaries in Fig. 2 show log Al_t and pH conditions that were effective in bench-scale tests for the settled removal of fulvic acid from Lake Drummond VA [24, 26, 28]. The PACl had 50% basicity and no sulfate. The polyaluminum sulfate (PAS) had 50% basicity and the same sulfate to Al(III) ratio as in alum. The results demonstrate that the addition of sulfate increased the pH-width of the sweep-floc zone and that addition of pre-neutralized Al(III) coagulants resulted in CN zones that were much larger in terms of effective coagulant dose than was observed for alum. These results support the claim that it is easier to achieve and maintain effective CN treatment by using pre-neutralized coagulants. The broader zones for destabilization are also useful when coagulant demand changes due to rapidly changing water quality. It is also notable that the CN zones for PAS and PACl extended outside the limits of solubility for $Al(OH)_3(am)$. This probably indicates slow dissolution of either the Al(III) polymers or the solid phases that are formed from these polymers, and this will be discussed below.

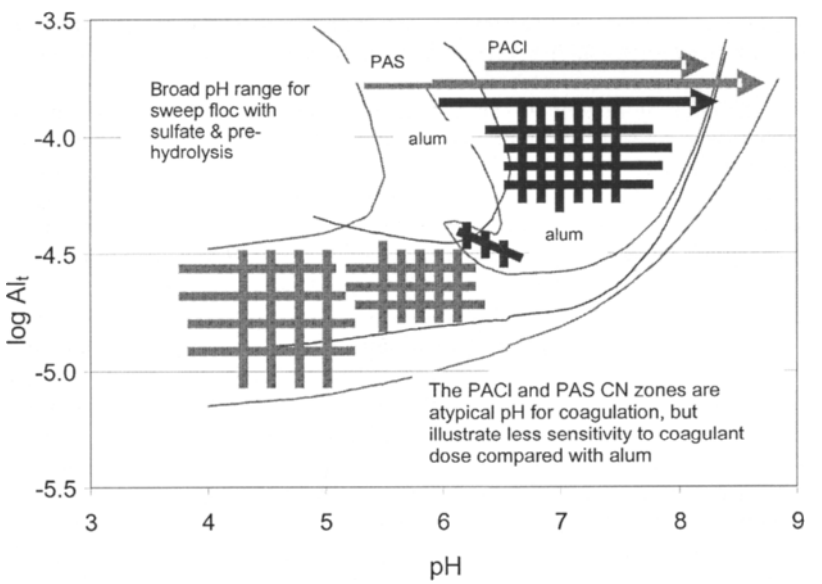


Fig. 2. Coagulant dose and pH conditions that resulted in effective removal of contaminants by charge neutralization (CN) or by sweep floc for alum, PACl, and PAS. The pH range for sweep floc is indicated by the arrows at the top of the figure. The effective range for CN is indicated by the cross-hatched regions at log $Al_t < -4.4$. The overall zones of effective removal by settling are shown by the solid-line boundaries

Specific problems may be encountered when using certain coagulants. Fe coagulants have lower pH and are much more corrosive than Al coagulants due to lower pK_a values for Fe(III), oxidizing power of Fe(III), and the catalytic effect of chloride at the low pH of ferric coagulants. Ferric salts often contain excess acidity, sometimes required to avoid precipitation of ferric hydroxide during storage. These aspects of Fe(III) coagulants can lead to severe corrosion problems if a facility that was designed for an Al coagulant is converted to use an Fe coagulant. Fe coagulants are sometimes by-products from steel-making or TiO_2 production. Quality control procedures and clear policies for acceptance or rejection of coagulants should be enforced no matter which commercial coagulant is used, but this is especially important for Fe coagulants if they are by-products. Alum is most frequently produced from alumina, but if it is produced directly from bauxite then greater emphasis should be placed on quality control [29].

Most synthetic organic coagulants are straight-chain, cationic polymers. The cationic charge density of organic coagulants is considerably higher than for $Al(OH)_3(s)$ or $Fe(OH)_3(s)$. As a result, organic coagulants offer the benefit of reduced sludge volume. However, dosing control is much more difficult for a pure organic coagulant than for inorganic coagulants like alum or ferric chloride. Also, organic coagulants are usually less effective than inorganic coagulants for the removal of NOM or other dissolved contaminants [24, 30]. As a result, complete substitution of organic coagulants for inorganic coagulants is rare. Several coagulant manufacturers blend organic coagulants with inorganic coagulants. These products can work well for charge-neutralization processes and still provide removal of NOM. The purchaser should demand complete product specifications and training in diagnosis of problems and identification of correct coagulation conditions, otherwise the buyer may become dependent on the vendor for assistance in selection of product and coagulation strategies.

Doses of coagulants are often based on the empirical formula and expressed in mg L^{-1} and this frequently causes confusion when comparing coagulant doses for different products or preparations. The formulas that were provided for alum and ferric chloride are the most common representations of these coagulants when they are purchased from vendors of commodity chemicals. However, the dose of alum can be expressed either as mg L^{-1} of dry alum or as mg L^{-1} of liquid alum (~50% water by mass). Ferric chloride can also be expressed as $FeCl_3$ or $FeCl_3 \cdot H_2O$ or other empirical formulas. This has been a problem in terms of purchase and application of coagulants; it has also been a problem in the scientific literature, where incorrect conclusions have sometimes been drawn about the relative benefit of one coagulant over another [27]. It is most reasonable to compare coagulants in terms of molar (or millimolar) concentrations of Al or Fe added. The volume of sludge generated, the consumption of alkalinity, and other indicators of performance are related to mols of coagulant metal rather than to mg L^{-1} concentrations. It is also easier to diagnose the coagulation mechanism if coagulant doses are expressed in mM rather than mg L^{-1}.

Table 3-1 of the U.S. EPA *Enhanced Coagulative and Enhanced Precipitative Softening Guidance Manual* (EPA 815-R-99-012) [31] described "Coagulant Dosage Equivalents" for six coagulants: dry filter alum, reagent grade alum, hydrated ferric chloride, ferric chloride, hydrated ferric sulfate, and ferric sulfate. The purpose of the EPA table is to provide guidance for performing jar tests using laboratory chemicals rather than bulk coagulants. The information in the table refers to dry chemicals that can be found in the laboratory, rather than to commercially available coagulants. However, manufacturers and vendors provide sufficient information (including specific gravity and w/w concentration of some form of the coagulant)

so that it is always possible to calculate relative doses based on the $\%Fe_2O_3$ or Al_2O_3. There can be other misunderstandings; liquid alum in the U.S. is usually sold in terms of the "dry" alum content. On the other hand the %Al or Al_2O_3 in poly-aluminum chloride is usually expressed relative to the wet weight of the product. Material Safety Data (MSD) sheets for coagulants that contain Al should state the $\%Al_2O_3$.

Characteristics of the most commonly used inorganic coagulants are given in Table 1. Negative values for basicity mean that the coagulant contains free acidity, i.e. excess HCl or H_2SO_4. Free acidity is present in most Fe(III) salts due to the method of preparation and to prevent precipitation of Fe(III) (hydr)oxide. Free acidity may be added to alum in order to decrease pH after addition of the coagulant and thereby increase the sorptive removal of NOM.

Table 1
Typical characteristics of coagulants (*eq mL* represents the mL of coagulant needed to obtain the same Al(III) or Fe(III) dose as 1.0 mL of alum)

Coagulant	Typical composition	% basicity	% Al_2O_3 or Fe_2O_3	Molarity of metal	eq mL
Liquid alum	$Al_2(SO_4)_3 \cdot 14H_2O$	0%	8.3%	2.17	1.0
Dry alum	$Al_2(SO_4)_3 \cdot 14H_2O$	0%	17.1%	-	-
Acid alum	$Al_2(SO_4)_3 \cdot 14H_2O$ plus H_2SO_4	-13 to -97%	4.4 to 8.3%	1.24 to 2.09	1.0-1.5
PACl (50% basic)	$Al(OH)_{1.5}(SO_4)_x(Cl)_y$	50%	6 to 10.25%	1.55 to 2.71	0.8-1.4
PACl (70% basic)	$Al(OH)_{2.1}(SO_4)_x(Cl)_y$	70%	10.5%	2.71	0.8
Aluminum Chlorohydrate	$Al(OH)_{2.46}(SO_4)_x(Cl)_y$	82%	23.6%	6.2	0.35
Alum-polymer blends	$Al_2(SO_4)_3 \cdot 14H_2O$ & cationic polymer	0%	<8.3%	-	-
Ferric Chloride	$FeCl_3 \cdot 6H_2O$	Negative (free acid)	18.7 to 22.1%	3.2 to 3.5	0.62-0.68
Ferric Sulfate	$Fe_2(SO_4)_3$	-0.5% typical	14.3 to 18.6%	2.6 to 3.7	0.59-0.83
Polyferric coagulants	$Fe_2(SO4)_{2.7}(OH)_{0.6}$	0 to 14%	14.3 to 18.6%	2.6 to 3.5	0.62-0.83

All of the pre-neutralized coagulants will gel if diluted. Vendors of these coagulants caution against dilution of stored coagulant or the use of chase water. Use of finished water to chase alum can also result in formation of precipitates, and this can result in poor performance of the coagulant and can even result in completely clogged supply lines. In general, coagulants should be fed at full strength to the rapid mix location. Gelling can also occur due to mixing of coagulants. Vendors of PACl recommend that all tanks, lines, and pumps should be rinsed with clean water before contact with the new coagulant. The new coagulant should

be circulated through pipes and pumps, to avoid clogging due to dilution or mixing with incompatible chemicals. All storage tanks, pipes, and pumps should be cleaned prior to changing coagulants.

3. HYDROLYSIS AND PRECIPITATION REACTIONS

Al(III) and Fe(III) can form thermodynamically stable monomeric, polymeric,and precipitated (hydr)oxy species. The identification and quantification of dissolved species is based on a range of qualitative and quantitative methods of analysis, including solubility measurements at pH extremes, or potentiometric titrations followed by least-squares modeling to predict the dominant soluble species and formation constants [12, 32, 33]. Spectroscopic techniques have been used for qualitative and quantitative analysis of dissolved and solid Al species. For Al there is general agreement that monomeric species $Al(OH)_n^{3-n}$ form with n from 1 to 4, although some suites of thermodynamic constants do not include the n=3 monomer. Several polymeric species have been proposed for Al(III), but the dimer, trimer, and Al_{13} are most consistently reported as the stable thermodynamic species [12, 34-36]. Some reported constants for hydroxyl complexation or precipitation of Al(III) are presented in Table 2, where log $K_1^* = \{AlOH^{+2}\}\{H^+\}\{Al^{+3}\}^{-1}$, log $\beta_2^* = \{Al(OH)_2^+\}\{H^+\}^2\{Al^{+3}\}^{-1}$ (etc.), and log $K_{sp}^* = \{Al^{+3}\}\{H^+\}^{-3}$ in the presence of the identified solid phase. There are disagreements about the species and the best thermodynamic constants; individual constants should not be selected from different sets of constants since the accuracy in prediction of the degree of hydrolysis (n) resides in the suite of constants rather than in individual constants [12, 32, 33].

Table 2
Selected hydrolysis constants for Al(III) reported as log K* (first hydrolysis constant), log β* (subsequent cumulative constants), and log K_{sp}* for solubility constants

Species	n	Ref [12]	Ref [37]	Ref [38]	Ref [39]	Ref [34]
$AlOH^{+2}$	1	-4.97	-4.97, -5.01#	-5.00	-4.997	-4.99
$Al_2(OH)_2^{+4}$	1	-7.7			-7.694	-7.7
$Al_3(OH)_4^{+5}$	1.33	-13.94	-13.9		-13.888	-13.9
$Al(OH)_2^+$	2	-9.3	-9.3,-9.24#	-10.1	-10.894	(-9.3)
$Al_{13}(OH)_{32}^{+7}$	2.42	-98.73	-98.7			-98.8
$Al(OH)_3$(aq)	3	-15.0	-15.0, -15.1#	-16.8	-16.691	(-15.0)
$Al(OH)_4^-$	4	-23.0	-23.0, -19.82#	-22.7		-23.0
$Al(OH)_3$(am)	3	-	10.85#			
α-$Al(OH)_3$(gibb)	3	8.5	8.5, 8.05#			8.5

*symbol after K or β means the reaction is written using H^+
#based on ΔG^0 values, p. 990 ref [37]

The core-and-links model has been used to describe the formation of a range of meta-stable polymers of Al(III). According to this model, colliding Al monomers can produce a dimer, and then colliding dimer and monomer can produce a trimer and so on [36, 40]. Except

for the dimer and trimer, the larger polymeric species are usually assumed to be unstable intermediates for formation of precipitates rather than thermodynamically stable polymeric species. The core and links model often invokes the formation of rings that contain six Al connected by (hydr)oxy linkages, and it is proposed that these hexamers then coalesce to form the planar structure of crystalline precipitates of Al such as gibbsite or bayerite.

In contrast to the intermediate polymeric species that are invoked by the core-and-links model, the formation and thermodynamic stability of Al_{13} has been extensively documented [41-46].

Table 3
Selected hydrolysis constants for Fe(III) reported as log K* (first hydrolysis constant), log β* (subsequent cumulative constants), and log K_{sp}* for solubility constants

Species	n	Ref [12]	Ref [37]	Ref [47]	Ref [39]
$FeOH^{+2}$	1	-2.19	-3.05	-2.02	
$Fe_2(OH)_2^{+4}$	1	-2.95	-2.91		-2.894
$Fe_3(OH)_4^{+5}$	1.33	-6.3	-5.77		-6.288
$Fe(OH)_2^+$	2	-5.67	-6.31	-5.75	
$Fe(OH)_3$(aq)	3	<-12	-13.8	-15	
$Fe(OH)_4^-$	4	-21.6	-22.7	-22.7	
FeOOH(am)	3	2.5	3.96, 3.00[#]	3.2	
α-FeOOH (goeth)	3	0.5	-1.69[#]		0.491
α-Fe_2O_3 (hem)$^+$	3		-1.92[#]		-0.709

*symbol after K or β means the reaction is written using H^+
$^+$based on $½Fe_2O_3+3H^+ = Fe^{+3} +1½H_2O$
[#]based on ΔG^0 values, p. 996 ref [37]

Some studies have reported that Al_{13} can form in soils [48] and upon addition of alum or aluminum chloride to water with sufficient alkalinity to result in near-neutral pH [40, 49]. Al_{13} is considered the key reactive species in most PACl and ACH coagulants [13, 24]. The formation of Al_{13} is favored by high $[Al^{3+}]$ and pH, but slightly higher pH results in precipitation of aluminum (hydr)oxide. Most of the recipes for laboratory or commercial production of PACl or ACH involve mixing a basic material into an acidic Al(III) solution such that precipitated phases are momentarily supersaturated before mixing is complete, and then the precipitated phase dissolves, leaving Al_{13} [12]. The Al_{13} structure contains the same hexamers of octahedral Al on each side of the central tetrahedral Al that are proposed to form during the core-and-links progress towards a precipitate and which form the planar structure of crystalline Al(III) (hydr)oxides. The pre-formation of these precursors to precipitates has been used to rationalize the improved effectiveness of PACl or ACH (relative to alum) when there is poor mixing or cold water; the formation of effective coagulant species from alum is dependent on the chemistry of the dilution water, the intensity of mixing, and the rates of competing reactions.

Fe(III) also forms monomeric species $Fe(OH)_n^{3-n}$ with n from 1 to 4, although some authors do not include the n=3 monomer (see Table 3). Several polymeric species have been proposed for Fe(III), but only the dimer and trimer are consistently reported as stable species [12]. Studies show that true nm-sized particles can be formed from Fe(III) salts for very low n values [50, 51]. Fast titration of $FeCl_3$ with $NaHCO_3$ produced 3 to 7 nm particles that were meta-stable, i.e. the particles did not precipitate after weeks of storage [51]. Slower titration with NaOH resulted in >50 nm particles [50]. There was no indication that polymers (except for dimers and trimers) were produced in any of these studies. Other studies have demonstrated that 2 to 3 nm particles can persist for years [52], especially for pH far from the pH_{zpc} where aggregation is electrostatically inhibited. This is analogous to the charge restabilization zone in conventional coagulation diagrams [46, 53, 54]. Commercial pre-neutralized Fe(III) coagulants have very low n values and likely consist of electrostatically stabilized nanoparticles of FeOOH(s) [55]. The stability and the effectiveness of poly-iron coagulants is improved by rapid neutralization [51] resulting in very small particles that retain a relatively large cationic charge. It is also reported that neutralization of $Fe_2(SO4)_3$ results in precipitation of H-jarosite that dissolves upon dilution and with the increase in pH that occurs when the coagulant is dispersed into raw water. Flocculation of the nano-particles of Fe(III) or Al(III) occurs upon increase in pH, related to the decrease in specific surface charge [56].

Crystalline Al(III) or Fe(III) oxides do not form during coagulation processes or during storage of most coagulated solid phases. Inhibition of crystallization could be due to "poisoning" of the reaction due to sorption of NOM onto the surface of the amorphous precipitate. The initial production of a more soluble and amorphous precipitated phase rather than the "thermodynamically" stable phase is a specific demonstration of Ostwald's Rule, which states that the more super-saturated the initial condition the more likely that the amorphous and more soluble phase will form, especially for high interfacial tension between water and the crystalline phases [37, 57]. Interfacial tension results in positive Gibbs' free energy, and crystal growth to stable particles is only spontaneous after a critical nucleus size has been exceeded. Predicted increases in solubility as a function of the number of Fe or Al atoms in the precipitate are shown in Figure 3 for hematite and gibbsite, based on published interfacial tensions [37] and assuming zero interfacial tension between amorphous aluminum or ferric (hydr)oxide and water. The assumption of zero interfacial tension for the amorphous phases means that the critical nucleus size is infinitely small, consistent with previously cited reports that the diameters of primary particles in flocs of amorphous aluminum or ferric (hydr)oxide are only a few nanometers. These characteristics of the amorphous precipitates will be invoked to postulate a qualitative model for the coagulation process.

The meta-stability of amorphous Al(III) and Fe(III) (hydr)oxides results in significantly higher solubility than would be predicted for large crystals of more stable precipitates such as hematite and gibbsite. This is important with respect the formation of Al_{13} since the concentration of Al_{13} at a given pH is proportional to $[Al^{+3}]^{13}$. This effect is illustrated in Fig. 4, which shows that Al_{13} becomes the dominant dissolved species for a broad range of acidic pH and high total dissolved Al(III) when $Al(OH)_3$(am) is the solid phase, but Al_{13} is never a dominant species in the presence of gibbsite. This figure also shows pH and Al(III) concentrations typical for alum (left-most circle), a PACl (middle circle), and a PAS preparation (right-hand circle); Al_{13} is "stable" in these coagulants for these materials if crystalline phases do not precipitate. There are many anecdotal reports that polyaluminum coagulants have gelled when stored in tank cars that are heated by ambient conditions; this is consistent with the shift of $Al(OH)_3$(s) solubility to lower pH values with increase in temperature, e.g., [58] or with the precipitation of crystalline solids at higher temperature.

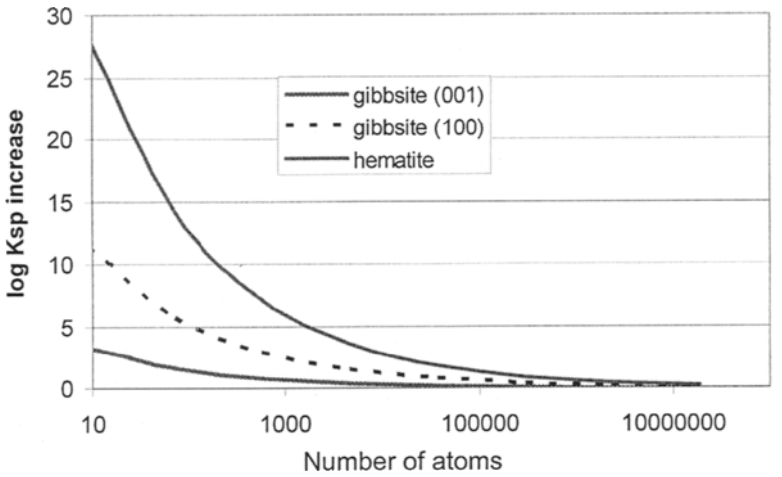


Fig. 3. Increased solubility of hematite and gibbsite as a function of number of Fe(III) or Al(III) atoms per particle due to interfacial tension with water. This results in the initial precipitation of amorphous Fe(III) and Al(III) (hydr)oxides that can persist over times significant relative to space times of the pertinent systems

High concentrations of counter-ions also affect solubility and colloidal characteristics of Al(III) and Fe(III) (hydr)oxides. Sulfate results in a broader pH range for formation of sweep floc from Al(III) salts [24]. Sulfate and other anions have a significant impact on pre-neutralized coagulants that contain Fe(III). Fukushi and Sato [59] reported that the solubility of nano-particles of hydrous ferric oxide (HFO) in sea water [60] was a function of $\{Fe^{+3}\}\{H^{+}\}^{-2.86}$ rather than $\{Fe^{+3}\}\{H^{+}\}^{-3.00}$ due to the effects of counter ions on surface charge and the incorporation of counter-ions in the bulk structure of the colloids. This results in greater tendency to form HFO nano-particles, decreased thermodynamic driving force for re-crystallization to hematite or other more stable phases, and could be significant in determining speciation in pre-neutralized Fe(III) coagulants.

4. RATES OF HYDROLYSIS AND A QUALITATIVE MODEL FOR COAGULATION REACTIONS

The speciation of Al(III) and Fe(III) in coagulants and in treated waters may depend on the rates of competing reactions and on mixing processes as well as on thermodynamic constants. Rates of reactions are discussed in this section. Some reported or calculated $t_{½}$ values are shown in Table 4.

Conversion from one Al(III) or Fe(III) monomer to another (e.g., Al^{+3} to $AlOH^{+2}$) is very fast [53]; reaction rates are diffusion-limited for some reactions. Therefore when a coagulant is added to a water that contains some alkalinity, the pH of the Al(III) or Fe(III) environment rises rapidly and monomers with n=2 or 3 (e.g., $Al(OH)_2^{+}$ or $Al(OH)_3^{0}$) become the major monomeric species in less than 10^{-3} s.

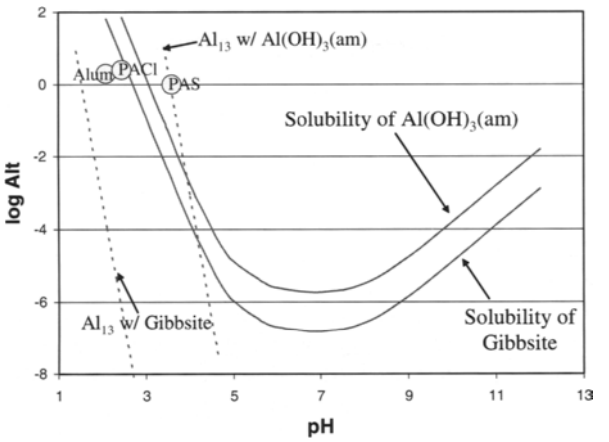


Fig. 4. Solubility of $Al(OH)_3$(am) and gibbsite considering only monomeric species [12]. The dashed lines represent additional solubility due to Al_{13} [12]. Al_{13} is a dominant soluble species in equilibrium with $Al(OH)_3$(am) but is insignificant in equilibrium with gibbsite

Table 4
Rates of formation and dissociation of Al(III) and Fe(III) hydroxyl species

Reaction	Expected $t_{½}$	Ref.
Formation of Al(III) monomeric species	$<10^{-3}$ s	61
Water exchange rate on Al^{+3}	<1 s	37
Water exchange rate on $Al(OH)_2^+$	$<10^{-4}$ s	est.
Al(III)-NOM complexation	<1 s	62
Al_{13} formation	minutes	12,63
Proton-catalyzed dissociation of Al_{13} at pH 5	hundreds of hrs	64
Phthalate dissociation of Al_{13}, equimolar, pH 5	10 hr	65,66
$Al(OH)_3$(am) precipitation upon dilution of a coagulant	<1 s	24
$Al(OH)_3$(s) crystallization in presence of NOM	Months to years	-
$Al(OH)_3$(s) dissolution, pH 5, 10 mM citrate, surface area	5 days	67
Water exchange rate on Fe^{3+}	0.003 s	37
Water exchange rate on $Fe(OH)_2^+$	$<10^{-7}$ s	37
Fe(III) dimerization at 10^{-4} M $Fe(OH)_3^0$	$<10^{-3}$ s	68
$Fe(OH)_3$(am) precipitation upon dilution of a coagulant	$<10^{-3}$ s	-
Fe_2O_3(s) crystallization in presence of NOM	Months to years	-

Dimerization is the first step leading to homogeneous precipitation of a solid phase and occurs by the following steps: (1) formation of an outer-sphere complex with formation constant K_0; (2) release of a water molecule with the rate constant k_{exch}, usually the rate determining step; and (3) rapid formation of an inner-sphere bond [69, 70]. Reported values for K_0 and k_{exch} for formation and then release of a water from the outer-sphere complex $Fe(H_2O)_5OH^{+2}\cdots Fe(H_2O)_5OH^{+2}$ are 1.2×10^{-3} M^{-1} [68] and 1×10^{5} s^{-1}[70]. K_0 and k_{exch} increase exponentially with a decrease in the charge on the combining species, e.g., K_0 and k_{exch} are approximately 2×10^{-2} M^{-1} and 10^{7} s^{-1} for $Fe(H_2O)_4(OH)_2^{+}\cdots Fe(H_2O)_4(OH)_2^{+}$. The rate of dimerization will increase exponentially as the pH increases.

The rate of dimerization increases with concentration of the monomeric Fe(III) species, as well as with pH. During rapid mix, the interface between droplets of coagulant and the bulk water phase provides a good environment for dimerization and subsequent precipitation. Al(III) or Fe(III) concentration is high (diffusion from the coagulant phase) and alkalinity diffuses from the raw water phase [71]. The preferential precipitation of the more soluble $Al(OH)_3$(am) or $Fe(OH)_3$(am) in lieu of bulk phases such as gibbsite and hematite was described earlier in this chapter; the high degree of supersaturation in the interface also provides a good location for precipitation of amorphous phases.

Formation of the Al_{13} polymer occurs within minutes [12, 43, 63], however dissolution of the Al_{13} polymer is slow [64]. The slow dissolution of Al_{13} might explain the effectiveness of PACl, PAS, and ACH outside the log Al_t and pH zones in which Al_{13} and $Al(OH)_3$(am) are thermodynamically stable, as shown in Fig. 3. Addition of more than equimolar concentrations of bidentate chelating ligands such as phthalate or oxalate accelerates the dissolution of Al_{13} but $t_{½}$ is still long compared to hydraulic residence times for coagulation and separation processes [65].

The rate data that are summarized in Table 4 can be used to construct a qualitative model for the coagulation process. When Al(III) salts are added to water, the conversion of Al^{3+} to $Al(OH)_2^{+}$ and other monomeric Al species is very fast with $t_{½}$ $<10^{-3}$ s [61]. Similar rates are expected for Fe(III) salts. The $t_{½}$ for formation of dimer and trimers is also expected to be $<10^{-3}$ s for Al(III) when the pH is high enough so that zero valent and monovalent monomers dominate and water exchange rates are fast. These rates are faster for Fe(III) due to higher water exchange rates (k_{exch}). As discussed above, precipitation may occur at rates similar to dimerization. Some investigators have used the term polymeric species [72] for these small particles but the experimental methods and physical descriptions make it clear that the "polymers" contained up to hundreds of Al(III) atoms. If formation of the primary particles is very rapid, then the thermodynamic driving force for further precipitation might be immediately reduced. Then primary particles can grow to larger flocs (fractal aggregates) by coagulation processes. If conditions result in the formation of highly charged particles (either positive or negative) then coagulation will not occur; this is analogous to the coagulation conditions of under-dosing (negative particles) or charge restabilization (positively-charged particles).

This qualitative model is in agreement with observations that Al and Fe hydroxides that are formed during water treatment are fractal aggregates [66, 73]. The primary particles within flocs of $Al(OH)_3$(am) that are formed during coagulation processes are between 3 and 10 nm in diameter [70, 74]. The amorphous fractal precipitates are very stable in the presence of NOM, sometimes persisting over decades of observation [66].

5. REMOVAL OF PARTICLES AND NATURAL ORGANIC MATERIALS

The primary objective of water treatment is production and distribution of water that is free from disease-causing chemicals and microorganisms. The secondary objective is to produce and distribute water that is appealing to all of our senses. Bacteria, viruses, asbestos fibers, *Giardia*, and other turbidity-causing particles are health concerns as well as aesthetic problems due to light-scattering. Arsenic, trace metals, and NOM are dissolved contaminants that directly or indirectly are health concerns but can also cause problems with color or taste in the finished water.

Removal of particulate or dissolved contaminants in conventional treatment requires the use of coagulants. Most turbidity-causing particles in raw waters are negatively charged due to inherent charge on the particles (e.g., isomorphic substitution in clays) or due to the sorption of potential-determining-ions at the particle surface, e.g., acidic functional groups at the surface of microorganisms or hydrous metal oxides. Sorption of NOM results in net negative charge on nearly all particles in raw waters [75]. The resulting electrostatic repulsion between particles inhibits aggregation and accounts for the persistence of small particles in natural waters. In the absence of successful coagulation, these particles will pass through clarifiers and rapid filters. The U.S. surface water filtration rules addressed this problem, i.e. some protected waters that had very low turbidity still contained sufficient infectious organisms to result in diffuse or wide-spread incidents of disease, and now must be treated with coagulation and defined particle removal operations.

NOM, and especially humic materials, must be removed from many waters due to reactions with disinfectants to produce disinfection by-products. *Enhanced coagulation* involves modifications in the coagulation procedure to enhance removal of NOM. This usually involves an increase in coagulant dose and sometimes a decrease in pH. NOM consists of a variety of organic materials. The focus for enhanced coagulation is on the dissolved organic carbon (DOC) component of NOM, since particulate NOM can be removed by conventional coagulation. Sorption models for the removal of NOM usually divide DOC into nonsorbable and one or more sorbable components [e.g., 76, 77]. The nonsorbable fraction cannot be removed during coagulation [78] and therefore the cumulative sorption isotherms intersect the X-axis at a non-zero value of $DOC_{remaining}$ (also called C_e) as shown in Figure 5. Sorption isotherms for each of the sorbable fractions usually take the form:

$$DOC_{removed}/D = a \times b \times DOC_{remaining}/(1+b \times DOC_{remaining}) \quad (1)$$

where DOC is in mg L^{-1}, D is the coagulant dose expressed as the metal, a is the maximum sorption capacity for the DOC onto the metal hydroxides produced by the coagulant, and b is the equilibrium constant for sorption.

Models have either adjusted the maximum sorption capacity, a, to reflect change in pH [76] or taken the acidity of DOC into account and presumed that the ionized fraction could not adsorb [77], also resulting in decreased sorption with an increase in pH. All models require calibration for the specific water to be treated, and calibration values for a, b, fraction in each NOM category, and effect of pH can change seasonally for the same water supply. It has often been demonstrated that humic materials with high specific ultraviolet absorbance (SUVA = absorbance at 254 nm / total organic carbon, $m^{-1}mg^{-1}L$) and high apparent molecular weight are most effectively removed by coagulation processes [e.g., 79, 80]. The sorption models that have been cited are empirical due to the poorly defined and continuously distributed properties of humic materials and other components of NOM [81]. More

sophisticated steric and electrostatic models have also been developed to describe the sorption of humic materials onto hydrous metal oxides [82-85] but direct application to water treatment is perhaps unjustified at this time.

If NOM is sorbed to coagulated particles, then the procedures and considerations for subsequent removal are the same as for other particles, i.e. sweep floc or charge neutralization must be employed in order to destabilize particles prior to removal by clarification or rapid filtration. It has sometimes been reported that the optimal conditions for removal of DOC and turbidity do not overlap. This observation is usually due to different analytical procedures (turbidity is based on a settled sample and DOC is performed after filtration through a membrane filter) and sometimes due to slightly deteriorated removal of turbidity at the lower pH imposed for enhanced sorption of NOM.

Coagulant dose and pH conditions for destabilization of particles by CN and sweep-floc mechanisms were identified in Figs. 1 and 2. Jar tests are commonly used to determine the appropriate dosage of coagulant for sweep floc conditions. Jar tests are appropriate when the flocs are large enough to observe and settle and when a clarifier is used in front of a rapid filter. The use of conventional jar testing is questionable, however, if clarification is not an essential part of the treatment process. Jar tests can be modified to provide information that will be useful for design and operation of conventional processes.

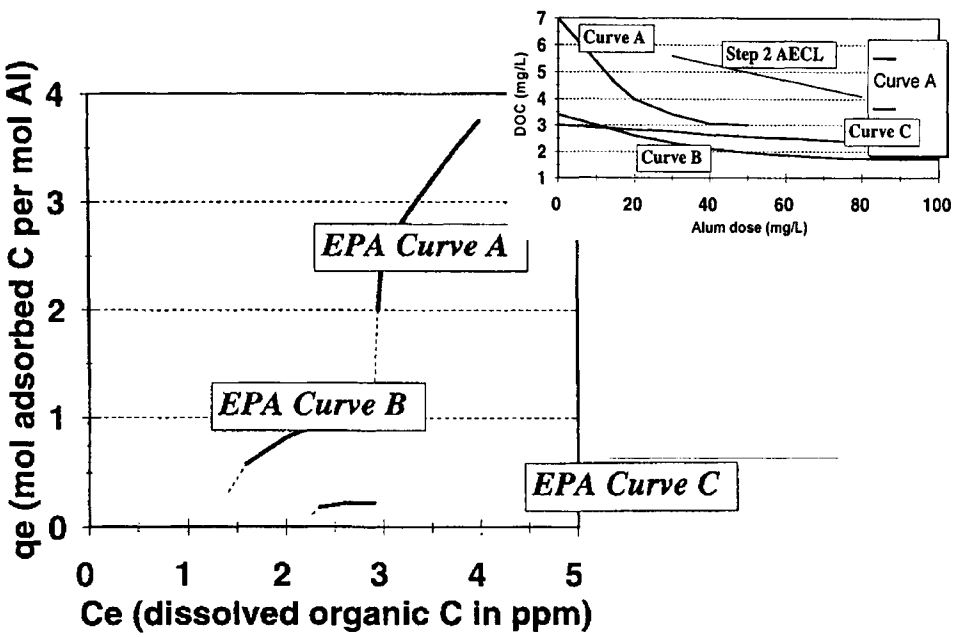


Fig. 5. Transformation of Fig. IX-1 of the U.S. EPA, Enhanced Coagulation and Enhanced Precipitative Softening Guidance Manual [31] for Step 2 enhanced coagulation (inset, from the initial rule notification) into adsorption isotherms. The rule requires addition of coagulant if the utility has not achieved compliance with the DBP rule and if at least 0.3 mg L^{-1} DOC can be removed for each additional 10 mg L^{-1} of alum (or equivalent dose of an alternative coagulant). Step 2 AECL in the inset shows the slope for 0.3 mg DOC per 10 mg alum

For example, jar tests can be used to estimate changes in settling velocity due to changes in coagulant, dose, pH, mixing, temperature, and other variables. Square jars or baffled beakers may be preferable to round beakers because more of the power from the paddle is converted to turbulence, and less power is lost to rotational velocity and a vortex. Some square jars are equipped with a sampling tap that is located 10 cm beneath the surface of the water. If samples are taken over time (e.g., at 0, 1, 2, 5, 10, and 20 min) then the removal of original turbidity can be plotted versus apparent settling velocity. Some utilities take advantage of the short residence time in a flash mixer and run combination full-scale and bench-scale tests. This is done by changing the full-scale dose for a few minutes (residence time in most rapid mix chambers is less than 30 sec) and taking a discharge sample that can be used for observation of the floc and for measurement of removals as a function of surface loading rate (settling velocity). Hudson [86] was a champion of using creative sampling and analysis techniques in conjunction with jar testing.

CN treatment often results in low coagulant doses so that settleable flocs are not formed. In that case diagnosis of appropriate coagulant dose and pH cannot be made using classic jar test experiments. Zeta meters and streaming current monitors are effective in identifying CN conditions when settleable flocs are not produced; these tools are more efficient than jar tests for identification of appropriate CN conditions even if flocs are formed. Prior to the availability of these electrokinetic techniques, it was often advised that CN treatment of relatively clean waters was difficult to achieve and that sweep floc was preferable for such waters. With modern diagnostic tools those conclusions are no longer valid; CN is usually the best way to treat relatively clean, low alkalinity waters.

The net negative charge on particulates in most raw waters, and the coagulant dose is typically determined by the concentration of NOM. For example, dissolved or sorbed humic materials at neutral pH have approximately 0.01 equivalent of negative charge per g of NOM. The most highly charged clays have a negative charge as high as 0.001 equivalents per g of clay. Thus, for equal concentrations of NOM and total suspended solids (TSS), the minimum coagulant demand for the NOM would be at least 10 times the coagulant demand for the TSS [25].

Plant operators should be aware of the coagulation mechanism at their facility and the diagnostic tools that can be used to improve treatment or to respond to changes in water quality. As an example, the city of Milwaukee WI changed from alum to PACl during the winter of '92-93. This resulted in a decrease in coagulant dose from ~30ppm alum to ~5ppm PACl. It is likely that the destabilization mechanism was changed from sweep floc to CN but jar tests (rather than electrokinetic tests) were retained for evaluation of appropriate coagulation conditions. Spring '93 brought unusual weather conditions, culminating with a significant freeze between two thaws. These conditions are often associated with rapidly deteriorating water quality due to run-off, sieving of contaminants behind a slush dam followed by release during a thaw, and the effects of cold water on the position of the $Al(OH)_3(am)$ solubility diagram. In addition, Milwaukee suffered some critical equipment and chemical failures. Hundreds of thousands were sickened by cryptosporidiosis [87]. It is nearly impossible to tell today whether the system failed in part because of overdosing or underdosing. It is possible the problems could have been correctly diagnosed and treated if zeta meters or streaming current monitors had been used.

A strategy for application of zeta meters for control of CN is provided in Fig. 6. Fig. 6 shows ZP versus pH for no coagulant addition (lowest line) and then for three increasing doses of alum. ZP is a representation of the electrical potential at the slipping plane, i.e. at a small distance from the water-particle interface [88]. ZP < -5 to -10 mV usually results in

underdosing and failure to remove contaminants during conventional treatment. CN produces aggregates that can be removed by conventional processes. Very positive ZP results in overdosing and restabilization. The charge and potential on the clay or oxide particles was negative for pH>1 in the absence of coagulant. The lowest coagulant dose resulted in CN from pH~5 to 5.7. Addition of higher coagulant doses produced two CN zones and an overdosing or restabilization zone. The second CN region, at higher pH, usually merges into the sweep floc zone.

It is relatively simple to develop Fig. 6 for a raw water. If alum is the coagulant, then adjust the pH to an acidic value less than required for precipitation of $Al(OH)_3(am)$, add the desired alum dose, measure ZP, and then incrementally add base to adjust pH and measure ZP with each pH adjustment. This procedure usually gives results similar to those obtained using separate beakers for each pH value; starting with a basic pH and incrementally adding acid does not give good results. As with any bench-scale test, the proof is in application at plant-scale. Since hydraulic retention time is very short in most rapid mix devices, an alternative to bench-scale testing is to momentarily change the plant-scale coagulant dosing and then take samples for pH and ZP from the rapid-mix tank or just downstream from an in-line mixer.

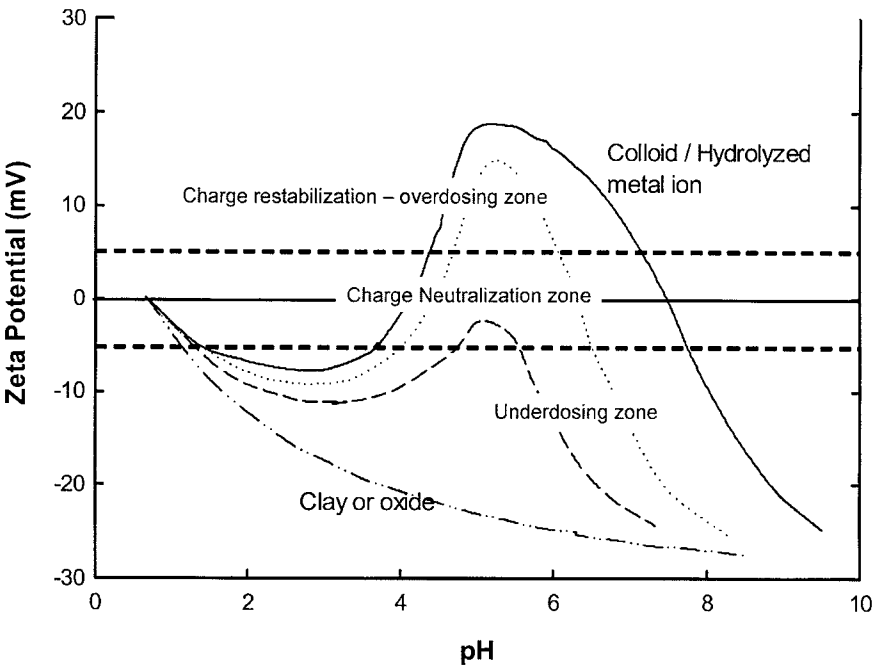


Fig. 6. Effects of pH and coagulant dose on destabilization of particles. The lowest line represents particles without addition of alum. The dashed, dotted, and top solid line represent increasing doses of alum

An application of ZP diagnostic procedures is shown in Fig. 7. Altoona PA receives acid mine drainage (AMD) as part of the water supply. They were using over 10ppm of alum. This caused excessive head loss in the rapid filters (direct filtration operation), excessive solids generation, and money spent on alum and caustic. The plant was equipped with a zeta meter and began using the strategy that is described by Fig. 6. NOM was low (TOC<2ppm) so coagulation at a higher pH was acceptable. The bench-scale results shown in Fig. 7 showed that coagulant dose could be reduced to 2ppm alum for pH~7.2. This coagulation strategy was successfully implemented at plant-scale. Altoona carried the CN logic a step further. Fe(II) in the raw water was oxidized with ozone prior to the rapid mix tank. The plant operators discovered they could operate the facility with no alum addition, as long as the pH was adjusted to achieve ZP~0 mV. The bench-scale tests should be repeated regularly and especially when there is a change in water quality or temperature, or a new batch of coagulant or other change in plant operation.

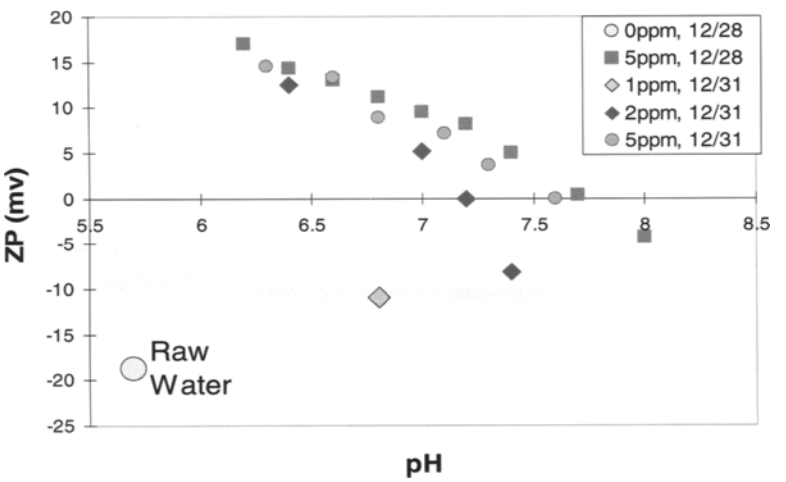


Fig. 7. ZP versus pH and alum coagulant dose for Altoona, PA. Bench-scale work was focused on the zone of decreasing ZP at pH above 6. Coagulant dose was successfully changed from ~10ppm alum to 2 ppm alum and pH 7.2

Streaming current (SC) is an electrokinetic measurement that can be used on-line at plant-scale to identify the appropriate coagulant dose for CN conditions. Coagulated water is passed through the instrument and coagulated particles are adsorbed onto the surfaces of a piston and cylinder. The clearance between piston and cylinder is small enough so that the reciprocating action of the piston transfers counter-ions. This results in an alternating current that is rectified to a SC value. SC is adaptable to on-line and plant-scale operation. SC values are not absolute and depend on pH and other factors. Therefore ZP and SC are complementary techniques; bench-scale ZP measurements can be used to calibrate the plant-scale SC values. Once calibrated to a set-point, SC and flow meters can be used to automatically control coagulant dosing.

A plot of SC versus coagulant dose is analogous to an acid-base titration curve; there is usually a rapid change in SC value with incremental coagulant dose at dose values that just result in good removal of contaminants, similar to the rapid change in pH at the end point of

an acid-base titration. This is illustrated in Fig. 8; these results were obtained by using a SC bench-scale, but similar data (or at least a portion of the curve) could be obtained by instantaneous changes in plant-scale coagulant dose as described above for ZP.

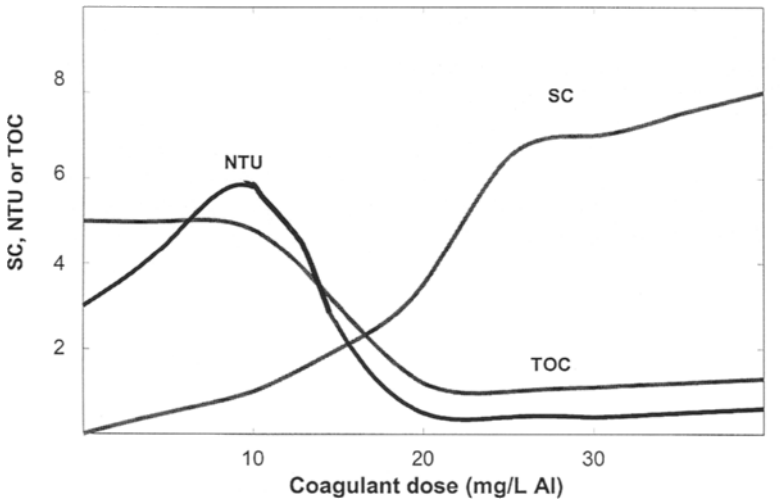


Fig. 8. Changes in streaming current value (SC), turbidity (NTU), and total organic carbon (TOC) with an increase in alum dose. SC often changes rapidly at the most appropriate dose for treatment

6. SUMMARY

Coagulation is an ancient technology [89]. It has been employed as a crucial part of modern water treatment for a century [90] yet there are still many unanswered questions about the nature of particles and NOM in raw water and the chemical and physical processes that occur when coagulants are added to raw water. These uncertainties are compounded by incomplete and sometimes contradictory knowledge about the hydrolysis of Al(III) and Fe(III), the nature of solid phases that are formed when salts of those metals are added to alkaline water, and the rates of parallel reactions that occur in complicated chemical systems. However, our understanding of coagulation processes continues to improve and we have access to an increasing variety of coagulants and of treatment process options. In this chapter an attempt was made to summarize some of the available information about the nature of coagulants and the reactions that occur when these coagulants are added to water that contains dissolved and particulate contaminants, and to provide qualitative and semi-quantitative models (usually empirical) that might assist in understanding and managing coagulation processes.

REFERENCES

[1] A. Amirtharajah and C.R. O'Melia, In Water Quality and Treatment (F.W. Pontius, ed.), McGraw-Hill, NY, 1990.
[2] F.J. Mangravite, J. - Am. Water Works Assoc., 67 (1975) 88.
[3] D.F. Lawler, C.R. O'Melia and J.E. Tobiason, In Particulates in Water: Advances in

Chemistry Series 189 (M.C. Kavanaugh & J.O. Leckie, eds.) Washington, D.C., 1980.
[4] K.Y.-J. Choi and B.A. Dempsey, Water Res., 38 (2004) 4271.
[5] V. Runkana, P. Somasundaan and P.C. Kapur, J. Colloid Interface Sci., 270 (2004) 347.
[6] K.K. Das and P. Somasundaran, Colloid Surf. A, 223 (2003) 17.
[7] K.K. Das and P. Somasundaran, J. Colloid Interface Sci., 271 (2004) 102.
[8] J. Roussy, M. Van Vooren, B.A. Dempsey and E. Guibal, Water Res., 39 (2005) 3247.
[9] N. Narkis and M. Rebhun, Water Sci. Technol., 36 (1997) 85.
[10] Kline & Co. Inc., Coagulants and Flocculants North America 1977, Little Falls, NY, 1997.
[11] A. Amirtharajah and K.M. Mills, J. - Am. Water Works Assoc., 74 (1982) 210.
[12] C.F. Baes Jr. and R.E. Mesmer, The Hydrolysis of Cations, Wiley-Interscience, NY, 1976.
[13] B.A. Dempsey, In Aquatic Humic Substances, Influence on Fate and Treatment of Pollutants (I. H. Suffet and P. MacCarthy, eds.), ACS Symp. Ser., 219, 1989.
[14] J.M. Dulko, US Patent No. 6 036 935 (2000).
[15] A. Masion, J.Y. Bottero, F. Thomas and D. Tchoubar, Langmuir, 10 (1994) 4349.
[16] Y.-H. Shen and B.A. Dempsey, Environ. Int., 24 (1998) 899.
[17] K.N. Exall and G.W. vanLoon, Water Res., (2003) 3341.
[18] J.-P. Boisvert and C. Jolicoeur, Colloid Surf. A, 5 (1999) 161.
[19] J.D. Murphy, R. Schuffenbecker and H. Suty, US Patent No. 5 348 721 (1994).
[20] M. Kvant and K. Stendahl, US Patent No. 5 182 094 (1993).
[21] D. Haase, N. Spiratos and C. Joliceour, US Patent No. 5 149 400 (1992).
[22] Y. Aiba, T. Furumori, S. Shinpo and K. Funabiki, US Patent No. 3 929 666 (1975).
[23] H. de Hek, R.J. Stol and P.L. de Bruyn, J. Colloid Interface Sci., 64 (1978) 72.
[24] B.A. Dempsey, R.M. Ganho and C.R. O'Melia, J. - Am. Water Works Assoc., 76 (1984) 141.
[25] B.A. Dempsey, CRC Crit.Rev. Environ. Control, 14 (1984) 311.
[26] B.A. Dempsey, H. Sheu, T.M.T. Ahmed and J. Mentink, J. - Am. Water Works Assoc., 77 (1985)73.
[27] J. DeWolfe, B.A. Dempsey, M. Taylor and J.W. Potter, Guidance Manual for Coagulant Changeover, AwwaRF, 2003.
[28] B.A. Dempsey, Production and Utilization of Polyaluminum Sulfate, AWWARF, 1994.
[29] H. Elliott and B.A. Dempsey, Land Application of Water Treatment Sludges, AwwaRF, Denver,CO, 1990.
[30] J.K. Edzwald, W.C. Becke and S.J. Tambini, J. Environ. Eng., 113 (1987), 167.
[31] U.S. EPA, Enhanced Coagulation and Enhanced Precipitative Softening Guidance Manual, EPA 815-R-99-012, 1999.
[32] G.H. Nancollas and M.B. Tomson, Guidelines for the Determination of Stability Constants, IUPAC Pergamon Press, 1980.
[33] L.-O. Öhman and S. Sjöberg, Coord. Chem. Rev., 149 (1996) 33.
[34] R.M. Smith and A.E. Martell, Critical Stability Constants, Plenum Press, NY, 1974.
[35] P.M. Bertsch, In: The Environmental Geochemistry of Aluminum (G. Sposito, ed.), pp. 88-115. CRC Press, Boca Raton, FL, 1989.
[36] C. Brosset, Acta Chem. Scand., 8 (1954) 299.
[37] W. Stumm and J.J. Morgan, Aquatic Chemistry, Wiley-Interscience, NY, 1996.
[38] D.K. Nordstrom and H.M. May, In: The Environmental Geochemistry of Aluminum (G. Sposito, ed.), pp. 29-53. CRC Press, Boca Raton, FL, 1989.
[39] R.M.Smith, A.E. Martell, and R.J. Motekaitis, NIST Standard Reference Database 46, version 7 NIST, Gaithersburg, MD, 2003.
[40] S. Bi, C. Wang Q. Cao and C. Zhang, Coord. Chem. Rev., 248 (2004) 441.
[41] G. Johansson, Acta Chem. Scand., 14 (1960) 771.
[42] D.L. Teagarden, J. Pharm. Sci., 70 (1981) 758.
[43] J.Y. Bottero, D. Tchoubar, J.M. Cases and F. Fiessinger, J. Phys. Chem., 86 (1982) 3667.
[44] P.M. Bertsch, Soil Sci. Soc. Am. J., 51 (1987) 825.
[45] H.-J. Liu, J.-H. Qu, C.-Z. Hu and S.-J. Zhang, Colloid Surf. A, 216 (2003) 139.
[46] L. Allouche and F. Taulelle, Inorg. Chem. Comm., 6 (2003) 1167.

[47] X. Liu and F.J. Millero, Geochim. Cosmochim. Acta, 63 (1999) 3487.
[48] D. Hunter and D.S. Ross, Science, 251 (1991) 1056.
[49] J.Y. Bottero, M. Axelos, D. Tchoubar, J.M. Cases, J.J. Fripiat and F. Fiessinger, J. Colloid Interface Sci., 117 (1987) 47.
[50] K.A. Gray, C. Yao and C.R. O'Melia, J. - Am. Water Works Assoc., 87 (1995) 136.
[51] T.-K. Liu and E.S.K. Chian, J. Colloid Interface Sci., 284 (2005) 542.
[52] U. Schwertmann, J. Friedl and H. Stanjek, J. Colloid Interface Sci., 209 (1999) 215.
[53] W.P. Cheng and F.H. Chi, Water Res., 36 (2002) 4583.
[54] D. Wang, Z. Luan and H. Tang, J. - Am. Water Works Assoc., 95 (2003) 79.
[55] N.J.D. Graham and J. Jiang, US Patent No. 5 785 862 (1998).
[56] B. Lo and T.D. Waite, J. Colloid Interface Sci., 222 (2000) 83.
[57] J.W. Morse and W.H. Casey, Am. J. Sci. 288 (1988) 537.
[58] D.J. Pernitsky and J.K. Edzwald, J. Water Supply – Aqua, 52 (2003) 395.
[59] K. Fukushi and T. Sato, Environ. Sci. Technol., 39 (2005) 1250.
[60] H.B. Byrne and Y.-R. Luo, Geochim. Cosmochim. Acta, 64 (2000) 1873.
[61] L.P. Holmes, D.L. Cole, and E.M. Eyring, J. Phys. Chem., 72 (1968) 301.
[62] C.H. Langford and T.R. Khan, Can. J. Chem., 53 (1975) 2979.
[63] J.W. Akitt, J. Chem. Soc., Dalton Trans., (1972) 604.
[64] G. Furrer, M. Gfeller and B. Wehrli, Geochim. Cosmochim. Acta, 63 (1999) 3069.
[65] A. Amirbahman, M. Gfeller and G. Furrer, Geochim. Cosmochim. Acta, 64 (2000) 911.
[66] E. Molis, F. Thomas, J.Y. Bottero, O. Barrès and A. Masion, Langmuir, 12 (1996) 3195.
[67] M. Dietzel and G. Böhme, Geochim. Cosmochim. Acta, 69 (2005) 1199.
[68] H. Wendt, Inorg. Chem., 8 (1969) 1527.
[69] R.G. Wilkins, The Study of Kinetics and Mechanism of Reactions of Transition Metal Complexes, Allyn and Bacon, Inc., NY, 1974.
[70] F.M.M. Morel and J.G. Hering, Principles and Applications of Aquatic Chemistry, Wiley Interscience, NY, 1993.
[71] M.M. Clark, R.M. Srivastava and R. David, Environ. Sci. Technol., 27 (1993) 2181.
[72] J.D. Hem, Trace Inorganics in Water (R.F. Gould, ed.), Adv. Chem. Ser., 73, 1968.
[73] D. Li and J. Ganczarczyk, Environ. Sci. Technol., 223 (1989) 1385.
[74] D.W. Schaefer, R.F. Shelleman, K.D. Keefer and J.E. Martin, Physica, 140A (1986) 105.
[75] K.A. Hunter and P.S. Liss, Nature, 282 (1979) 823.
[76] T. Tseng and M. Edwards, J. - Am. Water Works Assoc., 91 (1999) 159.
[77] G. Kastl, A. Sathasivan, I. Fisher and J. Van Leeuwen, J. - Am. Water Works Assoc., 96 (2004) 79.
[78] J.A. Davis and R. Gloor, Environ. Sci. Technol., 15 (1981) 1223.
[79] P. Bose and D.A. Reckhow, J. Environ. Eng. – ASCE, 124 (1998) 803.
[80] P.C. Chiang, E.E. Chang and C.H. Liang, Chemosphere, 46 (2002) 929.
[81] E.M. Thurman, Organic Geochemistry of Natural Waters, Kluwer Acad., Boston, 1985.
[82] T. Saito, L.K. Koopal, W.H. van Riemsdijk, S. Nagasaki and S. Tanaka, Langmuir, 20 (2004) 689.
[83] K. Kaiser, Org. Geochem., 34 (2003) 1569.
[84] G.U. Balcke, N.A. Kulikova, S. Hesse, F.D. Kopinke, I.V. Perminova and F.H. Frimmel, Soil Sci. Soc. Am. J., 66 (2002) 1805.
[85] K.K. Au, A.C. Penisson, S.L. Yang and C.R. O'Melia, Geochim. Cosmoschim. Acta, 63 (1999) 2903.
[86] H.E. Hudson, Jr., Water Clarification Processes: Practical Design and Evaluation, Van Nostrand Rheinhold Co., NY, 1981.
[87] K.R. Fox and D.A. Lytle, J. - Am. Water Works Assoc., 88 (1996) 8794.
[88] R.J. Hunter, Zeta Potential in Colloid Science, Academic Press, NY, 1981.
[89] American Water Works Association, The Quest for Pure Water, AWWA, 1981.
[90] A.P. Folwell, Water-Supply Engineering: The Designing and Constructing of Water Supply Systems, 3rd ed., John Wiley & Sons, NY, 1917.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 3: Floc formation and floc structure

John Gregory

Department of Civil and Environmental Engineering, University College London, Gower Street, London WC1E 6BT, UK

1. PARTICLE COLLISIONS AND AGGREGATION

1.1. Basics

Aggregation (coagulation or flocculation) of particles involves two essential steps:

- collisions between particles
- attachment of colliding particles

The first of these depends on particle size, concentration and transport mechanisms. Whether or not attachment occurs on collision depends crucially on interactions between particles, or their *colloid stability*, as discussed in the reference chapter on particle stability by Charles O'Melia These interactions are usually of very short range and have very little influence on the transport of particles. For this reason it is possible to treat transport and attachment separately, so that we can think of aggregation rate in terms of a *collision frequency* and a *collision efficiency factor*, α. The latter is the fraction of collisions that result in attachment. When there is no repulsion between particles every collision should be effective and $\alpha = 1$. However when particles have a high degree of colloid stability, then the collision efficiency is very low. The collision efficiency can be increased by dosing suitable additives (coagulants and flocculants) and this is often called *destabilisation*. Thus, the transition from stable particles to aggregates or flocs involves two steps: destabilisation and collisions. These are illustrated schematically in Fig. 1.

In this Chapter, we shall be mainly concerned with the collision process; in particular with the collision frequency and the nature of the aggregates formed.

1.2. Collision frequency and aggregation rate

A suspension may contain many types of particles at different concentrations. If we consider just two types of particle i and j, with number concentrations N_i and N_j, then the frequency of i-j collisions is governed by a second-order rate equation:

$$J_{ij} = k_{ij} N_i N_j \tag{1}$$

In this expression, J_{ij} is the number of collisions occurring between i and j particles in unit time and unit volume of suspension and k_{ij} is a second-order rate coefficient, which depends on a number of factors (see below).

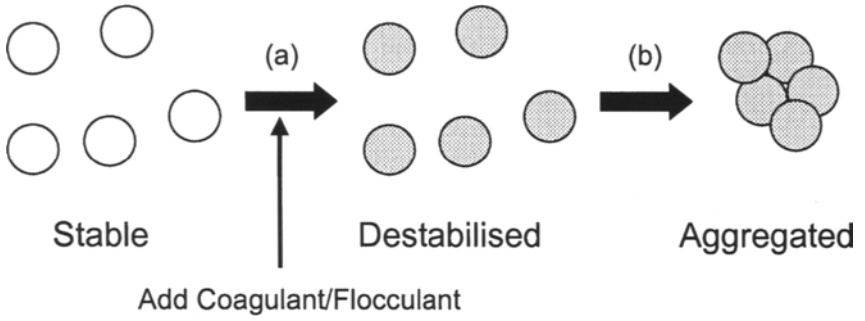


Fig. 1. Showing (a) destabilisation and (b) collisions of particles to give aggregation

In order to relate this to aggregation rate, it is convenient to consider the case where all the particles are identical and to restrict attention to the very early stages of the process, where collisions are nearly all between single (*primary*) particles. Each collision then leads to the loss of two primary particles and the formation of one doublet, or a net loss of one particle. The rate of decrease of the *total* particle concentration, N_T (initially equal to the concentration of primary particles, N_1) can then be shown to be given approximately by:

$$\frac{dN_T}{dt} = -\alpha \frac{k_{11}}{2} N_T^2 = -\alpha k_a N_T^2 \tag{2}$$

where k_{11} is the rate coefficient for the collision of primary particles and k_a is the aggregation rate coefficient (note that this is just half of k_{11}).

Since the aggregation rate depends on the square of the particle concentration and this decreases according to Eq. (2), it follows that the rate will decline as aggregation proceeds.

In order to progress further, we need to consider the collision rate coefficient *kij*. This depends greatly on the collision mechanism. There are three important ways in which particles may be brought into contact:

- Brownian diffusion (leading to perikinetic aggregation)
- Fluid motion (leading to orthokinetic aggregation
- Differential sedimentation

These are shown schematically in Fig. 2 and will be discussed in the following sections.

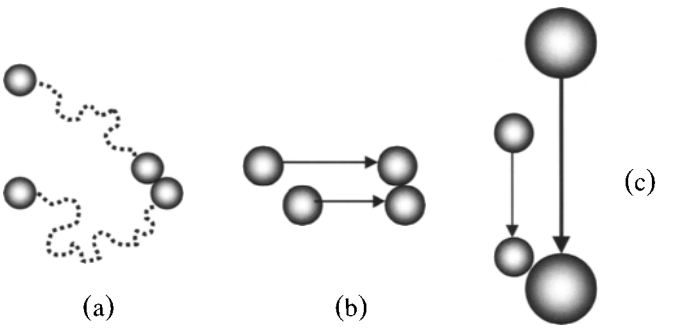


Fig. 2. Collisions of particles by (a) brownian diffusion, (b) fluid motion and (c) differential sedimentation

1.3. Brownian diffusion - perikinetic aggregation

All particles in water undergo random movement as a result of their thermal energy. This is known as *brownian motion* and it can be an important means of bringing about particle collisions. Smoluchowski [1] calculated the collision rate between spherical particles and the result can be expressed as a collision rate coefficient, as defined by Eq. (1):

$$k_{ij} = \frac{2k_B T}{3\mu} \frac{(d_i + d_j)^2}{d_i d_j} \tag{3}$$

where k_B is Boltzmann's constant, T is the absolute temperature, μ is the fluid viscosity and d_i and d_j are the diameters of the particles.

For particles that are not too different in size, the size term in Eq. (1) is approximately constant, with a value of 4. With this assumption, the rate coefficient becomes:

$$k_{ij} = \frac{8k_B T}{3\mu} \tag{4}$$

It is noteworthy that this form does not involve particle size and depends only on physical constants. For aqueous dispersions at 25^0C, the value of k_{ij} is 1.23 x 10^{17} m^3 s^{-1}. Strictly, this value only applies to particles of equal size, but is acceptable when the particle diameters do not differ by more than a factor of about 2. For particles of different size, the rate coefficient is always greater than that given by Eq. (4).

It turns out that perikinetic aggregation is only important in practice for particles that are quite small (typically less than 1 μm). This is because the reduction in particle number as aggregation proceeds gives a large decrease in collision rate, which is not compensated by an increased collision rate coefficient (since this does not depend greatly on particle size). In water treatment practice perikinetic aggregation is never sufficient to give the large aggregates (flocs) that are needed. The rate of floc formation can be greatly increased by some form of fluid motion.

1.4. Fluid motion - orthokinetic aggregation

The first treatment of orthokinetic collision rate was also by Smoluchowski [1], for spherical particles in a fluid undergoing *uniform, laminar shear*, as in Fig. 2. The fluid velocity varies linearly with distance in only one direction, *z*, and the gradient *du/dz* is the *shear rate*, *G*. Such conditions are never encountered in practice, but this simple model provides a convenient starting point for a discussion of orthokinetic aggregation. By assuming that particles move along rectilinear paths until collision, it is possible to show that the orthokinetic collision rate coefficient for i and j particles is:

$$k_{ij} = \frac{G}{6}(d_i + d_j)^3 \tag{5}$$

There is a striking difference between this expression and that for the perikinetic case, Eq. (3) - the very important effect of particle size. As we saw, the perikinetic collision rate is only slightly dependent on particle size, whereas the orthokinetic rate depends on the *cube* of the particle size. This is the reason why orthokinetic aggregation is very much more significant for larger particles and of great practical importance. Orthokinetic aggregation will be discussed further in Section 2.

1.5. Differential sedimentation

When particles differ in size and/or density, they will settle at different rates. Larger and denser particles will settle faster and can collide with more slowly settling particles. This can give a significant collision rate, especially for larger particles or aggregates. The appropriate rate coefficient can easily be calculated, assuming spherical particles and using Stokes law for their sedimentation rate. For particles of equal density, the collision rate coefficient is:

$$k_{ij} = \left(\frac{\pi g}{72 \mu}\right)(\rho_s - \rho_L)\,(d_i + d_j)^3 (d_i - d_j) \tag{6}$$

where g is the gravitational acceleration, ρ_S is the density of the particles and ρ_L the density of the suspending fluid.

It is clear from Eq. (6) that the collision rate depends greatly on the size, the difference in size between the colliding particles and the density difference between particles and fluid.

1.6. Comparison of rates

If we consider first only equal spherical particles, diameter d, the differential settling rate is zero and the ratio of the orthokinetic and perikinetic rate coefficients can be calculated from Eqs. (4) and (5):

$$\frac{k_{ortho}}{k_{peri}} = \frac{G\mu d^3}{2kT} \tag{7}$$

For a shear rate of 10 s^{-1} (corresponding to quite mild agitation), the collision rate coefficients are equal when the particle diameter is about 1 μm. For higher shear rates and, especially, for larger particles, the orthokinetic rate becomes very much larger.

The rate coefficients for the three collision mechanisms described above are compared in Fig. 3. The rate coefficients have been calculated for collisions between a particle of diameter of 2 μm and a second particle with diameter varying from 0.01 μm to 30 μm. The shear rate, G, is assumed to be 50 s^{-1} and the particle density 2 g cm^{-3}. The fluid is assumed to be water at 25^{0}C. There are several important features of these results:

The perikinetic rate passes through a minimum when the particle diameters are equal. Around the minimum, the rate coefficient is approximately independent of particle size, as in Eq. (3). However, when the sizes differ significantly the rate coefficient can be greater than the 'constant' value by an order of magnitude or more.

The differential settling rate is zero for equal particles, because they both settle at the same speed and do not collide. However, when the second particle is larger than about 2 μm this mechanism becomes very important.

The orthokinetic rate becomes greater than the perikinetic when the second particle diameter exceeds about 0.1 μm and becomes overwhelmingly larger when d is greater than about 2 μm.

Of course, these conclusions would be modified if different values for the fixed particle diameter, density and shear rate were chosen. However, it can be reasonably assumed that, for particles larger than a few μm in agitated suspensions, the perikinetic collision rate will be negligibly small.

2. HYDRODYNAMIC EFFECTS

All of our previous discussion of aggregation rates has assumed that, for fully destabilised suspensions (α = 1), all collisions lead to permanent attachment of particles. In fact, in a viscous fluid such as water, the approach of particles can be significantly hindered as the gap between them narrows. At close approach it becomes increasingly difficult for water to be 'squeezed out' of the gap. In the case of brownian collisions, the effect is to reduce the effective diffusion coefficient of approaching particles and this can typically reduce the predicted collision rate by a factor of around two.

For orthokinetic collisions hydrodynamic effects can be much more significant. The assumption that particles follow rectilinear paths until the moment of collision no longer holds. Particles will actually tend to move *around* other particles (in 'curvilinear' paths), which considerably reduces the chance of attachment. For hard spheres, especially of unequal size, effective collision efficiencies can be very low and it has been argued (e.g. Lawler [2]) that the simple Smoluchowski result, Eq. (5), grossly overpredicts actual aggregation rates. However real aggregates are not hard spheres, but porous, fractal objects (see 3.1) and it is likely that hydrodynamic effects are much less significant than previously supposed.

Because of hydrodynamic effects, the quantitative comparison of collision rates in Fig. 3 may not be entirely reliable. However, in practice, the relative importance of the different collision mechanisms and the effect of particle size are probably still broadly as shown.

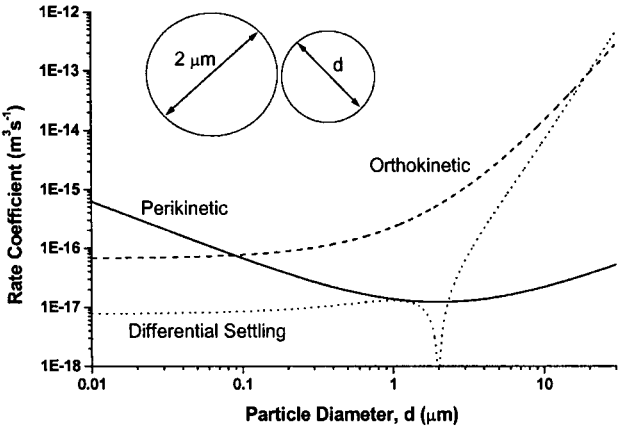


Fig. 3. Comparison of rate coefficients for three collision mechanisms (see text)

3. PRACTICAL ASPECTS OF ORTHOKINETICS FLOCCULATION

3.1. Role of solids concentration

From the orthokinetic rate coefficient in Eq. (5), it is possible to derive the flocculation rate for equal particles in terms of the rate of decrease of total particle number concentration, as in Eq. (2):

$$\frac{dN_T}{dt} = -\frac{2}{3}\alpha N_T^2 G d^3 = -\alpha k_a N_T^2 \tag{8}$$

This expression only applies in the very early stages of the flocculation process, but is still helpful for exploring practical consequences.

The presence of a d^3 term in Eq. (8) implies that the volume of particles plays a significant part. The volume fraction of particles in a suspension of equal spheres is simply the volume of each particle multiplied by the number per unit volume:

$$\phi = \frac{\pi d^3 N_T}{6} \tag{9}$$

Combining this with Eq. (8), and assuming that every collision is effective ($\alpha = 1$) gives:

$$\frac{dN_T}{dt} = -\frac{4G\phi N_T}{\pi} \tag{10}$$

If the volume fraction of particles remains constant during flocculation, which seems a reasonable assumption at first sight, then Eq. (10) implies a *first-order* rate process. While it is true that the volume of *primary* particles remains constant, this is not usually the case for aggregates (flocs), which have an effective volume greater than that of their constituent

particles (see 3.2). Nevertheless, it is worthwhile to consider the consequences of a 'pseudo' first order rate law.

Assuming that the shear rate G and volume fraction ϕ remain constant, then Eq. (10) can be integrated to give:

$$\frac{N_T}{N_0} = \exp\left(\frac{-4G\phi t}{\pi}\right) \tag{11}$$

where N_0 is the initial concentration of primary particles.

The exponential term in Eq. (11) contains the dimensionless group $G\phi t$, which plays a very important role in determining the extent of flocculation. In principle, for a given value of this number, the same degree of flocculation should occur, for any combination of the individual terms. For instance, doubling the shear rate and halving the time should have no effect on the overall process. In practice, high shear rate can have a detrimental effect, for instance as a result of floc breakage (see 4.1), so this conclusion may be misleading.

The dimensionless term Gt, sometimes called the *Camp number*, is of great importance in the design of practical flocculation units. However, the volume fraction ϕ is equally significant. In waters of fairly low turbidity the particle concentration may be too low to give acceptable rates of flocculation and there are well-established practical ways around this problem.

Hydrolysing coagulants are widely used in water treatment, often under conditions where hydroxide precipitation and 'sweep flocculation' occur (see previous chapter). The formation of a bulky, amorphous precipitate of metal hydroxide can give a substantial increase in the volume fraction of suspended particles and hence, according to Eq. (11), a significant enhancement of flocculation rate. This effect is clearly evident in the results shown in Fig. 4 [3]. These are from a modified jar test procedure in which kaolin clay suspensions were flocculated with aluminium sulphate ('alum') at two different concentrations and at pH 7. The process was continuously monitored by an optical technique, which gives a semi-empirical 'Flocculation Index' (FI). The FI value is closely related to floc size. At the lower alum concentration (10 μM Al), the kaolin particles are destabilised by adsorption of cationic hydrolysis products and collisions occur at a rate determined by the particle concentration and the effective shear rate. This gives a fairly low flocculation rate and the ultimate FI value is rather low. At the higher concentration (80 μM Al) hydroxide precipitation and 'sweep flocculation' occur. It is clear that this process is considerably more rapid than for the lower Al concentration, as evidenced by a steeper rise in the FI value. However, there is a noticeable lag time before significant flocculation begins. It is believed that this is related to the time needed for the precipitate particles to grow to an appreciable size. At still higher alum concentrations this lag time is reduced and may be eliminated altogether.

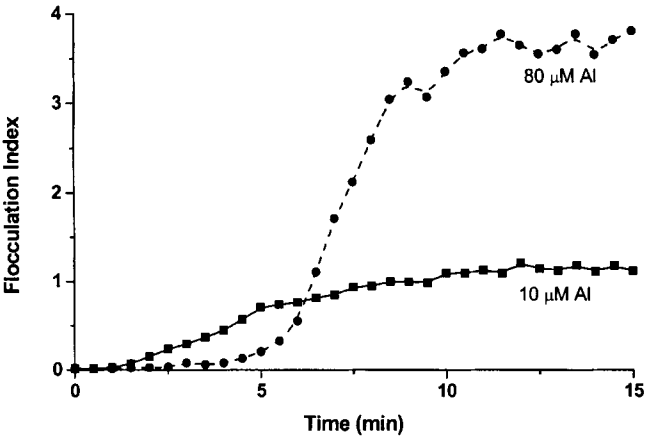


Fig. 4. Flocculation of kaolin suspensions at two dosages of alum

It is also evident from Fig. 4 that the FI value (and hence floc size) reaches a limiting value after a certain time. This is probably due to a dynamic balance between floc growth and breakage, giving a steady state floc size distribution. Floc strength and floc breakage will be discussed further in Section 4.

The role of solids concentration in determining flocculation rate can be exploited in practice by the use of certain types of flow-through units such as flocculator-clarifiers. These are operated in such a way that incoming water passes through a region where previously flocculated solids are held, for instance in a *floc blanket*. The solids content in the floc blanket is very much higher than in the raw water, giving a large increase in flocculation rate.

3.2. Non-uniform shear

No real flocculation processes occur under conditions of uniform laminar shear, as assumed in the Smoluchowski treatment, leading to Eq. (5). In fact, in nearly all cases, flocculation takes place under turbulent conditions. One way of dealing with this difficulty was proposed by Camp and Stein [4], on the basis of dimensional analysis. They calculated a mean or effective shear rate, $\overline{G}$, from the power input per unit mass of suspension, ε:

$$\overline{G} = \sqrt{\frac{\varepsilon}{\nu}} \tag{12}$$

where ν is the *kinematic viscosity* (= μ/ρ, where ρ is the density of the suspension).

This average value can then be inserted in place of G in Eq. (5), giving:

$$k_{ij} = \frac{1}{6}\left(\frac{\varepsilon}{\nu}\right)^{1/2}\left(d_i + d_j\right)^3 \tag{13}$$

Although this result is very similar to a more rigorous result for particle collisions in isotropic turbulence [5], the agreement is probably fortuitous, since it is known that the simple

averaging procedure of Camp and Stein is not applicable to the highly complex nature of turbulent flow [6].

Turbulent flow is characterised by eddies of various sizes. The largest of these are comparable in size to the vessel or impeller. The energy in these eddies cascades through eddies of decreasing size and eventually, below a certain length scale, the energy is dissipated as heat. The *Kolmogorov microscale* separates the inertial range, where energy transfer occurs with little dissipation, from the viscous subrange, where energy is dissipated as heat. The Kolmogorov microscale, l_K, depends on the kinematic viscosity of the fluid, ν, and the rate of energy dissipation per unit mass, ε:

$$l_K = \left(\frac{\nu^3}{\varepsilon}\right)^{1/4} \tag{14}$$

For typical values of average shear rate under practical flocculation conditions in water treatment ($\overline{G} \approx$ 50-100s^{-1}), the Kolmogorov microscale is of the order of 100-150 μm. For particles smaller than l_K, the collision rate coefficient should be reasonably well approximated by an expression of the form of Eq. (13), but for larger particles, transport by eddies in the inertial range is important. In this case, the collision rate shows a different dependence on power input (proportional to $\varepsilon^{2/3}$, rather than $\varepsilon^{1/2}$) [7].

3.3. Mixing conditions

A flocculation process involves dosing a coagulant and then allowing collisions to occur (see Fig. 1). In a typical water treatment jar test procedure, coagulant dosing is accompanied by a fairly brief period of 'rapid mixing' in order to distribute the additive uniformly. There then follows a much longer period of 'slow mixing' in order to promote particle collisions (orthokinetic aggregation). At the plant scale coagulant is usually added with some form of 'flash mixing', followed by flocculation in a stirred tank or under conditions where hydraulic effects provide the required velocity gradients.

Both the rapid and slow mixing stages can have important influences on the effectiveness of the flocculation process. Rapid mixing should not be too intense or prolonged, since the high shear rate involved can result in poorer floc growth in the subsequent slow mixing stage [8]. The results in Fig. 5 are for kaolin suspensions (50 mg L^{-1}) flocculated with commercial polyaluminium chloride (PACl) at a dosage (3.4 mg L^{-1} as Al) where sweep flocculation is the dominant mechanism. The flocculation was monitored in the same way as for the results in Fig. 4 and the Flocculation Index (FI) is plotted against time. The PACl was added after 60 s and rapid mixing (400 rpm) was applied for between 5 and 60 s. There then followed a period of slow stirring (50 rpm) during which FI values reached plateau values. The largest increase in FI occurs for 10 s of rapid mix and longer times give significantly smaller flocs. It is noteworthy that in the 60 s rapid mix case a plateau value of FI is apparent before the slow stirring begins. This shows that small flocs are formed in this stage, but floc growth is severely limited by the high shear. During the slow stirring phase further floc growth is possible, but to smaller limiting sizes than for shorter rapid mix times.

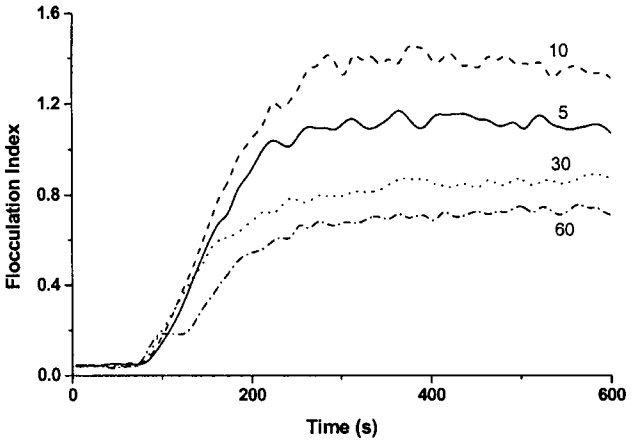


Fig. 5. Flocculation of kaolin suspensions with PACl for different rapid mix times. (Figures on curves show rapid mix times in seconds)

The effective shear rate in the slow mixing stage also needs to be carefully controlled. If the shear rate is high, the flocculation rate can be increased, according to Eq. (10). However, floc growth is restricted by the applied shear, giving a limiting floc size (see Figs. 4 and 5) and, for large flocs, quite low shear rates must be applied. The question of floc strength and limiting floc size will be considered further in 4.1.

As a compromise, a process known as 'taper flocculation' may be used, where the flocculating suspension passes through different zones with decreasing shear rates [9]. In this way, fairly rapid floc growth can occur in the early stages and large flocs can be formed in the low-shear regions.

4. FLOC STRUCTURE – FRACTAL AGGREGATES

4.1. Fractals

It is known that, in most cases, particle aggregates have a self-similar, *fractal* structure. The concept of fractals was brought into widespread use by the work of Mandelbrot [10] and has found very broad application. Typical fractal objects have a branching structure, with a similar pattern at all levels of observation (hence the term 'self-similar').

For aggregates, it is most useful to think in terms of *mass fractals*, for which the mass of an object, M, scales as its length, L, raised to some power, d_F, the fractal dimension. If log M is plotted against log L, then a straight line should be obtained, with a slope d_F. In principle, the fractal dimension for an object can vary from 1 (indicating a linear form) to 3 (for a solid, non-fractal object). A value of $d_F = 2$, indicates that a projection of the aggregate onto a plane should appear as a completely-filled shape. The higher the fractal dimension, the more compact is the aggregate. This has very important implications for floc density (see 3.3). Real aggregates can have fractal dimensions from around 1.7 up to about 2.6 or higher.

There are essentially two ways in which aggregates can form:

- by addition of single particles to growing aggregates (particle-cluster aggregation)
- by collisions of pre-formed aggregates (cluster-cluster aggregation).

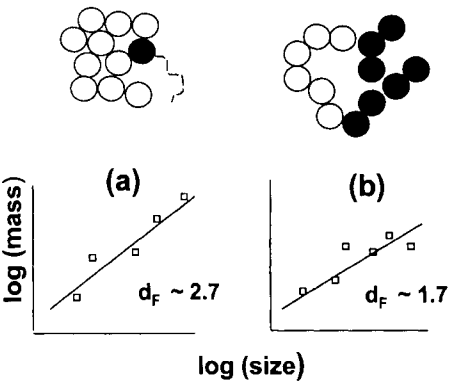


Fig. 6. Formation of fractal aggregates by (a) particle-cluster and (b) cluster-cluster collisions

These are illustrated schematically in Fig. 6. It should be clear from this picture that particle cluster aggregation will give more compact aggregates, since single particles have a greater chance of penetrating into the interior of the cluster before contacting another particle and sticking. When two clusters collide, it is likely that particle-particle contacts occur before any significant inter-penetration has occurred. The corresponding log M vs log L plots are also shown and in this case the fractal dimensions for the two cases are 1.7 and 2.7.

Another factor that can influence aggregate fractal dimension is the collision efficiency, α, discussed earlier (1.1). If particles have been fully destabilised, so that $\alpha = 1$, then a particle approaching an aggregate will stick on first contact, leading to rather open structures with low fractal dimension. In the literature this is often called *diffusion-limited aggregation (DLA)*, since most work on fractal aggregates has been for the perikinetic case. DLA generally leads to aggregates with fractal dimensions in the region of 1.7-1.8. For lower colloid stability, a particle makes several contacts before adhering, so that it is more likely that some degree of penetration into the aggregate interior will occur. This process may be known as *reaction-limited aggregation (RLA)* and typically gives fractal dimensions of about 2.1 [11].

In water treatment, flocculation by applied fluid shear is much more important and this tends to give aggregates (flocs) of rather higher d_F than the diffusion-controlled case. This may be partly due to some restructuring, as a result of floc breakage and re-growth. Generally, higher shear rates lead to more compact flocs and this could explain some observed effects of rapid mixing conditions on subsequent floc formation. With prolonged mixing at high speed there is opportunity for small, compact flocs to form. When the mixing speed is reduced, these 'micro-flocs' are allowed to grow further, but the resulting flocs will be smaller than those formed directly from primary particles. Fractal dimensions of aggregates are discussed further in a later chapter by Bruce Jefferson and Peter Jarvis.

4.2. Collisions of fractal aggregates

The Smoluchowski treatment of aggregation kinetics is based on the assumption that colliding particles and aggregates are solid spheres. Even if the primary particles are spheres, aggregates are likely to rather open, fractal structures, as we have seen. This will significantly affect the diffusion coefficient and the collision radius of aggregates. For brownian collisions, these effects tend to cancel out, so that the assumption of solid spheres may not be too bad.

However, for the orthokinetic case, the increase in effective aggregate size, especially if the fractal dimension is quite low, can give a great increase in the aggregation rate. Instead of Eq. (5), the following expression for the collision rate coefficient applies:

$$k_{ij} = \frac{G d_0^3}{6} (i^{1/d_F} + j^{1/d_F})^3 \tag{15}$$

where d_0 is the diameter of the primary particles and the diameter of an i-fold aggregate has been assumed to be given by:

$$d_i = d_0 i^{1/d_F} \tag{16}$$

This follows directly from the definition of mass fractal dimension, since the aggregation number i is directly proportional to the aggregate mass.

The corresponding Smoluchowski expression, Eq. (5) is based on the assumption of 'coalesced spheres', with $d_F = 3$. In this case, the increase in aggregate size with aggregation number is quite slow (a 10-fold increase in diameter for 1000-fold aggregates). For lower fractal dimensions, aggregate size increases more rapidly, which can give a very significant increase in aggregation rate [12]. It also follows that the assumption of a constant volume fraction, leading to Eq. (11), is not justified for fractal aggregates. The effective floc volume will increase substantially as flocs grow larger.

Another point, already mentioned in 1.7, is that hydrodynamic effects become much less important for fractal aggregates.

4.3. Density of fractal aggregates

A very important consequence of the fractal nature of aggregates is that floc density decreases with increasing floc size. This effect has been known at least since the 1960s [13], well before the emergence of fractal concepts.

The effective, buoyant density of an aggregate in water, ρ_E, is simply:

$$\rho_E = \rho_A - \rho_W = \phi_S(\rho_S - \rho_W) \tag{17}$$

where ρ_A, ρ_L and ρ_W are the densities of the aggregate, water and the solid particles respectively and ϕ_S is the volume fraction of solid *within the aggregate* (not to be confused with the volume fraction of particles in a suspension). For an aggregate consisting of k identical primary particles, it can be shown that ϕ_S is given by:

$$\phi_S = k^{d_F - 3/d_F} \tag{18}$$

If the effective floc density (for instance measured by a sedimentation technique [14, 15]) is plotted against the floc size on a log-log-scale, then the data often show a linear decrease, with a characteristic slope as shown schematically in Fig. 7. This behaviour implies a relationship of the form:

$$\rho_E = Bd^{-y} \tag{19}$$

where B and y are empirical constants.

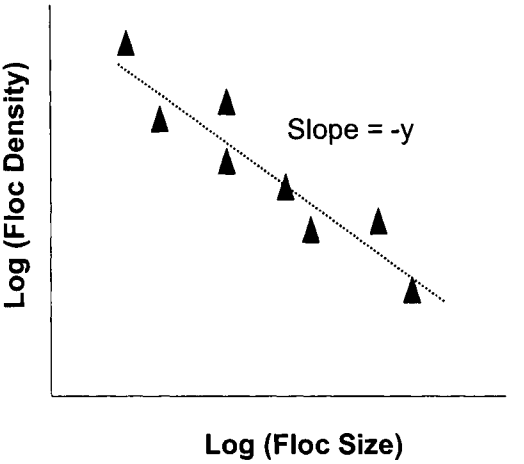


Fig. 7. Variation of floc density with size

It is not difficult to show that the exponent y is simply related to the fractal dimension:

$$d_F = 3 - y \tag{20}$$

Early measurements of floc density for systems of practical interest, including sewage sludges and hydroxide flocs [14], gave values of y in the range 1 - 1.4, corresponding to fractal dimensions of 2 - 1.6, in line with those obtained by other methods. However, it should be pointed out that experimental plots of log (floc density) vs. log (floc size) nearly always show a large degree of scatter, so that the derived slopes may not be entirely reliable.

The values of y just quoted imply a very significant decrease in density with increasing size. For instance, with $y = 1.1$ ($d_F = 1.9$), a 10-fold increase in floc size gives a nearly 13-fold decrease in effective density.

Another consequence is that, for flocs of comparable size and fractal dimension, the smaller the primary particles the lower the density. This follows from Eq. (18) and the fact that the smaller the primary particles, the larger the aggregation number, k, for a given aggregate size. Fig. 8 shows calculated densities of flocs composed of primary particles with two different diameters - 10 nm and 1 μm. The smaller size is typical of initial particles formed on precipitation of hydrous metal oxides and so is relevant to 'sweep flocculation'. The density of the primary particles is assumed to be 2500 kg m^{-3} and the fractal dimension of the aggregates is taken as 2.1. Note that the smallest floc size considered is 1 μm, which is the same as the primary particle size for the larger particles. In this case the effective density (in water) is 1500 kg m^{-3}. However, in the other case, a 1 μm aggregate (containing nearly 16,000 primary particles) has a much lower effective density and a solid volume fraction less than 2%. For 100 μm aggregates of 10 nm primary particles, the solid volume fraction is around 0.025%. This explains why hydrous oxide precipitates are so voluminous - they typically contain more than 99% water.

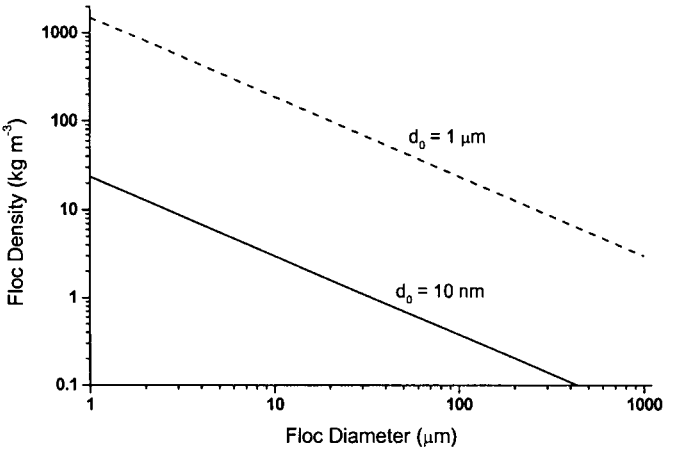


Fig. 8. Effective densities of flocs with different primary particle sizes (see text)

5. FLOC STRENGTH AND FLOC BREAKAGE

5.1. Limiting floc size

In practice, floc strength is an extremely important property. Breakage of flocs can have a serious effect on the effectiveness of solid-liquid separation processes. Nearly all flocculation processes are carried out in some form of turbulent flow and it is almost inevitable that some floc breakage will occur under these conditions.

Empirically, it is usually found that flocs grow only to a certain limiting size (see Fig. 4), which depends on their strength and the effective shear rate or energy dissipation, ε [16]:

$$d_{max} = C\varepsilon^{-n} \tag{21}$$

where C and n are empirical constants.

It is generally thought that the limiting size represents a balance between the formation and breakage of flocs. It will be seen that, in many cases of practical interest, floc breakage is irreversible to some extent (4.3), so that the concept of a dynamic equilibrium may be open to doubt. Nevertheless, the limiting floc size, under given shear conditions, gives a very useful indication of floc strength. Another approach is to subject pre-formed flocs to a sudden increase in shear (for instance by increasing stirring rate) and observe the decrease in floc size.

5.2. Mode of floc breakage

Analytical approaches to floc strength and floc breakage, especially under turbulent conditions, are quite complex and there is still no generally accepted theoretical picture (see e.g. Bache [17]). In turbulence, the Kolmogorov microscale, defined earlier in Eq. (14), is very important. Flocs larger than this are subject to pressure differences and normal forces across the floc, which lead to *rupture* of flocs into fragments of comparable size. Flocs smaller than the turbulence microscale are thought to be *eroded* by viscous forces, giving

many particles much smaller than the original floc. Floc rupture and erosion are shown pictorially in Fig. 9. The exponent n in Eq. (21) depends on the mode of breakage, being rather larger for rupture than for erosion. However, values are difficult to check experimentally and may depend on the specific nature of the system being considered. Under typical water treatment conditions, flocs tend to be of the same order of size as the microscale and the theoretical picture is still far from clear.

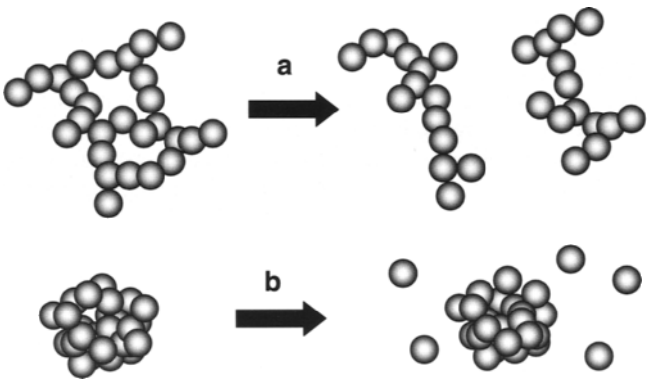


Fig. 9. Showing breakage of flocs by a) rupture and b) surface erosion

Fundamentally, the force, f_0, required to rupture a floc must depend on the number of particle-particle bonds broken and the strength of these bonds. The number of bonds broken depends on the cross-sectional area at the plane of rupture and it seems reasonable to define floc strength, σ, as the ratio of rupture force to the cross-sectional area:

$$\sigma = \frac{4f_0}{\pi d^2} \tag{22}$$

The floc strength expressed in this way is a force per unit area (or pressure) and values from a few Nm^{-2} up to around 1000 Nm^{-2} have been reported, depending on the system. This means that, for typical flocs of around 100 μm in size, the actual force required for breakage would lie somewhere in a range from nN - μN.

The number of particle-particle bonds per unit area must depend on the density of the floc and hence on the fractal dimension. For open, low density structures (low d_F), there will be few bonds per unit area and so the flocs should be relatively weak. More dense flocs should have higher strength. It may be that denser flocs are more prone to surface erosion, rather than splitting (see Fig. 9), although there is little experimental evidence on this point. Furthermore, as aggregates grow, they become less dense, and the strength, as defined by Eq. (22) should decrease. This would be another factor determining the maximum floc size.

Another influence is the size of the primary particles. Generally, the smaller the primary particles, the stronger the floc [16]. This can be explained by the larger number of particles within an aggregate of a given size and hence the larger number of particle-particle contacts per unit area.

The strength of bonds between particles is extremely important in determining floc strength. When particles are coagulated by simple salts only van der Waals attractive forces

are operative, giving rather weak aggregates. Hydrolysing metal salts and, especially, polymeric flocculants can give considerably stronger flocs. High molecular weight polymers acting by the 'bridging' mechanism can give very strong (and hence large) flocs.

In summary, floc strength is greater for:

- denser flocs
- smaller primary particles
- stronger inter-particle attraction

5.3. Reversibility of floc breakage

Although it has long been recognised that flocs may be quite fragile objects, which may be broken at increased shear rates, it is often assumed that this breakage is reversible, so that broken flocs can re-form when the shear rate is reduced. However, for flocs typical of those produced in water treatment plants, breakage may not be easily reversed. This effect has been known for some time (e.g. Francois [18]), and has more recently been the subject of systematic studies [8, 19, 20].

Fig. 10 shows the results [20] of modified jar test experiments with continuous monitoring of Flocculation Index, in the same way as in Figs. 4 and 5. Kaolin clay suspensions (50 mg L^{-1}) were flocculated with aluminium sulphate ('alum') and 2 commercial polyaluminium chloride products (PACla and PAClb). All were added at an equivalent dosage of 3.4 mg L^{-1} as Al. After a short rapid mix period (10s at 400rpm), flocs were formed by stirring at 50 rpm for 10 minutes. The stirring speed was then increased to 400 rpm (an increase of effective shear rate by a factor of more than 20). The stirring speed was maintained at 400 rpm for 10 seconds and then restored to 50 rpm.

It is clear from Fig. 10 that the three coagulants give different plateau values of FI, indicating significant differences in floc strength, in the order PAClb > PACla > Alum. On increasing the stirring speed, the FI shows a sudden and very rapid decrease in all cases, showing extensive floc rupture, rather than surface erosion. When the stirring speed is restored to 50 rpm some increase in FI occurs, but only to a limited extent. This shows that for all these coagulants, floc breakage is only partially reversible. Although the three coagulants show quantitative differences in floc growth, breakage and re-growth, the *relative* degrees of breakage and re-growth are reasonably similar. This indicates that the mechanisms of flocculation are broadly similar.

The time for which the high shear rate is applied (the floc breakage time) has a significant effect, as shown in Fig. 11. Here the breakage time is varied from 10 s to 5 min, but other conditions are the same as for Fig. 10. The coagulant is alum (3.4 mg L^{-1} Al), but very similar results are found for the PACl samples. It is clear that increased breakage times do not give smaller flocs – breakage seems to be practically complete after 10s. However, extended breakage times give less re-growth after the stirring speed is reduced. This is especially apparent for the 300s case.

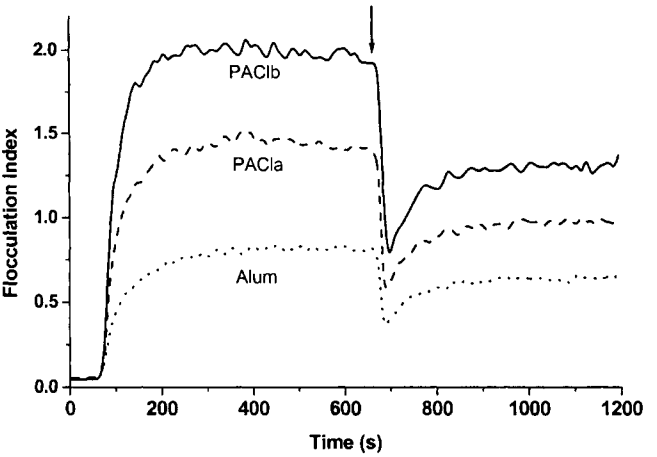


Fig. 10. Floc growth, breakage and re-growth with three coagulants, all at 3.4 mg L^{-1} as Al. Arrow shows increase of stirring speed (floc breakage)

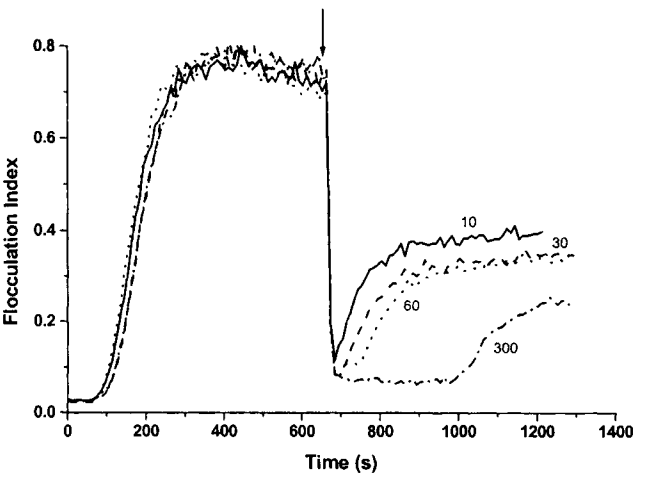


Fig. 11. Flocculation, breakage and re-growth with alum (3.4 mg L^{-1} Al) for different breakage times. The figures on the curves are the times (s) that the stirring speed was held at 400 rpm

Another important effect is that of rapid mix time [8]. We have already seen in Fig. 5 that rapid mix time has a very significant effect on floc growth. However, the breakage and re-growth of these flocs follows roughly the same pattern, irrespective of rapid mix time. The results in Fig. 12 are from the same experiment as Fig. 5, using PACIb at 3.4 mg L^{-1} Al, but in this case floc breakage (10s at 400 rpm) and re-growth (at 50 rpm) are also shown. In all cases flocs break to about the same FI value and there are only minor differences in FI after re-growth. This means that, although flocs formed after longer periods of rapid mixing are smaller, they break to a lesser extent and show more complete re-growth.

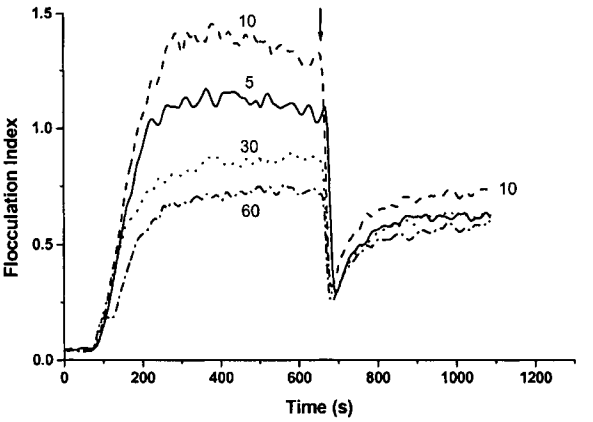


Fig. 12. Showing the effect of rapid mix time (400 rpm at times shown on curves) on floc growth, breakage and re-growth

The irreversible floc breakage found for hydrolysing coagulants is not observed for some other additives. For instance, the cationic polyelectrolyte polyDADMAC shows almost complete re-growth of flocs after breakage [19]. Fig. 13 shows the results of tests under the same conditions as for Fig. 10, for PAClb and polyDADMAC, except that in the latter case the period of floc growth at 50 rpm was extended to 30 minutes, because of the longer time needed to reach the plateau FI value.

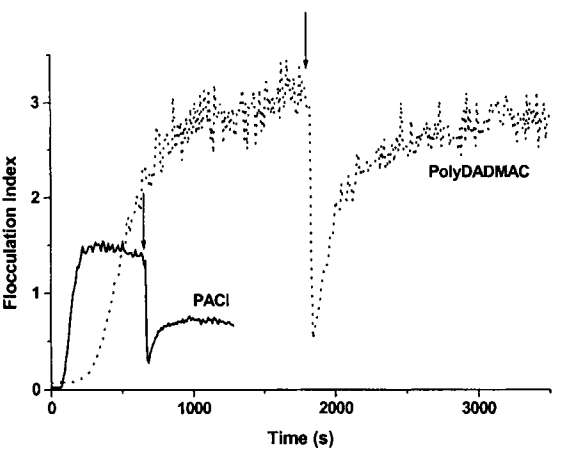


Fig. 13. Formation, breakage and re-growth of flocs with PAClb and polyDADMAC

Several important points emerge from Fig. 13. The FI value achieved by polyDADMAC is considerably higher than with PACl, indicating that the former gives larger and stronger flocs. The slower onset of flocculation with polyDADMAC is explained by the relatively slow adsorption of this flocculant. Also it can be seen that the rate of flocculation is lower with polyDADMAC, even though the flocs eventually grow larger. The explanation is essentially the same as that given earlier in relation to Fig. 4. The cationic polyelectrolyte destabilizes the particles by charge neutralization, but no new particles are produced, so the

collision rate remains quite low compared to the 'sweep floc' case. In both cases, there is a sudden and very rapid decrease in FI when the stirring speed is increased, as before. However, the re-growth behaviour is quite different for the two additives. With the polyelectrolyte, re-growth of flocs occurs to give nearly the same FI value as before breakage, indicating almost complete reversibility of floc breakage. By contrast, with PACl, as shown previously, the broken flocs re-grow only to a limited extent.

Reversible floc breakage would be expected when the interaction between particles is of an essentially physical nature, such as van der Waals or electrical attraction. The cationic polyelectrolyte in this case is of high charge and fairly low molecular weight, and is believed to act by an 'electrostatic patch' mechanism, giving electrostatic attraction between positive and negative 'patches' on the surfaces of neighbouring particles. There is no obvious reason why such bonds should not re-form after breakage. Particles coagulated by simple salts are held by van der Waals attraction and complete reversibility of floc breakage is found.

The irreversible floc breakage has been found for a wide range of coagulants based on aluminium and iron (III) [19], including pre-hydrolysed forms such as PACl. It is very likely that, under most practical conditions, the effect of these additives depends on the precipitation of an amorphous metal hydroxide, which binds impurity particles together ('sweep flocculation'). Although the details are not clear, it is possible that the breakage of 'sweep flocs' involves breaking chemical bonds within the hydroxide precipitate, and that this is responsible for the fact that broken flocs do not completely re-form.

Since floc breakage is quite likely to occur under water treatment conditions, the irreversibility is likely to have important practical consequences. However, this aspect has received rather little attention so far.

REFERENCES

[1] M. Smoluchowski, Z. Phys. Chem., 92 (1917) 129.
[2] D.F. Lawler, Water Sci. Technol., 10 (1993) 165.
[3] J. Duan and J. Gregory, Adv. Colloid Interface Sci., 100-102 (2003) 475.
[4] T.R. Camp and P.C. Stein, J. Boston Soc. Civ. Eng., 30 (1943) 219.
[5] P.G. Saffman and J.S. Turner, J. Fluid Mech., 1 (1956) 16.
[6] L.R. Spielman, in K.J. Ives (ed.) The Scientific Basis of Flocculation, Sijthoff and Noordhoff, Alphen aan den Rijn, 1978, pp. 63-88.
[7] J.L.Cleasby, J. Environ. Eng. ASCE, 110 (1984) 875.
[8] M.A. Yukselen and J. Gregory, J. Chem. Technol Biotechnol, 79 (2004) 782.
[9] K.J Ives and O. Hoyer, Water Sci. Technol., 37(10) (1998) 67.
[10] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco, 1982.
[11] P. Meakin, Adv. Colloid Interface Sci., 28 (1988) 249.
[12] Q. Jiang and B.E. Logan, Environ. Sci. Technol., 25 (1991) 2031.
[13] A.L Lagvankar and R.S. Gemmel, J. - Am. Water Works Assoc., 60 (1968) 1041.
[14] N. Tambo and Y. Watanabe, Water Res. 13 (1979) 409.
[15] C. Fargues and C. Turchiuli, Trans. IChemE., 81A (2003) 1171.
[16] K. Mühle, in B. Dobias (ed.) Coagulation and Flocculation, Marcel Dekker, New York, 1993, pp. 355-390.
[17] D.H. Bache, Chem. Eng. Sci., 59 (2003) 2521.
[18] R.J. Francois, Water Res., 21 (1987) 1023.
[19] M.A. Yukselen and J. Gregory, Int. J. Miner. Proc., 73 (2004) 251.
[20] J.Gregory, Water Sci. Tech., 50 (2004) 163.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 4: Practical application of fractal dimension

Bruce Jefferson and Peter R Jarvis

School of Water Sciences, Cranfield University, Bedfordshire, MK43 0AL

1. INTRODUCTION

An understanding of the physical characteristics of particles that exist or are generated in water treatment is important in determining the efficiency, operation and robustness of the separation processes used to remove them. Considerable attention has been traditionally placed on measuring and understanding the significance of particle size and to a lesser extent charge on these processes. However, consideration of aggregate structure has been less well explored due mainly to a lack of appropriate theories to describe complex random structures. The major change occurred in the mid 1970s with the publication of Mandelbrot's [1] work on fractal geometry, which provided the required framework from which all the following fractal work is based. The key feature was the identification that aggregate structures exhibit a self similarity such that their character appears constant at different scales of magnification. Many natural objects exhibit this property such as cauliflowers, trees and lungs. Perhaps the best known example of a natural fractal structure is the coastline of an island such as Great Britain. When the length of the coastline is measured with a fixed length stick a relationship is observed between the size of the measuring stick and the measured property (perimeter of the coastline). This occurs as smaller sized sticks can resolve details that the larger sticks will miss such as bays and inlets (Fig. 1).



Fig.1. Coastline of Britain measured by two different length measuring sticks

The relationship can be defined by a power law such that:

$$X \propto R^{d_f} \tag{1}$$

where X is the measured property, R is a linear measure of the size and d_f is a measure of the scaling and level of rugosity of the structure known as the fractal dimension. For Euclidean objects, the dimensional value of d_f will be 1 for a straight line, 2 for a two dimensional planar shape and 3 for a compact three dimensional shape. Fractal objects take non-integer values of d_f and are therefore said to show non-Euclidean dimensionality. Values approaching 3 for a three dimensional floc therefore indicate a high degree of compaction whilst values approaching 1 indicate a very loose and open structure. The fractal dimension can therefore give important structural information of floc compaction and the space filling nature of the aggregate.

The fractal dimension of aggregates can be determined based on a number of measurable properties. However, they all use the same basic concept in that a geometric power law scaling relationship applies between each dimensional geometry: mass or volume for three dimensions, projected surface area for two dimensions, perimeter for one dimension and the characteristic length scale of the aggregate. In the case of aggregates, the most common is the relationship between the mass of particles and the radius of gyration of the aggregate:

$$M \propto R_g^{d_{f3}} \tag{2}$$

where the radius of gyration (R_g) is defined as the standard deviation of the particle from its centre of mass. The term d_{f3} is the 3D fractal dimension.

Similar expressions can be written for both area and perimeter:

$$A \propto L^{d_{f2}} \tag{3}$$

$$P \propto L^{d_{f1}} \tag{4}$$

where A and P are the projected surface area and the perimeter respectively and d_{f2} and d_{f1} are the 1D and 2D fractal dimensions respectively. The subject of fractal dimension was introduced in the previous chapter, including examples of how such knowledge can be applied to the description of aggregated structures, and will not be expanded on here. Instead this chapter will address the issues of measurement and application to discuss the role fractal dimension plays in current and future studies of aggregated systems.

2. EXPERIMENTAL METHODS

The techniques commonly used to measure structural properties of aggregates can be categorised as based on either:

- image analysis
- scattering
- settling

2.1. Image analysis

Imaging is one of the oldest methods to characterise particles and provides one of the most versatile as it offers methods for understanding morphology as well as size. Consequently, it has been widely used to measure fractal dimension [2- 4] especially for simulated aggregates where experimental techniques are not possible.

The premise of the technique is to either generate (simulated) or capture (real) a projection of the image which is then analysed in terms of its morphological character (length, area, circularity, perimeter etc.). Measurements of d_{f1} or d_{f2} are generated from power law expressions between the perimeter or projected area and the characteristic length scale. The need for projected images means analysis is limited to d_{f1} and d_{f2} as values greater than 2 are not possible. A number of techniques can be employed to provide a determination of the measured property as a function of the length scale but the most commonly used are:

- structured walk
- box counting
- enclosing sphere

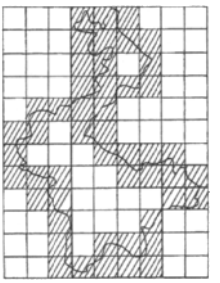
The perimeter of objects is commonly measured by either stepping round the edge of an object with a set of virtual dividers (structured walk) or overlaying a grid onto the image (box counting) (Fig. 2). In both cases the characteristic length of either the divider or box is changed and the resultant perimeter measured. In the case of projected area an enclosing circle, or sand box, approach is most common whereby the mass contained within a circle of increasing radius is the plotted against the radius and the d_{f2} calculated. In all three cases a log-log plot of P or A against L will generate a straight line plot whose gradient is equal to the fractal dimension -1 (Fig. 2).

The main requirement for 2D fractal dimension analysis using image analysis is for the image to be of suitable quality for commercial software packages to be able to distinguish the floc from the background [5]. In practice this often requires considerable image correction prior to fractal analysis and works best with flocs that show good contrast with their background and are not translucent. Image capture has been performed from a range of microscopes including transmission, scanning and optical such that a range of magnifications are possible. Common to all the techniques is the need for some sort of sample preparation and care is thus required to avoid altering the character of the aggregates being analysed.

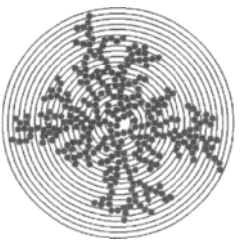
The above approach can be extended to 3D images in the case of simulated aggregates. An example is given in Fig. 3 which shows the structured walk and enclosing circle (or sphere) of 2D and 3D aggregates simulated by contacting single particles onto an aggregate with varying degrees of diffusion in the trajectory of the particles movement and a 100% sticking probability upon contact with the aggregate [6]. The simulations visually show that more open structures are generated when diffusion dominates the motion of particles. However, measurement of the fractal dimension in both 2D and 3D shows relatively little difference between the extremes of no diffusion and 100% diffusion. To illustrate in the case of 3D enclosing sphere measurements the d_f values decreased from 2.7 to 2.45 indicating an insensitivity in the technique compared with visual observation which can be common in practical situations. Recent research has attempted to link 2D fractal dimension to 3D values with reasonable success especially for inorganic primary particles. This has not been found to hold for organic particles as they contain greater irregularity in the perimeter of the primary particle [7].



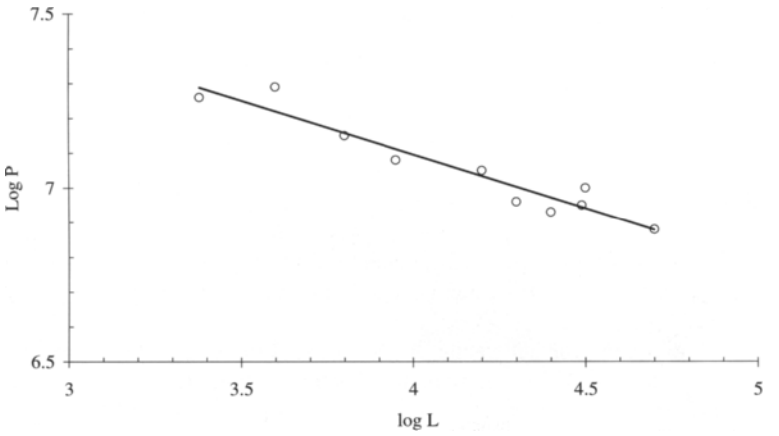
(a) Structured walk



(b) Box counting



(c) Enclosing sphere

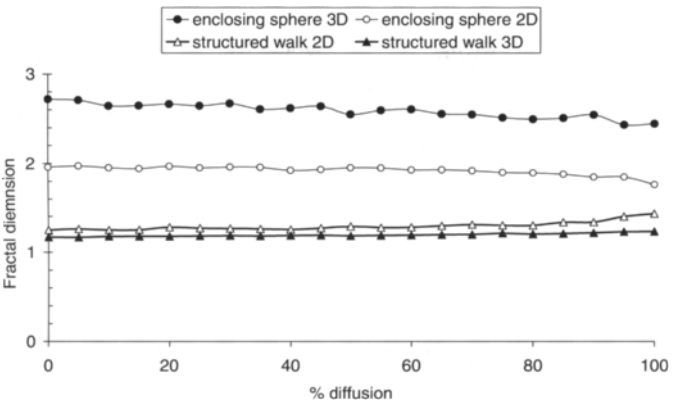


(d) Example plot of perimeter against divider length. $d_{f1} = 1.3$

Fig. 2. Examples of analysis approaches for d_{f1} and d_{f2}

2.2. Scattering

When irradiated, particles emit secondary or sympathetic radiation the intensity of which changes with the scattering angle. Consequently, measurement of the intensity of the radiation at different angles provides a method of determining the fractal dimension if enough is known about the scattering properties of the material contained within the aggregate. The approach has been demonstrated with X-Rays and neutrons but static light is the most common approach used due to its ease, cost and the fact that the longer wavelength of light permits measurements of aggregates in the micron size range and so is of most importance to water treatment applications.



(a) Summary of fractal dimension of the simulated 2D and 3D aggregates as a function of the % diffusion induced agglomeration.

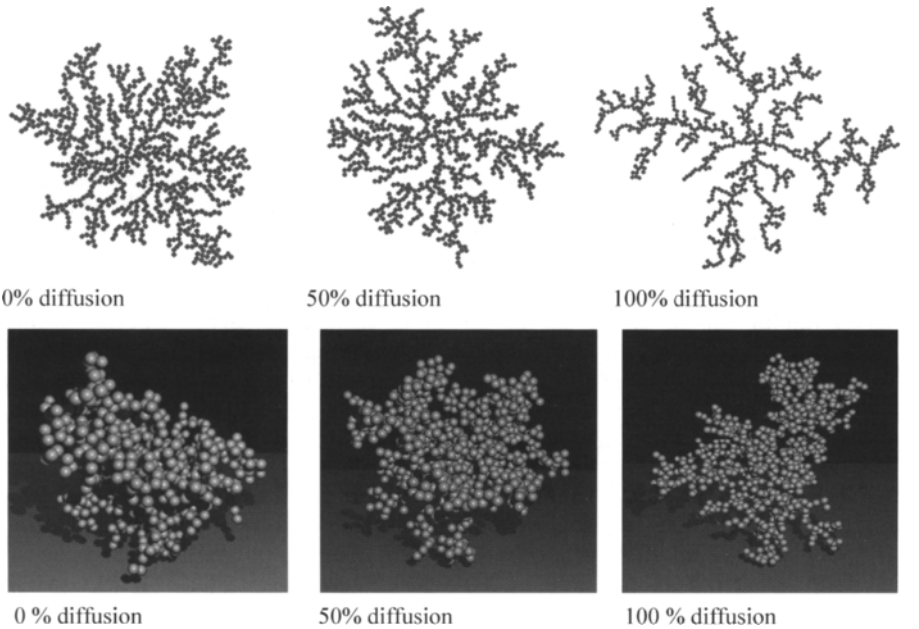


Fig. 3. Example results from simulated aggregated measured by enclosing sphere and structured walk approaches

The approach is based on light scattering theory, detailed elsewhere [8, 9] and requires a number of assumptions to hold:

(1) The primary particles that make up the aggregate are uniform in shape and size.
(2) The refractive index of the aggregate material is low so that the wavelength of the incident light does not become shortened.

(3) Light is only scattered once as it passes through the suspension of aggregates before hitting the detector. Multiple scattering should be minimised by ensuring the concentration of particles is low

The resulting description comprises of a power law expression between the intensity of the scattered radiation, I(Q), and the wave number [Q]:

$$I(Q) \propto Q^{-d_{f3}} \tag{5}$$

where Q is estimated by:

$$Q = \frac{4\pi n \sin(\theta/2)}{\lambda} \tag{6}$$

Where n is the refractive index of the suspending medium, θ is the scattered angle, λ is the wavelength of the radiation in a vacuum. The above approach is only valid when the wave number is much bigger than the primary particle and much smaller than the overall size of the aggregate:

$$\frac{1}{R_{agg}} \rangle\rangle Q \rangle\rangle \frac{1}{R_{part}} \tag{7}$$

This is because when $1/Q$ approaches the size of R_{agg} the relationship is affected by the edges of the aggregate identified by the Guinier region where the intensity becomes independent of the wave number (Fig. 4); Sorensen [8] recommends that d_f only be determined in the region where $QR_{agg}>5$. Similarly, when $1/Q$ approaches the size of R_{part}, light is mainly scattered by the primary particles and becomes independent of Q as indicated by the Porod region; therefore the fractal dimension should only be measured between these limits.

The above approach is based on Rayleigh-Gans-Debye (RGD) theory which is only valid when the primary particles satisfy a condition of independent scattering. This is normally considered appropriate when:

$$|m-1| << 1 \tag{8}$$

$$(2\pi n/\lambda)L|m-1| << 1 \tag{9}$$

where m is the material refractive index and L is the length of the scattering body.

To date the majority of systems investigated by this technique has been well defined mono dispersed primary particles such as latex, aluminium oxide and hematite for which the above approximations are valid. Extension of this to the more complex flocs commonly encountered in water treatment is more difficult. The primary particles are often a mixture of materials of unknown properties contained in a poly dispersed system. A notable exception to this is activated sludge as the very low refractive index of the bacteria in the floc allows the assumptions in RGD theory to be met, and consequently this technique is becoming more widely used for investigating the structure of wastewater flocs.

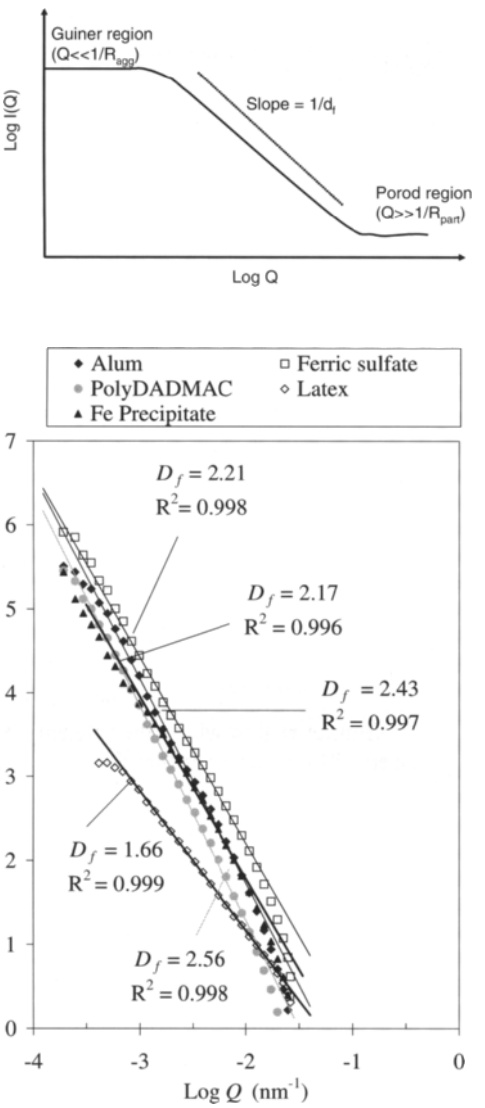


Fig. 4. (a) Characteristic light scattering plot and (b) example plot for NOM flocs

Perhaps the most important limitation is the scale of investigation. The power law relationship breaks down beyond small flocs sizes within the distribution. Reported limits are 50-100 μm for kaolin, 70 μm for activated sludge and 5 μm for NOM flocs. Given that the median flocs size of the different systems is likely to be between 400-800 μm the approach appears to principally analyse the micro flocs within the distributions rather than the flocs across the whole size range.

2.3. Settling rate

Monitoring of the settling behaviour of aggregates is a well established practice due to its direct relevance to sedimentation processes. The fractal nature of aggregates tends to influence settling behaviour in two ways:

(1) an increase in drag compared to a sphere of equal mass due to the increase in projected area
(2) a reduction in drag by advection of fluid through the aggregate.

Determination of fractal dimension from settling velocity is based on a modified Stoke's law expression to account for the two impacts:

$$v_s = \frac{d_f^{D_f-1} 4kg}{3A(\beta)\mu} \tag{10}$$

where v_s is the terminal settling velocity (m s^{-1}), k is a proportionality constant (kg m^{-D}), ρ_l is the density of the liquid (kg m^{-3}), $A(\beta)$ is a correction factor for advection through the floc, μ is the viscosity of the suspending medium (m s^{-1}) and g is acceleration due to gravity (m s^{-1}). Consequently, the slope of a log-log plot of floc settling velocity against size will therefore yield the fractal dimension, as d_f = slope +1.

The approach above only applies for floc with Reynolds numbers less than one and when the flocs fall in an isolated manner at their terminal settling velocity. These conditions are likely to hold for most practical cases although care is required when d_f values of less than 2 are recorded, as porosity effects are likely to significantly influence the results [10]. Practical experiments involve the videoing of settling flocs as they pass by a fixed CCD camera (Fig. 5). Care is required to ensure: temperature equalisation between the sample and the settling column, that the flocs are transferred into the column without breakage, only individual flocs are in view at any instant and that when recorded the flocs are in the focal plane of the camera. Any and all of these issues are common and can seriously impact on the quality of the recorded data, making the technique quite difficult and very tedious for extensive measurement. Practical experience indicates that a minimum of 80 flocs must be measured for representative results, which takes an experienced operator between 1 and 3 days to measure.

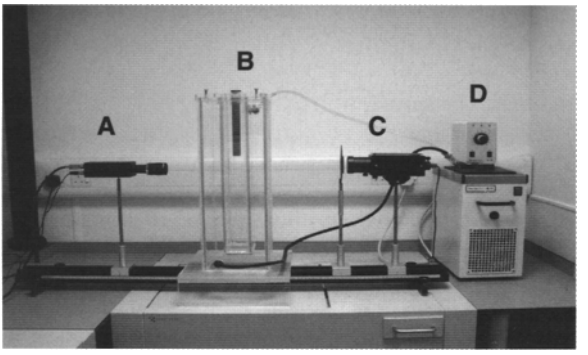


Fig. 5. Settling column apparatus. A = CCD camera linked to PC and image analysis software; B = the central settling column contained within a sealed water bath; C = lamp for background lighting; D = temperature controlled-circulator for the water bath

2.4. Comparison of techniques

As outlined in the above descriptions the three main methods of fractal analysis are based on very different premises, and have different strengths and weaknesses which require consideration when selecting the most appropriate technique (Table 1).

Table 1
Comparison of the techniques

Technique	Advantages	Disadvantages
Light scattering	- Rapid, non-intrusive method - Lends itself well to dynamic, online analysis - Very good for analysis of small aggregates with an open structure and low refractive index - Takes a large number of readings from many aggregates in a few seconds	- Not good for poly dispersed aggregates made from many primary particles - Choosing an appropriate model for scattering behaviour can be difficult - Results affected by contamination from dust etc. - Power-law relationship breaks down at large floc size
Settling	- Best for measuring fractals of compact flocs - Cheap and simple - Not prone to contamination - Good for aggregates of made from a number of different primary particles	- Time consuming - Finding an appropriate drag coefficient is difficult - Can get non-random orientation of falling aggregates - Careful regulation of settling column required
Image analysis	- Best for large, open aggregates - Not prone to contamination - Examination of single flocs allows detailed information on variation in floc structure within a sample	- Time consuming - Requires well defined, high contrast images for accurate analysis – which flocs generally aren't

Light scattering works well for small, open flocs that have low refractive indices, in fact the type of flocs that the other systems are not as effective at measuring. The low refractive index makes observation difficult for both settling and image analysis. Whilst this can be overcome with the use of electron microscopes, the cost and the uncertainties derived from sampling means it is seldom used. Perhaps the most important benefit of light scattering techniques is the speed of analysis which enables the technique to be used on line for kinetic and growth evolution experiments. An example is given for the cyclical breakage and regrowth of latex coagulated in NaCl (Fig. 6). The aggregate reached a d_f of 1.67 after the initial growth phase and compacted after each breakage event to a d_f of 2.05 which returned back to the original level of compaction upon returning to the original shear level. This demonstrates the reversibility of aggregation by simple double layer compression; a result only able to be observed by this technique. The major difficulties with scattering are the uncertainty in the models used and the upper limit on the size of aggregates included. When aggregates are dense and the primary particles are not tiny, scattering is affected by shadowing, scattering interactions and multiple scattering, which invalidates the RGD model [9]. The result is that such systems are not commonly used for flocs generated from complex waters, such as those containing NOM or algae, and are instead used principally in laboratory studies of more fundamental concerns.

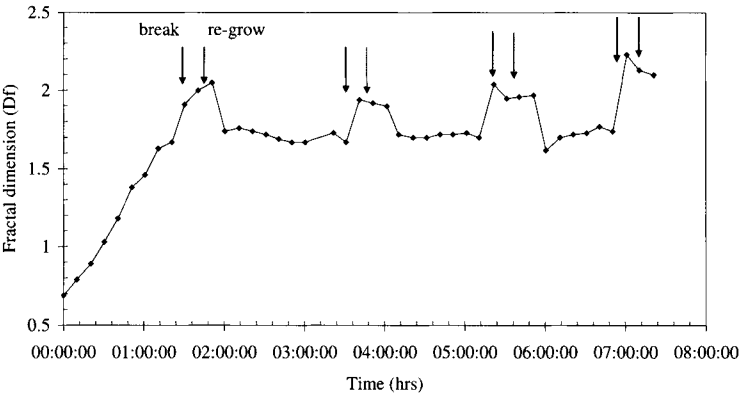


Fig. 6. The change in fractal dimension with time of latex flocs coagulated in 1 M NaCl for cyclical breakage and re-growth regimes

Image analysis works best with large images that provide a high level of contrast to the background. The major difficulties relate to the sample preparation required to generate the image and, whilst confocal laser microscopy has to a degree reduced this, it remains the major limitation of the technique. The benefits of the technique are that individual images are analysed one at a time providing information of the structural detail of individual flocs as well as the whole distribution. This also permits the removal of spurious aggregates eliminating to a degree the problems of contamination that are common with light scattering.

Settling is the most reliable technique for aggregates with a high fractal dimension. The technique is good with large aggregates and is suitable for polydispersed systems. These advantages make the technique appropriate for studying the types of aggregates generated at water treatment works. The major limitations are the experimental difficulties rather than analytical and the fact that it must be conducted off line. In particular, individual flocs must be measured in sufficient number to represent the distribution, no temperature gradients can exist and care is needed in transferring the flocs to the test equipment to avoid break up and lateral movement out of the focal plane. It is important to report the technique used in measuring the fractal dimension as each technique will provide a different answer for the same aggregate. To illustrate, in a comparative study of activated sludge the measured value of the d_f was 1.3 from settling measurements and 2.06 from light scattering [11]. Similarly, comparison of the aggregates formed during the coagulation of NOM with different coagulants revealed d_f values of 2.25, 2.42 and 2.5 for ferric, alum and polyDADMAC when using light scattering and 1.75, 1.99 and 2.25 when using settling experiments (Fig. 7). Interestingly, the range of values measured was 0.25 for scattering and 0.5 for settling suggesting the latter was a more suitable technique in this case.

3. APPLICATIONS

The application of fractal dimension in science and engineering crosses many disciplines from being able to identify psychiatric disorders from magnetic resonance images of cerebral cortical surfaces in medicine [12], to the detection, location and depth of cracking in structural supports; from an increase in complexity of a vibrational signal [13] through to the modelling of settlements from ancient civilisations in archaeology [14].

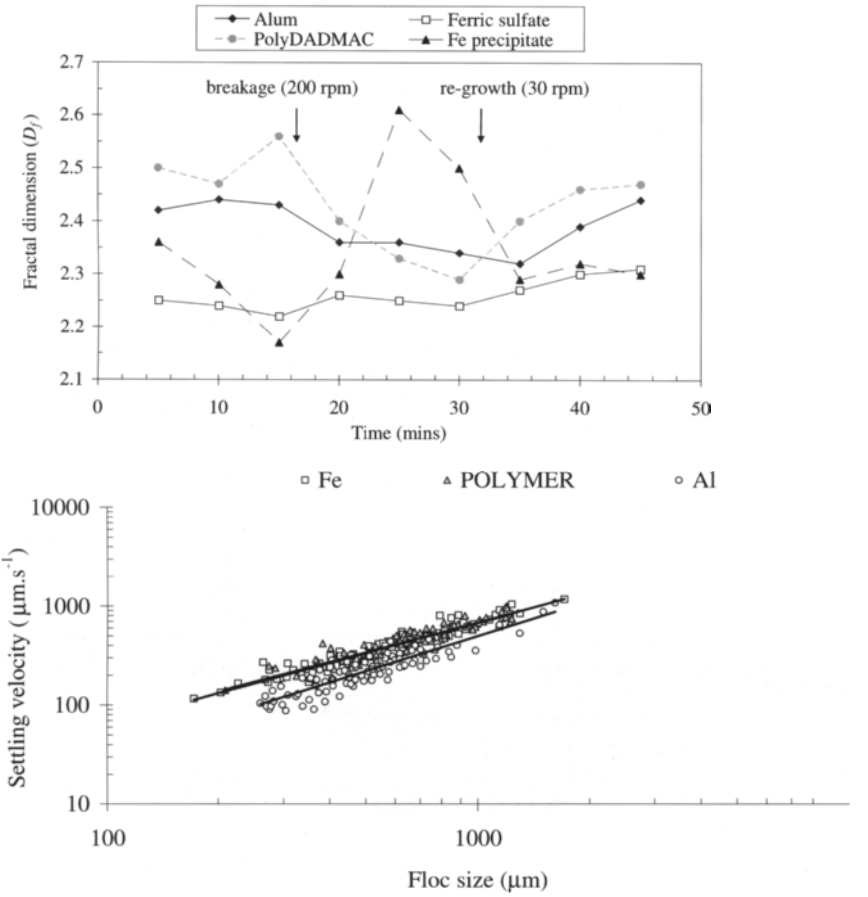


Fig. 7. d_f of NOM flocs measured by either scattering or settling

Water treatment investigations to date have focussed on understanding structural changes to flocs during coagulation and flocculation especially with idealised particles or simulated aggregation. To a much lesser extent some work has attempted to make exploitable links between the fractal properties of aggregates and the performance of downstream processes such as filtration and membranes. Table 2 summarises some recent publications that involve fractal dimension to demonstrate the level of variation observed and the implications that were attributed to it if any. Comparison of the studies reveals reported changes in d_f of between 0.05 and 0.53 with the majority between 0.2 and 0.4. In general, discussion of the errors associated with the measurement techniques is uncommon but it is not inappropriate to assume that changes of less than 0.1-0.2 are difficult to clearly discern experimentally. Detailed error analysis is thus clearly required in the future to assist in our understanding of the importance of the observed changes especially as in most cases the measurement is effectively an average value for the distribution of aggregates in the system.

Table 2
Recent publications relating to the application of fractal dimension measurements

System description	Method of obtaining fractal dimension	Range of fractals measured	Δd_f	How fractal changed	Practical implications of results	Reference
The effect of fractal dimension on membrane filtration for haematite flocs	Light scattering (3D)	1.83 → 2.25	0.42	df increased as mixing speed increased	Specific cake resistance affected by fractal dimension for small flocs only (~10 μm)	[23] Lee *et al.* (2003)
Change in floc structure and adsorption with NOM and PAC	Light scattering (3D)	2.0 → 2.40	0.4	df decreased when reservoir water rich in NOM diluted by a third.	Organic:coagulant ratio important in determining optimum floc removals.	[17] Ho and Newcombe (2005)
The effect of increasing the ratio of organic matter on organo-ferric floc structure	Settling (3D)	1.78 → 2.20	0.42	As the ratio of organic in the floc decreased from 3.8 to 0 (by mass), the df increased	Organic:coagulant ratio important in determining optimum floc removals.	[18] Jarvis *et al.* (2005)
The change in fractal dimension of activated sludge floc with source wastewater	Light scattering (3D)	1.96 → 2.49	0.53	df changed from source to source dependent on wastewater being treated.	Flocs with high fractal dimension stronger and more resistant to shear.	[25] Wilen *et al.* (2003)
The effect of polymer addition on: 1) kaolin sludge 2) activated sludge	Light scattering (3D)	Kaolin: 1.68 → 1.98 Activated sludge: 1.85 → 2.12	0.3 0.27	Addition of increasing concentrations of polymer reduced df and increased floc size		[11] Wu *et al.* (2002)
The effect of polymer addition on: 1) kaolin sludge 2) activated sludge	Settling (3D)	Kaolin: 1.86 → 2.31 Activated sludge: 1.31 → 1.55	0.45 0.24	No observable trend relating to polymer concentration and df		[11] Wu *et al.* (2002)
Change in floc structure from charge neutralisation to sweep flocculation for alum-particle flocs.	Image analysis (2D): $A \propto P^{\frac{2}{D_f}}$	Lake water: 1.96 – initial 1.84 – CN 1.65 – SF Clay suspension 1.89 –initial 1.81 – CN 1.77 – SF	 0.14 0.31 0.08 0.12	df lower for sweep flocculation than charge neutralisation	Mechanistic understanding of floc formation. Floc size was more important than fractal dimension for residual turbidity	[5] Chakraborti *et al.* (2000)

Table 2 contd.
Recent publications relating to the application of fractal dimension measurements

System description	Method of obtaining fractal dimension	Range of fractals measured	Δd_f	How fractal changed	Practical implications of results	Reference
Change in floc structure from charge neutralisation to sweep flocculation for alum-particle flocs	Image analysis (2D): 1) $A \propto P^{\frac{2}{D_f}}$ 2) $A \propto L^{D_f}$	Image analysis: 1) 1.31 – SF 1) 1.08 – CN	0.23	No consistent trend between sweep flocculation and charge neutralisation for df across 3 techniques.		[26] Kim *et al.* (2001)
		2) 1.48 – SF 2) 1.53 – CN	0.05			
	Light scattering (3D)	Light scattering 2.20 – SF 1.84 – CN	0.36			
Differentiation of microbial flocs and granules	Image analysis (2D): 1) $A \propto L^{D_f}$ 2) Box counting	Power law 2.14 – Granules 1.84 – Flocs	0.3	df higher for granules than floc aggregates		[2] Bellouti *et al.* (1997)
		Box counting 1.95 – Granules 1.90 – Flocs	0.05			
Change in fractal dimension of activated sludge flocs during cyclical shearing	Light scattering (3D)	2.45	0	df unchanged during breakage at high shear and re-growth at reduced shear		[27] Chaignon *et al.* (2002)
Change in fractal dimension of latex flocs during cyclical shearing	Light scattering (3D)	1.67 → 2.05	0.38	df increased during breakage shear but returned to original during re-growth phase		[28] Jarvis *et al.* (2005)
Change in fractal determination of Fe flocs during shear	Light scattering (3D)	2.52 → 2.71	0.19	df increased as shear rate increased.	Flocs re-structure and become more compact after breakage in higher shear rates.	[29] Jung *et al.* (1996)
Determination of fractal dimension for different coagulation mechanisms	Light scattering (3D)	1.82 – DLA 2.06 – RLA	0.18	df increased as the coagulation mechanism changed from DLA to RLA	Mechanistic understanding of floc formation	[30] Tang (1999)
Effect of different coagulants on sewage sludge	Settling (3D)	1.67 – PAC 1.87 – $FeCl_3$ 1.90 – $Fe(SO_4)_3$	0.23	df changed with coagulant type. Flocs with lower D_f had increased phosphate adsorption	Nutrient adsorption maximised by low fractal dimension	[31] Smoczynski and Wardzynska (1996)
Effect of polymer dose on sewage sludge flocs	Light scattering (3D)	1.78 → 2.23	0.45	As the polymer dose was increased, the df decreased.	Flocs with low fractal dimension had improved de-watering characteristics and lower membrane fouling	[32] Waite (1999)

Table 2 contd.
Recent publications relating to the application of fractal dimension measurements

System description	Method of obtaining fractal dimension	Range of fractals measured	Δd_f	How fractal changed	Practical implications of results	Reference
Change in fractal dimension with aggregation time for charged silica particles	Light Scattering (3D)	1.79 → 1.88	0.09	Increase in df as flocs grow		[15] Kim and Berg (2000)
Change in fractal dimension with a change in primary particle size for charged silica particles	Light Scattering (3D)	2.26 → 2.64	0.38	Decrease in df as the primary particle size changed		[15] Kim and Berg (2000)
Effect of polymer dose on drinking water sludge flocs	Settling (3D)	1.06 → 1.77	0.71	df increased with the addition of polymer, however further increases in polymer did not significantly change the df		[33] Zhao (2004)
Effect of dual polymer dosing on alumina flocs	Settling (3D) Light scattering (3D)	Settling: 1.65 → 1.77 Light: 1.85 → 2.12	0.12 0.27	df increased with the addition of dual polymer	Floc properties significantly improved through the addition of dual cationic/anionic polymer	[34] Glover *et al.* (2000)

Without question the most successful application of fractal dimension has been in improving the understanding of agglomeration from a mechanistic point of view and how this influences the structures that form. Early work related to the influence of attachment efficiency by comparing structured formed under diffusion limited aggregation (DLA) [no repulsive barrier] to those under reaction limited aggregation (RLA) [repulsive barrier]. Flocs formed in the DLA regime tend be more open structures than those under RLA as the barrier to agglomeration effectively permits restructuring through requiring multiple contact before a particle, or cluster, sticks to a growing floc. Experimental determinations usually involve coagulation of idealised particles such as latex, hematite or alumina with simple salts such as NaCl or $CaCl_2$. DLA flocs have d_f values of around 1.7-1.8 compared to 2.2-2.3 for RLA. Similarly, d_f value of 1.8 and 2.5 are reported for the difference between particle-cluster and cluster-cluster agglomeration from simulated agglomeration. In stirred systems, significant restructuring takes place such that d_f values of 2.2 are common, almost half way between the two extremes.

Chakraborti *et al* [5] investigated the coagulation of lake water with alum and measured the perimeter fractal dimension by image analysis. The fractal dimension was seen to decrease from 1.84 to 1.65 as the mechanism changed form charge neutralisation to sweep flocculation. The decreased fractal dimension in the sweep floc zone corresponded to a reduction in residual turbidity commonly observed for this type of system. However, the size of the flocs was also very different such that the observed differences were more likely to be related to size rather than structure. Indeed, similar work by Kim *et al* [15] showed the reverse effect

when using scattering techniques. Flocs formed under charge neutralisation had a fractal dimension of 1.84 compared to 2.2 during sweep flocculation.

In a more applied field, measurement of fractal dimension has been used as a indicator of floc character to enable comparison between different chemicals and operational regimes. Comparison of the flocs formed during the coagulation of natural organic matter has shown an ability to control structure, to a degree, by appropriate choice of chemical. Alum was seen to produce the most compact flocs even compared with polymers with a d_f, based on settling, of 2.2 compared to 1.7 and 1.99 for iron and polyDADMAC respectively [16]. For any given choice of coagulant the structure then becomes a function of the dose ratio used. For instance, both Ho and Newcombe [17] and Jarvis *et al* [18] have shown that flocs become more open as the dose ratio increases. The observations were commensurate with an overall decrease in the floc size and an increase in floc strength. A traditional method of altering floc character is through the use of polymers. For instance, addition of 2 $mg.L^{-1}$ of an anionic polymer to alum sludge increased the fractal dimension, measured by settling, from 1.06 to 1.71. However, addition of more polymer made no difference to the structure beyond the dose resulting in a d_f or 1.77, even with 20 $mg.L^{-1}$ of polymer. Similarly, comparison of alumina aggregates formed with either salts or polymers showed more compact structures formed in the case of the latter identified by a fractal dimension of 2.12 (scattering) compared to 1.75 with the salt.

An area of particular growth has been in the measurement of fractal dimension for wastewater applications. In part this is due to the fact that scattering techniques are appropriate due to the lower refractive index of the particles. However, it has also become more important due to recycling of wastewater and the need for better understanding of flocs to control the release of colloids into reuse sources. Recent comparisons of different wastewaters gave d_f values ranging from 1.96-2.49. Flocs with d_f values > 2.4 were observed to have strength values more than double those with a $d_f < 2$. The potential impact due to changing structure has been related to phosphate uptake and dewatering. For instance, Smoczynski and Wardzynska [19] have shown P removal to drop from 94 to 87 % as the d_f increase from 1.67 to 1.9, based on settling measurements.

One commonly held expectation is that as flocs become more open they contain more bound water and so are more difficult to dewater. Experimental evidence with incompressible solids has shown a relationship between d_f, as measured by image analysis of the filter cake, and specific cake resistance and permeability. However, the observed changes in d_f were small (2.135-2.142) and these corresponded to changes in other operating variables such as the applied pressure [6]. Similarly, some recent work has demonstrated the bound water content of synthetic particle coagulated with alum and ferric to increase from 0.2 to 0.4 as the d_f, as measured by settling, increased from 2.45 to 2.6. The more compact alum flocs settled faster and lead to sludges containing about 20 % more bound water than those formed with ferric [20]. However, alternative evidence suggests changes in dewatering performance are due to variations in floc size for water sludges [21] or physical properties such as charge, hydrophobicity and viscosity for sewage sludges, rather than compactness of the flocs [22].

A growing area of application related to this is in understanding membrane fouling, especially of porous dead end systems where cake layer fouling tends to control the operation of the plant. For instance, Lee *et al* [23] reported micro filtration of hematite-NOM flocs and showed that the fractal dimension was more important for small (~10 μm) flocs than large (>40 μm). For small flocs the cake resistance doubled as the d_f value increased from 1.9 to 2.1, as measured by scattering. In the case of larger flocs no change in cake resistance was observed with changing d_f. During the experiments the d_f was seen to alter as a function of the amount of organic material in the floc and the method by which it was added. Similar trials

with submerged membrane systems have shown that flocs with a lower fractal dimension have lower specific resistance and higher compressibility index leading to a higher permeability, or lower fouling rate [24].

The use of fractal dimension is now widespread due to its relatively easy measurement and promise of potential insight but questions remain as to exactly what role it plays. It is clear that it is very useful in providing new insights into agglomeration, especially in regard to idealised particles and simulations but its extension to particles encountered in real waters is less convincing. Perhaps its greatest weakness/flaw is in the paucity of a distinct link to the operation of downstream processes. A significant problem is that the change in fractal dimension often occurs commensurately with changes in other variables and a clear picture of the potential influence of fractal dimension remains unclear. Critics of the use of fractal dimension often exploit this lack of clear evidence to its practical significance to label it as an academic concept rather than a practical tool. Whilst sympathy must be given to this viewpoint at the current time sufficient evidence does exist to show its use as at least a diagnostic tool. Perhaps therefore, it is more appropriate to consider the discovery of its true potential as one of the challenges for the next generation of colloid scientists.

REFERENCES

[1] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco, 1982.
[2] M. Bellouti, M.M. Alves, J. M. Novais and M. Mota, Water Res., 31 (1997) 1227.
[3] C.P. Cousins and J. Ganczarczyk, Water Qual. Res. J. Can., 33, 4, (1998) 565.
[4] R.K. Chakraborti, K.H. Gardner, J.F. Atkinson and J.E. van Benschoten, Water Res., 37 (2003) 873.
[5] R.K. Chakraborti, J.F. Atkinson and J.E van Benschoten, Environ. Sci. Technol., 34 (2000) 3969.
[6] S.T.H. Brock, Fractal Dimensions and their Relation to Filtration Characteristics. PhD thesis, Loughborough University, 2000.
[7] C. Lee and T.A.. Kramer, Adv. Colloid Interface Sci., 112 1-3 (2004) 49.
[8] C.M. Sorensen, Aerosol Sci. Technol., 35 (2001) 648.
[9] G.C. Bushell, Y.D. Yan, D. Woodfield, J. Raper and R. Amal, Adv. Colloid Interface Sci., 95 (2002) 1.
[10] J. Gregory, Filtration Sep., 35 (1998) 367.
[11] R.M. Wu, D.J. Lee, T.D. Waite and J. Guan, J. Colloid Interfacial Sci., 252 (2002) 383.
[12] T.H. Ha, U.Yoon, K.J. Lee, Y.W. Shin, J-M Lee, I.Y. Kim, K.S. Ha, S.I Kim and J.S. Kwon, Neurosci. Lett., 384 1-2 (2005) 172.
[13] L.J. Hadjileontiadis, E. Douka and A. Trochidis, Mechanical Syst. Signal Proc., 19 3 (2005) 659.
[14] C.T. Brown and W.R..T. Witschey, J. Archaeological Sci., 30 (2003) 1619.
[15] A.Y. Kim, and J.C. Berg, J. Colloid Interface Sci., 229 (2000) 607.
[16] B. Jefferson, P. Jarvis, and S.A. Parsons Proc. 11th International Gothenburg Symposium on Chemical Treatment of Water and Wastewater, Orlando, Florida, 2004.
[17] L. Ho and G. Newcombe, Water Res., 39 (2005) 3668.
[18] P. Jarvis, B. Jefferson and S.A. Parsons, Environ. Sci. Technol., 39 (2005) 8919.
[19] L. Smoczynski and R. Wardzynska, J. Colloid Interface Sci., 183 (1996) 309.
[20] C. Turchiuli and C. Fargues, Chem. Eng. J., 103 (2004) 123.
[21] C.C. Wu, C. Huang and D.J. Lee, Colloids Surf. A, 122 (1997) 89.
[22] B. Jin, B-M.Wilen and P. Lant, Chem. Eng. J., 98 (2004) 115.
[23] S.A. Lee, A.G. Fane, R. Amal and T.D. Waite, Sep. Sci. Technol. 38 4 (2003) 869.
[24] M-H. Cho, C-H. Lee and S. Lee, Water Sci. Technol., 51 (2005) 143.
[25] B-M Wilen, B. Jin and P. Lant, Water Res., 37 (2003) 3632.

[26] S-H, Kim, B-H. Moon and H-I. Lee, Microchem. J., 68 2-3 (2001) 197.
[27] V. Chaignon, B.S. Lartiges, A. El Samrani and C. Mustin, Water Res., 36 (2002) 676.
[28] P. Jarvis, B. Jefferson and S.A. Parsons, Environ. Sci. Technol., 39 (2005) 2307.
[29] S.J. Jung, R. Amal and J. A. Raper, Powder Technol., 88 1 (1996) 51.
[30] S. Tang, Colloids Surf. A, 157 (1999) 185.
[31] L. Smoczynski and R. Wardzynska, J. Colloid Interface Sci., 183 (1996) 309.
[32] T.D. Waite, Colloids Surf. A, 151 1-2 (1999) 27.
[33] Y.Q. Zhao, Separation and Purification Technology, 35 3 (2004) 175.
[34] S. M.Glover, Y-D. Yan, G. J. Jameson and S. Biggs, Chem. Eng. J., 80 1-3 (2000) 3.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 5: Coagulation and flocculation with organic polyelectrolytes

B. A. Bolto

CSIRO Manufacturing & Infrastructure Technology, PO Box 56, Highett, Victoria 3190, Australia

1. INTRODUCTION

The main applications of organic polyelectrolytes in potable water production are in coagulation and flocculation, and in the dewatering of treatment plant sludges. The water production processes are usually followed by sedimentation and filtration, although with only slightly contaminated waters the sedimentation step may be omitted. Flotation is an option instead of sedimentation, especially for algae laden waters. The sludges obtained from the various separation processes have very high water contents and must be further concentrated to minimise transportation costs; polymers have a role in this sludge conditioning.

The impurities present in the source water can be in the form of suspended material, such as clay, silica, microbial cells or algae, and as dissolved and colloidal natural organic matter, and as dissolved salts. Raw water processing normally involves physicochemical procedures, based on coagulation and flocculation of suspended solids and colloids, and the adsorption of soluble material on solid substrates such as metal hydroxide flocs. The focus in this Chapter is on the use of soluble polymers in coagulation and flocculation processes.

Polymers have been utilized in coagulation/flocculation processes for water purification for at least four decades [1]. In comparison with alum, some of the advantages flowing from the use of polymers in water treatment are:

- lower coagulant dose requirements
- a smaller volume of sludge
- a reduction in the ionic load of the treated water
- avoidance of the presence of aluminium ions
- cost savings of up to 25-30% [2, 3].

Polymers are especially beneficial in coping with the problems of slow settling flocs in low-temperature coagulation or in treating soft coloured waters, where they improve settleability and increase the toughness of flocs [4]. The capacity of a treatment facility may be more than doubled with the formation of larger and stronger flocs, the rate of solid and water phase separation can be significantly increased, and the dosage of other chemicals lowered. Also, the range of waters that can be treated is wider.

There are disadvantages of course, with higher costs in particular situations and environmental factors being the main concern. There is a greater sensitivity to incorrect dosage, with turbidity and natural organics removal less efficient in some instances [3]. Degradation in the presence of chlorine or other disinfectants covered in a recent report [5] is reviewed here.

With a few notable exceptions [6] there is not a great deal of published information on the relationship between polymer structure and treatment performance in drinking water production; that is, on the influence of molecular structure on coagulation/flocculation, on the rates of both precipitation and sedimentation, on product water quality and on the solids content of the final sludge.

The types of impurities present in poor quality water supplies are first briefly outlined.

2. NATURAL IMPURITIES IN WATER

Impurities may be dissolved compounds as well as insoluble particles, and may be of organic or inorganic origin [7]. Some of the more commonly found natural components containing organic material are, in decreasing size order, zooplankton, phytoplankton, bacteria, viruses, clay-humic acid complexes, humic acids, proteins, polysaccharides, fulvic acids, and very small species such as fatty acids, carbohydrates, amino acids and hydrocarbons. They are formed by the biological degradation of organic life substances [8], and include highly coloured compounds. Inorganic salts of natural origin are also present to some degree.

2.1. Dissolved organic compounds

Dissolved compounds, defined as that which will pass through a membrane having pores of 0.45 μm size when measured as dissolved organic carbon (DOC), have levels in the range 0.1-115 mg L^{-1}, with 5.75 mg L^{-1} being reported as a global average for streams [9]. DOC poses a problem for the water treatment industry for a number of reasons. Apart from the aesthetic problems of colour, taste and odour, its presence poses a health hazard because of the formation of potentially carcinogenic compounds when the water is disinfected with chlorine or chloramine - the problem of disinfection by-products (DBPs). As well, DOC exacerbates the deterioration of the microbiological water quality in distribution systems, fouls membranes and ion-exchange resins, interferes with the oxidation of dissolved iron and manganese to insoluble easily removed forms, and encourages corrosion [10].

Although humic substances, encompassing fulvic and humic acids, constitute about 50% of the total DOC of a typical surface water, the proportion can be as high as 80% [11]. Humic substances are unique and troublesome materials in that they have quite variable properties, in terms of acidity (pK_a 3-5), molecular weight (several hundred to ten thousand) and molecular structure (mostly phenolic and carboxylic acid functionalities, but also alcohol, quinone, ether, ester and ketone groups). They behave as negatively charged colloids or anionic polyelectrolytes at natural pH levels and have surface-active properties, so that they can interact with particulate material and also, via their hydrophobic aromatic and aliphatic regions, with non-polar pollutants such as pesticides and polychlorinated biphenyls. Humic substances are often present as stable complexes with metal ions. The subtleties of their structure have been further detailed recently [12]. These variable properties influence reactivity, which as mentioned changes spatially and temporally. If the smaller charged organic molecules are first removed from raw water by ion exchange, as carried out in one full-scale plant [13], a subsequent alum clarification stage is greatly facilitated: larger flocs are formed that settle three times more rapidly, far less organics are left in the product water, and only 25% of the original alum dose is required in a conventional clarification process [14].

The composition and effects of DOC on water treatment are covered more fully in the NOM reference chapter later in this book.

2.2. Insoluble impurities

Suspended particulate matter and minute particles that cause turbidity are an important component of all natural waters and this is discussed in detail in a reference chapter. The charge on the particles is controlled by an adsorbed layer of natural organic matter, as well as by the salinity and the concentration of divalent cations in the water [15]. Humic substances can adsorb onto the particles via surface metal cations. The surface charge potential of the particles is an important parameter influencing coagulation and adsorption behaviour. It can be monitored via particle microelectrophoresis, and in natural systems is invariably negative, irrespective of the nature of the primary particle [16]. The coating of organics has a strong impact on the amount of coagulant required and the rate of coagulation, slowing the rate markedly at low salinities, but having less of an effect as the salinity increases [17].

3. COAGULATION AND FLOCCULATION MECHANISMS

The main challenge in removing turbidity and organics from water supplies is to cope with impurities that are negatively charged at natural pH levels, and have formed a stabilised dispersion. Initially it is a matter of destabilising the dispersion and coagulating the contaminants. This is done by adding positively charged species in appropriate quantities to neutralise the charge on the impurities. A flocculation step involving a polymer is then used to bring together the particles so that larger flocs are produced. In the terminology used here, the terms coagulation and flocculation are not synonymous. Coagulation is used in the sense of describing the process whereby the original dispersion is destabilised by overcoming the forces which maintain the stability, and flocculation to describe the process whereby the destabilised particles join together to form larger agglomerates [18]. As well as the addition of an inorganic salt followed by a polymer to achieve these effects, the two processes can be achieved by the addition of only a single polymer.

3.1. Coagulation by charge neutralisation with metal salts

Coagulation with hydrolysable metal salts is a long-standing technology, the mechanism of which has been thoroughly investigated [19, 20]. Charge neutralisation is brought about by the addition of aluminium or iron salts to form soluble hydrolysed cations and metal hydroxide flocs, the relative amounts formed being dependant on the operating pH.

Originally utilised as alum in the case of aluminium, the coagulant can be applied also as the polynuclear cation via preformed poly(aluminium chloride) or the corresponding sulfate. The salt residual is then reduced and it is found that the usage of coagulant can be up to 30% lower [21]. The large inorganic species so formed is effective in neutralising the negative charge on suspended clay particles and natural organic material [22]. The polymeric forms are better for cold waters and can be employed over a wider pH range than alum, but at greater cost.

3.2. Flocculation via polymer bridging

The small flocs produced by coagulation with metal salts can be built up into larger agglomerates by subsequent treatment with a polymer, to form larger particles that have more rapid rates of sedimentation. This involves polymer bridging, in which polyelectrolyte bound to a floc particle has looped and dangling chains that can attach to further particles nearby. Long chain polymers that do not have a high level of charge give best results. One polymer chain adsorbs on two or more particles via an electrostatic or non-electrostatic driving force [23, 7]. It takes place when the surface is not completely covered, as in the initial stages of the

process or when the polymer dose is low. It occurs immediately after addition of the polymer, and is very dependent on mixing conditions [24].

The nature of the surface charge on flocs formed by coagulation with a metal salt will depend on the dose of metal salt used. The net charge is normally close to zero, either slightly positive or slightly negative. In the first case an anionic polymer will increase the size of the flocs by interaction with the positive sites on the floc surface [25]. Bridging will occur when the adsorbed chains interact with another floc.

Polymer characteristics normally favouring bridging are a low charge density (CD) and a high molecular weight (MW), with a minimum of 800,000 Daltons advocated. The CD maximum generally ranges from 5-15%, although levels as high as 45% have been effective in some instances [26, 27]. Higher CD polymers have improved shear resistance, the transition occurring at a CD between 10 and 20%. Re-flocculation is incomplete at low CDs, but more or less complete at a CD of >12%. Bridging after shearing is also very dependent on chain length, with higher MW polymers giving better performance.

3.3. Coagulation by charge neutralisation with cationic polymers

Organic polymers may be used as primary coagulants as well as in the more traditional flocculation step described in Chapters 2 and 3. The polymer acts as a destabilising agent via a charge neutralisation/precipitation mechanism, as well as an agent for floc growth. At the optimum polymer dosage for adsorption of a polyelectrolyte on an oppositely charged surface, the surface charge is nearly eliminated so that only weak electrostatic repulsion remains. The best turbidity and colour removal occurs with particles of net zero charge, which is when the mobility or zeta potential of the particles is close to zero [28].

Charge neutralisation takes place at low and high surface coverage, depending on the charges on the polymer and on the surface. Here the polymer has a passive role, merely acting as a multivalent ion, so the process is strictly coagulation [24]. Polymer characteristics that favour charge neutralisation are substantial doses of a high CD, low MW polymer. Thus in jar tests on synthetic and natural waters of low turbidity and moderate to high colour, high CD cationic polymers are effective in removing the organics responsible for colour, and in lowering the production of trihalomethanes after chlorination [29, 30].

When the polymer dose is low, a mechanism known as the electrostatic charge patch model has been proposed [23]. The polymer adsorbs where there are domains of opposite charge on the floc to that of the polymer, leading to a local excess of charge associated with the polymer, but leaving patches of the opposite charge on the floc [7]. Direct electrostatic attraction between the patches on different particles results via domains of opposite charge, leading to flocculation. Thus when a cationic polymer adsorbs on a negatively charged surface, patches of positive surface are formed which may attach to negative patches on another particle, with no bridging chains being involved.

Patch formation, like bridging, takes place when the surface is not completely covered, as in the initial stages of adsorption or with a low polymer dose. As it takes place immediately after addition of the polymer, it is very dependent on mixing conditions. It is a flocculation phenomenon, as the polymer has an active role in creating patches that link particles together. A characteristic of this model is that flocs may reform after floc breakage. The patch charge model is favoured by low doses of a high CD, high MW polymer. A low ionic strength is preferable. Thus a patch charge mechanism is proposed when highly charged polymers adsorb strongly onto silicate mineral particles and silica, resulting in a low surface coverage [31].

Fundamental colloid science studies of the flocculation of particles in very clean systems by well defined polyelectrolytes have been carried out over the years, covering for

example latices, cellulose, silica, iron oxide and clays, as reviewed earlier [7]. Generally there is a balance between two mechanisms: charge neutralisation with highly charged polymers, and inter-particle bridging when the polymer has a low content of ionic groups [32]. There are particular issues of great practical importance, such as effects on reaction rates [33] and the size and strength of flocs [34], which are found to be dependent on the CD of the polymer, its MW and the ionic strength.

4. POLYMER TYPES AVAILABLE

Many reviews of polymeric flocculants have been published over the years, covering synthesis, applications and the range of commercial organic flocculants available [35-38]. Polyelectrolytes are best broadly classified according to the nature of the charge they possess; that is, whether they are cationic, anionic or non-ionic. Other crucial parameters are CD and MW. The CD is generally expressed as a percentage of ionic groups (both those that are charged irrespective of pH and those that can become charged under certain pH conditions) relative to all the groups in the polymer. The terminology high, medium and low will be used, as outlined in Table 1. Further details on CD will be presented later.

Table 1
Classification of polymeric flocculants by molecular size and charge

Amount	MW (Dalton)	CD (mol %)
High	ca. 10^7	50-100
Medium	10^5-10^6	ca. 25
Low	10^4-10^5	ca. 10

4.1. Cationic polyelectrolytes

The many varieties of cationic polymers available have been reviewed in detail [7]. The molecular structures of the ones most commonly used are shown in Fig. 1. Generally they possess quaternary ammonium groups that have a formal positive charge present irrespective of the pH level, giving rise to the name strong electrolyte polymers. Weak electrolyte versions that acquire cationic properties in acidic media are available. Some natural products or their derivatives, such as chitosan, are also in use.

4.1.1. Poly(diallyldimethyl ammonium chloride)

Polymerisation of diallyldimethylammonium chloride produces poly(diallyldimethyl ammonium chloride) or polyDADMAC, a water soluble polymer. The polymer is of low to medium MW and contains five-membered pyrrolidinium units as established long ago [39], although some commercial literature still erroneously suggests that six-membered rings prevail. It should be noted that for the cationic polymers listed in Fig. 1 the counterion has been omitted. PolyDADMAC is much used in the water industry. Higher MW copolymers have been made with acrylamide [40]. The random copolymer has practical application in sludge conditioning.

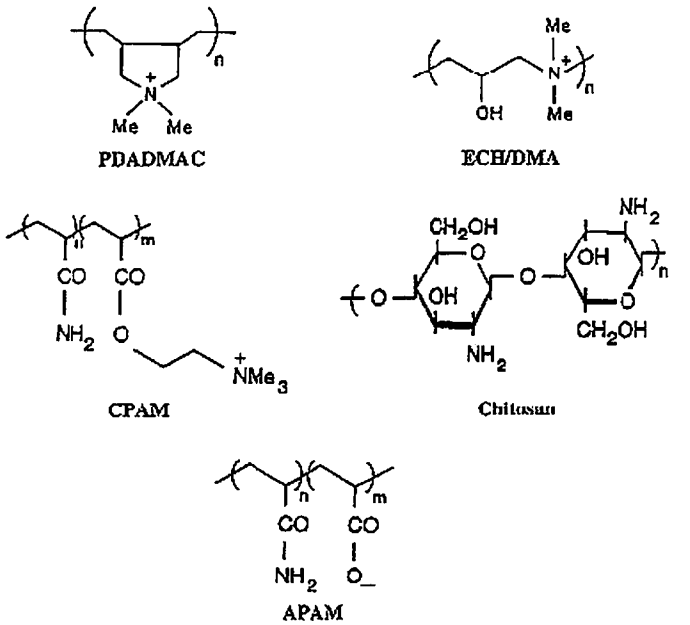


Fig. 1. Structures of the cationic polyelectrolytes polydiallyldimethylammonium chloride (PDADMAC), polymers from epichlorohydrin and dimethylamine (ECH/DMA), cationic polyacrylamide (CPAM), chitosan and anionic polyacrylamide (APAM)

4.1.2. Epichlorohydrin/dimethylamine polymers

Epichlorohydrin forms polymers with ammonia and primary and secondary amines, the preparation of which has been thoroughly reviewed [41]. The reaction of epichlorohydrin with a secondary amine such as dimethylamine produces a low MW linear polymer, denoted here by ECH/DMA, in which all the nitrogens are in the quaternary ammonium form, making the commonly used generic term of 'polyamines' a misnomer.

4.1.3. Cationic polyacrylamides

Random copolymers of acrylamide and the cationic ester acryloyloxyethyltrimethyl-ammonium chloride, formed by quaternisation of dimethylaminoethyl acrylate with methyl chloride, are extensively used in the water industry. An extensive survey of this category of polymers has been made [35, 42]. The methacrylate analogue is also commercially available. The cationic content in the cationic polyacrylamides or CPAMs can be 10-100%.

In the commonly available polymers, hydrolysis of the ester groups and consequent loss of cationic charge has been investigated and found to be CD and pH dependent, with hydrolysis increasing under more alkaline conditions:

$$\text{Pol-COOCH}_2\text{CH}_2\text{NMe}_3^+ + \text{OH}^- \rightarrow \text{Pol-COO}^- + \text{HO-CH}_2\text{CH}_2\text{NMe}_3^+$$

Earlier studies [43] suggested that some degradation occurred even at pH 6 for polymers having a CD of 24%, with a half life of 24 h at pH 7 and 0.25 h at pH 8.5. The polymer was stable in a pH 4 environment. There is no evidence of hydrolysis of acrylamide units up to pH

8.5. More recent work on ester hydrolysis on less ionic polymers having a CD of 6% gives a half life of 22 months, but the pH level is not specified [44]. For a CD of 30% the process is also very slow, especially for pure solutions, whereas salt or pH levels above 8 causes an ease of degradation that is not observed for the polymer of 100% charge [45]. Apart from loss of cationic sites, there is a change in the chain conformation on hydrolysis because of the anionic carboxylates formed, which reduces the chain extension and make the polymer less efficient as a flocculant. In the case of the homopolymers, the corresponding methacrylate is less vulnerable to hydrolysis [46].

4.1.4. Natural cationic polymers

There are several naturally occurring polymers that have inherent cationic properties or can be modified to yield a cationic polyelectrolyte. The most prominent of these is chitosan, a partially deacetylated chitin which can be considered as a 1:4 random copolymer of N-acetyl-α-D-glucosamine and α-D-glucosamine. The commercial product is of medium MW and has a CD that is pH dependent and can be as high as 80% [47]. The use of chitosan in water purification has been extensively reviewed [48].

4.1.5. Charge densities of cationic polyelectrolytes

Table 2 outlines the CD of various cationic polymers in mol% and meq/g of polymer, the latter being a theoretical value calculated from the molecular formulae. As there are other components present in the polymers such as co-monomers to increase chain length or to create chain branching, as well as initiator fragments, this is only an approximation that is useful for comparison purposes.

Table 2
Charge densities of commonly used cationic polyelectrolytes

Polymer	Molecular formula	CD, mol%	CD, meq g^{-1}
PDADMAC	$C_8H_{16}N\ Cl$	100	6.2
ECH/DMA	$C_5H_{12}ON\ Cl$	100	7.3
CPAM	$C_8H_{16}O_2N\ Cl$	100	5.2
CPAM	$(C_8H_{16}O_2N\ Cl)_{0.5}\ (C_3H_5ON)_{0.5}$	50	3.8
CPAM	$(C_8H_{16}O_2N\ Cl)_{0.25}\ (C_3H_5ON)_{0.75}$	25	2.5
CPAM	$(C_8H_{16}O_2N\ Cl)_{0.1}\ (C_3H_5ON)_{0.9}$	10	1.2
Chitosan	$C_6H_6O_4\ HCl$	100*	5.2

4.2. Anionic polyelectrolytes

By far the most common anionic polymers are those containing weakly acidic carboxylic acid groups, where the CD will depend on pH.

4.2.1. Anionic polyacrylamides

High MW carboxylic acid polymers based on polyacrylamide are extensively employed as flocculating agents in the water and other process industries, where a low CD is the general rule. Copolymers of the structure included in Fig.1 can be prepared either by copolymerisation of acrylamide and acrylic acid or its salts, or by polymerisation of acrylamide followed by partial hydrolysis [41]. The former route gives a roughly random copolymer, whereas some clustering of anionic groups can occur in the alkaline hydrolysis approach. The CD can be determined by potentiometric titration of the copolymers. Anionic polyacrylamides, or

APAMs, containing varying proportions of acrylamide co-monomer are listed in Table 3 as regards CD, calculated on a theoretical basis in meq g^{-1} of polymer.

Table 3
Charge densities of commonly used anionic polyacrylamides

Molecular formula	CD, mol%	CD, meq g^{-1}
$C_3H_3O_2$ Na	100	10.6
$(C_3H_3O_2 Na)_{0.75}$ $(C_3H_5ON)_{0.25}$	75	8.5
$(C_3H_3O_2 Na)_{0.50}$ $(C_3H_5ON)_{0.50}$	50	6.1
$(C_3H_3O_2 Na)_{0.25}$ $(C_3H_5ON)_{0.75}$	25	3.3
$(C_3H_3O_2 Na)_{0.1}$ $(C_3H_5ON)_{0.9}$	10	1.4

4.3. Non-ionic polymers

The use of synthetic polymers like polyacrylamide in the water industry dates back some four decades, but some polymers of natural origin have been used for centuries.

4.3.1. Polyacrylamide

Synthetic polymers that are often quoted as non-ionic actually contain, as in the case of polyacrylamide, some 1-3% of anionic groups arising from the hydrolysis of amide groups under the preparative conditions [37]. Polyacrylamide (PAM) having less than 1% hydrolysis can be prepared by careful attention to the monomer concentration, pH, temperature and initiator, with a redox type preferred [41]. ^{13}C nuclear magnetic resonance spectroscopy has been found to be the most reliable method for determining the degree of hydrolysis of concentrated polyacrylamide samples [49].

4.3.2. Natural polymers

Tannins have received attention [50]. The purification of drinking water with macerated seeds from the horseradish tree *Moringa oleifera* is also of interest [51].

5. APPLICATIONS IN DRINKING WATER TREATMENT

The focus here is on coagulation and flocculation. With slightly contaminated waters the usual sedimentation step after these stages can be omitted, as in direct filtration. Flotation sometimes replaces sedimentation, notably when there is a high level of algae. Both separation processes yield sludges of very high water contents, so that sludge conditioning with polymers and subsequent dewatering are necessary to minimise transportation costs.

Waters of low turbidity and moderate to high colour are ideal candidates for coagulation with a cationic polyelectrolyte. Avoiding the use of an inorganic coagulant has the advantages that a smaller amount of sludge is produced, the system is less pH sensitive, there is less dissolved salt added [52, 53] and there is no aluminium residual. On the other hand, there can be concerns about residual polymer and monomer in the product water, and these are discussed later. The level of residual polymer is minimised when the turbidity is high [54].

5.1. Primary coagulation in drinking water treatment

5.1.1. Conventional sedimentation and filtration

In the production of drinking water a cationic polyelectrolyte of high CD and a low to medium MW such as polyDADMAC can be used instead of a metal salt as the primary coagulant. ECH/DMA polymers are also effective. The CD rather than the MW is important in selecting the optimal conditions, and the initial periods of both rapid and slow mixing are critical in the formation of flocs [6]. There are fewer examples of polymer-only coagulation in conventional coagulation/sedimentation/filtration plants compared with polymer use in conjunction with a metal salt, because of the high polymer dose that would be required. Thus a combination of 7 mg L^{-1} of alum, 0.26 mg L^{-1} of the high CD cationic polymer ECH/DMA and 0.01 mg L^{-1} of neutral polyacrylamide as filter aid successfully treated water of turbidity 21-28 nephelometric turbidity units (NTU) [55]. A number of workers have discussed the turbidity limits when alum is the coagulant, and this has been reviewed recently in a consideration of highly turbid waters, covering a range of 5-200 NTU [56]. It was noted that the alum dose required for effective coagulation and filtration could be the limiting factor, with 12-15 mg L^{-1} as Al suggested as the upper limit. Another example of the use of ECH/DMA polymers of low MW (10,000 and 50,000 Daltons) as primary coagulants showed that the longer chained polymer was more effective for raw waters of higher turbidity [57]. Pilot work indicated a 50-80% saving of poly(aluminium chloride) using 0.4-1 mg L^{-1} of polymer for waters of average and high turbidity. The removal efficiency for organics was also enhanced.

The capacity of a filter plant that has been operated with inorganic coagulant may be increased by polyelectrolyte addition. Thus an increased flow of 25% was achieved by adding 0.3 mg L^{-1} of cationic polymer after ferric chloride dosing, the level of which was lowered by more than a third [58]. This was accompanied by a filter cycle time increase of 67%, and an improvement in the product water quality. The cost of the polymer was completely recovered by the decreased outlay on ferric chloride.

In laboratory studies it has been demonstrated that when particles are present, as would normally occur in natural waters, the performance of quaternary ammonium polymers of high CD and MW is improved [59]. By adding kaolinite when polyDADMAC was used there was a small gain, with 16% better removal of colour in one case, but there was little change in the removal of natural organic matter (NOM), as measured by UV absorbance. The commonly used mixes of alum and polymer are a convenient way to have suitably reactive particles present; for one of the waters studied a 67% reduction in the alum dose was possible by adding 1 mg L^{-1} of polyDADMAC, to give even better performance than alum alone at the optimum dose.

In similar studies of cationic polymer as the sole coagulant for the removal of NOM, a highly charged polyDADMAC of the highest possible MW was found to be the most effective of commercially available polymers, with the effectiveness diminishing for polyDADMACs of lower MW [60]. A cationic polyacrylamide of high charge also performed well, taking out double the amount of NOM compared with a low CD polymer of the same MW. Chitosan gave reasonable results, despite its low CD and MW, suggesting that a different mechanism prevails for this type of polymer. Generally, organic polymers did nearly as well as alum for the waters investigated, taking out 86-100% of the colour that alum does. The uptake of NOM after fractionation into four fractions based on hydrophobic and hydrophilic properties was also determined [61]. Alum was best for removal of humic and fulvic acids as measured by UV absorbance, although a cationic polymethacrylate or polyDADMAC could perform well.

Of the high CD polymers, those with the least polar structure were the most effective in removing the hydrophobic fractions. Neutral organic compounds were a very minor component and there was little effect (where data could be obtained). For proteins and other charged hydrophilic compounds, alum then polyDADMAC were the best performers for one source.

5.1.2. Direct filtration

The settling stage normally employed in conventional treatment plants is kinetically inefficient. In direct filtration, where some time is allowed for floc growth to achieve optimum size, there is no such sedimentation step, which makes for lower capital costs. Cationic polyelectrolytes have a distinct advantage over the metal salts that are commonly used in this procedure, because they effect charge neutralisation without the formation of additional solids in the form of a metal hydroxide precipitate. The flocs from inorganic coagulants take up much of the space in the filter media, rapidly causing a pressure drop [62]. Polymers give the advantages of a deformable floc, extended filter runs, and decreased sludge volume from filter backwashing. Polymer MW does not seem to be a major variable. The initial mixing intensity and the mixing time are the most important parameters in determining the particle size of the floc formed prior to filtration [63].

For waters high in organics a substantial polymer dose is needed, but compared with alum usage the filter run can be more than double in length [28]. If it is found that direct filtration of waters containing large amounts of humic substances with cationic polymers as the sole flocculant is not economic because of the high dose required, conventional metal ion coagulants can be used with a lower dose than is normal, to minimise the amount of precipitate. This is followed by a high CD cationic polymer such as polyDADMAC, used in charge neutralisation mode rather than causing particle bridging to give large flocs, although the floc size is still increased [64]. Such a mode of operation is in extensive use.

Cationic polymers can coagulate freshwater algae at doses of 1-10 mg L^{-1} [65]. However, the high salinity of marine systems inhibits the process. Direct filtration experiments have demonstrated that the dose of cationic polymer must be more than that necessary to overcome the effects of exocellular matter released by the algae, when good cell removal is possible [66]. A high CD polymer is preferred. Employing an inorganic coagulant for this purpose can disrupt the cell membrane and release compounds that give rise to taste and odour in the product water [67]. With blue green algae toxins are produced that can be a health hazard. They can be removed by oxidation or by adsorption onto activated carbon.

5.1.3. Dissolved air flotation

In dissolved air flotation a proportion of the treated water is recycled through a pressurised air saturation system, and the air-saturated water then released into the water to be treated [68]. Suitable polymers can redress any floc shearing brought about by overly vigorous contact of air bubbles with the particles [69].

5.2. Polymers as coagulant aids

A major use of organic polymers in water treatment is as a coagulant aid to bridge the coagulated particles formed when an aluminium or iron salt has been used as the primary coagulant. An appropriate polyelectrolyte can increase floc size; a strong dense floc of regular shape is preferred. The large aggregates that form then settle more rapidly. The particles produced by the inorganic salt generally have a slightly positive or slightly negative surface charge, depending on coagulation conditions and the dose of metal salt. With flocs of slightly

positive character an anionic polyacrylamide of low or medium CD and high MW is appropriate. The use of polymers in this way results in a substantial lowering of the alum dose required, a 40-60% reduction being possible [33, 68]. Thus for a water containing 5 mg L^{-1} of humic acid, a dose of 75 mg L^{-1} of alum will remove only 20% of the humics, but adding an anionic polymer in conjunction with 10 mg L^{-1} of alum will give a 95% reduction [25]. An excess of polymer can cause re-dispersion of the impurities. When the coagulated solids have a slightly negative charge, a cationic polymer of low CD and high MW, such as a cationic polyacrylamide, is effective in forming larger flocs [28].

Processes have been optimised for organics removal in enhanced coagulation systems so that the production of harmful disinfection by products is minimised [70]. These systems are also used for the elimination of particles to the greatest possible extent to ensure the absence of pathogenic organisms like *Giardia* and *Cryptosporidium* in the product water [71]. This requires the best combination of inorganic salts as coagulant and polymer as flocculant, the best type of polymer, the optimum concentration ratio, and optimum process conditions [72].

5.3. Recycling of filter backwash waters

In arid areas conservation of water resources can be achieved by recycling spent filter backwash waters, and this is an economic necessity in other locations also. There is a concern that such recycling can compromise product water quality because contaminants can be concentrated to a level beyond the plant's multi-barrier treatment capability, with *Cryptosporidium* being a particular worry. A major survey of the situation and a pilot plant study of treatment options prior to recycling have been completed recently [73, 74]. *Cryptosporidium* has been detected in backwash waters at a range of levels, one result being as high as 9-33 million per 100 litres [75]. In a survey of 34 water treatment plants it was found that *Cryptosporidium* levels were up to 61 times and *Giardia* levels 16 times higher in the backwash water than in the original raw water [76]. In the UK it has been recorded that a backwash water contained *Cryptosporidium* at 1M/100L and supernatant water decanted from a settling tank 100K/100L [77]. The most recent study puts the significantly higher levels of protozoa in backwash water than in raw water at 21 times for *Cryptosporidium* and 16 times for *Giardia* [74]. Similar trends have been observed for DOC and other contaminants.

In a survey of the 335 water treatment plants in the US that recycle spent filter backwash water, the average generation of spent waters was found to be 2.5% [73]. Most of the plants (88%) use surface water as their source, with 83% of these recycling to the head of the plant and only 2% to just before the filters. Further treatment is provided by 65% of the plants before reuse of the backwash water, the type of treatment varying from site to site. Options include sedimentation with or without added coagulants, dissolved air flotation with and without polymer, oxidation, conventional filtration and membrane filtration. Pilot studies of various methods to establish appropriate strategies showed that 0.5 mg L^{-1} of a cationic polyacrylamide of very high MW and medium CD removed 99.6% of the turbidity after filtration, and a similar dose of an anionic polyacrylamide removed 99.4% [73, 74]. The polymers were selected as the best options after jar tests on two cationic, three anionic and three non-ionic polymers. Similar performance was obtained with 15-20 mg L^{-1} of ferric chloride, but treatment effectiveness was lost soon after a steady state was achieved. Ferric chloride and a CPAM in combination, however, gave stable long-term treatment with the same turbidity result as for CPAM alone. Ferric chloride was superior for DOC removal. Generally, treatment was much better when polymer was added in both sedimentation and dissolved air flotation, which gave equivalent performances. In other work on dissolved air flotation, treatment of a high degree in a very cost effective manner was achieved, a treated water

turbidity of 1 NTU being easily obtainable when the original backwash water turbidity was in excess of 50 NTU [78]. Chemical requirements were low, with no primary coagulants being added and only a single low dose (0.1-0.5 mg L^{-1}) of polymer being required to bind the floc particles and form agglomerates that were suitable for flotation. The optimum polymer type was site specific.

In a full-scale test there was a dramatic effect of adding polymer to a side-stream sedimentation plant. Polymer addition lowered the settled turbidity by 50%, with the addition of only 0.1 mg L^{-1} of polymer [74]. Capital costs for a range of processes have been estimated; sedimentation plus polymer was US$480,000-580,000 per million US gallons per day (mgd), but to return the oocyst level to less than the raw water concentration, dissolved air flotation plus polymer would be necessary. This would require capital of US$590,000-700,000 per mgd.

The recycled water may have some influence on the treatment process. Polymer residues at the μg L^{-1} level will see the raw water before addition of any coagulants of either the metal salt or organic polymer variety. In a turbid water some floc building will result, and this may affect the nature of the final flocs, making them lighter and more feather-like, so that settling rates may become slower and the final water content of the settled sludge may be increased.

5.4. Sludge thickening

Sludge properties are very dependent on the additives utilised, especially the amount of inorganic coagulant. The emerging technology of coagulant recovery will do much to reduce the volume of sludge for disposal [79, 80].

A degree of concentration is normally required to reduce sludge transport costs. The various methods available are summarised in Table 4 for water treatment plant sludge [68]. This application consumes more polyelectrolytes than any other in the water treatment area. Polymers are used to give large dense flocs that result in a more rapid settling of sludges and also clearer supernatants, which are recycled. In static settling the resulting strengthened flocs settle to a slightly larger volume, but in raked continuous thickeners they can be compressed without being broken, thus producing a much thicker material. Polymers can also improve dewatering characteristics during centrifugation and filtration. Typically, in centrifugation 1.5-3 kg of polymer is used per tonne of dry solids [81]. The high shear forces present during centrifugation make the use of polymers essential.

Table 4
Concentration of sludges from water treatment plants [68]

Conditioning method	Concentration in, % solids	Concentration out, % solids
Batch settlement	0.03-0.2	1-3
Continuous thickening		
without polymer dosing	0.03-0.2	2-3
with polymer dosing	0.03-0.2	5-10
Centrifuging	1-5	12-17
Filter pressing	1-10	20-25

For sludge treatment generally, polymers of low or medium CD and high MW such as cationic or anionic polyacrylamides give best performance, with the charge type depending on the origin of the sludge. Thus for a water treatment plant sludge where a high dose of alum and/or cationic polymer is employed, the sludge particles should generally have a positive surface charge, so an anionic polymer will be appropriate, although neutral polymers such as polyacrylamide are also employed.

Such an alum sludge in a typical gravity filter operation at a solids loading of 25 kg m^{-2} day^{-1} would have 1.5-2 % solids in the underflow without polymer use, and 3-4 % solids when an anionic polyacrylamide is present [52]. By way of contrast in sewage treatment, sludge from an activated sludge plant will have a negative surface charge, so that there a cationic polymer should be the better performer. The topic has received much attention recently [82].

The mixing or pumping of sludges can result in an increased difficulty of dewatering because of their sensitivity to shear. When a medium CD, high MW anionic polyacrylamide is employed on alum sludge, the dewatering response is dependent on both the shear and the mixing time [83]. The polymer requirements increase with longer mixing times and at higher shear because of the disaggregation of the sludge. A higher polymer dose is then needed to re-agglomerate the particles created by excessive mixing conditions.

A recent development has been the synthesis of high MW cationic polyacrylamides that have been crosslinked [84]. Since such a polymer is much less deformable it cannot spread over the surface of the particle, so that its charge cannot be neutralized by that particle. Some charge remains for flocculation or re-flocculation, to give the polymer unique dewatering characteristics. In sludge conditioning massive agglomeration takes place that is initially capable of some breakdown with subsequent shear. Ultimately equilibrium is reached and the floc structure assumes a stable state. The floc size can be 30% larger with a crosslinked polymer than that obtained with the corresponding linear polymer, and the reduction in size on applying shear about 10% instead of more than 90%. Re-flocculation can take place, the overall effect being the production of a very stable floc. Crosslinked polymers have been most successful in dewatering sludges via centrifugation, when a cake of higher solids content offers significant cost savings. Other applications include dewatering via belt presses and the sedimentation of slurries subject to high shear. Another advantage of these materials is that overdosing does not cause the restablisation of the slurry, since adsorption of the polymer onto the surface of a particle cannot result in it occupying all of the active sites [84]. The outcome is that there are increased throughput rates, a higher solids content of centrifuge cakes, and cleaner centrates.

Practical difficulties can arise in operating a polymer coagulated system that has been designed for alum treatment [3]. The clarifier hoppers have a side wall angled at 40-50°, inadequate for the use of polymer, which tends to stick to the walls. Rat-holing can also be a problem.

6. PRACTICALITIES IN THE USE OF POLYELECTROLYTES

There are a number of practical issues that need particular attention when using polyelectrolytes for water treatment. Equipment designed for storing, mixing and feeding polymers takes into account the unique physical characteristics of polymers, and has been much discussed elsewhere [68, 85, 86] Mixing is of paramount importance [87], beginning with initial high energy mixing during the polymer wetting stage in the case of a solid polymer, especially for a polymer of very high molecular weight, to a high rate of polymer dispersion in the reaction stage, but not of such intensity as to cause polymer fracturing [88].

Items considered here are selection of polymer type and methods of controlling and monitoring the dosage. The major concerns with polyelectrolytes, apart from costs, are the environmental impact and toxicity to aquatic life. This has given rise to investigations into polymer toxicity, the analysis of residual polymer in the product water and in wastes from water treatment processes and polymer degradation. A recent concern has been over the production of disinfection by-products from the reaction of residual polymers with disinfectants.

6.1. Polymer selection

There is a complex interrelationship involving polymer structure, MW, CD, dose, mixing conditions, amount and type of impurity particles and organic matter [34, 60, 89]. The CD and MW of cationic polymers affect the rate of adsorption, with the CD being important in determining the optimum dosage when a charge neutralisation mechanism prevails, as the amount of positive charge adsorbed by impurities is about the same irrespective of CD [90]. High MWs are essential when the reactions are via a bridging mechanism. Confirmation of the optimum type and dose of polymer is best determined by jar tests, by the same method employed for inorganic coagulants such as alum or iron salts [52, 68]. For drinking water, typical doses are 1-10 mg L^{-1} for polymers used as primary coagulants, but only 0.1-0.2 mg L^{-1} for polymers used as coagulant aids. The maximum amount is determined by the health limit or cost, if it is lower than the optimum dose.

6.2. Monitoring systems

Under or overdosing can have very significant detrimental effects. Underdosing causes high turbidity and colour levels in the final water; overdosing can result in the re-dispersion of impurities and carryover of polymer from a clarifier into the filtration stage, blinding the filter so that an increase in backwash frequency occurs, accompanied by a lower water yield [91]. Avoiding overdosing helps minimise the amount of sludge formed, consequently reducing land disposal costs. Hence a considerable effort has gone into ways of monitoring polymer feed during operation of a water treatment process.

The most successful method of controlling polymer use is by means of a streaming current detector (SCD). The movement of counter ions beyond the shear plane of surfaces within the detector gives rise to the streaming current [92]. The fluid motion carrying the ions results from piston reciprocation within a closed cylinder that is in contact with a sampled flow. The output from the SCD is related to the zeta potential or electrophoretic mobility of the solid particles. It is therefore possible to directly follow the reduction of the negative surface charge on the particles in the water, and to select the optimum dose, which occurs when the charge is neutralised. Automated control for drinking water production is then achievable when polymers are used as primary coagulants [93]. The technique has also been successfully applied to sludge conditioning, optimal sludge flocculation being obtained at near neutral streaming current [92]. Rheological characteristics have been applied as well, especially on solid residuals in the mineral industry [94]. Dewatering of such sludges has received considerable attention [95, 96].

Another promising method is based on measurements of fluctuations in the intensity of light transmitted through a flowing suspension [97]. In some cases quantitative information on floc size can be derived, but the main use of the technique is for sensitive monitoring of the state of aggregation. A fibre optic flocculation sensor has also been devised [98], and optimising flocculant demand by a laser light diffraction method that follows the floc size distribution has been advocated [99].

A charge titration unit has been designed that automatically titrates the negative charge carriers in raw waters with a cationic polymer [100]. It is capable of computing the required coagulant dose faster and more accurately than manual titration. The dose is automatically determined and transmitted on-line from the unit to the dosing pumps.

6.3. Polymer toxicity

The normally used anionic and nonionic polymers are of low toxicity generally, but cationic types are more toxic, especially to aquatic organisms. Concerns about contaminants have led Japan and Switzerland not to permit the use of polyelectrolytes in drinking water treatment, whilst Germany and France have set stringent limits. The health significance of possible contaminants has been reviewed in detail [101]. The monomers are more toxic than the polymers [102]. Limits on the level of monomer are strictly controlled, especially with acrylamide products, where as a general rule the maximum allowable content of free acrylamide is 0.025%, and the residue in drinking water is limited to 0.5 $\mu g\ L^{-1}$ [103]. For polyDADMAC the monomer content limit is 0.5% in Europe and 2% in the USA [104, 105]. For drinking water production, the National Sanitation Foundation has recommended maximum doses for frequently used commercial polymers in the USA. The limits are generally < 50 mg L^{-1} for polyDADMAC, < 20 mg L^{-1} for ECH/DMA polymers and < 1 mg L^{-1} for polyacrylamides, irrespective of the polyacrylamide charge type. The polyDADMAC figure is based on a maximum carryover of polymer into the product water of 50 $\mu g\ L^{-1}$.

In determining the toxicity of polymers to aquatic species there are different parameters compared to non-polymeric additives [106]. Synthetic polymers tend not to be readily absorbed by organisms, and their toxicity may be substantially altered by key aquatic components [107]. Cationic polymers are rated at moderate to high toxicity, and are markedly more toxic to aquatic organisms than anionic or non-ionic polymers [108]. The cationic types are detrimental to fish because of mechanical gill blockage that causes suffocation [109, 110]. This is greatly reduced by the addition of solids such as clays that are normally present in various forms in receiving waters and sediments, where they can strongly adsorb the cationic polyelectrolyte. Humic acids also affect the toxicity, reducing it by an order of magnitude at humic acid levels of 5 mg L^{-1} [111].

A study of the toxicity of cationic polyacrylamides of various MW and CD towards daphnia and minnows has shown that the toxicity is greatest for polymers of high CD, almost irrespective of chain length [112]. At lower CD the higher MW polymers are more toxic in the case of daphnia, but the actual toxic level of 0.2 mg L^{-1} is well above the likely polymer concentration in the final product water. The toxic level for minnows is 10 times greater. For anionic polyacrylamides the most toxic are those of longer chain length, but they are at least 100 times less toxic than the cationic versions. Cationic polymers drastically disrupt yeast cells at a polymer dose of *ca.*70 mg L^{-1}, while anionic or non-ionic ones do not [113]. The higher the hydrophobicity of the polycation the greater the disruption.

A summary of the toxicity of polymers to freshwater organisms indicates that fish are more sensitive to cationic polymers, but algae are sensitive to anionic polymers because of the chelation of nutrient metal cations [114]. This effect can be offset by the addition of Ca^{++}. The presence of humic substances or clays can markedly reduce the bioavailabilty and hence toxicity of the polymers, and this must be taken into account in any risk assessment of environmental damage resulting from the presence of polymer in surface waters.

6.4. Residual polymer

It is essential to determine the ultimate fate of polymers used in the treatment process, plus that of any impurities present in the original polymer, to see what quantities are present not only in the product water, but in recycled backwash water and sludge. The problem was reviewed long ago [115], when by the use of thin layer chromatography it was found that as well as polymer and monomer, other contaminants appeared to be present in the product water obtained by using a CPAM of the day. In the analysis of an APAM of low CD and high MW (Fig. 1), separation by size exclusion chromatography and fragmentation by flash pyrolysis followed by gas chromatography/mass spectra have been employed [116]. Size exclusion chromatography was also used in a study of residual non-ionic polyacrylamide in an alum sludge [117]. A number of impurities were present at very low levels in water treated with the commercial product, including the monomers acrylamide and sodium acrylate, and also hydroxypropionitrile and traces of isobutyronitrile from the initiator. The sensitivity was 10-20 $\mu g\ L^{-1}$.

A number of approaches of following residual polymer has been studied. To confirm that all added polymer remains attached to particles that are removed prior to water use, it is necessary to analyse for polymer residues in solution. This can be done by colloid titration against a polyelectrolyte of opposite charge, using dyes or fluorescent compounds as indicators [118-120]. The method is not especially sensitive, 0.5-1 mg L^{-1} being the lower limit. A variation on this theme for cationic flocculants is to follow turbidity when tannic acid is used to precipitate the polymer [121]. A similar method measures the light absorbance at 680 nm following progressive additions of a standard solution of an o,o'-dihydroxyazo compound that participates in an association reaction with the polymer [122].

In a review of methods available to determine CPAMs, fluorescence spectroscopy was selected as the most promising analytical method [123]. Further method development is needed to achieve the desired detection limit of 50 $\mu g\ L^{-1}$, and it was felt that there should be more effort by polymer manufacturers to develop suitable methods. Size exclusion chromatography was also considered, but adsorption of polymer to the stationary phase made for uncertainties. Another approach is based on a standard clay test developed long ago [124], where a calibration curve of polymer versus turbidity can be used to measure the content of residual polymer when its sample is tested on the suspension. However, the sensitivity is not high [125]. A review of 17 groups of methods for determining polyacrylamides used in the petroleum industry has been published [49]. The lowest detection limits were 10 and 20 $\mu g\ L^{-1}$, for SEC and fluorescence spectrometry respectively.

A number of workers have followed the course of flocculation with specially synthesised ^{14}C tagged polymer [116]. It has been used to show that dissolved organic matter reacts with the polymer before the turbidity particles [126], and that hydrolytic degradation of polyacrylamide is negligible under normal operating conditions [127]. The final destination of a polymer used in a particular treatment process has been shown to be on solid particles [128]. A fluorescently-labelled cationic polyacrylamide has been made by including choline in a Hofmann reaction on polyacrylamide [129].

More recently, polyDADMACs have been made that contain 1-2% of an amine-functional monomer. These copolymers were reacted with a flurophore so that a fluorescent tag was formed within the polymer structure [54]. Residual polyDADMAC could be detected fluorometrically at concentrations below 100 $\mu g\ L^{-1}$. Measurable amounts of residual polymer were always present in the treated water, especially at doses above or below the optimum. Turbidity particles or alum flocs lowered these levels.

6.5. Polymer degradation

The hydrolysis of the ester links in cationic polyacrylamides is known to be pH and CD dependent, as already mentioned, and is more facile as the pH is increased. Amide groups generally are much more hydrolytically stable. Degradation of polymers by oxidation under disinfection conditions is reported separately below.

The biodegradation of synthetic polymers is extremely slow as most of the structures utilised are resistant, although amide groups are susceptible [130], as are the ester links in CPAMs [131]. Partial cleavage of the latter under both aerobic and anaerobic conditions has been confirmed [132]. The poly(acrylic acid) formed can be degraded in the natural environment, but only material of very low MW (below 4000) seems to be affected [133].

Natural polymers based on polysaccharides and proteins are readily degraded at the ether and amide sites respectively [130]. Synthetic analogues of natural polymers that are easily degraded, such as poly(aspartic acid) and acrylic acid grafts on polysaccharides, are seen to have a future as biodegradable reagents in scale control [134, 135].

6.6. Disinfection by-products

6.6.1. From chlorine

Cationic polymers in common use in the water industry, such as polyDADMAC and ECH/DMA, generally produce little in terms of disinfection by-product formation when the normal levels of polymer utilised in water treatment are exposed to conventional amounts of chlorine [136]. This is illustrated in Fig. 2. However, there is a considerable quantity of trihalomethanes (THMs) formed in the case of CPAMs, albeit after 20 days' exposure to an initial 20 mg L^{-1} of chlorine.

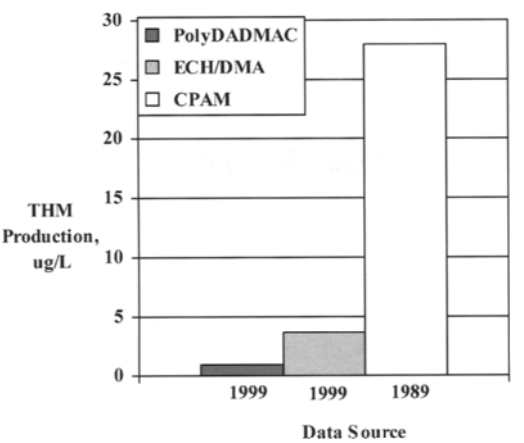


Fig. 2. THM formation from the chlorination of the cationic polymers PDADMAC, ECH/DMA [137] and CPAM [131] at polymer levels of 2, 5 and 10 mg L^{-1} respectively

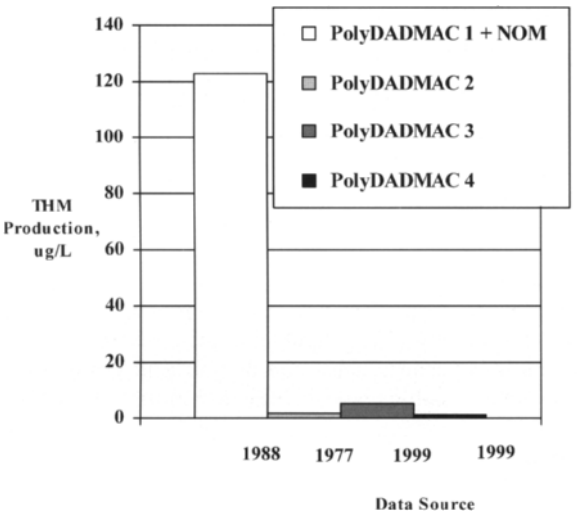


Fig. 3. THM formation from the chlorination of PDADMAC 1 in a water containing NOM with COD 29 mg L^{-1} [138], PDADMAC 2 [139], PDADMAC 3 [140] and PDADMAC 4 [137] at polymer levels of 0.5, 10, 10 and 2 mg L^{-1} respectively

More work is warranted on CPAMs, particularly with regard to the hydrolysis of these polymers as a function of CD, and the reactivity of the hydrolysis products towards chlorine. This is especially so because of the age of the data, as polymer purity and stability has improved much in the interval. Measurement techniques are also much better. The polyDADMAC results do not show this sensitivity to the date of the study. Relative to the amounts formed from NOM in the raw water, THM production is indeed trivial, as shown for polyDADMAC in Fig. 3.

A recent concern is the reported formation of N-nitrosodimethylamine (NDMA), initially detected in the treated water of a small community where polyDADMAC and chlorine were employed [141]. NDMA is a known animal carcinogen of 1000 times the potency of THMs. In detailed exploratory experiments the authors showed that it was not formed when the reagents were present at concentrations normally used in drinking water treatment, to within the limit of detection of 10 ng L^{-1}. Later work on the formation of NDMA in the treatment process found that it was not related to polyelectrolyte use, as it could be formed when polymer was absent from the system [142].

For ECH/DMA there is a more variable production of THMs on chlorination, still in the insignificant range, as depicted in Fig. 4. Of some concern is the high level of AOX produced, at 192 μg L^{-1} versus 12 μg L^{-1} for polyDADMAC, indicating that there are many chlorinated compounds formed from ECH/DMA that have not been identified. The worst result is from one of the older polymers, although one of similar age parallels that of a more recent investigation [137].

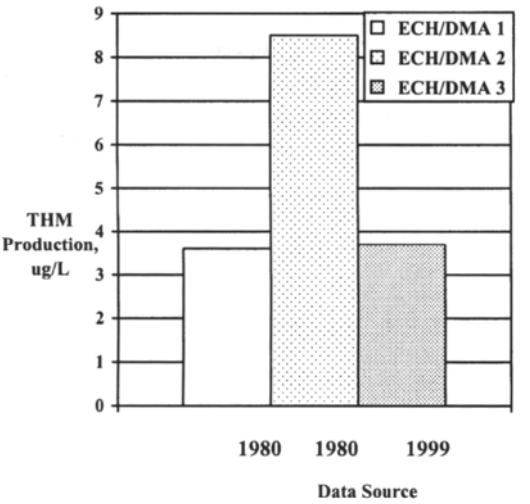


Fig. 4. THM formation from the chlorination of ECH/DMA 1 and ECH/DMA 2 [143], and ECH/DMA 3 [137] at polymer levels of 10, 20 and 5 mg L^{-1} respectively

Other polyacrylamides of the anionic and non-ionic type present less of a problem, with THM formation at trivial levels compared with the amounts produced by NOM [137, 138, 143, 144]. These polymers have contributions equivalent to the amount produced by the monomer present. Acrylamide monomer is a very potent precursor. Strict regulatory control on monomer and impurity content hence has to be maintained. All the results with NOM present show how essential it is that natural precursors be removed before chlorination. If treatment to remove NOM includes polymer use, it would be advantageous to have a more efficient method than those currently available for analysing the level of residual polymer in the product water. A simple method of tagging polymers is required.

6.6.2. From ozone

Although very destructive in its reaction with polymers at high doses, ozone produces mainly formaldehyde when used at levels met with in water treatment. At polymer levels of 10 mg L^{-1} the amount detected for polyDADMAC is 36 μg L^{-1}, and for ECH/DMA 62 μg L^{-1} [137]. Formic acid and lesser amounts of other aldehydes are produced. For APAMs the complete destruction of polymer can be achieved under forcing conditions. Normal doses result in the production of 49 μg L^{-1} of formaldehyde as well as other compounds. More work should be done on identifying the compounds formed. Mutagenicity tests on an ozonated APAM have proved negative. A biodegradation step should nevertheless be carried out after ozone treatment to remove these by-products.

6.6.3. From chlorine addition after ozone

Chlorination when there has been no removal of low MW material after ozone treatment can result in copious quantities of chlorinated by-products. Thus for polyDADMAC an extremely high AOX of 435 μg L^{-1} was obtained, admittedly following excessive doses of both reagents [145]. A major product was chloropicrin (197 μg L^{-1}), and chloroform, dichloroacetic acid and

formaldehyde were each produced at levels above 50 μg L^{-1}, as well as smaller amounts of many other chlorinated species and ketoacids. A biodegradation step such as that provided by biological activated carbon may be necessary prior to final chlorination.

6.6.4. From permanganate

This reagent has a profound effect on polymer degradation, judging from results for polyDADMAC, where there is a clear indication of chain shortening and loss of charge density [146]. The reaction between permanganate and polymer may contribute to the formation of disinfection by-products. Here formation of haloacetic acids and haloacetonitriles from the reaction of chlorine alone was more significant than THM formation. Similar results were obtained with permanganate present as well. Tests on a wider range of commonly used polymers are needed.

6.6.5. From light

Exposure to sunlight can have a significant effect, as an investigation of the reaction of polymers with chlorine or permanganate, or both, with and without prior irradiation with sunlight for an hour or with UV at 254 nm for 30 min has shown [5]. The effects on the formation of chlorinated compounds and polymer performance were determined. Following pretreatment of 0.5 mg L^{-1} of neutral or anionic polyacrylamides with chlorine (5 mg L^{-1}) or permanganate (0.8 mg L^{-1}), with or without exposure to one hour of sunlight, DBPs were measured after a further treatment with 10 mg/L of chlorine in the dark for 7 days. Sunlight combined with permanganate or chlorine caused an increase in the total THM and haloacetic acid (HAAs) formation potentials for APAM (Table 5). For PAM exposure to a single oxidant or to sunlight combined with an oxidant increased the DBP formation potential.

6.6.6. Effect on polymer performance

Damage to polymers from oxidants can be detrimental to the treatment efficiency of polymers [5]. It can be of serious practical significance, as chain cleavage or reduction in polymer charge is detrimental to plant performance. More research under conditions close to those met with in practice is essential on this topic, with regard to chlorine, ozone and permanganate. Disinfection procedures including UV exposure of a kaolinite dispersion of turbidity 2 NTU altered the final turbidity of the product water and the rate of filtration when PAM was employed, as shown in Table 6. A range of polymer doses was explored, from 0-1 mg L^{-1}, which determined the optimum dose. UV treatment resulted in the need for a higher dose of polymer, and any oxidant lowered the filtered water quality, which was much worse with UV and worst when chlorination followed UV. Filtration times were shortened on progressive degradation of the polymer. Filterability was worse with undegraded polymer as the longer, intact polymer chains actually lower permeability, and in the process result in a higher quality product water.

7. ECONOMICS OF USING POLYELECTROLYTES

There is not much current material published on costs. One evaluation of the performance and economics of 23 different organic polymers in treating a Texan turbid water source for drinking water production was aimed at identifying the polymer that reduced filter loading [147]. Increased turbidity removal by a polymer would cause such a reduction, increasing filter run times. Filter backwash savings greater than the cost of polymer were sought. A 5.4 hour increase in filter run time equated to an 18% reduction in the volume of backwash water

required, and a saving of US$4.32 per million gallons of product water. Depending on the cost of polymer and its performance at a dosage of 0.5 mg L^{-1}, the net savings varied from US$2.56 to $3.20 per million gallons. The two most cost effective polymers were not those of lowest cost, ranking 6 and 7 in that regard.

A South African experience on changing from alum to a polymeric coagulant showed a 30% decrease in the unit cost over a three-year period despite an inflation rate of about 10% per annum [3]. Chemical costs in Rand/kL changed from 1.41 for alum and sodium hydroxide in 1987/88 to 1.11 for polymer in 1989/90. There was also a small reduction in power cost as the polymer dose was 10-20% of that for alum, allowing the use of smaller dosing pumps that required less power. As a further bonus the organic polymer was less aggressive than inorganic coagulant, resulting in reduced maintenance costs.

Table 5
DBP formation with and without sunlight exposure

	Disinfection by-product formation potential, μg DBP/ mg polymer				
Polyelectrolyte, DBP	No pretreatment	Sunlight, 1 hour exposure	Permanganate, 0.8 mg/L	Sunlight & permanganate, 0.8 mg/L	Sunlight & chlorine, 5 mg/L
APAM					
Total THMs	16	15	13	19	21
HAAs	11	11	10	13	16
PAM					
Total THMs	14	44	45	40	47
HAAs	36	46	39	39	45

Table 6
Performance parameters after various disinfection procedures

Further treatment	Optimum polymer dose, mg/L	Final turbidity at optimum polymer dose, NTU	Time to filter 100 mL, min
None	0.04	0.16	670
Cl_2	0.04	0.22	490
$KMnO_4$	0.04	0.39	530
UV	0.10	0.47	450
UV + Cl_2	0.10	0.64	400

8. CONCLUSIONS

The role of polymers in water treatment is very well established, with myriad examples of the benefits of polymer use in conventional sedimentation and filtration, and in direct filtration, mostly arising from the lower solids production. The influence of variations in the details of

the chemical structure of the polymer on performance has as yet only been investigated superficially. To illustrate the potential in this regard it has been found recently that of high CD polymers, those with the least polar structure are the most effective in removing hydrophobic organic matter.

Polymer toxicity does not seem to be a problem, as the normally used anionic and non-ionic polymers are of low hazard generally, although cationic types are more toxic, especially to aquatic organisms. Strict limits on the amounts that can be used for drinking water treatment can prevent environmental damage resulting from the presence of polymer in receiving surface waters, as applies also to the maximum permissible carryover of polymer into product water. The monomers used in polymer manufacture are more toxic than the polymers, but rigorous limits on the level of monomer are maintained, especially with regard to acrylamide.

Issues of current relevance include:

- Polymer use in cleaning up filter backwash waters, especially with regard to the recycling of *Cryptosporidium* oocysts
- Better methods for analysis of residual polymer in product water
- Reactions of polymers with oxidants used as disinfectants, to form DBPs.

It would be advantageous to have a more efficient method than currently available for analysing the level of residual polymer in the product water. A simple method of tagging polymers is required to quantify the amount of such material present.

In reactions with chlorine, there is minimal DBP formation from polymers if normal levels of polymer and chlorine are used in a post-chlorination mode. The commonly used polymers are not the principal precursors, except for cationic polyacrylamides, on which more work is needed. Acrylamide monomer is a potent precursor, but is not of concern as a source of DBPs when its presence in the polymer is strictly controlled. With ozone, harmful by-products are formed mostly from monomers and polymer impurities. Profound shortening of the polymer chain occurs, which has an impact on water treatment performance, a point that requires further investigation. Destructive reactions are accelerated in the presence of UV, whatever the chemical oxidant present. Work so far has revealed that because of chain shortening and the loss of polymer charge, the damaged polymer results in shorter filter times and a reduction in the efficiency of turbidity removal. More quantitative research is necessary on this aspect.

REFERENCES

[1] S. Kawamura, J. - Am. Water Works Assoc., 68 (1976) 328.
[2] D. Rout, R. Verma and S. K. Agarwal, Water Sci. Technol., 40(2) (1999) 137.
[3] D.J. Nozaic, S.D. Freese and P. Thompson, Water Sci. Technol.: Water Supply, 1(1) (2001) 43.
[4] D. Faust and O.M. Aly, Chemistry of Water Treatment, Butterworths, Boston, 1983.
[5] A.D. Levine, B.A. Bolto and D.R. Dixon, Reactions of Polyelectrolytes with other Water Treatment Chemicals, AwwaRF Report, AWWA Research Foundation, Denver, 2004.
[6] R.-J. Leu and M.M. Ghosh, J. - Am. Water Works Assoc., 80(4) (1988) 159.
[7] B.A. Bolto, Progr. Polym. Sci., 20 (1995) 987.
[8] E.M. Thurman, Organic Geochemistry of Natural Waters, Martinus Nijhoff-Dr Junk, Dortrecht, 1985, pp. 3, 30
[9] S. Boggs, D.G. Livermore and M.G. Seitz, Rev. Macromol. Chem. Phys., C25 (1985) 599.
[10] H.-J.Huang and H.-H. Yeh, Proc. Water Technol. Conf., Am. Water Works Assoc., Denver, 1993.

[11] J. Lawrence, Analysis of Trace Organics in the Aquatic Environment, B. K. Afghan and A. S. Y. Chau (eds.), CRC Press, Boca Raton, 1989.
[12] F. H. Frimmel, G. Abbt-Braun, K. G. Heuman, B. Hock, H.-D. Lüdemann and M. Spiteller (eds.) Refractory Organic Substances in the Environment, Wiley-VCH, Weinheim, (2002).
[13] M. Bourke and M. Slunjski, Water, 26(6) (1999) 17.
[14] D. B. Bursill, P. T. Hine and J. Y. Morran, Proc. 11th Fed. Conv. Aust. Water & Wastewater Assoc., Sydney, Australia, (1985)197.
[15] R. Beckett, Surface and Colloid Chemistry in Natural Waters and Water Treatment, R. Beckett (ed.), Plenum, New York, 1990
[16] R. Beckett and N. P. Le, Colloids and Surfaces, 44 (1990) 35.
[17] R. J. Gibbs, Environ. Sci. Technol., 17 (1983) 237.
[18] P. R. Hutchison and T. W. Healy, Surface and Colloid Chemistry in Natural Waters and Water Treatment (R. Beckett, ed.), Plenum, New York, 1990
[19] W. Stumm, Chemistry of the Solid-Water Interface, Wiley, New York, 1992.
[20] M. Rebhun and M. Lurie, Water Sci. Technol., 27(11) (1993) 1.
[21] R. Nilsson and M. Kvant, Recycling in Chemical Water and Wastewater Treatment, H. H. Hahn, R. Klute and P. Balmer (eds.), Schriftenreihe des ISWW, Universität Karlsruhe, 1986
[22] P.C. Singer and G.W. Harrington, Proc. Water Technol. Conf., Am. Water Works Assoc., Denver, (1993).
[23] J. Gregory, Flocculation, Sedimentation and Consolidation, B.M. Moudigil and P. Somasundaran (eds.), Engineering Foundation, Washington, 1985
[24] N.G. Hongeveen, M.A. Cohen Stuart and G.J. Fleer, Colloids Surf., A., 117 (1996) 77.
[25] J.K. Edzwald, J.D. Haff and J.W. Boak, J. Environ. Eng. Divn., 103 (1977) 989.
[26] T.K. Wang and R. Audebert, J. Colloid Interface Sci., 119 (1987) 459.
[27] L. Eriksson, B. Alm and P. Stenius, Colloids Surf., A., 70 (1993) 47.
[28] J.K. Edzwald, Organic Carcinogens in Drinking Water, N. M. Ram, E. J. Calabrese and R. F. Christman (eds.), Wiley, New York, 1986.
[29] G.L. Amy and P. Chadik, J. - Am. Water Works Assoc.., 75 (1983) 527.
[30] D. Levine, L.M. Mercurio and J.R. Carman, Chemical Water and Wastewater Treatment IV, H. H. Hahn, E. Hoffmann and H. Odegaard (eds.), Springer, Berlin, 1996.
[31] D.P. Parazak, C.W. Burkhardt, K.J. McCarthy and M.P. Stehlin, J. Colloid Interface Sci., 123 (1988) 59.
[32] G. Durand-Piana, F. Lafuma and R. Audebert, J. Colloid Interface Sci., 119 (1987) 474.
[33] S.Y. Lee and J. Gregory, Water Supply, 9 (1991) 11.
[34] M.M. Ghosh, C.D. Cox and T.M. Prakash, J. - Am. Water Works Assoc.., 77(3) (1985) 67.
[35] M.F. Hoover, J. Macromol. Sci.- Chem., A4 (1970) 1327.
[36] L.K. Schwoyer, Polyelectrolytes for Water and Wastewater Treatment, W.L.K. Schwoyer (ed.), CRC Press, Boca Raton, 1981.
[37] L.B. Luttinger, Polyelectrolytes for Water and Wastewater Treatment W. L. K. Schwoyer (ed.), CRC Press, Boca Raton, 1981.
[38] D.A. Mortimer, Polym. Int., 25 (1991) 29.
[39] J.E. Lancaster, L. Bacchei and H.P. Panzer, J. Polym. Sci., B14 (1976) 549.
[40] C. Wandrey and W. Jaeger, Acta Polym., 36 (1984) 100.
[41] N. Vorchheimer, Polyelectrolytes for Water and Wastewater Treatment W. L. K. Schwoyer (ed.), CRC Press, Boca Raton, 1981.
[42] W. Baade, D. Hunkeler and A.E. Hamielec,. J. Appl. Polym. Sci., 38 (1989) 185.
[43] R. Aksberg and L. Wagberg, J. Appl. Polym. Sci., 38 (1989) 297.
[44] T. Smith-Palmer, N. Campbell, J.L. Bowman and P. Dewar, J. Appl. Polym. Sci., 52 (1994) 1317.
[45] F. Lafuma and G. Durand, Polym. Bull., 21 (1989) 315.
[46] P. van de Wetering, N.J. Zuidam, M.J. van Steenbergen, O.A.G. van der Houwen, W.J.M. Underberg and W.E. Hennink, Macromol., 31 (1998) 8063.
[47] R.A.A. Muzzarelli, Chitin, Pergamon Press, Oxford, 1977.

[48] E.R. Pariser and D.P. Lombardi, Chitin Sourcebook: A Guide to the Research Literature, Wiley, New York, 1989.
[49] K.C. Taylor and H.A. Nasr-El-Din, Pet. Sci. Eng., 12 (1994) 9.
[50] D.M. Rice, J.H. Denysschen and G.J. Stander, Evaluation of a Tannin-base Polyelectrolyte as a Coagulant for Turbid Waters, CSIR Research Report 223, Pretoria, South Africa, 1964.
[51] A. Ndabigengesere, K.S. Narasiah and B.G. Talbot, Water Res., 29 (1995) 703.
[52] R.M. Schlauch, Polyelectrolytes for Water and Wastewater Treatment W. L. K. Schwoyer (ed.), CRC Press, Boca Raton, 1981, p. 91.
[53] E.A. Vik and B. Eikebrokk, Aquatic Humic Substances, I.H. Suffet and P. MacCarthy (eds.), Am. Chem. Soc., Washington, 1989.
[54] D.M. Bennett, B.A. Bolto, D.R. Dixon, R.J. Eldridge, N.P. Le and C.S. Rye, Chemical Water and Wastewater Treatment VI, H. H. Hahn, E. Hoffmann and H. Odegaard (eds.), Springer, Berlin, 2000.
[55] G.S. Logsdon, D.G. Neden, M.D. Ferguson and S.D. LaBonde, J. - Am. Water Works Assoc.., 85(12) (1993) 39.
[56] G.S. Logsdon, D.G. Neden, M.D. Ferguson and S.D. LaBonde, Proc. Water Quality Conf., Am. Water Works Assoc., Denver, 1993.
[57] S.H. Lee, W.S. Shin, M.C. Shin, S.J. Choi and L.S. Park, Environ. Technol., 22 (2001) 653.
[58] J.M. Reuter and A. Landscheidt, Pretreatment in Chemical Water and Wastewater Treatment, H. H. Hahn and R. Klute (eds.), Springer-Verlag, Berlin, 1988.
[59] B.A. Bolto, D.R. Dixon, R.J. Eldridge and S.J. King, Water Res., 35 (2001) 2669.
[60] B.A. Bolto, D.R. Dixon, R.J. Eldridge and S J. King, Chemical Water and Wastewater Treatment V, H. H. Hahn, E. Hoffmann and H. Odegaard (eds.), Springer, Berlin, 1998.
[61] B.A. Bolto, G. Abbt-Braun, D.R. Dixon, R.J. Eldridge, F. Frimmel, S. Hesse, S.J. King and M. Toifl, Water Sci. Technol., 40 (1999) 71.
[62] L. Coccagna, Water, Wastewater, and Sludge Filtration, S. Vigneswaran and R. Ben Aim (eds.), CRC Press, Boca Raton, 1989, p. 57.
[63] G.E. Jackson, CRC Crit. Rev. Environ. Control, 11 (1980) 1.
[64] M. Rebhun, Z. Fuhrer and A. Adin, Water Res., 18 (1984) 963.
[65] B. Bilanovich, G. Shelef and A. Sukenik, Biomass, 17 (1988) 65.
[66] J. Haarhoff and J.L. Cleasby, J. Environ. Eng., 115 (1989) 348.
[67] S.L. Kenefick, S.E. Hrudey, H.G. Peterson and E.E. Prepas, Water Sci. Technol., 27 (1993) 433.
[68] T. Hall and R.A. Hyde, Water Treatment Processes and Practices, Water Research Centre, Swindon, UK, 1992.
[69] J.G. Walzer, Polyelectrolytes for Water and Wastewater Treatment, W.L.K. Schwoyer (ed.), CRC Press, Boca Raton, (1981).
[70] M.E. Tryby, R.J. Miltner and R.S. Summers, Proc. Water Quality Technol. Conf., Am. Water Works Assoc., Denver, 1993.
[71] H. Bernhardt, Proc. Regional Conf., Internat. Water Supply Assoc., Zurich, 1994.
[72] N. Narkis, B. Ghattas, M. Rebhun, and A.J. Rubin, Water Supply, 9 (1991) 37.
[73] H. Arora, G. Digiovanni and M. Lechevallier, J. - Am. Water Works Assoc., 93 (2001) 100.
[74] D.A. Cornwall, M.J. MacPhee, N.E. McTigue, H. Arora, G. DiGiovanni, M. LeChevallier and J.S. Taylor, AwwaRF Report, AWWA Research Foundation, Denver, 2001.
[75] J.B. Rose, A. Cifrano, M.S. Madore, C.P. Gerba, C.R. Sterling and M.J. Arrowood, Water Sci. Technol., 18(10) (1986) 233.
[76] M. LeChevallier, W.D. Norton and R.G. Lee, Appl. Environ. Microbiol., 57 (1991) 2617
[77] J.S. Colbourne, Proc. Am. Water Works Assoc. Water Quality Technol. Conf., Denver, 1989.
[78] A. Eades, B. J. Bates and M.J. Macphee, Water Sci. Technol., 43(8) (2001) 59.
[79] N. Anderson, B. Bolto, C. Chin and L. Kolarik, 15th Fed. Conv. Aust. Water & Wastewater Assoc., Sydney, Australia, 1993.
[80] D. Petruzzelli, A. Volpe, A.C. Di Pinto and R. Passino, Reactive Polym., 45 (2000) 95.
[81] C. Alt, Water, Wastewater, and Sludge Filtration, S. Vigneswaran and R. Ben Aim (eds.), CRC Press, Boca Raton, 1989.

[82] S.K. Dentel, From Sludge to Biosolids, L. Spinosa (ed.), IWA, London, 2002.
[83] J.T. Novak and N. Bandak, J. - Am. Water Works Assoc.., 86(11) (1994) 84.
[84] A. Mohammed, S. Weir and G. Moody, Filtr. & Sepn., 37(8) (2000) 24.
[85] R.J. Chamberlain, Polyelectrolytes for Water and Wastewater Treatment, W.L.K. Schwoyer (ed.), CRC Press, Boca Raton, 1981.
[86] C.F. Lockyear, P.J. Jackson and J.H. Warden, Polyelectrolyte Users' Manual, Technical Report TR 184, Water Research Centre, Swindon, UK, 1983.
[87] A. Amirtharajah and S.C. Jones, Chemical Water and Wastewater Treatment IV, H. H. Hahn, E. Hoffmann and H. Odegaard (eds.), Springer, Berlin, 1996.
[88] J.P. Scott, P.D. Fawell, D.E. Ralph and J.B. Farrow, J. Appl. Polym. Sci., 62 (1996) 2097.
[89] M. Lurie, and M. Rebhun, Water Sci. Technol., 36(4) (1997) 93.
[90] J. Gregory and S.Y. Lee, J. Water Supply: Res. Technol.-AQUA, 39 (1990) 265.
[91] C.R. Veal, Ultrapure Water, 7(4) (1990) 20.
[92] S.K. Dentel, K.M. Wehnes and M.M. Abu-Orf, Chemical Water and Wastewater Treatment III, R. Klute and H. H. Hahn (eds.), Springer, Berlin, 1994.
[93] W. Barron, B.S. Murray, P.J. Scales, T.W. Healy, D.R. Dixon and M. Pascoe, Colloids Surf. A., 88 (1994) 129.
[94] S.K. Dentel, M.M. Abu-Orf and C.A. Walker, Chem. Eng. J., 80 (2000) 65.
[95] R. Hogg, Int. J. Miner. Process., 58 (2000) 1.
[96] J.B. Farrow, P.D. Fawell, R.R.M. Johnston, T.B. Nguyen, M. Rudman, K. Simic and J.D. Swift, Chem. Eng. J., 80 (2000) 149.
[97] J. Gregory and D.W. Nelson, Colloids Surf., 18 (1986) 175.
[98] D.Bartelt, W.Horn, Geiger and G. Kern, Prog. Colloid Polym. Sci., 95 (1994) 161.
[99] S. Lartiges, J.Y. Bottero, C. Democrate and J.F. Coupel, J. Water Supply: Res. Technol.-AQUA, 44 (1995) 219.
[100] H. Bernhardt and H. Schell, J. Water Supply: Res. Technol.-AQUA, 45 (1996) 19.
[101] R.D. Letterman and R.W. Pero, J. - Am. Water Works Assoc., 82[11] (1990) 87.
[102] J. Criddle, A Review of the Mammalian and Aquatic Toxicity of Polyelectrolytes, Report NR 2545, Foundation for Water Research, Medmenham, UK, 1990.
[103] P.J. Norman, Personal communication, Ciba Specialty Chemicals, Leeds, UK, (2002).
[104] BSI Standards, Chemicals used for treatment of water intended for human consumption – Poly(diallyldimethylammonium chloride). BS EN 1408:1998, British Standards Institute, London, (1998).
[105] NSF International, Certified product listings, ANSI/NSF Standard 60, 2001. See www.nsf.org.
[106] B.R. Vitvitskaya, A.A. Korolev, I.N. Skachkova, G.A. Savonicheva, S.G. Sergeev, and O.L. Nokova, Gig. Sanit., 3 (1988) 66.
[107] J.D. Hamilton, K.H. Reinhert and M.B. Freeman, Environ. Sci. Technol., 28 (1994) 187.
[108] W.S. Hall and R J. Mirenda, J. Water Poll. Control Fed., 63 (1991) 895.
[109] K.E. Biesinger and G.N. Stokes, J. Water Poll. Control Fed., 58 (1986) 207.
[110] G.A. Cary, J.A. Mahon and W.J. Kuc, Environ. Toxicol. Chem., 6 (1987) 469.
[111] M.S. Goodrich, L.H. Dulak, M.A. Freedman and J.J. Lech, Environ. Toxicol. Chem., 10[4] (1991) 509.
[112] S.S. Timofeeva, A.M. Beim and A.A. Beim, Khimiya i Teknologiya Vody, 16 (1994) 72.
[113] T. Narita, R. Ohtakeyama, M. Matsukata, J.P. Gong and Y. Asada, Colloid Polym. Sci., 279 (2001) 178.
[114] WRc, A review of polyelectrolytes to identify priorities for EQS development, R&D Technical Report P21, Foundation for Water Research, Marlow, UK, 1998.
[115] V. Goppers and C.P. Straub, J. - Am. Water Works Assoc., 68 (1976) 319.
[116] P. Black, F.. Birkner and J. J. Morgan, J. Colloid Interface Sci. 21 (1966) 626.
[117] H.E. Keenan, E.N. Papavasilopoulos and D.H. Bache, Water Res., 32 (1998) 3173.
[118] L.K. Wang and W. . Shuster, Ind. Eng. Chem., Prod. Res. Dev., 14 (1975) 312.
[119] D.P. Parazak, C.W. Burkhardt and K.J. McCarthy, Anal. Chem., 59 (1987) 1444.
[120] H. Tanaka and Y. Sakamoto, J. Polym. Sci. Part A: Polym. Chem., 31 (1993) 2687.

[121] Y.A. Attia and J. Rubio, Br. Polym. J., 7 (1975) 135.
[122] K. Kirie, T. Zaitsu and C. Igarashi, Jpn. Kokai Tokkyo Kohu JP 86 86,656 assigned to Ebara-Infilco Company Ltd. (1986).
[123] N.J. Howes and G. Harper, R&D Technical Report No. E47, Environment Agency, Bristol, UK, 1998.
[124] H. Burkert, GWF-Wasser/Abwasser-Munich, 111 (1970) 282.
[125] N.S.C. Becker, N.A. Booker, A. Davey, S.R. Gray, R. Jago and C. Ritchie, Chemical Water and Wastewater Treatment , H. H. Hahn, E. Hoffmann and H. Odegaard (eds.), Springer, Berlin, 2000, p. 223.
[126] N. Narkis and M. Rebhun, J. Water Poll. Control Fed., 55 (1983) 947.
[127] F. Hollander, P. Somasundaran and C.C. Gryte, J. Appl. Polym. Sci., 26 (1981) 2123.
[128] N.J. Anderson, B.A. Bolto, R.J. Eldridge and M.B. Jackson, React. Polym., 19 (1993) 87.
[129] H.Tanaka and L. Odberg, J. Polym. Sci. A: Polym. Chem., 27 (1989) 4329.
[130] D. Satyanarayana and P.R. Chatterji, Rev. Macromol. Chem. Phys., C33 (1993) 349.
[131] T. Soponkanaporn and R. Gehr, Water Sci. Technol., 21 (1989) 857.
[132] L.L. Chang, D.L., Raudenbush and S.K. Dentel, Water Sci. Technol. 44[2-3] (2001) 461.
[133] R.W. Lenz, Adv. Polym. Sci., 107 (1993) 1.
[134] K.C. Low and L.P. Koskan, Abstracts of Papers, 206th National Meeting, PMSE 124, Am. Chem. Soc., Washington, 1993.
[135] W. Dezinger, H. Hartmann, C. Gousetis and K.H. Buechner, Ger. Offen. DE 4,038,908 assigned to BASF A.-G., (1990).
[136] B. A Bolto, J. Water Supply: Res. Technol.-AQUA, 55(8) (2005) 531.
[137] M. Fielding, J. Hutchison, D.M. Hughes, W.H. Glaze and H.S. Weinberg, AwwaRF Report, AWWA Research Foundation, Denver, 1999.
[138] L.P. Alekseeva and Ya.L. Khromchenko, Soviet J. Water Chem. Technol., 10 (1988) 140.
[139] K.L.E. Kaiser and J. Lawrence, Science, 196 (1977) 1205.
[140] E.E. Chang, P.C. Chiang, S.H. Chao and C.H. Liang, Chemosphere, 39 (1999) 1333.
[141] P. Child, G. Kaar, D. Benitz, P. Fowlie and R. Hong-You, Proc. Fourth National Conf. on Drinking Water, Ontario, 1991.
[142] D.B. Jobb, R.B., Hunsinger, O. Meresz and V.Y. Taguchi, Proc. Fifth National Conf. on Drinking Water, Winnipeg, 1992.
[143] M.A. Feige, E.M. Glick, J.W. Munch, D.J. Munch, R.L. Noschang and H.J. Brass, Water Chlorination: Environmental Impact and Health Effects, R. L. Jolley, W. A. Brungs, R.B. Cumming and V.A. Jacobs (eds.), Ann Arbor Sci. Pub., Ann Arbor, Michigan, 1980.
[144] J. A. Mallevialle, A. Bruchet and F. Fiessinger, J. - Am. Water Works Assoc., 76(6) (1984) 87.
[145] P. Stockham and J. Morran, Proc. WaterTECH Conference, Aust. Water Assoc., Sydney, 2000.
[146] D. Levine, R.L. Swickley, J.R. Carman, C.M. Ballif, M.S. Beal and M.E. Owens, Proc. Am. Water Works Assoc. Water Quality Technology Conf., Denver, 1998.
[147] J. Tarquin, G.B. Haan and D.Rittmann, Water Eng. Manage., 145[1] (1998) 27.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 6: Dissolved air flotation in drinking water treatment

James K. Edzwald

Department of Civil and Environmental Engineering, University of Massachusetts, Amherst, MA 01003, USA

1. INTRODUCTION

Dissolved air flotation (DAF) was first applied to drinking water treatment in the late 1960s in Scandinavia and South Africa. The process is particularly efficient in removing low density particles and flocs. Hence, DAF is a good clarification process for treating supplies of low to moderate turbidity, and those containing algae and natural color. It has found increasing favor over sedimentation processes for treating these type supplies, and it is now widely used world-wide. Fig. 1 shows a schematic process diagram for a DAF drinking water plant. Coagulation and flocculation are required pretreatment processes. Chemical coagulation is necessary to produce particles or flocs with low repulsive forces between particles and air bubbles so that bubble attachment to flocs may occur. A brief flocculation period is necessary to obtain high collision rates between flocs and bubbles. Air is dissolved in a recycle flow by adding air under pressure in a saturator. Recycle rates (defined as the recycle flow (Q_r) divided by the plant through-put flow (Q_o)) are about 8 to 12%. Saturator pressures are between 400 and 600 kPa. The recycle flow is injected through nozzles into the front section of the DAF tank, called the contact zone, producing air bubbles of 10 to 100 μm and giving a milky appearance to the water; hence the name *white water*.

The DAF reactor has two functions, to provide collisions or contact opportunities between air bubbles and flocs and to provide removal of the floc-bubble aggregates. Conceptually, DAF is thus modeled as a two step process as depicted in Fig. 2. In the contact zone, floc particles are brought into contact with the bubble suspension with the purpose of forming floc-bubble aggregates. The flow moves to the second part of the reactor, called the separation zone, where rising aggregates are collected in the float layer at the top of the tank and the treated water or subnatant is withdrawn from the bottom of the tank. The objectives of this chapter are four-fold. First, bubble properties (size and density), bubble suspension concentrations (mass, number, and volume), and bubble characteristics (separation distances and rise velocities) are summarized. Second, the contact zone model is summarized. Model variables are discussed beginning with the role of pretreatment coagulation chemistry and flocculation for successful bubble attachment and collisions of flocs with bubbles. The importance of bubble size and concentrations is also addressed. Third, separation zone modeling is summarized. Rise velocities of floc-bubble aggregates are compared to separation zone theory and hydraulics. Finally, some applications are presented on the removals of algae and protozoa pathogens by DAF, and on recent trends of designing DAF plants with short flocculation times and high hydraulic loadings.

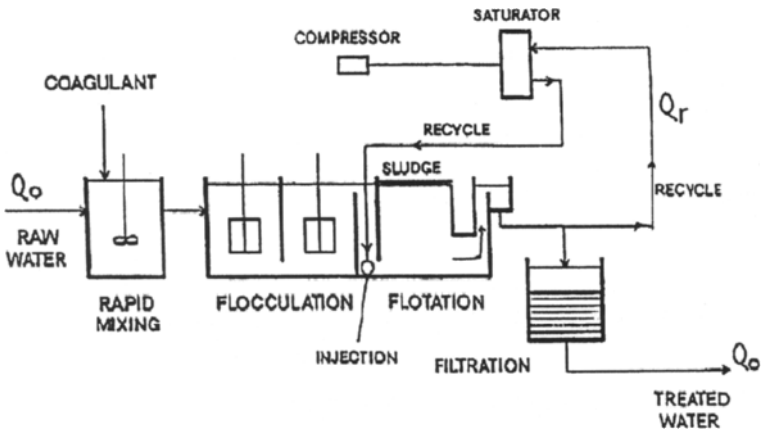


Fig. 1. Process schematic of a DAF drinking water treatment plant

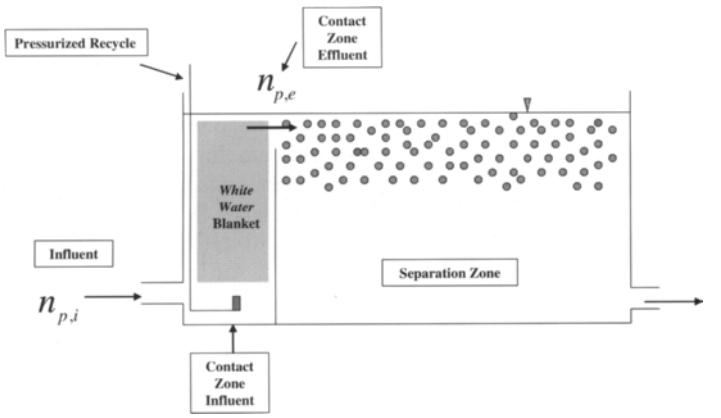


Fig. 2. Schematic diagram of a DAF tank showing the contact and separation zones

2. BUBBLES

2.1. Bubble properties

Small air bubbles are formed instantaneously after injection of the pressurized recycle flow into the contact zone influent at the bottom of the tank– see Fig. 2. The bubble sizes formed depend on the saturator pressure, the injection device (nozzle type or needle valve), and the recycle flow rate. The injection devices are usually nozzles, which produce a more uniform bubble size distribution. The pressure difference across the injection device is the primary factor affecting bubble size. Saturator gauge pressures can range from 400 to 600 kPa, but 500 kPa is typical. Hence for a saturator pressure of 500 kPa and with very low pressure loss in the recycle piping, the pressure difference is about 475 kPa for nozzles placed at a depth of 2 to 2.5 m (i.e., under a water pressure of 20 to 25 kPa).

A size distribution of bubbles from to 10-100 μm or greater is produced with generally lower mean sizes found for higher saturator pressures. A mean value of 60 μm is a good

estimate for contact zone modeling for DAF saturator pressures of about 500 KPa [1]. It is assumed that the bubble size is maintained throughout the contact zone depth even with the decreasing hydrostatic pressure [2]. Larger bubble sizes are found in the separation zone, and this is discussed in Section 4.

The density of air bubbles is a fundamental property needed for calculating bubble volume concentrations and bubble rise velocities. The density of air bubbles depends on pressure, temperature, and humidity. While the pressure decreases from about 125 kPa at the bottom of the contact zone to approximately 101.3 kPa at the top, for purposes of using densities to calculate bubble rise velocities, the lower atmospheric pressure of 101.3 kPa is used. Moist air densities are presented in Table 1 for conditions of 100% humidity and with the dew point temperature equal to the water or air temperature [3].

Table 1
Moist air densities and Henry's Law equilibrium constants

Temperature (°C)	ρ_b (kg m^{-3})	H^* (kPa mg^{-1} L)
0	1.290	2.722
5	1.266	3.075
10	1.241	3.459
15	1.218	3.825
20	1.194	4.184
25	1.171	4.537
30	1.146	4.858

*[c]alculated based on air composition of 78.1% N_2, 21.0% O_2 and accounting for Ar at 0.9%; molar mass of air of 28.95 g $mole^{-1}$ [4]

2.2. Bubble suspension

The air bubble suspension in the contact zone is characterized first in terms of the mass concentration (C_b). Henry's Law calculations are used to determine the concentration of air dissolved in recycle water leaving the saturator (C_r). Table 1 presents Henry's constants for air for water temperatures of 0 to 30 °C. It is imperative to understand that 100% dissolution of the air in the saturator is difficult because air is a mixture of gases with a higher solubility of oxygen in water than nitrogen leading to a nitrogen rich air in the saturator limiting air dissolution, and because of kinetic limitations of air transfer. Therefore, the actual air concentration released in the contact zone is less than 100% based on assuming complete solubility of air in the recycle water. The pressure loss in the recycle piping can also have a minor impact on the air released. In practice the overall efficiency (e) of air delivery for the contact zone varies from 60 to 90% depending on the specific saturator and recycle system. Packed tower saturators are more efficient than unpacked, and can achieve about 90% efficiency [5].

The mass concentration of air released in the contact zone (C_b) can be calculated from Eq. (1). Eq. (1) is obtained from a mass balance accounting for the air in the recycle flow (C_r), the saturation air concentration (C_s) in water for the pressure in the contact zone (101.3 kPa is used), the recycle ratio (R), and k which accounts for any air deficit concentration in the influent flocculated water to the DAF tank. The two main variables in Eq. (1) for controlling C_b are the saturator pressure, which affects C_r, and the recycle ratio (R). C_b is controlled primarily in practice through R, since for a particular plant the saturator pressure is not

changed very much. Volume (ϕ_b) and number (n_b) concentrations are then calculated from Eqs. (2) and (3).

$$C_b = \frac{[e(C_r - C_s)R - k]}{1 + R} \quad (1)$$

$$\varphi_b = \frac{C_b}{\rho_b} \quad (2)$$

$$n_b = \frac{\phi_b}{(\pi d_b^3 / 6)} \quad (3)$$

Concentrations are summarized in Table 2 for 10% recycle for low and high saturator pressures of 400 and 585 kPa, respectively, and for low and high overall efficiencies (e) of 70 and 90%. A bubble size of 60 μm was assumed for calculating n_b in Eq. (3). These calculations provide reasonable estimates of low to high air bubble concentrations for the contact zone. Mass concentrations (C_b) range from 5 to 11 mg L^{-1}, while the bubble volume concentrations (ϕ_b) range from 4000 to 9000 ppm and number concentrations (n_b) between 40 and 80x10^6 bubbles L^{-1}. Compared to particle floc volume and number concentrations, these are quite high allowing for good opportunity for collection of floc onto air bubbles and lowering floc density. This is discussed further in Sections 3.2-3.4. The high bubble number concentrations yield separation distances (d_{sd}) between air bubbles of about 200 μm yielding a *white water* blanket of bubbles in the contact zone that is modeled much like a filter.

Table 2
DAF tank contact zone mass, volume, number concentrations of air bubbles and separation distances between bubbles for10 percent recycle (20 °C, bubble diameter of 60 μm)

Saturator pressure (kPa) and efficiency (%)	C_b (mg L^{-1})	Φ_b (ppm)	n_b (bubbles L^{-1})	d_{sd} (μm)
585 at 90%	11.2	9400	83x10^6	169
585 at 70%	8.2	6900	61 x10^6	195
400 at 90 %	7.6	6370	56 x10^6	201
400 at 70%	5.4	4540	40 x10^6	233

The classical Stokes equation is used to calculate air bubble rise velocities for bubble sizes up to 100 μm. The bubbles have a spherical shape [6], and Reynolds numbers are less than 1 indicating laminar flow for the streamlines of flow around the rising bubble.

$$v_b = \frac{g(\rho_w - \rho_b)d_b^2}{18\,\mu_w} \quad (4)$$

Eq. (4) shows that v_b depends greatly on bubble size (d_b^2), on the density difference ($\rho_w - \rho_b$), and on the water viscosity. A distinctive feature of flotation is the large density difference of almost 1000 kg m^{-3} that acts as the driving force for the bubble rise velocity. The main effect of water temperature is on water viscosity, lower temperature increases the viscosity reducing the rise velocity. Fig. 3 shows the effects of the main variables of bubble size and water

temperature on rise velocity. Bubbles of 10 μm or less have rise velocities < 1 m h^{-1}. Bubbles in the contact zone with an average size of 60 μm have a rise velocity of 7 m h^{-1} at 20 °C reduced by 36% in cold water (4 °C) to 4.5 m h^{-1}. Larger bubbles of 100 μm exist in the separation zone as discussed in Section 4.2. These bubbles have much greater rise velocities of 12.5 m h^{-1} in cold water to 19.5 m h^{-1} in warmer waters.

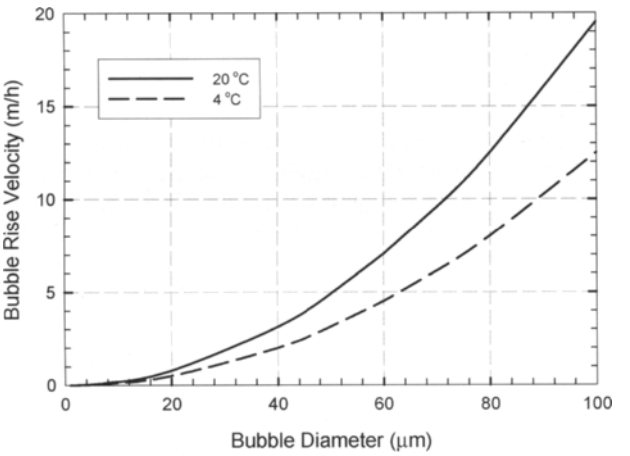


Fig 3. Bubble rise velocities as a function of bubble size for water temperatures of 4 and 20 °C

3. COLLISIONS AND ATTACHMENT OF PARTICLES WITH BUBBLES

3.1. Modeling framework

It is useful to distinguish between the sizes of bubbles and particles for dispersed air flotation versus dissolved air flotation processes. Dispersed air flotation processes commonly used for mineral separation from ores involve large bubbles and large particles, sizes of millimeters or fraction thereof. The adhesion or attachment of these large bubbles to particles is described as a three phase phenomena in which contact angles are measured and used to quantify adhesion. High contact angles indicate strong bubble attachment and hydrophobic particle surface conditions. Dissolved air flotation, on the other hand, involves much smaller bubbles and particles, sizes of micrometers. Here, small particles attach to small bubbles but do not involve contact angles. In fact, Derjaguin *et al.* [7] refer to this as *contactless* flotation meaning that at close distances between particles and bubbles there are forces of interaction, either causing attachment or providing a repulsive energy barrier preventing particle-bubble attachment. This idea of *contactless* flotation for the attachment step is incorporated into the contact zone model described next.

The model framework considers two steps: 1) collision opportunities and 2) particle attachment to the bubbles. The first step involves mechanisms of how particles are transported from the bulk water to the vicinity of the bubble surface. As particles approach the bubble surface, interaction forces can prevent attachment or facilitate attachment if the forces are attractive. The key equations of the contact zone model are summarized next. Discussion

follows using the model variables to gain insight into how DAF contact zone performance is affected by pretreatment and by bubble size and concentration.

3.2. Contact zone model

The DAF reactor has two functions, to provide collisions opportunities between flocs and air bubbles and to provide removal of the floc-bubble aggregates. These two functions are carried out in the contact and separation zones, respectively, as depicted in Fig. 2. The contact zone model considers bubbles in the *white water* blanket within the contact zone as collectors of particles using the single collector efficiency approach to account for particle transport to the bubble interface. The *white water* blanket of bubbles exists in the contact zone at a dynamic steady state at a high bubble concentration (> 10^7 bubbles L^{-1}, Table 2). This dynamic steady state is maintained by continuous injection of bubbles from the recycle flow and output at the exit of the contact zone. Details of the model are found in two papers of Edzwald and co-workers [8, 9]. The model has been critically reviewed recently by Haarhoff and Edzwald [1].

Air bubbles in the *white water* blanket act as collectors of particles or flocs, and the dimensionless particle transport coefficient (η_T) is used to account for the total collision efficiency of a single bubble or collector due to the various transport mechanisms. Collisions may occur by Brownian diffusion (η_D), by fluid flow or interception (η_I), and by settling of flocs (η_S) onto bubbles. Eqs. (5-7) describe the individual single collector efficiencies.

$$\eta_D = 6.18 \left[\frac{k_b T}{g(\rho_w - \rho_b)}\right]^{2/3} \left[\frac{1}{d_p}\right]^{2/3} \left[\frac{1}{d_b}\right]^2 \tag{5}$$

$$\eta_I = \left(\frac{d_p}{d_b} + 1\right)^2 - \frac{3}{2}\left(\frac{d_p}{d_b} + 1\right) + \frac{1}{2}\left(\frac{d_p}{d_b} + 1\right)^{-1} \tag{6}$$

$$\eta_S = \left[\frac{(\rho_p - \rho_w)}{(\rho_w - \rho_b)}\right]\left[\frac{d_p}{d_b}\right]^2 \tag{7}$$

The interception expression (Eq. (6)) is more complicated than the classical equation used to model granular media filtration. If $d_p/d_b << 1$ (small flocs relative to bubble size), the classical, approximate, equation (Eq. (8)) can be used. The latter equation demonstrates directly the dependence of η_I on $(d_p/d_b)^2$, however, in general Eq. (6) should be used.

$$\eta_I = \frac{3}{2}\left(\frac{d_p}{d_b}\right)^2 \tag{8}$$

Conceptually, the *white water* blanket of bubbles acts like a filter. Interestingly, the separation distance between bubbles of ~ 200 μm (Table 2) is not unlike the pore space between filter grains in a filter bed. The single collector efficiency parameter used to model collisions of particles with bubbles has been used successfully to model removal of particles in water filtration [10] and for other types of flotation processes [11, 12].

The total single collector efficiency is incorporated into a kinetic expression for collision of particles with bubbles [1], and a steady state mass balance yields the contact zone

model performance equation.

$$\left(1 - \frac{n_{p,e}}{n_{p,i}}\right) = \text{Efficiency} = \left[1 - exp\left(\frac{-\frac{3}{2}\alpha_{pb}\eta_T\phi_b v_b t_{cz}}{d_b}\right)\right] \quad (9)$$

Eq. 9 describes the efficiency of removal of particles or flocs onto bubbles within the contact zone. The model is instructive in that it identifies important variables affecting contact zone performance. These include the particle-bubble attachment efficiency (α_{pb}), the single collector efficiency (η_T), bubble size (d_b), bubble rise velocity (v_b), bubble volume concentration (Φ_b), and the contact zone detention time (t_{cz}). Table 3 provides a summary of these variables. The variables are categorized in terms of how they are affected by pretreatment processes and by flotation tank design and operation.

Table 3
Contact zone model variables

Variable	Dependence	Comments
	Pretreatment	
α_{pb}	Interaction forces between particles and bubbles	Coagulation chemistry including dose and pH critical in maximizing particle attachment to bubbles
η_T	$d_p^{-2/3}$: Brownian Diffusion d_p^2: Interception & Settling	Flocculation increases d_p. Desire flocs of 10s of microns
	Flotation tank	
d_b	d_b^{-1} in Eqn (9) and d_b^{-2} for η_T dependence.	Smaller bubbles yield better performance. Bubble size set mainly by the pressure difference across the nozzle; influenced also by the type of nozzle. Range of 10 to 100 μm; mean of 60 μm
η_T	Primarily floc size ($d_p^{-2/3}$ & d_p^2) and bubble size (d_b^{-2})	For mean bubble size of 60 μm, high η_T achieved for flocs of 10s of microns
Φ_b	Recycle flow and saturator pressure.	Controlled mainly by the recycle flow or ratio. Values as high as 8000-9000 ppm insure large volumes of bubbles for collection and floating of particles
v_b	d_b^2	Describes motion of bubble relative to the water flow in the contact zone. Lower value means bubbles reside for longer period increasing opportunity for collection of particles
t_{cz}	Depth and flow rate through the contact zone.	Contact zone detention times need to exceed 1-2 min

3.3. Pretreatment chemistry and flocculation

Pretreatment coagulation affects the particle-bubble attachment efficiency (α_{pb}), and pretreatment flocculation affects the particle size (d_p) of flocs in the influent to the flotation tank contact zone. Both pretreatment steps have been described in earlier Chapters, 2 and 5, and both are important to DAF. Model predictions (Eq. (9)) are presented next for typical DAF tank operation followed by discussion of the fundamentals.

One approach to evaluating α_{pb} is an empirical viewpoint. α_{pb} can have values between 0 (no collisions lead to attachment) and 1 (all collisions result in attachment). In this empirical approach, α_{pb} depends on coagulation pretreatment chemistry (coagulant type, dosage, and pH). Fig. 4 shows the contact zone efficiency as a function of particle or floc size for two cases of α_{pb}: 0.5 for good coagulation and 0.01 for poor coagulation. Model predictions in Fig. 4 show that for poor coagulation the contact zone efficiency is very low (< 5%) for particles or flocs < 10 μm, and that high efficiencies are not obtained unless flocs approach 100 μm. For good coagulation (favorable attachment, α_{pb} of 0.5), floc particle sizes of tens of microns result in high efficiencies (> 80%). The model predictions also show a minimum in the contact zone efficiency for particles of 0.9 μm. This minimum at ~ 1 μm is analogous to granular media filtration efficiency. This is because the physics of particle transport is the same for the two processes. Particle transport by Brownian diffusion is the controlling mechanism for particles < 1 μm while particle transport by interception controls for particle sizes > 1 μm; transport by settling is not significant for floc particles with densities of 1100 kg/m^3. In summary, floc particles of tens of microns have high η_T values yielding good contact zone removal efficiencies (Fig. 4). An important model finding then is that floc particles of tens of microns should be prepared by the pretreatment flocculation process for the influent to DAF tanks.

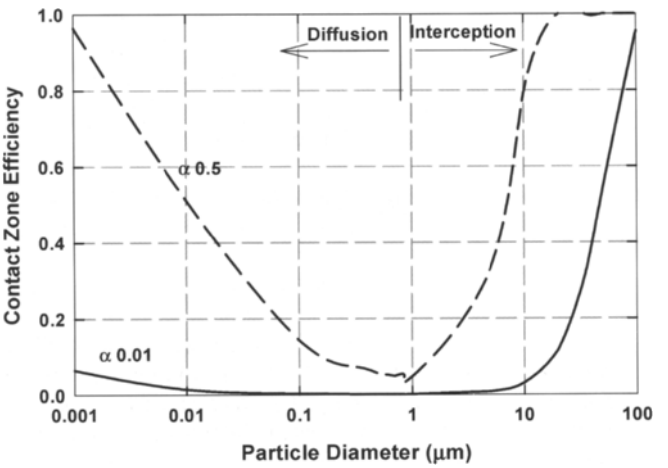


Fig. 4. Contact zone efficiency as a function of particle or floc size for good (α_{pb} = 0.5) and poor (α_{pb} = 0.01) coagulation for conditions of t_{cz} = 3 min, ρ_p = 1100 kg m^{-3}, T = 293 °K, d_b = 60 μm, and Φ_b = 8000 ppm (Source: Adapted from Haarhoff and Edzwald [1])

The empirical approach to the evaluation of α_{pb} has had some success. Haarhoff and Edzwald [1] reported on empirically determined α_{pb} values between 0.5 and 1 for optimum alum coagulation characterized by pH around 6 and dosing yielding flocs of zero or no charge. Shawwa and Smith [13] found α_{pb} values of 0.35-0.55 for good coagulation conditions, and Schers and Van Dijk [14] found α_{pb} values of 0.2-1 from data for six DAF plants in The Netherlands.

As particles are transported from the bulk solution to close distances at the bubble surface, forces between the bubble and particle affect attachment, as depicted in Fig. 5. Derjaguin *et al.* [7] have described this bubble-particle interaction for small particles according to classical colloidal particle interactions. These forces occur at separation distances of tens of nanometres. These forces include electrostatic forces from overlapping of electrical double layers (charge repulsion), London-van der Waals forces, and a hydrophobic force. Another force affecting bubble-particle attachment is hydrodynamic retardation. Excluding the hydrophobic force, Han [15] and Leppinen [16] have considered the above forces to model collisions and attachment of particles to bubbles. Next, some discussion is presented of these fundamental factors affecting the particle-bubble attachment efficiency, α_{pb}.

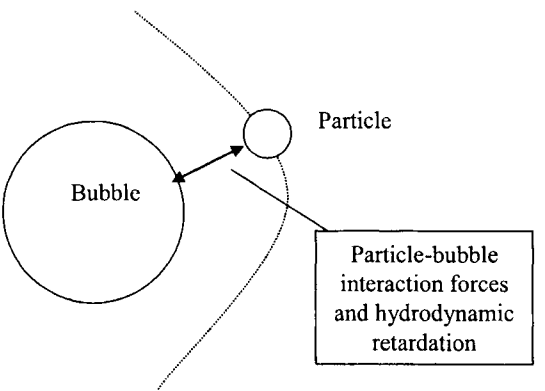


Fig. 5. Attachment affected by particle-bubble interaction forces and hydrodynamic retardation

Air bubbles carry a negative charge, even in distilled water or distilled water containing simple non-hydrolyzing salts. Negative zeta potentials have been reported in the range of -15 to -25 mV [7, 17, 18]. Dockko and Han [18] reported zeta potentials of about – 25 mV at pH 6 to 8 for bubbles of about 30 μm. They found the isoelectric point at about pH 2.5. Unlike inorganic and organic particles in waters, air bubbles do not contain surface functional groups that ionize or complex metals leading to charged surface groups; however, they can obtain a charge by three mechanisms. First for the case of bubbles in clean water, their charge is attributed to a higher concentration of small anions close to the bubble surface compared to the larger hydrated cations. Second, water supplies may contain negatively charged surfactants that accumulate at the bubble surface causing a negative charge. Third, the addition of positively charged polymers or metal coagulants that precipitate positively charged particles can accumulate at the bubble interface reducing its charge or even reversing its charge.

London-van der Waals forces between dissimilar particles – i.e., solid particle and air bubble – may be attractive or repulsive. Lu [19] in considering the London dispersion force for a non-polar bubble and a particle in water makes a case for a repulsive interaction. Others, however, indicate that the net DLVO force (double layer repulsion and London-van der Waals) may be attractive or repulsive [7, 17] with the double layer repulsion due to interaction of negatively charged particles and bubbles playing a key role preventing particle attachment to air bubbles.

Ducker *et al.* [17] made force measurements and found that the hydrophobic force is strong at distances exceeding those associated with double layer interactions, and it is the primary force explaining attachment of hydrophobic particles to air bubbles.

Another force that can hinder collisions, and thus attachment, is hydrodynamic retardation. This is the force that causes a deviation in the particle trajectory as the particle approaches the bubble surface due to resistance to motion from thinning of the viscous water between the particle and bubble. The single collector collision efficiency equations presented above, for interception and settling (Eqs. (6) and (7)), do not account for hydrodynamic retardation, and show a basic dependence of η_I and η_S according to d_p^2, although the interception dependence is a little more complicated than that. The exponent on d_p should be less than 2 if hydrodynamic retardation affects η. Collins and Jameson [20] found dependence according to $d_p^{1.5}$ indicating an effect from hydrodynamic retardation. Their experiments were undertaken using polystyrene particles of 4-20 μm collected by bubbles of about 50 μm, a system where η_I (interception) dominates. On the other hand when settling (η_S) was the main transport mechanism, Reay and Ratcliff [12] found dependence according to $d_p^{2.05}$.

In summary, coagulation is essential in water treatment to reduce repulsive charge interactions between particles or flocs and bubbles. Favorable attachment (high α_{pb} values) of particles to bubbles requires reduction in the repulsive charge interaction between particles and bubbles. Flocs with zero or very low zeta potentials should be produced through coagulation. Under these conditions, attractive forces can prevail (London-Van der Waals or hydrophobic) leading to attachment. Coagulant chemicals are used to obtain favorable attachment so α_{pb} depends on coagulation conditions (type, dosage, and pH). In some waters hydrophilic colloids may exist that are resistant to bubble attachment, but proper coagulation can alter the colloid surface properties so that attachment to bubbles is favored.

3.4. Bubble size, bubble concentration, and contact time

Eq. (9) shows the variables affecting contact zone performance. Important design and operating variables associated with the flotation tank and recycle system are bubble size (d_b), bubble volume concentration (Φ_b), and the contact zone detention time (t_{cz}) as summarized in Table. 3.

Overall, the contact zone efficiency depends on d_b^{-1} – see Eq. (9) and Table 3. Smaller bubbles provide for better performance. Bubble size also affects η_T with higher single collector efficiencies for smaller bubbles. The principle that smaller bubbles are better for the contact zone agrees with observations that DAF is more efficient than dispersed air flotation. Bubble size of say 60 μm is characteristic of DAF in contrast to bubbles of about 1 mm for dispersed air flotation.

Model predictions in which the contact zone detention time is varied (t_{cz}) are made using Eq. (9). The model assumes plug flow. Haarhoff and Edzwald [1] examined the hydraulics of the contact zone and the effects of dispersion, and found that plug flow provides reasonable predictions of contact zone efficiency. To evaluate the effect of bubble concentrations, two cases of Φ_b are examined. In one case Φ_b was set at 9000 ppm, this

represents high bubble concentrations (Table 2) that are achieved with recycle ratios (R) of about 10-12% and saturator pressures of 500-600 kPa. Φ_b was set at a much lower value of 3000 ppm, which may result from low recycle ratios such as 4% or combinations of lower R and lower saturator pressures.

Model predictions for removing 20 μm flocs are presented in Fig. 6 showing the dependence of contact zone efficiency on the detention time for the two bubble volume concentration cases. Poor efficiency occurs for low bubble volume concentrations unless high contact zone detention times are used. Low bubble volume concentrations do not provide sufficient bubble volume for efficient collection of flocs. Low bubble volume concentration of 3000 ppm means the air bubble mass (C_b) and number concentrations (n_b) are also low, corresponding to values of 3.6 mg L^{-1} and $25x10^6$ bubbles L^{-1}, respectively. Fig. 6 shows that for a bubble volume concentration of 9000 ppm, the contact zone efficiency is high and insensitive to detention times exceeding 1.5 min. Contact zone detention times used in practice are 2 to 4 min. The theoretical predictions support design practice, and show that shorter times lead to poor performance. They also show there is no benefit to increasing the detention time where Φ_b is about 9000 ppm. A Φ_b of 9000 ppm corresponds to C_b and n_b concentrations of 10.7 mg L^{-1} and $70x10^6$ bubbles L^{-1}, respectively. DAF systems are usually designed to release about 10 mg L^{-1} of air so a good benchmark value for Φ_b is 9000 ppm.

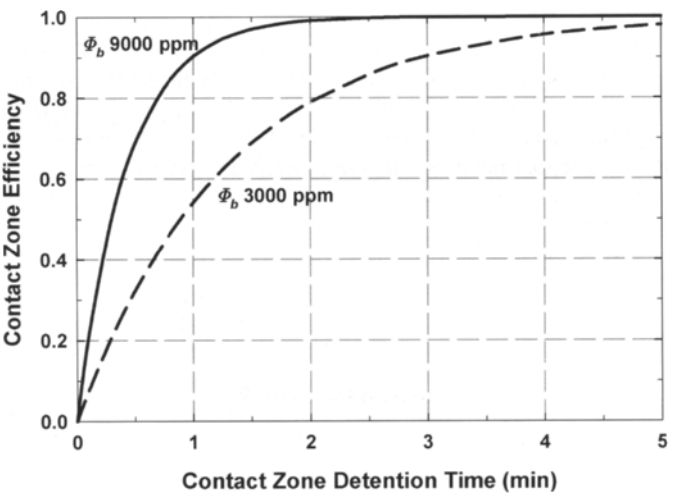


Fig. 6. Contact zone efficiency as a function of detention time for high (Φ_b = 9000 ppm) and low (Φ_b = 3000 ppm) bubble volume concentrations for floc size of 20 μm for conditions of α_{pb} = 0.5, ρ_p = 1100 kg m^{-3}, T = 298 °K, and d_b = 60 μm (Source: Adapted from Haarhoff and Edzwald [1])

4. RISE VELOCITY OF FLOC-BUBBLE AGGREGATES

4.1. Model framework

To calculate the rise velocity of floc-bubble aggregates requires the determination of the aggregate size and density. The aggregate equivalent diameter (d_{fb}) produced from N bubbles of size d_b attached to a floc particle of size d_p is estimated by:

$$d_{fb} = \left(d_f^{\,3} + N d_b^{\,3}\right)^{1/3} \tag{10}$$

The floc-bubble aggregate density ρ_{fb} is estimated from the weighted average of the air bubble and floc densities:

$$\rho_{fb} = \frac{\rho_f d_f^{\,3} + N\rho_b d_b^{\,3}}{d_f^{\,3} + N d_b^{\,3}} \tag{11}$$

Next, a modified Stokes equation is used to estimate the floc-bubble aggregate rise rate:

$$v_{fb} = \frac{4g(\rho_w - \rho_{fb})d_{fb}^2}{3K\mu_w} \tag{12}$$

K accounts for floc shape and the drag force. Since the bubbles are spherical, the shape of the aggregate depends on floc size [1]. Model predictions of rise velocity are presented below for a mean bubble size of 100 μm. For small flocs relative to the bubble size, the aggregate is approximately spherical and $K = 24$ so that Eq. (12) simplifies to the classic Stokes expression of Eq. (4). For flocs much larger than the bubble size, the aggregate shape approaches the floc shape and $K = 45$. For the calculations presented below it is assumed that K varies gradually from 24 at a floc size at or below 40 μm to 45 at a floc size at or above 170 μm.

4.2. Rise velocity predictions

Bubble sizes in the separation zone are more likely larger than in the contact zone due to a decrease in water pressure in the upper portion of the separation zone, and perhaps, due to coalescence. Leppinen and Dalziel [21] measured bubble sizes in the separation zone of full-scale DAF tanks and found that a large fraction of the bubbles had sizes near 100 μm. Consequently, a bubble size of 100 μm is used for aggregate rise velocity predictions for the separation zone, in contrast to the 60 μm bubble diameter used in Section 3 for the contact zone model. Eq. (12) shows that the rise velocity is greatly influenced by aggregate size (d_{fb}) and the density difference (ρ_w-ρ_{fb}); the latter is called, conceptually, the driving force for flotation. The density of the aggregate must be less than the water density or flotation cannot occur. Eq. (12) is valid for laminar flow of bubble motion through the water or for Reynolds number < 1. This was confirmed except for rise velocities of aggregates with multiple bubbles attached to flocs ≥160 μm, where Reynolds number values of 1.0 to 1.5 were found. Still, these values are reasonable for use in Eq. (12).

Fig. 7 presents aggregate rise velocity and density difference predictions for three cases of bubble attachment: 1) one bubble per floc, 2) the maximum number of bubbles that can occupy a floc surface, and 3) half the maximum number. Multiple bubble attachment to large flocs is possible. The maximum number of bubbles that can attach depends on the floc area and bubble area (Tambo *et al.* [22] and Haarhoff and Edzwald [1]). Thus, one 100 μm bubble attaches to flocs of 70 μm or less and then 2 bubbles attach for flocs of 80 μm with the number of bubbles attaching per floc increasing with floc size reaching 12 bubbles for flocs of 200 μm.

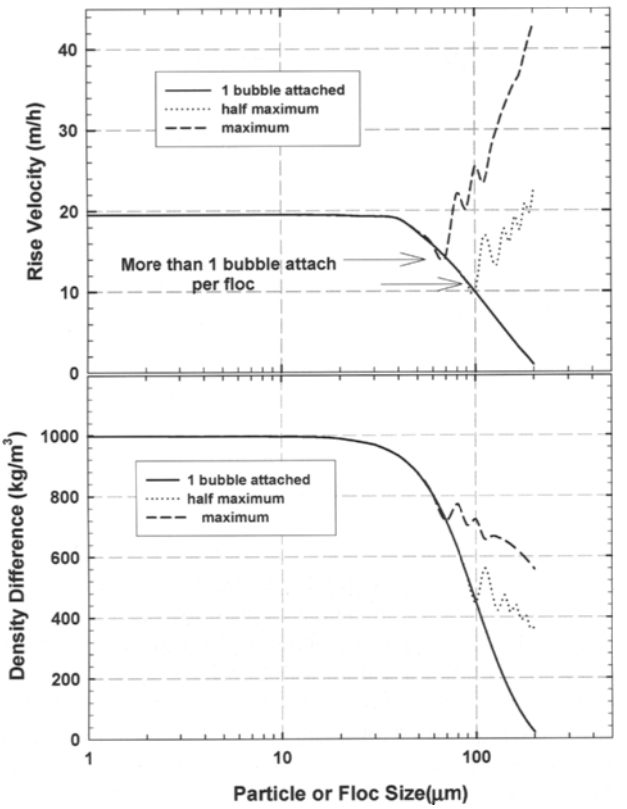


Fig. 7. Aggregate rise velocity and density difference for 1 bubble and multiple bubble attachment of 100 μm bubbles, initial floc density of 1100 kg m^{-3}, and temperature of 20 °C

Fig. 7 provides some interesting insights beginning with the limiting case of allowing one bubble of 100 μm attached per floc. For small flocs of 50 μm or less, the density difference is ≥870 kg m^{-3} (close to the maximum value, which is ρ_w) and so the aggregate rise velocity with one bubble attached to small flocs is approximately the rise velocity of a 100 μm bubble without attached floc – compare the nearly 20 m h^{-1} rise velocity here with that in Fig. 3. Density differences and rise velocities decrease for single bubble attachment to flocs greater than 50 μm. For example, for 100 μm floc or $d_f/d_b = 1$, the density difference is < 500 kg m^{-3}, and the rise velocity is about 10 m h^{-1}. If we limit one bubble attachment to 200 μm flocs or $d_f/d_b = 2$, then the density difference is far too low and the rise velocity decreases to < 1 m h^{-1}. Fig. 7 also shows the effect of multiple bubble attachment per floc on aggregate rise velocity for two cases: half the maximum number of possible bubbles that can attach (2-6 bubbles) and maximum bubble attachment (4-12 bubbles). For floc sizes of 100-200 μm, multiple bubble attachment per floc can yield fairly high aggregate density differences of 400-700 kg m^{-3} producing aggregate rise velocities of 15-20 m h^{-1} for the case of the half the maximum number of bubbles per floc and as high as 20-40 m h^{-1} for the maximum bubble number attachment case. It is important to note that if a water plant keeps floc size small at 50 μm or less, then aggregate rise velocities with single bubble attachment can be as high at

about 20 m h^{-1} as those for large flocs with several bubbles attached per floc; for example 200 μm flocs with 6 air bubbles attached.

4.3. Separation zone

The separation zone of the DAF tank (see Fig. 2) allows for gravity removal of free bubbles and floc-bubble aggregates. The *white water* suspension exits the contact zone by passing over the baffle separating the two zones. Bubbles and aggregates are collected at the top of the tank as floated sludge, and the clarified water (subnatant) exits the separation zone from the bottom of the tank. In practice, separation zone performance and design are based on the Hazen theory that the rise velocities of the bubbles (v_b) and aggregates (v_{fb}) must exceed the separation zone hydraulic loading.

$$v_b \geq v_{hl} = \frac{Q}{A_{sz}} \tag{13}$$

$$v_{fb} \geq v_{hl} = \frac{Q}{A_{sz}} \tag{14}$$

The separation zone area (A_{sz}) should be used above to calculate the DAF separation zone hydraulic loading (v_{hl}), but it is usually about 90% of the tank area including the contact zone. Consequently in practice, nominal DAF tank loadings are reported in terms of the gross footprint area. These may range from 5 to as high as 40 m h^{-1} for recent high rate designs. This trend to high rate systems is discussed in Section 5.2.

Using the simple Hazen concept, Fig. 3 indicates that for the low hydraulic loading of 5 m h^{-1}, free bubbles of about 50 μm or greater will rise to the top of the separation zone and are removed. For hydraulic loadings > 20 m h^{-1}, it is predicted that all free bubbles less than about 100 μm are not separated and will exit the DAF tank. In a similar analysis, Fig. 7 indicates that DAF will remove aggregates consisting of small flocs attached to 100 μm bubbles or larger flocs with multiple bubble attachment for hydraulic loadings up to 20 m h^{-1}.

There is a discrepancy between these predictions and experience. Practical experience and design practice shows that DAF tanks can operate at 20 m h^{-1} and even higher hydraulic loadings with good removal of aggregates and with few bubbles in the tank effluent. A reasonable explanation for the discrepancy between the simple Hazen theory and practice has to do with the hydraulics of flow through the separation zone [1]. Hazen theory assumes plug flow through the separation zone, but this does not occur. DAF tanks, especially with higher hydraulic loadings exhibit stratified flow in which the flow moves horizontally along near the top of the separation zone to the end and then returns in a horizontal flow structure in a layer below the surface [23, 24] before proceeding to the exit at the tank bottom. A simple illustration of this stratified flow is shown in Fig. 8. This flow pattern effectively triples the separation area for flotation thereby tripling the equivalent hydraulic loading for separation of free bubbles and aggregates per Eqs. (13-14). Although the flow patterns through DAF tanks vary with hydraulic loading, recycle, and tank geometry, this simple analysis explains why DAF tanks with hydraulic loadings > 20 m h^{-1} can be designed. Stratified flow effectively increases the separation zone area, and is analogous to inserting a tray (floor) in a sedimentation tank.

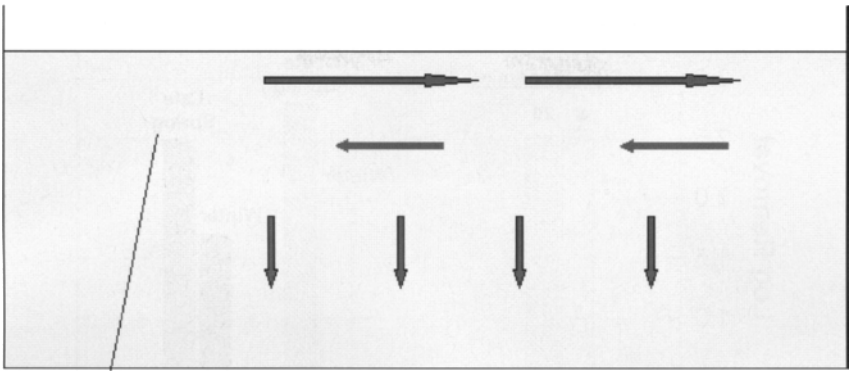


Fig. 8. Illustration of stratified flow in the separation zone

5. APPLICATIONS

5.1. Types of water supplies

DAF is particularly suited to water supplies with characteristics such as those containing low density particles and those yielding low density particles following chemical coagulation. Supply types include those containing 1) natural color, 2) algae, 3) low turbidity (< 10 NTU) and low TOC, and 4) moderate mineral turbidities (10-50 NTU). Those containing color or algae are obvious choices. High quality supplies with low turbidity and TOC are good candidates for DAF treatment compared to direct filtration treatment because DAF clarification provides an additional process for particle and pathogen removal, and compared to sedimentation because high quality source waters produce low density floc difficult to settle without extensive flocculation and often the use of floc aid polymers. DAF performance is also less affected by low temperature compared to sedimentation so supplies in regions that have cold waters are also a good choice for DAF. To generalize, DAF is a good choice for reservoir supplies and should be considered for river supplies that fit the water quality types described.

DAF has several advantages. It is more efficient than settling in removing low density particles, even at the much higher hydraulic loadings – compare 5 to 40 m h^{-1} for DAF to 0.5 to 1 m h^{-1} for conventional settling and 2 to 5 m h^{-1} for high rate settling with plates and tubes The greater removal of particles (turbidity) by DAF versus settling means lower particle loading to filters is consequently found, which means granular media filters can be designed at higher rates or longer filter runs (higher water production) are achieved. It is common to find DAF effluent turbidities < 1 NTU and lower than 0.3 NTU when coagulation is optimum. The overall footprint for DAF plants can be small: smaller floc tanks because of shorter flocculation pretreatment time, smaller DAF tank areas compared to sedimentation tanks, and smaller filter area if designed at higher rates as explained above. DAF tanks that accumulate floated sludge and remove by scrapping with flight paddle or brush systems produce sludge with higher solids content (1-5%) than settled sludges, reducing sludge treatment.

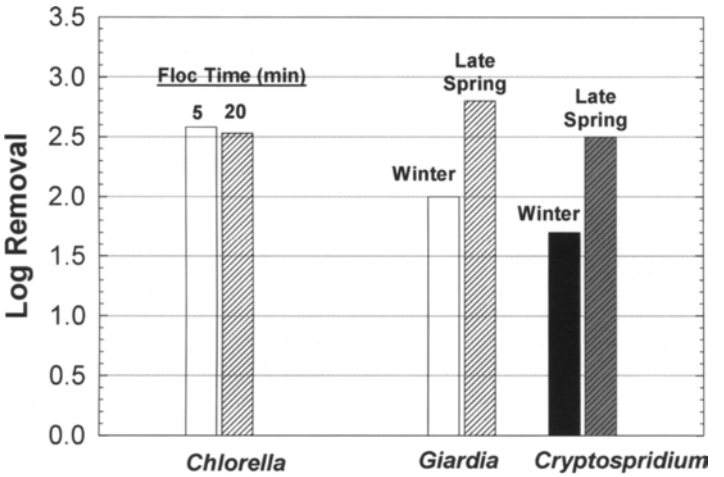


Fig. 9. Log removals by DAF for algae (*Chlorella vulgaris*) and two protozoa (*Giardia* and *Cryptosporidium*) (Data reported in Edzwald *et al.* [25, 26])

5.2. Algae, *Giardia,* and *Cryptosporidium*

Algae have low densities, near that of water, which enables them to be nearly buoyant and reside near the surface of reservoirs and maintain photosynthetic activity. If you consider that many algae have sizes of 1 to 20 μm, then Fig. 7 shows that by simply attaching one air bubble to each alga can produce high rise velocities (~ 20 m h^{-1}) and removal. DAF is thus a widely used method to treat supplies with algae in which removals of 90 to 99% are achieved [5]. Fig. 9 shows these high DAF removals for *Chlorella vulgaris* from a reservoir supply with alum as the coagulant at pH 6.5 [25]. The same data also illustrate that DAF is as effective with a short pretreatment flocculation time of 5 min versus 20 min in removing *Chlorella*.

Recent studies have demonstrated that DAF is more effective than settling in removing *Giardia,* and *Cryptosporidium* [26, 27]. Fig. 9 summarizes some data taken from Edzwald *et al.* [26]. These data were obtained from pilot plant studies conducted at a DAF hydraulic loading of 15 m h^{-1}, and for pilot runs conducted under cold (winter, 2-3 °C) and moderate (late spring, 13-14 °C) water temperatures. Higher log removals were obtained for the warmer spring water temperatures in accordance with theory about temperature effects on rise velocities, but generally log removals for both *Giardia* and *Cryptosporidium* were very good, about 2-log or higher compared to plate settling which was 1-log or less [26].

5.3. Developments in DAF

Developments in DAF plant design are summarized, especially those dealing with flocculation pretreatment and with DAF hydraulic loadings. Prior to about 1990, it was typical to design flocculation tanks much like flocculation ahead of sedimentation. Long flocculation tank detention times of 20 to 30 min with 2 to 3 stages of tapered mixing intensity were common. Using the principle obtained from the contact zone modeling presented in Section 3.3 that good performance can be achieved with small flocs of 10s of

microns, Edzwald and co-workers demonstrated this through laboratory [25] and pilot-scale DAF studies [28-30]. This led to the acceptance that floc tank design should be tailored to DAF plants, and that the flocculation tank effluent should produce flocs of 10s of microns, not flocs of 100s of microns as desired for settling processes. After 1990, flocculation times for many DAF plants have decreased to about 10 min, with at least one design at 5 min (Croton plant for New York City).

Again prior to about 1990, DAF tanks were sized for hydraulic loadings of 5 to 10 m h^{-1}. The considerable theoretical and applied research conducted on DAF in the last 25 years has led to a better understanding of the process. This knowledge and optimization of the process in design studies have pushed the technology so that much higher DAF hydraulic loadings are now used leading to very small footprint areas. In the 1990s, DAF plants were designed at 10 to 20 m h^{-1}. In the last 5 years, successful DAF operation at 20 to 40 m h^{-1} with a short flocculation time of 5 min was demonstrated through pilot-scale studies [30, 31]. A DAF manufacturer has ingeniously developed a porous plate that is placed at the bottom of the DAF tank for better flow distribution in the separation zone and for collection of the subnatant [24]. Several new plants have been built using this system, e.g., a facility in the USA (West Nyack, NY) designed at 30 m h^{-1} and one in Tampere (Finland) at about 40 m h^{-1}. In certain situations where land is scarce or expensive, it is possible to reduce total water plant footprint area by placing DAF directly over the granular media filter. It does limit the hydraulic loading,, because the filtration controls the overall process. However, even with this design facilities have been built with a hydraulic loading of 15 m h^{-1}.

ABBREVIATIONS

DLVO	Derjaguin Landau Verwey and Overbeek

LIST OF SYMBOLS

A	area in plan
C	mass concentration
d	size or diameter
e	percent of air released
g	gravitational constant of acceleration
k	air deficit concentration
k_b	Boltzmann's constant
K	shape factor
n	number concentration
N	number of bubbles per floc
Q	volumetric flow rate
R	recycle ratio
v	rise rate or hydraulic loading
t	detention time
T	absolute temperature
α	attachment efficiency factor
ϕ_b	bubble volume concentration
π	mathematical constant
η	single collector efficiency
μ	dynamic viscosity

ρ	density
subscript *b*	indicating *bubble*
subscript *cz*	indicating *contact zone*
subscript *D*	indicating *Brownian Diffusion*
subscript *e*	indicating *effluent*
subscript *f*	indicating *floc*
subscript *fb*	indicating *floc-bubble*
subscript *hl*	indicating *hydraulic loading*
subscript *i*	indicating *influent*
subscript *I*	indicating *Interception*
subscript *o*	indicating *plant through-put*
subscript *p*	indicating *particle*
subscript pb	indicating *particle-bubble*
subscript *r*	indicating *recycle*
subscript *s*	indicating *saturation*
subscript *sz*	indicating *separation zone*
subscript *S*	indicating *Settling*
subscript *T*	indicating *Total*
subscript *w*	indicating *water*

REFERENCES

[1] J. Haarhoff and J.K. Edzwald, J. Water Supply Res. Technol. - Aqua, 53 (2004) 127.
[2] T. Takahashi, T. Miyahara and H. Mochizuki, J. Chem. Eng. Jpn., 12 (1979) 275.
[3] J.A. Dean (ed.), Lange's Handbook of Chemistry, 13th Ed., McGraw Hill, NY, 1985.
[4] R.H. Perry and D.W. Green (eds.), Perry's Chemical Engineering Handbook, 6th Ed., McGraw Hill, NY. 1984.
[5] R. Gregrory, T.F. Zabel and J.K. Edzwald, Water Quality and Treatment, R.D. Letterman (ed.), 5th edition, McGraw Hill, NY, 1999.
[6] R. Clift, J.R. Grace and M.E. Weber, Bubbles, Drops, and Particles, Academic Press Inc., NY, 1978.
[7] B.V. Derjaguin, S.S. Dukhin and N.N. Rulyov, Colloid Surf. Sci., 13 (1984) 71.
[8] J.K. Edzwald, J.P. Malley and C Yu, Water Supply, 8 (1990) 141.
[9] J.K. Edzwald, Water Sci. Technol., 31 (1995) 1.
[10] K.M. Yao, M. T. Habibian and C.R. O'Melia, Environ. Sci. Technol., 5 (1971) 1105.
[11] L.R. Flint and W.J. Howarth, Chem. Engr., Sci., 26 (1971), 1155.
[12] D. Reay and G.A. Ratcliff, Can. J. Chem. Eng., 51 (1973) 178.
[13] A.R. Shawwa and D.W. Smith, Water Sci. Technol., 38, (1998) 245.
[14] G.J. Schers, and J.C. Dijk, Chemical Water and Wastewater Treatment II, R. Klute and H.H. Hahn (eds.), Springer Verlag, NY, 1992.
[15] M.Y. Han, J. Water Supply: Res. Technol. - Aqua, 51 (2002) 27.
[16] D.M. Leppinen, J. Water Supply: Res. Technol. - Aqua, 49 (2000) 259.
[17] W.A. Ducker, Z. Xu and J.N. Israelachvili, Langmuir, 10 (1994) 3279.
[18] S. Dockko and M. Y. Han, Water Sci. Technol., 50 (2004) 207.
[19] S. Lu, Colloids Surf., 57 (1991) 73.
[20] G.L. Collins and G.J. Jameson, Chem. Eng. Sci., 31 (1976) 985.
[21] D.M. Leppinen and S.B. Dalziel, J. Water Supply: Res. Technol. - Aqua, 53 (2004) 531.
[22] N. Tambo, Y. Matsui and K. Fukushi, World Congress of Chemical Engineering, Tokyo, 1986.
[23] M. Lundh, L. Jonsson and J. Dahlquist, Water Sci. Technol., 43 (2001) 185.

[24] H.J. Kiuru, Proc. 4th Int. Conference on Dissolved Air Flotation in Water and Wastewater Treatment, Helsinki, 2000.
[25] J.K. Edzwald and B.J. Wingler, J. Water Supply: Res. Technol. - Aqua, 39 (1990) 24.
[26] J.K. Edzwald, J.E. Tobiason, L.M. Parento, M.B. Kelley, G.S. Kaminski, H.J. Dunn and P.B. Gallent, J. - Am. Water Works Assoc., 92 (2000) 70.
[27] J.K. Edzwald, J.E. Tobiason, C.T. Udden, , G.S. Kaminski, H.J. Dunn, P.B. Gallent and M.B. Kelley, J. Water Supply: Res. Technol. - AQUA, 52 (2003) 243.
[28] J.K. Edzwald, J.P. Walsh, G.S. Kaminsky and H.J. Dunn, J. - Am. Water Works Assoc., 84 (1992) 92.
[29] M.T. Valade, J.K. Edzwald, J.E. Tobiason, J. Dahlquist, T. Hedberg and T. Amato, J. - Am. Water Works Assoc., 88 (1996) 35.
[30] J.K. Edzwald, J.E. Tobiason, T. Amato and L.J. Maggi, J. - Am. Water Works Assoc., 91 (1999) 41.
[31] T. Amato, J.K. Edzwald, J.E. Tobiason, J. Dahlquist and T. Hedberg, Water Sci. Technol., 43 (2000) 19.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 7: Membrane filtration processes and fouling

A.G. Fane[a], Wei Xi[b] and Wang Rong[b]

[a] School of Civil and Environmental Engineering
University of New South Wales,
Sydney NSW 2052.

[b]Institute of Environmental Science & Engineering
School of Civil and Environmental Engineering
Nanyang Technological University, Singapore

1. INTRODUCTION

Until relatively recently the major application of membrane technology to produce drinking water has been for desalination of saline waters by reverse osmosis. However since about 1990 technical advances in the lower pressure membrane processes coupled with more stringent water quality standards has seen these methods increasingly applied to the production of drinking water from 'traditional' non-saline sources. The major reasons favoring this trend are,

- membrane technologies are capable of removing a wide range of contaminants from particulates , including pathogens, to ionic species,
- the technology only requires modest energy usage (no phase change),
- membranes are modular allowing small (decentralized) to large (centralized) facilities,
- most importantly, the economics of water treatment by membranes has become increasingly competitive with conventional processes.

This chapter discusses the use of membrane technologies in drinking water production with the emphasis on non-saline water sources, such as surface and ground waters. Contaminants removed by the membranes include turbidity, pathogens, colloids and natural organic matter (NOM). The focus is on surface interactions and membrane fouling. Section 2 provides a brief description of the relevant membrane processes, membrane hardware and modes of operation. Section 3 is a brief overview of membrane interactions and fouling mechanisms and Section 4 provides a review of reported studies on fouling in water treatment with an emphasis on the role of NOM. In section 5 some of the techniques for fouling mitigation are discussed.

2. MEMBRANE TECHNOLOGY

2.1. Membrane processes

The pressure-driven membrane processes applied to water treatment make use of finely microporous membranes, usually polymeric, in the form of sheets, tubes or hollow fibres.

Table 1 summarizes the key features of these processes. For non-saline water treatment the methods used are microfiltration (MF), ultrafiltration (UF) and nanofiltration (NF). All three processes remove turbidity, bacteria and protozoa. Virus removal can be achieved by UF and NF. Substantial removal of NOM requires NF or low pressure MF or UF coupled with coagulants or adsorbents. Reverse osmosis (RO) is applied if the raw feed water is saline. Because of their different pore sizes and applications the various membrane processes respond in different ways to foulants in raw water, as discussed in 3.1 below.

Table 1
Membrane processes and characteristics for water treatment

Membrane process	Pore size (nm)	Species retained	Typical fluxes (litres/m^2hr)**
Microfiltration (MF)	100-1000	Bacteria, turbidity, particles	50-200
Ultrafiltration (UF)	5-50	Virus, colloids, macrosolutes	50-100
Nanofiltration (NF)	2-5	NOM, sugars, divalent ions	10-50
Reverse Osmosis (RO)	< 1*	Monovalent ions, low MW org.	10-30

* No detectable pores ** For water treatment applications

2.2. Modes of operation

There are important operational modes that need to be considered as these influence the type of module used (see 2.3) and the kinetics and degree of fouling; understanding the basic differences between the modes may help us interpret apparently conflicting data. Fig. 1 introduces these concepts, which are:

i) Crossflow: in this situation surface shear is applied as the feed moves across the surface of the membrane, usually pumped in tangential flow but it could be stirred (laboratory systems) or induced by bubbles or other agents (vibrations , rotations etc). The surface shear provided by crossflow operation encourages back-transport of retained species and this helps to minimize surface interactions and fouling. Crossflow is typically used for NF and RO and for feeds with high solids content in MF and UF. In theory this mode can achieve steady-state, and with very idealized systems it does so, but in practice some degree of fouling tends to occur, so that flux (Fig. 1(c)) or transmembrane pressure (TMP) (Fig. 1(d)) continue to change with time. It is important to note that 'batch concentration' with crossflow will not achieve steady state as the feed concentration increases, but polarization will be controlled to some extent by the crossflow.

ii) Deadend: operation without surface shear which allows retained species to accumulate at the membrane and then to be removed by intermittent backwash etc (Fig. 1 (a) and (b)). Thus deadend operation inevitably accepts the deposition of potential foulants but it tends to require less energy and is favored by feeds with low solids content. Nowadays it is frequently used in water treatment with low pressure (MF and UF) membranes.

iii) Constant pressure operation: as depicted in Fig. 1 (c) with a fixed, constant, pressure and crossflow the flux is initially high and drops rather rapidly to a pseudo 'steady state' as cake or retained species (particles or solutes) build up. For constant pressure and deadend operation (Fig. 1 (a)) the flux can effectively decline to zero if the cake is not backwashed. Laboratory studies have tended to favour experiments at constant pressure. It is obvious that with constant pressure operation the initial rate of fouling would be high, due to the high flux, but that eventually the rate of fouling can slow down considerably (although it may not reach a zero rate).

iv) Constant flux operation: is more typical of large-scale plant where the system is required to produce a given amount of permeate each day. Fig. 1 (d) shows that for crossflow operation the steady flux will lead to a gradual rise in the required TMP if fouling occurs. This cannot continue indefinitely and occasional cleaning will be required. For deadend mode, Fig. 1 (b) shows that the TMP steadily increases until backwash occurs and a new filtration cycle commences. With constant flux the initial rate of fouling can be relatively slow, but foulants continue to arrive at the same rate over time, unlike constant pressure operation.

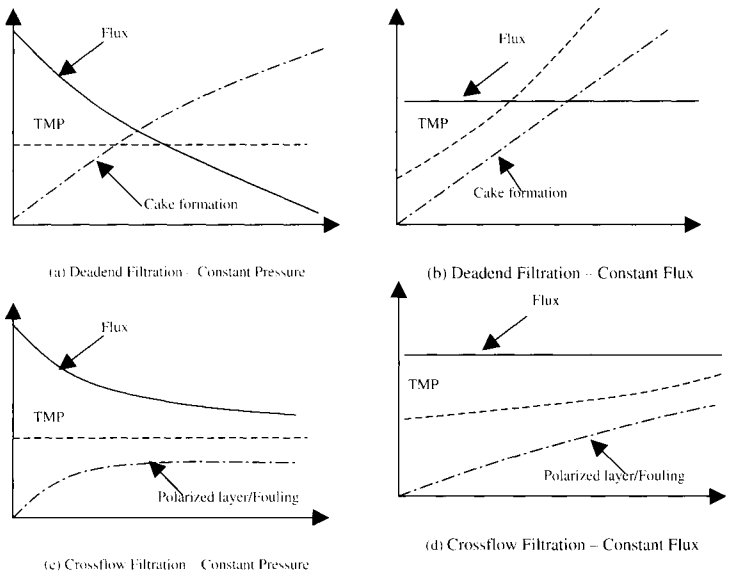


Fig. 1. Mode of operation: crossflow vs deadend; Constant pressure vs constant flux

2.3. Membrane modules

Membranes are housed in modules that serve to support the membranes and facilitate 'fluid management'. Table 2 lists the various modules and refers to their typical mode of operation. More detailed descriptions of module types and their characteristics can be found in Fane [1]. Crossflow modules are typically housed in a pressure vessel, such as the spiral wound module which is the standard for large scale NF and RO.

The modules commonly used in water treatment for MF and UF are hollow fibres contained in pressure vessels or submerged in tanks; tubular systems tend to have niche applications in water treatment.

Table 2
Modules types and characteristics for water treatment

Module type	Module features	Advantages	Disadvantages	Mode of operation
Tubular (C)*	Tubes (5-25mm) in series /parallel	Dirty feeds with no pretreatment	High energy cost. Low packing density.	Crossflow**
Hollow fibre (C)	Contained in a pressure vessel	Low energy. High packing d.	Pretreatment may be required.	Dead end or Crossflow***
Hollow Fibre (S)*	Submerged fibres in a tank	Low energy. High packing d.	Hydrodynamics and flow distribution may be poor.	Dead end
Spiral Wound (C)	Flat sheets and spacers wrapped in pressure vessel	High packing d. Good hydrodynamics	Pretreatment may be required. Moderate energy cost.	Crossflow

*C=Pressure Vessel, S=Submerged **Deadend possible ***Crossflow if turbidity rises

3. MECHANISMS OF FOULING AND POLARISATION

3.1. Definitions

Fouling is the deposition of material on or within the structure of membranes that is not readily reversed by simply releasing the pressure or by backwashing. Some foulants may be removed by physical or chemical cleaning but some fouling may be irreversible. The effect of fouling is to reduce membrane permeability (flux declines at constant pressure, or pressure rises at constant flux), and to alter solute retention (tends to increase for MF and UF, and decrease for NF and RO). Thus membrane fouling deteriorates membrane performance, increases operating cost and ultimately shortens membrane life.

The fouling effect can be described by Darcy's law as follows,

$$J_v = \frac{dV}{A\,dt} = \frac{\Delta p}{\eta (R_m + \sum R_i)} \qquad (1)$$

J_v is the volumetric flux of permeate,
V the total volume of permeate,
A the membrane surface area;
Δp the pressure drop imposed across the fouling layer and membrane,
η the viscosity of the permeate,
R_m the resistance of clean membrane,
R_i the various additional resistances during the filtration process as a result of the polarization (see below) and subsequent fouling.

The pressure drop, Δp, is frequently referred to as the transmembrane pressure or TMP. It is often convenient to lump the deposit resistances into a 'cake' resistance, R_c, so Eq. (1) is written,

$$J_v = \frac{\Delta p}{\eta (R_m + R_c)} \qquad (2)$$

Eqs. (1) and (2) apply to all modes of operation (crossflow, deadend, constant pressure or flux). For conditions where there is a significant osmotic pressure the driving force in Eqs. (1) and (2) becomes $(\Delta p - \Delta \pi)$.

Polarisation is the accumulation of retained species adjacent to the membrane and is known as concentration polarisation (solutes) or particle polarisation (particles). By definition, polarization is a reversible phenomenon (for example by stopping flux) and once it becomes irreversible it is fouling. Due to concentration polarization the concentration at the membrane surface (C_W) will be greater than in the bulk feed (C_B). The well-known film model provides a useful relationship for C_W in crossflow operation,

$$C_W = C_B \exp(J_v / k) \qquad (3)$$

Where J_v is flux and k is the boundary layer mass transfer coefficient, which increases with crossflow and decreases with molecular weight. This equation is important as it shows that the surface concentration, C_W that can determine fouling, is very dependent on the J_v/k ratio. For constant pressure crossflow tests, the ratio J_v/k will not be constant and this may complicate interpretation.

For particles and colloids in the > 0.5 μm range crossflow produces shear-induced 'back transport' from the membrane; the effect decreases as size decreases and small particles tend to deposit first. Colloids < 0.1 μm tend to behave as Brownian species with back transport decreasing as size increases. Around 0.1 to 0.2 μm the mechanisms crossover and colloids of this size tend to have the lowest 'backtransport' and become foulants. This could be modified if electrokinetic interactions are significant, and this is dependent on the materials and solution chemistry (pH, ionic content etc).

For deadend operation 'polarisation' is inevitable and since it usually applies to MF and UF the polarization is particulate, forming a cake. However, with mixed feeds, like raw water, fine colloids, macro and microsolutes can pass into the cake and either be partially retained, modifying the cake, or pass into and through the membrane.

3.2. Fouling mechanisms and interpretation

Fouling can take three 'generic' forms (depicted in Fig. 2),

- pore restriction: a reduction in pore diameter due to adsorption/deposition. Small changes in d_{pore} have a significant effect as membrane permeability varies with $[d_{pore}]^4$
- pore plugging/blocking: a loss of pore density due to species, such as colloids or aggregates, blocking the entrance or internal passageway of pores
- cake formation/surface deposition of individual particles, aggregates and precipitates. This includes deposition favored by the raised concentration at the membrane surface, such as inorganic scale formation, crosslinking of hydrogels, and aggregation of colloids.
- biofouling: a type of surface deposition due to bacterial adhesion and biofilm growth, where the cake comprises bacterial colonies and expressed biopolymer.

The biofouling mechanism is more prevalent in wastewater processing, due to the availability of nutrients. The other 3 mechanisms can be found in water treatment, with (i), (ii) and (iii) occurring in MF and UF, either in sequence or simultaneously, and (iii) most likely for NF as the very small pore size tends to lead to surface deposition. For MF and UF membranes, which have a distribution of pore sizes around a mean, the effect of colloids and macrosolutes may vary locally.

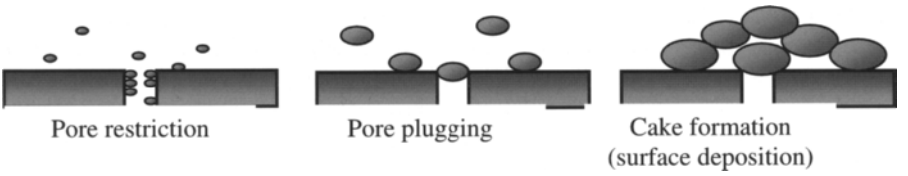


Fig. 2. Schematic of fouling mechanisms

During operation a membrane may go through several fouling stages. Fig. 3 depicts the flux history of a membrane operated under constant pressure conditions. For constant flux conditions the history would be TMP change which would tend to mirror Fig. 3. Four phases are possible. Phase I is a possible change during water flux testing of the new membrane. Phase II is the initial flux drop when put into use, and this will be due to concentration polarization, pore plugging and adsorptive interactions. Phase III is the longer term operation where flux will usually still continue to decline, and phase IV is the recovery phase during membrane cleaning. The measured water fluxes, $J_{W1,2,3}$ will provide some evidence of fouling (see below).

Fouling data are presented in various ways which include the following:

- Flux decline ratio (J_v / J_o), where J_o is the initial flux at the start of the filtration process. This ratio is illustrative but potentially misleading, as it is not linearly related to fouling resistance. Typically J_v / J_o drops sharply and then changes more slowly, but this could still correspond to a steady increase in fouling resistance. When membranes are compared, based on J_v / J_o, a membrane with an initially high flux will 'look' more fouling prone than one with an initially low flux.
- Fouling resistance (R_c) can be obtained from flux and TMP data by applying Darcy's Eq. (2). This is a better measure of analysis and comparison of different membranes and protocols. Comparing resistances after washing and cleaning allows estimation of 'easily reversible' and 'irreversible' foulant.
- Transmembrane pressure (TMP) - for constant flux processing the parameters dTMP/dt or dTMP/dV are useful and direct measures of the fouling effect (TMP is proportional to (R_m+R_c) via Eq. (2)).

3.3. Foulant species

Any species in the feed water is a potential foulant and its impact will depend on its characteristics and concentration as well as membrane properties such as pore size, charge, and hydrophobicity, module properties, mode of operation (see 2.2) and applied flux.

The potential membrane foulants in raw waters include colloidal and suspended particles as well as microorganisms, dissolved macromolecular organic substances and inorganic compounds. Natural organic matter (NOM) with a broad spectrum of molecular weights and size distributions, functional groups and sub-structures has been shown to be the major foulant in the membrane filtration of natural waters [2-7]. Table 3 illustrates the range of species found in river water by Croue et al. [7], based on the fractionation protocol of Leenheer et al. [8]. The term NOM covers the particulate colloids (POM), 'dissolved' colloids and macrosolutes (DOM), and the humics. Since NOM is from environmental origin its composition is highly variable both in space and time, which makes comparisons rather difficult. One of the major groups in NOM are humic substances, which include a complex mixture of organic compounds such as humic acids, fulvic acids and other hydrophilic

compounds. As noted above humic substances have been found to vary in different water sources. In general, humic substances have characteristic lengths around 1nm to 10 nm and molecular weight from few hundreds to approximately 100,000Da. In Table 3 humics are indicated as < 3.5kD (based on Leenheer), but this is arbitrary and the literature refers to humics with molecular weights ranging from 0.2 to 100 kD. They contain both aromatic and aliphatic components with carboxylic and phenolic groups attached to the aromatic rings. As a result, humic substances can be negatively charged as well as having hydrophobic and hydrophilic (humic and non-humic) components [4-9]. The importance of the charge character is that it can lead to significant interactions with the membrane and with divalent cations, such as calcium which is usually present in raw water (see 4.2(iii)). As indicated in Table 3 the humics fraction also has neutral components.

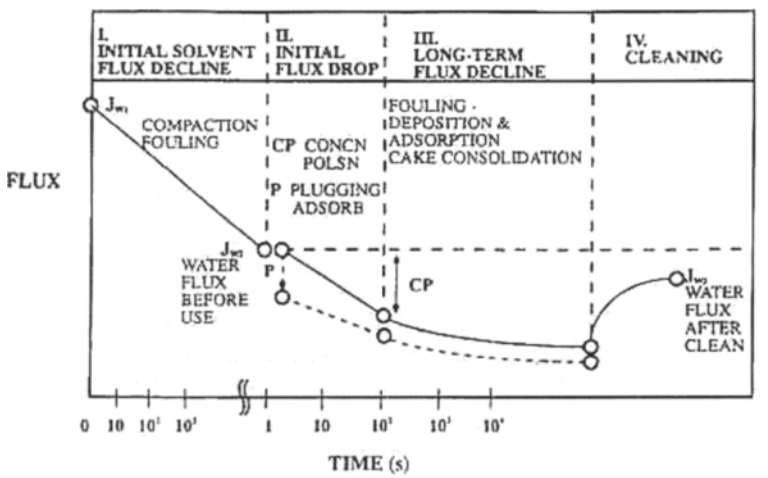


Fig. 3. Typical flux history during membrane fouling

Many studies have used the International Humic Substances Society (IHSS) standards for comparative purposes, but these presumably exclude the POM fraction. In actual raw water the POM fraction (< 1μm >0.45 μm) can be significant, and Croue et al. [7] found it to be 30-40% of the organic colloid fraction (> 3.5kD). It was also a very significant foulant (see 4.2(i) below).

4. REVIEW OF NOM MEMBRANE FOULING STUDIES

4.1. Mechanisms of NOM fouling

From the above it is obvious that NOM fouling of membranes is a complex process (for example [6-11]) with many influencing factors, not least being the variable nature of raw waters. This section reviews the extensive literature that has evolved over the past two decades. At the outset it should be noted that not one single component of NOM has been identified as the chief culprit for membrane fouling, and there are apparent conflicts in the literature. However some components are clearly implicated. In order to aid the discussion Table 4 attempts to summarise a number of studies and gives details of the key features, such as membrane process, mode of operation etc. From Table 4 it is evident that the types of

NOM studied vary widely from 'standard' IHSS material, to local NOM samples to soil-derived NOM (Aldrich). It is also apparent that for MF the mode of operation in research studies is usually crossflow (stirred) and constant pressure, whereas the industry tends to operate at deadend and constant flux, with backwash. For UF and NF the preponderance of studies have used constant pressure, and again the industry favours constant flux (constant production rate).

Table 3
NOM and other foulant species in raw water (based on [7, 8])

Foulant species	Characteristic	Comment
Particles (POM)	Colloidal (< 1μm >0.45 μm)	Bacterial cell wall residues (polysaccharides/proteins)
Dissolved organic matter (DOM)	Colloidal and macromolecular (<0.45 μm > 3.5kD)	Bacterial cell wall residues (polysaccharides/proteins)
Humics (< 3.5KD)	Hydrophobic – neutral – charged Transphilic Hydrophilic – neutral – charged	Proteins, aminosugars, polysaccharides, polyhydroxyaromatics, organic acids
Inorganic ions	Divalent cations (particularly calcium)	Aggregates NOM species. Bridges/binds NOM to membrane surface.

As a general statement, NOM fouling on membranes can be caused by several interacting phenomena. For example, Seidel and Elimelech et al. [9] found that permeation drag controlled by the permeate flux and calcium binding to NOM were the major causes for the development of a densely compacted fouling layer on the membrane surface in the process of NOM nanofiltration. NOM fouling and filtration performance were governed by the *coupled* influence of chemical and hydrodynamic interactions. For these reasons researchers such as Amy et al. [4, 5] report fouling at controlled values of $J_v k^{-1}$ (see equation 3).

The study of Lahoussine-Turcaud et al. [10] on the ultrafiltration of natural water with different organic and inorganic macromolecules indicated that organic matter had a much greater effect on the flux decline than inorganic colloids. Similar results were reported by Bersillon et al. [11] in the analysis of the deposited cake formed on the membrane surface after natural water filtration. This heterogeneous process can be generally attributed to the various mechanisms discussed in 3.2 including pore restriction and plugging as well as cake/surface deposition. Fouling by NOM adsorption is proposed as a significant factor and is a consequence of the interactions between the organic compounds and membranes either at the membrane surfaces or within the pores. The interactions include chemical, physical and electrostatic effects. Chemical adsorption of organic compounds on the membrane materials is generally irreversible because of the relatively high strength of attachment, and recovering the permeate flux is difficult. However physical adsorption shows a weak attachment to the membrane surface and the flux decline due to this mechanism is reversible. The electrostatic interaction between the membrane materials and NOM may be either attractive or repulsive, depending on their charged functional groups. Ionic species, such as calcium may play a role.

Table 4
Summary of selected NOM membrane fouling studies

Membrane Process	Mode* of Operation	Feed properties and preparation	Factors** influence fouling	Proposed fouling species and mechanism	Ref.
1. MF	Xf/CP	Aldrich NOM solution	(i), (iii)	Aldrich NOM, cake formation, pore blocking, adsorption, calcium role	[14], [15], [35], [49]
	Xf/CP	Surface water, filter< 0.45μm	(i), (ii), (iii)	Hydrophilic neutral compounds, hydrophobic acids, transphilic acids, hydrophilic charged compounds colloidal material, adsorption, cake formation, calcium role	[25], [70]
	De/CP	River water	(i),(ii), (iii)	Colloidal material, NOM ,small, neutral,hydrophilic	[32],[61]
2. UF	De/CP	Surface water, river water , filter< 1μm ;	(i), (ii), (iii), (iv)	Organic colloids (hphilic) polysaccharides/proteins & aromatics;	[7], [8], [26]
	Xf/CP	River water, reservoir water<0.2μm, ground water, lake water; Aldrich NOM, IHSS NOM	(i), (ii), (iii), (iv)	Hydrophilic NOM, adsorption, cake formation; steric and electrostatic exclusion humic acid adsorption, concentration polarization, aggregate deposition	[5],[17], [21],[23], [27],[28], [31], [37], [42], [59]
	Xf/CF	Reservoir water, ground water	(i), (ii), (iii), (iv)	Hydrophilic NOM, hydrophobic NOM; cake layer formation	[16], [26], [39], [53], [61]
3. NF					
	Xf/CP	River water, reservoir Water, IHSS NOM; Aldrich humic acid	(i), (ii), (iii), (iv)	Hydrophilic NOM, hydrophobic NOM; adsorption, cake formation permeation drag, electrostatic double layer interactions calcium binding	[9], [18], [28],[34], [40],[42], [43],[44] [47]
	Xf/CF	Lake water river water	(ii),	NOM adsorption	[38], [71]

De = deadend, Xf = crossflow, CP = constant pressure, CF = constant flux
** (i) NOM properties (ii) membrane type (iii) solution conditions (iv) operating conditions.

As Table 4 shows there have been many studies of the mechanisms of membrane fouling by NOM. Mallevialle et al. [12] examined the fouling layer formed during the microfiltration (MF) and ultrafiltration (UF) of natural waters. They reported that NOM served as a 'glue' for the inorganic clays in the fouling layer, this is analogous to the proposed role of extracellular polymeric substances (EPS) in biofouling of membranes. Kalya et al. [13] pointed out that organic matter was a major component of the fouling layer deposited on a MF membrane after the filtration of eutrophic lake water.

The study of Yuan et al. [14, 15] on the microfitration of NOM indicated that large humic acid aggregates collected on the membrane surface initially and a cake layer of humic acid began to form with relatively little internal fouling within the membrane pores. Subsequent humic acid fouling was accelerated by this initial humic acid deposit. Intra- and intermolecular electrostatic repulsions were thought to be responsible for the increased fouling. Additionally, the solution properties such as pH and salt have significant effects on the fouling behavior. The repulsions were reduced at low pH and the presence of divalent calcium caused an increase in the adsorption and aggregation of humic acids. It should be noted that Yuan et al. used NOM derived from soils, rather than natural waters in most of their studies. It has been shown that NOM derived from soils contains some organic compounds of larger molecular weight than aqueous NOM, and there are also differences in functional groups, and hydrophobic /hydrophilic character between the NOM derived from soils and aqueous NOM.

In other studies, the mechanisms of ultrafiltration membrane fouling have been shown to be cake formation, pore restriction by adsorption and pore plugging. Crozes et al. [16] observed that both reversible and irreversible fouling could be found in the treatment of natural river water with UF membranes. Taniguchi et al. [17] examined the modes of NOM fouling during ultrafiltration and showed that cake formation was the dominant mechanism of UF fouling for large molecular weight cutoff (MWCO) ultrafiltration membranes. Cake formation was identified by a constant rate of increase in membrane resistance with permeate throughput and was independent of pore size over a range of 10kDa to1000 kDa MWCO. However, the lower MWCO membranes exhibited some irreversible fouling suggesting that low molecular weight species penetrated into the pore structure.

For NF membranes, several studies have shown that the combination of steric exclusion and electrostatic exclusion could be the main mechanisms for reducing NOM fouling. Hong and Elimelech [18] investigated the role of chemical and physical interactions in natural organic matter fouling of nanofiltration membranes. They recognized that the NOM fouling involved interrelationship (coupling) between the physical and chemical interactions. The rate of fouling was controlled by an interplay between permeation drag and electrostatic effects. Braghetta et al. [19] also indicated that neutralization of charge, electric double layer compression and the apparent shift in the conformation of NOM on neutralization are the main reasons for inducing NOM fouling in NF membranes. The influence of calcium, retained by NF, can also be important (see below).

4.2. Factors which influence NOM fouling

The nature of NOM such as size, hydrophobicity and charge density of species, the properties of the membrane such as MWCO, hydrophobic/hydrophilic character and surface charge (electrostatic repulsion), the chemistry of the water solution such as pH and ionic environment, and the operating conditions of the system all have impacts on membrane fouling [20-23]. The following sections highlight the effects of some of these factors.

4.2.1. Impact of NOM properties

Extensive studies have been conducted to evaluate the roles of NOM properties in fouling [for example, 21-26]. The relevant NOM properties include bulk NOM concentration, size and molecular mass distribution, the structure (aromatic/aliphatic, hydrophobic/hydrophilic), and charge density.

Although NOM rejection is anticipated to be mainly dependent on the NOM molecular sizes when other factors are excluded, NOM has a wide range of sizes and structures, and

since membranes normally have a pore size distribution, it is not straightforward to illustrate NOM rejection simply by size exclusion effects [24, 25]. It is commonly believed that a fraction of NOM with molecular weights smaller than the membrane pores can be adsorbed on the membrane surface and reduce the cross sectional area for permeation, while larger fractions can block pore entrances and contribute to cake or gel formation [23, 26].

Nilson and DiGiano [24] reported that only the large molecular weight fraction of NOM contributed to the formation of a fouling layer during NF filtration. Kalya et al. [13] found that compounds larger than 100 kDa were the major foulants in a MF system. Fan and Harris [25] also indicated that the higher molecular weight fractions of NOM (>30 kDa) were responsible for the greater flux decline on a hydrophobic MF membrane. In recent work on NOM removal by a charged hydrophobic UF membrane, Cho et al. [28] suggested that the larger size NOM compounds played a major role in influencing NOM fouling during the membrane filtration process. However, Maartens et al. [26] observed that a mixture of small and large organic compounds exhibit a greater fouling effect than a fraction of only large organic compounds for a UF filtration system. It is likely that all three mechanisms depicted in Figure 2 occur to some extent, and that internal fouling is more irreversible.

The hydrophobicity/hydrophilicity of NOM is also found to be one of the main factors in influencing membrane fouling in natural water treatment. A vast amount of research regarding this aspect has been presented in the literature in recent decades [for example, 21-28]. There is strong evidence that the hydrophobic components in NOM are mainly responsible for the fouling of hydrophobic membranes due to the interaction between the hydrophobic NOM and the hydrophobic membrane surface. For instance, Nilson and DiGiano [24] evaluated the filtration performances of hydrophobic NF membranes and found that the hydrophobic fraction of NOM significantly affects the flux decline. The hydrophobic NOM was more readily removed by the membranes than hydrophilic fractions. Jucker and Clark [27] demonstrated that humic macromolecules adsorbed more favorably onto hydrophobic membranes. The hydrophilic components were thought to impact water quality less than the hydrophobic components.

Cho et al. [28] suggested that flux decline resulting from hydrophobic NOM adsorption on a relatively hydrophilic membrane surface was greater than that with hydrophilic-NOM feed-water. In another study on the retention of NOM by a negatively charged hydrophobic UF membrane, Cho et al. [29] concluded that the retention of hydrophobic compounds of NOM was higher than for the hydrophilic compounds because of the combined hydrophobic interaction and electrostatic exclusion between NOM and the membrane surface.

The components of NOM can be divided into the following four fractions by pyrolysis-GC/MS [30]: polysaccharides, proteins, polyhydroxy-aromatics, and amino sugars. It was found that the fouling potential of these fractions could be ranked in the order of polyhydroxy-aromatics, proteins, polysaccharides and amino sugars. The polyhydroxy aromatics, which were thought to be the main foulants for negatively-charged NF membrane surfaces, are probably hydrophobic acids with phenolic groups, which exhibit no negative charge at a neutral pH. Base proteins were also found to be foulants for the NF membrane. Polysaccharides and amino sugars were other foulants and are derived from high polyhydric alcohols and have neutral characters. These observations are supported by a recent study by Croue et al. [7] who tested NOM fractions from river water on polyethersulfone UF membranes. They introduced the concept of a biopolymer mixture of polysaccharides, proteins and amino sugars which is made more fouling prone in the presence of polyhydroxy-aromatics.

A contrary view concerning the importance of NOM hydrophobicity has been provided by several researchers. Fan et al. [25] isolated NOM into four fractions; hydrophilic neutral, hydrophobic acids, transphilic acids, and hydrophilic charged. It was found that the order of the fouling potential of the individual fractions during microfiltration with PVDF membranes was hydrophilic neutral > hydrophobic acids > transphilic acids > hydrophilic charged. It was suggested that the hydrophilic neutral fraction exhibited the greatest fouling potential because of its size (>30kD) and due to interactions with calcium in the water. It was postulated that fouling was within the internal pore structure. Interestingly, hydrophobic PVDF membranes were significantly more prone to fouling than hydrophilic PVDF. Recently, Cho et al. [29] suggested that NOM hydrophobicity per se was not an influential factor except for extremely hydrophobic NOM-source water, as hydrophobic NOM comprises a relatively high fraction of ionizable hydrophobic acids. For typical NOM-source water, hydrophilic NOM was found to be the major membrane foulant, due to its non-charged character, thereby promoting easier adsorption onto the membrane surface. Lin et al. [31] studied the effects of the characteristics of fractionated humic acids on the performance of a negatively charged hydrophobic ultrafiltration membrane. Again it was observed that the hydrophilic fraction induced the most rapid flux decline despite little retention of DOC. Carroll et al. [32] also concluded that the major foulants were the hydrophilic neutral components of NOM while the fractions comprising humic acid and fulvic acids have less effect on membrane fouling in their study on the NOM fouling of a hydrophobic MF membrane treating a single water source. Recent results from the same group [33] using two different waters and polypropylene MF membranes found that the hydrophilic components were the major foulants for one water and the hydrophobic components were the major foulants for the other. In this study the authors also noted that short term tests did not correspond well to long term trials. The lack of a clear cut picture from the literature with regard to hydrophobic vs hydrophilic components of NOM may due to the variations in source waters and the molecular weights of the dominant foulants.

In some cases electrokinetic interactions could dominate [17, 18, 30]. For example, electrostatic repulsion may take place between a negatively charged hydrophobic membrane and NOM components with negative charge density. In addition the adsorbed species could change the properties of the membrane. For example, Yoon et al. [34] performed zeta potential measurements at different pH values (3.45-7.2) with and without humic acids. They verified that over the entire range of pH studied, the membrane was positively charged in the absence of humic acids. However, when the zeta potential was measured in the presence of humic acids, it became negatively charged for all pH values, except at pH 4.5. This phenomenon was attributed to the adsorption of negatively charged humic acids onto the membrane surface. Conversely if there is no net surface charge on the membrane the hydrophobic and hydrophilic neutrals could be easily adsorbed and this may explain the observed fouling and flux decline with these species.

Finally, the role of particulate organic matter (POM , < 1μm >0.45 μm, – see Table 3) should not be overlooked. Croue et al. [7] found these to be very significant foulants for their two river water samples. In their study the POM contributed from 63 to 86 % of the colloidal fouling; only 37 to 14% was due to the dissolved colloids. However it is possible that the effect of the POM was exaggerated in their experiments which used a constant pressure protocol that would have had a J_o in the region of 700 l/m^2hr, potentially giving rapid pore plugging and cake formation. Nevertheless POM, which may not be detected in standard DOC analysis (usually involves prefiltration at 0.45 μm), could be a major foulant, particularly for MF and UF in the absence of chemical addition (see 5.1).

4.2.2. Impact of membrane properties

Membrane properties include pore size and distribution, MWCO, hydrophobicity, surface/pore charge (electrostatic repulsion), as well as the morphology of the membrane surface. These characteristics play different roles that affect NOM filtration performance during various processes ranging from nanofiltration to microfiltration.

In the MF filtration of natural water, NOM fouling can result in flux decline and NOM retention, even though the large pores of the MF membranes are expected to cause minimal retention of the humic substances. This behavior is related to the permeability of the cake or 'gel' layer formed by organic/inorganic compounds on the membrane surface and to the effective size of the membrane pores, which become smaller due to the adsorption of dissolved compounds. Pore blockage caused by the physical deposition of large organic compounds on the membrane surface induces the initial NOM fouling. These aggregates appear to serve as nucleation sites for the subsequent deposition of humic acids on the membrane surface. For example, Yuan and Zydney [35] investigated the extent and mechanisms of humic acid fouling with a 0.16μm hydrophilic MF membrane. They observed that the flux decline was due to the convective deposition of large humic acid aggregates/particles on the membrane surface followed by the subsequent deposition of macromolecular humic acids, which was accelerated by the initial step of humic acid fouling. It was also found that the three classical filtration models, i.e., the pore blockage, pore restriction, and cake filtration models were unable to provide an accurate description of the experimental data over the full course of the microfiltration.

There are various opinions on NOM fouling during microfiltration. Nystrom et al. [36] found a dramatic flux decline during humic acid microfiltration through inorganic capillary membranes. This rapid decline in flux caused by the nearly complete blockage of the membrane pores was attributed to electrostatic interaction between the negatively charged humic acids and the positively charged inorganic membrane. This phenomenon is less likely with polymeric MF membranes that are more typically negative.

For UF membranes, the main mechanisms responsible for NOM fouling are the interactions between NOM and the membrane surfaces. These interactions include pore adsorption and plugging which are most likely to be the initial, and even longer term, fouling mechanisms for UF membranes. Therefore, the low molecular weight molecules smaller than the membrane pore sizes could lead to significant membrane fouling due to adsorption on or within the pores. Larger humic acid aggregates/particles would tend to form surface cake layers. Kulovaara et al. [37] reported that fouling of a 50 kD polysulfone membrane while treating a surface water was largely irreversible due to the formation of a strongly adherent deposit on the membrane surface. As for the fouling of ultrafiltration membranes by a lake surface water, Cho et al. [38] found that most of the flux decline was attributed to the weak NOM adsorption on the membrane surface along with a small contribution from gel layer formation, while concentration polarization was only important for polysulfone membranes with the largest pore size of 10 kDa. In their later study, Cho et al. [39] compared three types of UF membrane, in terms of their characteristics of flux decline, retention of NOM, and the adsorbed foulants, with two very different (relatively hydrophilic and relatively hydrophobic) NOM-containing source waters. They came to the conclusion that the humic components were either excluded by the membrane or passed through the pores; they were not selectively adsorbed (weak adsorption) onto the membrane surface. The adsorbed components, which could not be differentiated between surface and pore sorption, were predominantly neutral and basic NOM species.

In the NF filtration process, the mechanisms of NOM retention and fouling are believed due to a combination of steric and electrostatic exclusion, thus the hydrophobicity and surface charge of the NF membrane plays an important role in NOM fouling. Generally, hydrophilic NF membranes with negative charge exhibit high retention with low NOM fouling tendency (for example, Her et al. [40]). In contrast, Jucker and Clark [27] found that hydrophobic membranes were severely fouled by humic substances adsorption in their static (zero flux) adsorption analysis. More recently, Lee et al. [41] analyzed the relationship between the flux decline of NF membranes with NOM transport characteristics. They concluded that diffusion and convection contributed to NOM transmission. The membrane surface charge (in terms of zeta potential) was also found to be an important factor associated with NOM removal and membrane fouling by Braghetta et al. [19].

4.2.3. Impact of solution conditions

The chemical composition of source water, which includes pH, ionic strength, and the concentration of divalent ions such as Ca^{2+}, is another important factor influencing NOM fouling. Braghetta et al. [19] found a substantial flux decline at low pH and high ionic strength and attributed this observation to the conformational changes of NOM during NF fouling experiments with aquatic NOM. In contrast Ruohomaki et al. [42] observed that fouling was higher at neutral pH than at acid or alkaline pH when the pH effect on humic acid retention was evaluated using polyethersulphone membranes. The possible reason was shown to be the formation of compact humic acids under the conditions of lower or higher pH values. Nilson and DiGiano [24] found that flux decline was relatively significant at pH 4 and 7 but negligible at pH 10, probably due to the fact that the acidic components of NOM and the membrane have a higher negative charge density at higher pH values thus inhibiting NOM adsorption on the membrane surface.

Divalent cations are also believed to exacerbate NOM fouling of membranes. The mechanisms involved are aggregation of humics and acting as a 'chelating' bridge between negative membrane surfaces and negative humic species. The role of calcium in UF membrane fouling has been described by Jucker and Clark [27] and in NF fouling by, for example Schaefer et al. [43, 44] and Yoon et al. [34]. However, under some conditions the presence of calcium could be beneficial, as described by Schaefer et al. [45] for the addition of coagulant in a hybrid MF/UF removal processes. Calcium enhanced the coagulation to produce more filterable floc.

NOM would be expected to change conformation in different ionic environments. According to the polyelectrolyte solution theory, charged macromolecules undergo coil-to-globule conversion in response to changes in solution conditions. The roles of ion concentration and pH are schematically depicted in Fig. 4. When an ion salt is added to a NOM solution, the NOM solubility is likely to be affected. NOM could contract to small globules to adjust to the change in its environment.

The NOM transition could be expected in two ways: as a polyelectrolyte chain, the transition would pass through a series of intermediate conformations, which consist of chains bearing several sub-globules along the chain; the number of these globules would decrease with increasing the salt concentration. On the other hand, NOM is a complex mixture which consists of different macromolecules. These macromolecules would show different responses with increasing salt level. The coil-to-globule transition would be faster for some components [25, 46].

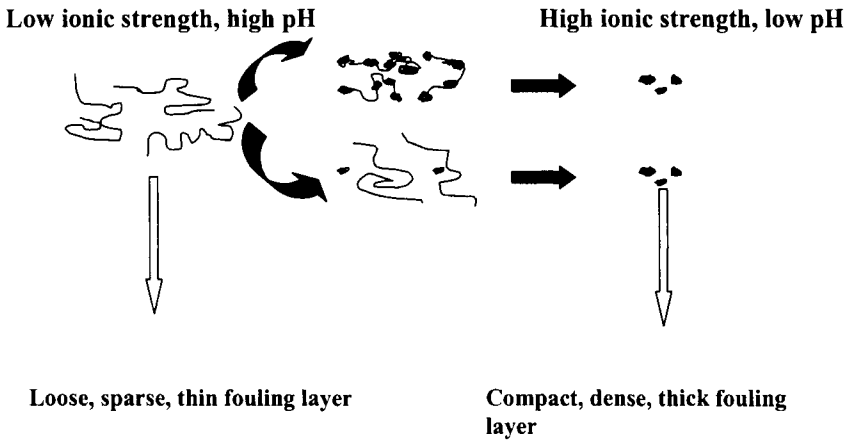


Fig. 4. Schematic depiction of the effect of solution conditions on the conformation of NOM macromolecules in the solution and on the membrane surface

The NOM transition could be expected in two ways: as a polyelectrolyte chain, the transition would pass through a series of intermediate conformations, which consist of chains bearing several sub-globules along the chain; the number of these globules would decrease with increasing the salt concentration. On the other hand, NOM is a complex mixture which consists of different macromolecules. These macromolecules would show different responses with increasing salt level. The coil-to-globule transition would be faster for some components [25, 46].

Unlike indifferent monovalent cations, divalent cations such as Ca^{2+}, Mg^{2+} interact specifically with NOM and form metal-humic complexes [16]. The complexation of divalent cations with NOM changes the electro-kinetic properties of the NOM. The extent of complex formation is strongly dependent on the solution chemistry and on the characteristics of the NOM. Wei et al. [46] demonstrated that increasing Ca^{2+} showed competing effects on the size of NOM. At relatively low concentrations of Ca^{2+} the size of the NOM was reduced, due to charge shielding. However as the concentration of Ca^{2+} increased there was an increase in NOM size, presumably due to aggregation or chelating between NOM and Ca^{2+}.

Clearly, the chemistry of the feed water affects the electrostatic properties of the dissolved organic compounds and the membrane and influences NOM interactions with the membrane surface. As a consequence the fouling potential of NOM may be changed.

4.2.4. Impact of operating conditions

As noted above, from Table 4 it is apparent that for MF, UF and NF the mode of operation in research studies is, with few exceptions, crossflow (pumped or stirred) and constant pressure. This differs from typical industrial operation which tends to use constant fluxes, with crossflow (NF or UF) or in deadend mode (MF or UF). There appears to be a need to develop standard fouling protocols for lab scale studies.

The relevance of controlling the crossflow /flux ratio has been pointed out in discussion of the J/k ratio in Eq. (3). Some studies (for example [5]) have examined the influence of the J/k ratio on fouling, and have shown that as this ratio is increased so is the rate of fouling. This points to the means of controlling fouling by limiting flux or improving the mass transfer coefficient k, typically increased by increasing crossflow. A similar approach was demonstrated in hollow fibre operation by Crozes et al.[16].

The strategy used for submerged MF or UF is to add some coagulant and then operate on a cycle at constant flux in deadend mode with regular backwash. The benefit of deadend is that with relatively low solids content the energy usage can be minimized. Choice of imposed flux is dictated by the frequency of backwash, irreversible fouling, and capital costs.

4.3. Modeling NOM fouling

In conjunction with the broad experimental studies, several researchers have explored theoretical approaches to describe the membrane fouling and flux decline associated with NOM in source water. A number of mathematic models have been developed with system parameters, material balance equations plus appropriate boundary conditions and flow dynamics, such as the modified fouling index, the membrane resistance-in-series model, the gel resistance model, and the fouling kinetic model etc [47-51]

Normally, the fouling effect is described by Darcy's law (Eqs.(1) and the fouling resistance R_c (in Eq. (2)) lumps together the various components of resistance. Ho and Zydney [50] developed a fouling model for the MF of macrosolutes where simultaneous pore blockage and cake formation were accounted for. It was assumed that the cake layer only formed over those regions of the membrane that have been blocked by the physical deposition of initial large aggregates. The model is different from most previous pore blockage models, since it also assumes that the aggregates might allow some fluid flow through the blocked pores. Based on these assumptions, R_c in Eq. (2) can be rewritten as:

$$R_c = (R_m + R_{co})\sqrt{1 + \frac{2f'\alpha_c \Delta P C_b t}{\eta (R_m + R_{co})^2}} - R_m \qquad (4)$$

where R_{co} is the resistance of the first layer of the humic acid deposit, C_b the bulk humic acid concentration, f' the fraction of humic acids in the bulk solution that contributes to the growth of the existing deposit, α_c the specific resistance of the cake layer, η the solution viscosity and t the time. The filtration flux through the fouled membrane was then a combination of the flow rates through the open and blocked pores:

$$\frac{J}{J_o} = \exp\left[-\frac{F_b \Delta P C_b}{\eta R_m} t\right] + \frac{R_m}{R_m + R_c}\left\{1 - \exp\left[-\frac{F_b \Delta P C_b}{\eta R_m} t\right]\right\} \qquad (5)$$

where J_0 is the initial flux through the clean membrane, R_c is evaluated from Eq. (4) and F_b is the pore blockage parameter ($m^2\ kg^{-1}$). The two terms on the right hand side of Eq. (5) are the flow rates through the open and blocked pores, respectively. The first term is derived using the classical pore blockage model.

This model was evaluated by measuring the flux decline in the microfiltration of an Aldrich humic acid solution through 0.2μm pore size membranes [49]. It was found that the theoretical results were in good agreement with experimental data of the filtration flux as a

function of time under different solution conditions. A possible limitation of this model is that it was assessed with data obtained with Aldrich NOM, a soil-based material which has a higher proportion of high molecular weight components. The constant pressure protocol was also applied.

In a study of NOM fouling kinetics in ultrafiltration systems, Tansel et al. [51] proposed a simple flux model based on the resistances-in-series approach to estimate the characteristic fouling time, τ. Assuming the overall resistance to flow is the combined result of two types of resistances namely a time-dependent resistance and a constant resistance. The relationship is;

$$\left(R_m + R_c\right)_{t=t} \Big/ \left(R_m + R_c\right)_{t=0} = \left(1 - X_f\right) + X_f \exp\left(t/\tau\right) \tag{6}$$

where X_f indicates the extent of fouling and τ is the fouling time constant. This relationship can be used to correlate the normalized flux data and predict the critical fouling time when the flux decreases to a specified fraction of the initial flux.

Braghetta et al. [19] also used the resistances-in-series model to describe the accumulation of NOM at NF membrane surfaces. They assumed that the membrane fouling associated with NOM occurred in three stages: concentration polarization, cake layer accumulation on the membrane surface and adsorption of NOM components on the membrane surface and/or within the pores. The summation term of R_i in Eq. (1) was then expanded to be a number of different resistances, as follows:

$$J_v(t) = \frac{\Delta p - \Delta \pi}{\eta(R_m + R_{cp} + R_e + R_c)} \tag{7}$$

where R_{cp}, R_e, R_c denoted the additional resistances during the filtration processes due to concentration polarization, internal pore fouling, and external deposition or cake/gel formation respectively.

A parameter, called the specific foulant mass resistance ϵ_s was also defined [19] to evaluate the value of R_{chem} associated with the amount of NOM removed from the membrane surface by chemical cleaning:

$$\varepsilon_s = \frac{\sum R_{chem}}{NOM \text{ recovered}} \tag{8}$$

where ϵ_s has units of $m^{-1}mg^{-1}$. The physical meaning of ϵ_s is related to the extent of superimposed resistance offered per unit mass of foulant associated with the membrane surface. According to Braghetta et al. [19], the specific foulant mass resistance ϵ_s can be used to quantify the extent of membrane fouling. Relatively high ϵ_s would suggest a densely compacted layer of foulant that superimposed a large resistance to filtration per unit mass of foulant. In contrast, a low ϵ_s would suggest a loosely connected layer of foulant that contributed little additional foulant-related resistance to the filtration. The parameter ϵ_s is obviously related to specific cake resistance α_c (resistance per mass load of membrane area).

5. TECHNIQUES TO REDUCE NOM FOULING

Preventing or reducing NOM fouling would significantly cut down the cost of the raw water treatment process, extend the membrane life and reduce the energy demand. By doing so, the additional costs associated with frequent cleaning operations and membrane replacement would also be reduced. Much effort has been made in this area and the techniques developed to minimize NOM fouling include hybrid membrane processes, the pretreatment of NOM raw water, the optimization of hydrodynamic parameters, the modification of the membrane surface and cleaning of the membrane system. These techniques are briefly reviewed below.

5.1. Hybrid membrane processes

The combination of coagulation/flocculation, or sorbents, such as powdered activated carbon (PAC) or an ion exchange resin (such as MIEX®) with membrane processes is a common practice in water and wastewater treatment. For example, coagulants such as alum, ferric salts and cationic polymers are often used prior to membrane units to assist removal of insoluble metals, proteins and bio-organics from industrial wastewater. Recently, the application of the hybrid membrane processes for NOM removal has received attention.

In general, the hybrid membrane processes can improve the removal of NOM and some other contaminants from the water, but their effects on membrane fouling are not always beneficial. Thus membrane fouling can be reduced or enhanced by adding particulate species. Vickers et. al. [52] concluded that without pretreatment, MF/UF can achieve only nominal (~10%) removal of NOM, whereas these processes can be used in combination with coagulation chemistry or PAC to achieve a higher NOM removal. Zhang et al. [53] investigated the behavior of UF processes to which powdered activated carbon (PAC), heated iron oxide particles (HIOPs), or (nonadsorbent) SiO_2 particles were added. The NOM removal efficiency was found to be in the order of PAC> HIOPs>SiO_2. They also studied the fouling tendency of these introduced particles and found that in the case of both PAC and SiO_2, increasing the dosage of solids led to a steady increase in fouling, but the opposite trend applied when HIOPs were added. Interestingly, in the absence of NOM, none of the solids fouled the membrane significantly. Lin et al. [54] reported that the use of PAC in a combined PAC/UF membrane system increased the membrane fouling even though the NOM removal efficiency increased. Schaefer et al. [45] also observed that, under certain conditions, the addition of ferric chloride to solutions containing humics can significantly increase the fouling of MF/UF membranes. Carroll et al. [32] found that small molecular weights, non-ionics, hydrophilic NOM that were poorly removed by coagulation were responsible for the fouling after the coagulation process.

To understand these phenomena, it is important to understand the interactions between the membrane, the NOM and the adsorbent particles. One explanation for the more severe fouling after the addition of the particles is that the NOM residue and the flocs form a 'composite' cake layer at the membrane surface, with NOM filling the interstices between particles. These cakes become compressed and distorted due to transmembrane pressure, forming a low permeability layer and blocking the pores on the membrane surface.

In order to reduce membrane fouling in the process of NOM removal, it is necessary to reduce the adsorption in the membrane pores and surface, increase cake porosity and reduce the interaction between NOM and the membrane surface. For example, the heated iron oxide particles (HIOPs) have a much higher tendency to bind with NOM than the membrane surface, which reduces the tendency of either the NOM or the HIOPs to interact with the membrane

surface and leaves enough porosity in the cake layer for water to reach the membrane surface [53]. It was noted by Carroll et al. [32] that the selection of coagulant would define the size, permeability and stability of flocs on the membrane surface and thus affect the NOM fouling. The filtration resistance of alum treated NOM was significantly less than that of the residual NOM, but this observation was not repeated for other coagulants because of the differences in the interactions between the NOM, membrane surfaces and coagulants. However it should be noted that alum may not be a panacea as it tends to slowly foul some MF membranes, and is not easily backwashed.

Recently Cabane et al. [55] reported that the aggregates produced under sweep floc conditions were more compressible than under charge neutralization conditions, resulting in compaction when the membrane filtration system was pressurized. In practice, this effect will depend on the imposed fluxes or allowed TMPs. Choi and Dempsey [56] revealed that NOM removal was effective and fouling was reduced by 'under-dosing' the coagulation (relative to conventional coagulation) in a UF membrane filtration process. The most appropriate under-dose conditions would likely depend on the nature of the raw water and on the operational requirements. This indicates that the control of water chemistry and optimization of operating conditions are important strategies to minimize NOM fouling.

MIEX® is an anion exchange resin capable of removing relatively low molecular weight negative organics from NOM. The combination of MIEX® with low pressure membranes is of interest [57], and can achieve good removal of TOC. It has been shown by Galjaard et al. [58] that using MIEX® to remove organics leads to very low fouling of an UF membrane which is hydrophilic and positively charged. Interestingly in this study the feed was chemically softened so the detrimental effects of calcium were obviated. The operating mode was deadend and constant flux (range, 50 to 100 L $m^{-2}hr^{-1}$).

5.2. Water chemistry control and optimization of operating conditions

The chemical composition of source water, which includes pH, ionic strength, and the concentration of divalent ions such as Ca^{2+}, is one of the main factors influencing NOM fouling. Therefore, the adjustment of the pH and ionic strength could help to reduce NOM fouling. Maartens et al. [59] showed that the pH of the raw water had a significant effect on the adsorption of NOM onto a polysulfone UF membrane surface. Since the lowest adsorption occurred at pH 7, adjusting the pH of the feed solution to 7 should minimize the membrane fouling caused by the interaction between NOM and the membrane surface. Hong and Elimelech [18] suggested that pretreatment was necessary to prevent NF membrane fouling by NOM. For hard water, it was evident that the removal of calcium and magnesium ions via a pretreatment step could substantially reduce NOM fouling. Cleaning with strong chelating agents such as EDTA to remove free and NOM-complexed calcium ions was the most effective way to eliminate the fouling layer and restore permeate flux.

It is also useful to optimize the operation conditions to minimize membrane fouling caused by the NOM. This involves consideration of the options for 'mode of operation' (section 2.2). Thus, for NF in spiral wound modules, and other systems with significant solids load, crossflow operation at a controlled flux (constant J/k) is recommended. The importance of this is illustrated by Hong and Elimelech [18] who investigated the distribution of the NOM fouling layer due to strong hydrodynamic shear at high cross flows. Although they did not operate at constant flux they found that NOM fouling was controlled by the interplay between fouling drag and electrostatic double layer repulsion. For hollow fibre ultrafiltration of surface water Crozes et al. [16] investigated the impact of flux (and TMP), crossflow velocity and backwash frequency on fouling. In their case they found an optimum flux of 75

l/h.m^2 with 0.9m s^{-1} crossflow velocity to limit fouling by particulates (turbidity). This 'optimum' would be specific to the module used and the condition of the feed. Some hollow fibre systems treating surface waters operate in deadend (no crossflow) for low turbidity feed and switch automatically to crossflow when turbidity increases. In all cases backwash would apply.

5.3. Membrane surface treatments

In general, hydrophilic membranes, with hydroxyl, amine or carboxylic acid groups, have been found to be less prone to fouling than hydrophobic membranes when treating natural water containing NOM. Unfortunately, most commercially available UF and MF membranes are made from relatively hydrophobic materials with low surface energies such as polypropylene, polysulfone, polyethersulfone and polyvinylidene fluoride as they offer better chemical, thermal and biological stabilities. However it is possible to combine the lower fouling potential of hydrophilic surfaces with the stability of hydrophobic membranes by surface modification [16, 60-66].

The successfully modified membranes should have a high degree of coverage to the internal pore channels and the separation surface with long-term stability and high membrane permeability. They should also minimize the impact on the bulk membrane properties. In addition, the costs of the modification technique should be compensated for by reduced costs of cleaning materials and operation (reduced fouling).

Surface modification is normally implemented through the adsorption of surfactants and soluble polymers onto the membrane surface to change the hydrophobicity and surface charge in order to enhance the membrane performance. Several surface modification techniques for commercially available UF membranes have been developed, which are summarized as follows:

5.3.1. Coating a hydrophilic thin layer onto the surface of the membrane

Forming a hydrophilic monolayer on a support membrane surface will give the hydrophobic membrane hydrophilic character, which is the so-called surface functionalization of membranes [64-66]. According to previous studies, this approach can be realized by several techniques such as the lamination of a pre-formed film, chemical dip coating, in-situ polymerization and electrophoretic coating, etc. For example, Fane and Fell [67] confirmed that a reduction in foulant adsorption by hydrophobic membranes can be achieved by forming a hydrophilic monolayer on the membrane surface, as it gave the membrane a hydrophilic character with fewer hydrophobic sites for foulant adhesion. Maartens et al [68] found that pre-coating membranes with hydrophilic surfactants significantly reduced the adsorption of foulant species in NOM and improved initial color retention of the membranes to above 96%, but the flux through the membranes was marginally reduced.

5.3.2. Initiating polymerization by radiation and grafting reactions onto a support membrane to impart hydrophilic properties

Photochemical modification has distinct advantages in simplicity, cost and breadth of application. This simple method increases membrane surface wettability and shifts the membrane pore size distribution to smaller sizes. The results of previous studies suggested that the selective UV excitation of a photo-initiator adsorbed onto the polymer surface causes a heterogeneous hydrogen abstraction and subsequent initiation by polymer mechanisms for polymer surface modification. In this way, the surface modification is achieved by applying a very simple, inexpensive graft polymerization technique. Recently, this method has been widely used to increase NOM retention and reduce NOM fouling.

Carroll et al. [69] studied the surface modification of a polypropylene hollow fibre membrane with charged and non-charged hydrophilic polymers. They showed that non-ionic and cationic hydrophilic grafts reduced the flux decline due to NOM fouling by up to 50% compared with the ungrafted polypropylene. In contrast, anionic hydrophilic grafts had initial flux increases up to 140% at high graft yields due to multi-valent ions in the natural water, although the pure-water flux was substantially lower than that of the ungrafted membrane. Kilduff et. al. [70] investigated photochemical modification of poly(ether sulfone) and sulfonated poly(sulfone) nanofiltration membranes for fouling control of natural organic matter, with N-vinyl-2-pyrrolidinone (NVP) grafting. They concluded that the fouling by NOM was reduced significantly after surface grafting, but solute retention was also reduced. Taniguchi et al. [71] revealed that photochemical grafting increased membrane surface wettability and shifted the membrane pore size distribution to smaller sizes, which consequently increased natural organic matter retention.

5.3.3. Deposition of a hydrophilic layer from a glow discharge plasma onto the membrane

Plasma-induced polymerization is another technique that has been successfully used for surface modification of polymeric membranes. The capability of plasma to alter the physical and chemical properties of polymeric surfaces without affecting the bulk properties (especially mechanical properties) of the base material is advantageous for the design and development of surface modified polymer membranes. With plasma treatment, specific surface chemistries can be created for increasing NOM retention and reducing membrane fouling. However, a limitation of plasma treatment is its temporal instability such as the gradual loss of surface chemical properties with time [72].

5.3.4. Chemical attachment of selected polymer and/or monomers

Concentrated emulsion polymerization and blending have been developed as one of the approaches for the preparation of hydrophobic-hydrophilic composite membranes. With this method two types of membrane structure have been developed: the asymmetric structure obtained by the phase inversion technique and the composite one formed by over-coating a porous substructure with a thin top layer acting as the barrier. The latter is made of an interfacial polymerized polyamide supported on a micro porous polysulfone and has gained importance because of the improved performance characteristics.

The choice of the specific surface modification techniques depends on the chemical structure of the given support membrane and the desired characteristics of surface modification required. Several patents have been issued to membrane manufacturers for modifying well-known polymers or precast porous films in this regard, such as photo-induced graft polymerization surface modification, surface modification by low-temperature plasma-

induced graft polymerization and direct microemulsion polymerization using polymerizable surfactants. For examples see references [73-75].

5.4. Membrane cleaning

The techniques mentioned above focus on reducing the membrane fouling caused by NOM. Nevertheless, membrane fouling inevitably occurs over a long period of operation, resulting in loss of membrane permeability, manifesting as flux decline or increase in required TMP. To recover the fouled membranes, hydraulic backwash and chemical cleaning is needed. Of course backwash is only applied to hollow fibre membranes. Cleaning agents include acid and caustic agents, as well as alkaline solutions. Many studies have been conducted to investigate the efficiency of these chemical cleaning agents and hydraulic wash procedures. A full survey is beyond the scope of this text.

An example is provided by Lee et al. [61] who investigated various cleaning strategies for the UF membranes used to treat source waters containing different NOMs. They found that high ionic strength cleaning (with 0.1 M NaCl) was more effective in promoting flux recovery for the membrane fouled with relatively hydrophilic NOM than acid, caustic or surfactants solutions. They also showed that increased cross-velocity and prolonging cleaning time influenced the cleaning efficiency of caustic solution, but did not influence the effectiveness of high ionic strength cleaning. Maartens et. al. [68] found that a 0.1% Triton X100 solution with NH_3 was successful to remove the NOM foulant on their UF membranes. The treated membrane flux increased even to levels higher than that of unfouled membranes. However, the chemical cleaning agent not only removed the foulants, but also absorbed onto the membrane surface and interacted with the membrane.

Interaction between the chemical cleaning agent and the membrane surface may not be good for the membrane system. Repeated chemical cleaning may affect the membrane life. Chemical cleaning is also typically cumbersome and requires shutdown of the unit being washed for several hours. This results in a reduction of the overall plant capacity, and produces a waste that may be difficult to dispose of. Chemical cleaning should thus be limited to a minimum.

6. CONCLUSIONS

The application of membrane technology to drinking water production continues to grow. However, the ability to control fouling by NOM and other species has not been perfected because of the complexity of the feed and the potential surface interactions. We have some clear insights, such as the benefit of hydrophilic membranes, the adverse effects of colloids and the detrimental role of calcium. The significances of other foulants are still moot. This is possibly due to local variations in NOM and also the lack of standard fouling protocols to make a fair comparison.

REFERENCES

[1] A.G. Fane, Module Design and Operation, Chapter 4 in A.I. Schaefer, A.G. Fane and T.D. Waite (Eds) Nanofiltration; Principles and Applications., Elsevier Science, ISBN 185617 4050, 2004.

[2] C. Jarusutthirak, Fouling and Flux decline of RO, NF and UF Membranes Associated with Effluent Organic Matter (EfOM) During Wastewater Reclamation Reuse, Ph.D Dissertation, University of Colorado, 2002.

[3] S. Hong, Natural Organic Matter and Colloidal Fouling in Crossflow Membrane Filtration, Ph.D Dissertation, University of California, Los Angeles, 1996.
[4] N. Her, G. Amy, H-R. Park and M. Song, Water Res., 38 (2004) 1427.
[5] G. Amy and J. Cho, Water Sci. Technol., 40 (1999)133.
[6] A.R. Costa and M. N.de Pinho, Desalination, 145 (2002) 299.
[7] G. Makdissy, J-P.Croue, G. Amy and H. Buisson, Water Sci. Technol.: Water Supply, 4 (4) (2004) 205.
[8] A. Leenheer, J-P. Croue, M. Benjamin, G.V. Korsin, C.J. Hwang, A. Bruchet and G.R. Aiken Comprehensive isolation of NOM from water for spectral characterizations and reactivity testing, in *Natural Organic Matter and Disinfection By-Products: Characterization and Control in Drinking Water,* ACS Symposium Scries 761, Edited by S.E. Barrett, S.W. Krasner and G.L. Amy, *American Chemical Soc.,* 68-83. 2000
[9] A. Seidel and M. Elimelech, J. Membrane Sci., 203, (2002) 245.
[10] V. Lahoussine-Turcaud, M.R. Wiesner and J-Y Bottero, J. Membrane Science, 52 (1990) 173.
[11] J.L. Bersillon, Fouling analysis and control, in: L. Cecille, J.C. Toussaint (Eds.), Future Industrial Prospects of Membrane Processes, Elsevier, Oxford, 1998.
[12] J. Mallevialle, C. Anselme and O. Marsigny, Effects of humic substances on membrane processes, in: I.H. Suffet, P. MacCarthy (Ed.), Aquatic Humic Substances: Influence on Fate and Treatment of Pollutants, ACS, Washington, DC, 1989.
[13] Y. Kalya, Y. Itoh, K. Fujita and S. Takizawa, Desalination, 106 (1996) 71.
[14] W. Yuan, A. Kocic and A.L. Zydney , J. Membr. Sci., 198 (2002) 51.
[15] W. Yuan and A.L. Zydney, Desalination, 122 (1999) 63.
[16] G.F. Crozes, J.G. Jacangelo, C. Anselme and J.M. Laine, J. Membr. Sci., 124 (1997) 63.
[17] M. Taniguchi, J.E. Kilduff and G. Belfort, Environ. Sci. Technol., 37 (2003) 1676.
[18] S. Hong and M. Elimelech, J. Membr. Sci., 132 (1997) 159.
[19] A. Braghetta, F. A. DiGiano and W. P. Ball, J. Environ. Eng., Nov (1998) 1087.
[20] K.L. Jones and C. R. O'Melia, J. Membr. Sci., 193 (2001) 163.
[21] M. R. Teixeira and M. J. Rosa, Desalination, 151 (2002) 165.
[22] J. C. Chen and A. Seidel, J. Environ. Eng., Oct. (2002) 967.
[23] W. Yuan and A. Zydney, Environ. Sci.Technol., 34 (2000) 5043.
[24] J.A. Nilson and F.A. DiGiano, J. - Am. Water Works Assoc., 88 (1996) 53.
[25] L. Fan, J.L. Harris, F.A. Roddick and N. A. Booker, Water Res., 35 (2001) 4455.
[26] A. Maartens, P. Swart and E.P. Jacobs, Desalination, 115 (1998) 215.
[27] B. Jucker and M. M. Clark, J. Membr. Sci., 97 (1994) 37.
[28] J. Cho, G. Amy, J. Pellegrino and Y. Yoon, Desalination, 118 (1998) 101.
[29] J. Cho, G. Amy and J. Pellegrino, J. Membr. Sci., 164(2000) 89.
[30] T.F. Speth, R.S. Summers and A.M. Gusses, Evaluation of membrane foulants from conventionally treated drinking waters, Natural Organic Matter Workshop Proceedings. 1996
[31] T. Lin, T. Lin and O.J. Hao, Water Res., 34(4) (2000) 1097.
[32] T. Carroll, S. King, S.R. Gray, B.A. Bolto and N.A. Booker, Water Res., 34(11) (2000) 2861.
[33] S.R. Gray, C.B. Richie and B.A.Bolto, Water Sci. Technol: Water Supply, 4(4) (2004) 189.
[34] S-H. Yoon, C-H. Lee, K-J. Kim and A.G. Fane, Water Res., 32 (1998) 2180.
[35] W. Yuan, and A. Zydney, J. Membr. Sci., 157 (1999) 1.
[36] M. Nystrom, K. Ruohomaki and L. Kaipia, Desalination, 106 (1996) 79.
[37] M. Kulovaara, S. Metsamuuronen and M. Nystrom, Chemosphere, 38 (1999) 3485.
[38] J. Cho, G. Amy and J. Pellegrino, Water Res., 33 (1999) 2517.
[39] J. Cho, G. Amy and J. Pellegrino, Desalination, 127 (2000) 283.
[40] N. Her, G. Amy, C. Jarusutthirak, Desalination, 132 (2000) 143.
[41] S. Lee, J. Moon, S-K. Yim, S-H. Moon and J. Cho, Desalination, 147 (2002) 237.
[42] K. Ruohomaki, P. Vaisanen, S. Metsanuuronen, M. Kulovaara and M. Nystrom, Desalination, 118 (1998) 273.
[43] A.I. Schäfer, A. Pihlajamäki, A.G. Fane, T.D. Waite and M. Nyström, J. Membr. Sci., 242(1-2) (2004) 73.

[44] A.I. Schafer, A.G. Fane and T.D. Waite, Desalination, 118, (1998) 109.
[45] A.I. Schäfer, U. Schwicker, M.M. Fischer, A.G. Fane and T.D. Waite, J. Membr. Sci., 171 (2000) 151.
[46] X. Wei, R.Wang, A.G. Fane and F.S. Wong, Water Sci.Technol.: Water Supply, 4(4) (2004)197.
[47] S. Lee, G. Amy and J. Cho, J. Membr. Sci., 240 (2004) 49.
[48] T. L. Champlin, Desalination, 131 (2000) 105.
[49] W. Yuan, A. Kocic and A.L. Zydney, J. Membr. Sci., 198, (2002) 51.
[50] C. Ho, and A.L. Zydney, J. Colloid Interface Sci., 232 (2000) 389.
[51] B. Tansel, W.Y. Bao and I.N. Tansel, Desalination, 129 (2000) 7.
[52] J.C. Vickers, M.A. Thompon and U.G. Kelkar, Desalination, 102 (1995) 57.
[53] M. Zhang, C. Li, M.M. Benjamin and Y. J. Chang, Environ. Sci. Technol., 37 (2003)1663.
[54] C.F. Lin, T.Y. Lin and O. J. Hao, Water Res., 34(4) (2000) 1097.
[55] B. Cabane, M. Meireles and P. Aimar, Desalination, 146 (2002) 155.
[56] K. Y-J Choi and B. A. Dempsey, Water Res., 38 (2004) 427.
[57] J. Morran, M. Drikas, C. Hepplewhite and C. Pelekani, MIEX and microfiltration – a winning alliance, Proceedings AWA 19th Convention, Canberra , April 2001;
[58] G. Galjaard, J.C. Kruithof and G.Raspati, Influence of NOM and membrane surface charge on UF membrane fouling, Proceedings of IWA Leading Edge Conference, Prague, June 2004;
[59] A. Maartens, P. Swart and E. P. Jacobs, J. Membr. Sci., 163 (1999) 5.
[60] H. Lee, G. Amy, J. Cho, Y. Yoon, S.-H. Moon and I.S. Kim, Water Res., 35 (2001) 3301.
[61] J. Meier-Haack, N.A.Booker and T. Carroll, Water Res., 37 (2003) 585.
[62] J. Marchese, M. Ponce, N.A. Ochoa, P. Pradanos, L. Palacio and A. Hernandez, J. Membr. Sci., 211 (2003) 1.
[63] S.H. Kim, S.-Y. Kwak, B.-H. Sohn and T.H. Park, J. Membr. Sci., 211 (2003) 157.
[64] J. Pieracci , J.V. Crivello, G. Belfort, J. Membr. Sci., 156 (1999) 223
[65] F. Hester, P. Banerjee and A.M. Mayes, Macromolecules, 32(1999) 1643.
[66] M. Li., C.H. Chew, W. K Teo and L.H. Gan, Langmuir, 12 (1996) 5863.
[67] G. Fane and C.J.D. Fell, Desalination, 62 (1987) 117.
[68] A. Maartens, P. Swart and E.P. Jacobs, J. Colloid Interface Sci., 221 (2000) 137.
[69] T. Carroll, N.A. Booker and J. Meier-Haack, J. Membr. Sci., 203 (2002) 3.
[70] J.E. Kiduff, S. Mattaraj, J.P. Pieracci and G. Belfort, Desalination, 132 (2000) 133.
[71] Taniguchi, J.E. Kiduff and G. Belfort, J. Membr. Sci., 222 (2003) 59.
[72] K.S. Kim, K.H. Lee, K. Cho and C.E. Park, J. Membr. Sci., 199 (2002) 135.
[73] J. Hosch and E. Staude, J. Membr. Sci., 121 (1996) 71.
[74] L Ling and Q. Chen, Desalination, 101 (1995) 51.
[75] S. Belfer, Y. Purinson, R. Fanstein, Y Radchenko and O. Kedem, J. Membr. Sci., 139 (1998) 175.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 8: Removal of natural organic material and algal metabolites using activated carbon

G. Newcombe

Australian Water Quality Centre, a partner in the Cooperative Research Centre for Water Quality and Treatment, PMB 3, Salisbury, South Australia, Australia

1. INTRODUCTION

The activated carbon adsorption process is of major importance to the international water industry, and its cost effective application is dependent on an understanding of the adsorbent and the processes influencing its use. This chapter gives an overview of the structure of activated carbon and the interfacial processes influencing the removal of naturally-occurring contaminants from drinking water. The chapter describes the application of activated carbon in drinking water treatment, and gives examples of removal of some important contaminants, natural organic material, taste and odour compounds and algal toxins, from drinking water.

2. WHAT IS ACTIVATED CARBON?

2.1. Production of activated carbon

Activated carbon can be created from a wide range of natural and synthetic organic raw materials. Some of the more common starting materials and activation methods are shown in Table 1. The production process involves carbonisation - heating in the absence of air to temperatures of about 800° C - then activation in the presence of an activating agent, usually steam, at temperatures between 850 and 1000° C [1]. Wood-based carbons are often carbonised and activated in the same step in the presence of a dehydrating chemical such as phosphoric acid, usually at lower temperatures than steam activation [1]. Carbons produced in this way are called chemically activated carbons. The carbonisation process removes most of the volatile components in the form of gaseous oxygen and hydrogen compounds. The remaining carbon atoms form graphite-like structures consisting of sheets of condensed aromatic ring systems. The sheets are held together by dispersion forces and may be flat or distorted into a more convoluted pore structure. The surfaces of these sheets form the basal planes of the carbon. The activation step further enhances the pore and graphite-plate structure, burning out any decomposition products remaining from the carbonisation process, and opening the pores. If the activation continues, walls between pores can be destroyed, leading to larger pores and decreased volume of very small pores [2]. There are a number of representations of the structure of activated carbon in the literature [1-3].

Table 1
Starting material, mode and extent of activation of seven activated carbons.

Carbon	Starting material	Activation method
C1	Coal	Steam
C2	Wood*	Chemical
C3	Wood*	Steam
C4	Coconut†	Steam, high level of activation
C5	Coconut†	Steam, low level of activation
C6	Wood	Chemical
C7	Lignite	Steam

*† same material

2.2. Physical properties of activated carbon

Activated carbon is a porous material with a very high surface area. For example, a level teaspoon of a good quality powdered carbon has an internal surface area larger than a standard soccer pitch. In contrast, the external surface area, that provided by the outside of the particles, is less than one hundredth of that value. The internal surface provides the majority of the sites for the target contaminants to adsorb upon.

The pores on activated carbon are categorised according to their size as follows-

Primary micro pores	< 0.8 nanometre (nm)
Secondary micro pores	0.8 - nm - 2 nm
Mesopores	2 nm -50 nm
Macropores	> 50 nm [4]

Micropores are generally considered to be slit-shaped, with either parallel walls, or, when the crystallites are at an angle, wedge-shaped. Transmission electron microscopy studies reported by Bansal et al. [2] indicate that the larger pores, mesopores and macropores, may be more cylindrical in shape. More recently, Daley et al. [5] used scanning tunnelling microscopy to investigate the pore structure of three activated carbon fibre materials, produced by the activation of a phenolic fibre precursor. They discovered micropores in the shape of elongated tubes up to 30 nm in length, which wound and twisted through the fibre. Both micro- and mesopores were found to be ellipsoidal in cross-section, and there was no evidence of slit-shaped pores. It is difficult to relate these results to the activated carbons usually used in water treatment, as the natural raw materials would have a significant impact on the pore structure of the resultant carbon. There have been other recent investigations of activated carbon structure [6-10].

Even small variations in the structure and chemical composition of the raw material, and activation conditions, can result in large differences in the physical properties of the finished product. Fig. 1 shows scanning electron micrographs of two types of activated carbon, carbons C5 and C6 from Table 1. The external structure of the carbons is very different, with C5 displaying a structure with few large openings onto the external surface, while C6 has an external surface with large openings. These differences are mainly a result of the structures of the raw materials.

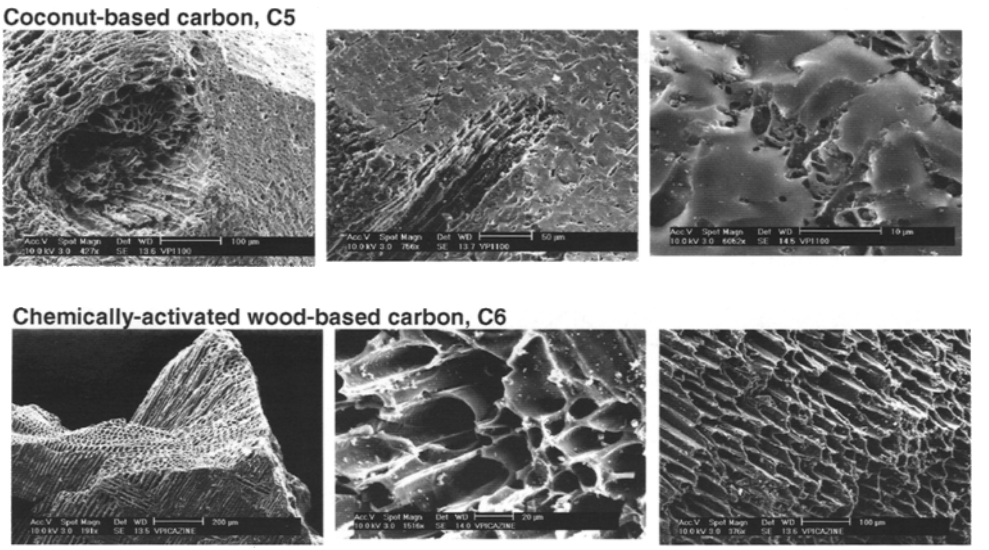


Fig. 1. Scanning electron micrographs of external activated carbon structure of coconut-based activated carbon, C5 and wood-based, chemically activated carbon, C6

The internal pore structure of activated carbon can be investigated using gas adsorption, usually nitrogen gas. Nitrogen adsorption isotherms are obtained over pressure ranges in the order of 10^{-5} to 1 relative pressure. There are numerous theoretical models that can be applied to the data to determine pore size distributions. The most commonly used currently is the Density Functional Theory (DFT) [11]. Fig. 2 shows the pore volume distributions of the activated carbons shown in Figure 1, obtained using nitrogen adsorption and DFT. The high volume of micropores of C5 is typical of a steam activated coconut or coal carbon, while the relatively low volume of micropores, but high volume of mesopores is typical of a chemically-activated wood-based carbon. The combination of different raw materials and methods and degree of activation results in a wide variety of very different activated carbons that are commercially available for drinking water treatment.

2.3. Chemical properties of activated carbon

Although the great majority of the surface is composed of carbon, all activated carbons contain some heteroatoms such as S, Cl, Na, K, P, Si, Al, Fe, depending on the impurities in the starting material and the method of activation. Newcombe et al. [12] reported results of an X-ray photoelectron spectroscopy (XPS) investigation into the surface heteroatoms of a large range of activated carbons and found that, for the carbons studied, coal-based carbons contained iron and silica, coconut carbons contained sodium, potassium and chloride, and chemically activated wood carbons contained phosphorus as a major impurity. The latter is not surprising, as phosphoric acid is the most common agent used for chemical activation of wood-based carbons. However, by far the most abundant elements present on the surface of activated carbon are carbon (approximately 80 to 98%) and oxygen (approximately 2 to 20%) [1]. Oxygen-containing functional groups are the most important in terms of activated carbon surface properties and as a result have been studied for over fifty years [13-17].

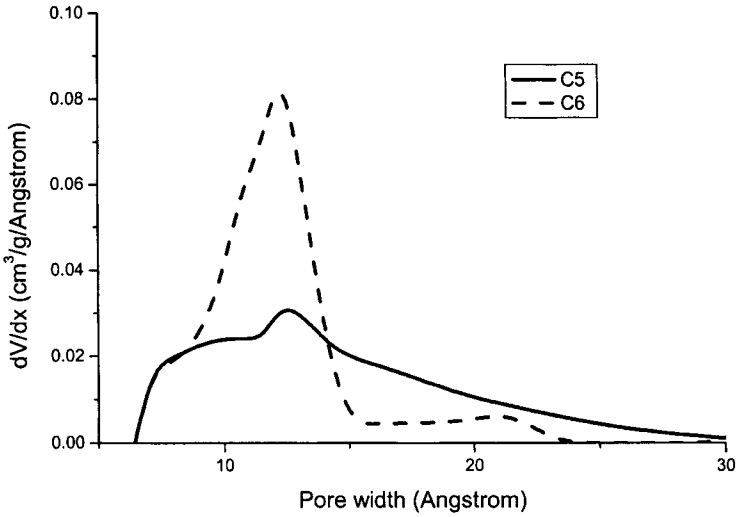


Fig. 2. Differential pore size distributions of two activated carbons, C5 and C6

Activated carbons have a surface charge in aqueous solution, which is usually attributed to oxygen surface groups. For example, proposed acidic groups include: carboxyl, phenol, lactones and quinones, and proposed basic groups include: chromene- and pyrone-type structures [1, 2, 3]. However, the existence of oxygen-containing basic sites has been the subject of some divergence of opinion in the literature, with little convincing evidence for the various proposed forms. In fact, most of the evidence indicates that carbons with low oxygen content are more likely to show basic properties [13, 18, 19]. Puri [20] and Monterra et al. [21] suggested that the main contribution to basic sites on the activated carbon surface is not from oxygen containing groups. As early as 1956, Studebaker et al. [13] suggested the basic sites on the activated carbon surface were due to Lewis bases, with the surface acting as an electron donor in the same manner as polycondensed aromatic hydrocarbon molecules. These sites do not involve oxygen, and oxygen groups would tend to localise electrons, lowering the basicity of nearby Lewis acid sites. The basic sites are generally thought to reside on the basal planes of the carbon surface. Excellent reviews describing the types of oxygen-containing functional groups present on the activated carbon surface can be found in [1,2,18] and comprehensive studies on the basic sites present on the surface are reported in [22-26].

Fig. 3 shows the variation of surface charge with pH for the 7 activated carbons listed in Table 1. Four of the carbons display similar charge behaviour whereas the other three (C3, C4 and C6) are markedly different. C6, a chemically-activated wood carbon, shows little evidence of positive surface sites, perhaps indicative of the absence of graphitic basal planes, however C2, also a chemically-activated wood carbon, shows significant positive character at pH <7. In contrast, C3, a steam activated wood carbon, shows little evidence of negative surface sites. This figure gives a good illustration of the variation in surface chemistry properties that can be expected in a range of activated carbons. A more detailed description of variations in surface chemistry, and the effects on adsorption is given in the following chapter.

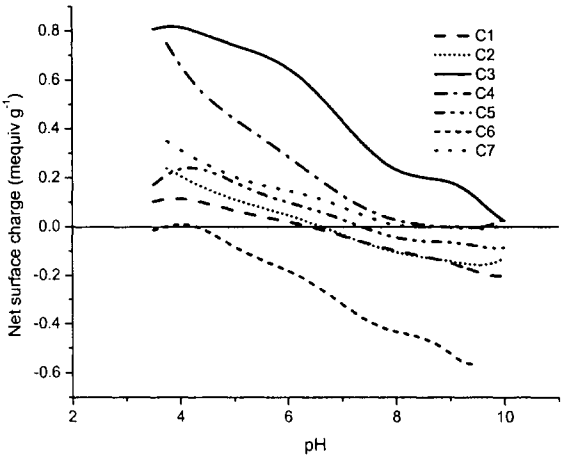


Fig. 3. Net surface charge vs. pH for 7 activated carbons.

2.4. Application of activated carbon for drinking water treatment

Two forms of activated carbon are commonly used in drinking water treatment, powdered activated carbon and granular activated carbon.

2.4.1. Powdered activated carbon

Powdered activated carbon (PAC) can be added before coagulation, during chemical addition, or during the settling stage, prior to sand filtration. It is removed from the water during the coagulation process, in the former cases, and through filtration, in the latter. As the name implies, PAC is in particulate form, with a particle size typically between 10 and 100 μm in diameter. One of the advantages of PAC is that it can be applied for short periods, when problems arise, then ceased when it is no longer required. With problems that may arise only periodically such as algal toxins or tastes and odours, this can be a great cost advantage. A disadvantage with PAC is that presently it cannot be reused and is disposed to waste with the treatment sludge or backwash water.

2.4.2 Granular activated carbon

Granular activated carbon (GAC) is used extensively in Europe and the United States for the removal of micropollutants such as pesticides, industrial chemicals and tastes and odours. It is also becoming more widely used in Australia, particularly for taste and odour removal and insurance against the possibility of a toxic algae bloom in the water source. The particle size is larger than that of PAC, usually between 0.4 and 2.5 mm. Granular activated carbon is generally used as a final polishing step, after conventional treatment and before disinfection. The advantages of GAC are that it provides a constant barrier against unexpected episodes of contamination of water sources, and the large mass of carbon provides a very large surface area. The disadvantage is that it has a limited lifetime, and must be replaced or regenerated when its performance is no longer sufficient to provide high quality drinking water. Replacing the GAC on a regular basis results in high on-going treatment costs. The regeneration process involves heating the carbon to very high temperatures to volatilise the adsorbed compounds.

The costs associated with this process are the capital costs of establishing a regeneration facility, or transport to an existing facility, energy costs, and the loss of some of the GAC through attrition.

Filtration through GAC is often used in conjunction with ozone. When used in conjunction with ozone it is sometimes called BAC, or biological activated carbon; however, this is a misnomer as all GAC filters function as biological filters within a few weeks to months of commissioning.

3. ADSORPTION PROCESSES

3.1. Overview of the adsorption process on activated carbon

Removal of contaminants by activated carbon is a complex process. Fig. 4 is a schematic representation of the major processes occurring during adsorption, which are largely diffusion related. Before its adsorption by activated carbon a molecule must diffuse-

1 to the particle surface from the bulk liquid
2 through the liquid surface film
3 through the pore structure of the carbon

finally being removed from solution at the adsorption site (step 4, Fig. 4)

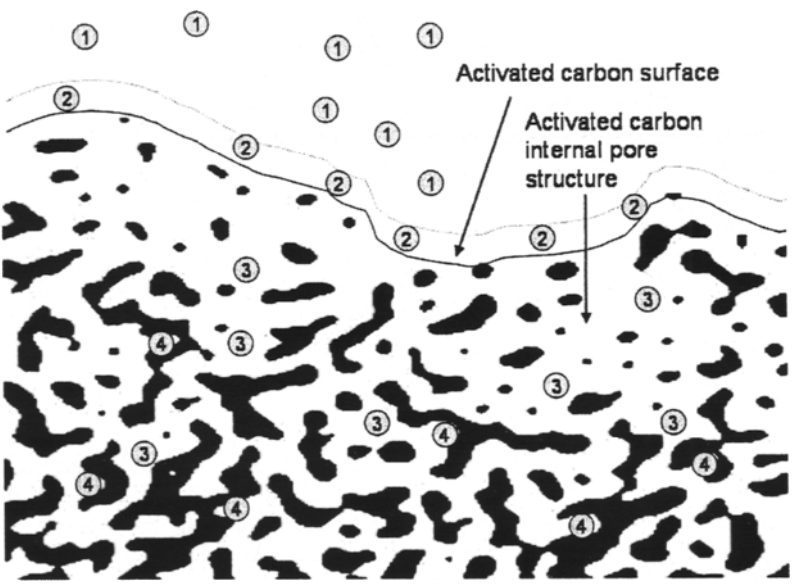


Fig.4. Representation of diffusion into the activated carbon structure

Processes 1 and 2 depend on the physical parameters of the system, for example mixing conditions for PAC, flow rates for GAC. Processes 3 and 4 are dependent on the activated carbon pore size distribution and surface chemistry. In general, the most favourable energy for adsorption is provided by pores slightly larger than the adsorbing molecule, as there are more contact points for the compound to adhere, and it fits "snugly" into the pore [27]. In water treatment another very important factor is how quickly the contaminant can reach a

suitable adsorption site. This is strongly influenced by the access to the internal structure through the pores on the external surface, as well as the structure and size of the "transport pores", those through which the contaminant must travel prior to reaching the adsorption site (ie step 3, Fig. 4).

3.2. Factors influencing adsorption onto activated carbon

Many different processes and mechanisms can play a part in adsorption from solution onto activated carbon. The type of adsorption interaction will depend on the properties of the surface, the adsorbate and the solution.

Factors influencing adsorption include:

- the pore size distribution of the carbon and the relationship between the size of the pores and the size of the adsorbate [27, 28]
- hydrophobic interaction between the graphitic surface of the carbon and hydrophobic parts of the adsorbing species [29, 30]
- electrostatic interaction between a charged carbon surface and a charged or polar adsorbate [31, 32]
- attraction between aromatic rings on the adsorbate and delocalised electrons on the activated carbon surface [33]
- hydrogen bonding between functional groups on the surface and those on the adsorbate [34]
- solubility of adsorbate [35, 36]
- the affinity of the solvent for the adsorbent [36, 37]
- the concentration of compounds competing for adsorption sites [38]
- the solution conditions, pH, ionic strength, temperature [1]

The particular combination of these factors existing in the system will govern the mechanism of adsorption, and whether the major influence is surface chemistry of the carbon, for example, or the volume of pores in the suitable size range for the solute.

In subsequent sections, and the following chapter, these factors will be explored using examples of specific contaminants of interest to the drinking water industry.

4. ADSORPTION OF NATURAL CONTAMINANTS IN DRINKING WATER TREATMENT

4.1. Adsorption of natural organic material

Detailed descriptions of the types of compounds present in NOM and the range of characterisation techniques commonly used can be found in the reference chapter by Ronald Beckett and James Ranville, and further information is given in [39-42].

4.1.1. Influences on the adsorption of natural organic material

Natural organic material is an issue in drinking water treatment as it creates increased demand for chemicals such as coagulants, ozone and chlorine. It also reacts with chlorine and ozone to produce a range of unwanted oxidation by-products. In conventional treatment processes coagulation/flocculation is used to remove a portion of the NOM. Activated carbon can also be a useful technique for the removal of NOM.

Clearly, with such a complex mixture many of the factors listed in section 3.2 will have a significant influence on the adsorption of NOM. However, many studies have shown that

the major influences are the pore size distribution of the carbon [27, 43-45] and the charges on the adsorbate and adsorbent [27, 31, 46, 47].

Fig. 5 gives an excellent illustration of the relative effects of these two factors. The figure shows the results of a series of adsorption isotherm experiments undertaken on 10 activated carbons at two pH values, 3 and 7. In this study two ultrafiltration fractions of concentrated reservoir NOM of nominal molecular weight 500-3000 and > 30000 g mol^{-1} were used. The pore volumes of all of the carbons were known, and the average hydrodynamic diameter of the NOM fractions was measured using flow field flow fractionation (FFFF) [48]. This knowledge allowed the estimation of the volume of pores that was accessible to each NOM fraction for each carbon. Adsorption of NOM was measured in terms of dissolved organic carbon (DOC) in mg L^{-1}. The DOC adsorbed at a solution concentration of 100 mg DOC L^{-1} was determined from the isotherm then plotted against available pore volume in Fig. 5. The 10 carbons were produced from a range of raw material and displayed the expected range of pore volume distributions, heteroatoms and surface charges [49]. It is therefore surprising to see such a consistent dependence of the adsorption of NOM, at pH 3, on the pore volume of the carbons. At this pH most of the charged groups on the NOM would be protonated [49], and therefore electrostatic interactions between the adsorbate and the adsorbent would be minimised. The clear linear relationship between the available pore volume and the adsorption of NOM at this pH suggests a negligible influence of surface chemistry, and the major influence is simply the size of the adsorbate compared with the size of the activated carbon pores. The adsorption at pH 7 is clearly different from that at pH 3, and does not show such a strong dependence on available pore volume. These results, and other work [50-52] strongly suggest that the adsorption at this pH is governed by strong electrostatic effects due to the relatively high negative charge on the NOM and the charge on the activated carbon surface. The electrostatic interactions influencing adsorption of NOM at pH 7 are surface-adsorbate interactions as well as lateral interactions between adsorbed charged molecules, particularly at high NOM surface concentration [52, 53].

Due to these electrostatic interactions, the adsorption of NOM at the pH range usually found in water treatment (6.5-8) cannot be accurately predicted using any commonly measured physical characteristic of the activated carbon such as pore volume or surface area. According to Fig. 5 as a general guide, when the volume of mesopores (2-50 nm) is low (as in most coconut- and coal-based carbons) the relative adsorption, or trend in NOM adsorption, can be estimated from pore volume data. As the volume of mesopores increases, the difference between the carbons becomes indistinct, as indicated by the plateau in NOM adsorption shown in Fig. 5.

The strong electrostatic influences present at neutral or higher pH can be reduced by increasing the ionic strength of the solution. Summers and Roberts [27] studied the effect of increased ionic strength on the adsorption of a commercial humic acid (HA) onto five activated carbons. They expected a decrease in adsorption on a positively-charged activated carbon with a screening of the electrostatic attraction between the surface and the carboxyl groups of the HA. The increase in adsorption that they observed was attributed to a decrease in the size of the molecules with the increase in salt concentration. Once this decrease in size was taken into account they maintained there was an appreciable decrease in adsorption. When they plotted the adsorption of the HA fraction on the five activated carbons in terms of the available surface area rather than per unit mass of carbon, they interpreted the ranking of the adsorbates in terms of the surface charge. As they used a commercial humic acid the results were perhaps not representative of NOM in drinking water sources which can be expected to have a much lower molecular weight.

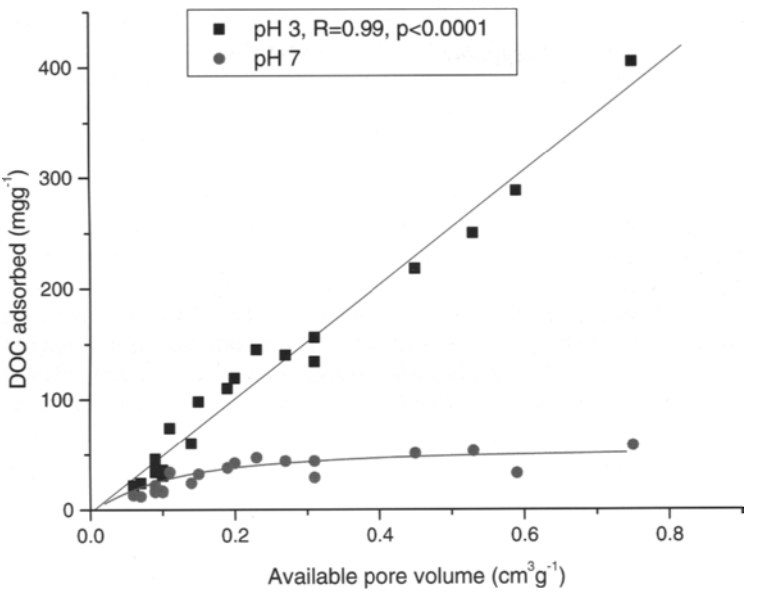


Fig. 5. Dissolved organic carbon adsorbed vs. available pore volume at pH 3 and 7 [50]

A more comprehensive study by Bjelopavlic et al. [52] investigated the adsorption of NOM on two activated carbons at five pH values. The carbons were chosen for their different surface properties, one had a point of zero charge (pH_{pzc}) of 7.5, and the other a pH_{pzc} of 5. The results showed that at pH 3, where electrostatic interactions were minimised, the adsorption was the same on both carbons, per unit volume of pores available for adsorption, and the addition of salt had no effect. At pH 4, 5, 7, and 9, an increase in ionic strength increased the adsorption of NOM at high surface concentration. The increase was more pronounced at high pH as the electrostatic repulsions between segments of the NOM are higher in that pH region, and are effectively screened by added salt. At low surface concentrations the direct surface-NOM interactions are predominant and electrostatic attraction, apparent on the carbon with the lower pH_{pzc}, was screened by added salt, thereby decreasing adsorption. The conversion from the "screening reduced" to "screening enhanced" adsorption regime took place where the negative charge from the adsorbed NOM balanced the positively charged sites on the activated carbon surface. At higher surface concentration adsorption takes place through non-electrostatic means, carboxyl groups are further from the surface and a strong lateral interaction is manifest. The lateral repulsion was screened by increased salt concentration and adsorption was enhanced.

4.1.2. Effect of NOM on adsorption of other contaminants

While naturally occurring problem compounds in drinking water are typically found in concentrations of ng or μg per litre, NOM is virtually always three to six orders of magnitude

higher in concentration, and always offers significant competition for adsorption sites. This usually substantially reduces the adsorption capacity for the target compound [44, 54-56].

An important issue in the understanding of the competitive effect is the determination of the composition and concentration of directly competing compounds. With a complex mixture such as NOM this is not a trivial exercise as the commonly utilised characterisation techniques such as DOC concentration and ultraviolet absorbance at 254 nm (UV254) are bulk parameters describing the total amount of carbon in NOM and the UV absorbing species respectively. These parameters give no information regarding the specific NOM components competing with a particular target contaminant, as this is likely to be only a small portion of the entire NOM "soup" [57]. Ideal adsorbed solution theory (IAST) was developed using thermodynamic considerations to describe multicomponent adsorption from solution onto activated carbon [58]. The model can be used to predict the adsorption of individual components in a mixture if the single solute isotherms are known [59]. Researchers have attempted to apply the model to describe the competition observed between NOM and microcontaminants [44]. This process isn't always successful, mainly due to the fact, as mentioned above, that competing NOM may only be a small portion of the total and the measurable parameter used to apply the IAST is usually the bulk DOC concentration.

The equivalent background compound (EBC) model can be used to obtain the concentration of competing compounds when the adsorption of the target compound is known in the presence and absence of the competing NOM [58, 60]. The EBC is not considered to be the entire NOM present in natural waters, as only an unknown portion of the NOM will compete. The model uses IAST to obtain the EBC adsorption parameters (Freundlich K, 1/n and initial EBC concentration) by the minimisation of the error between the experimental adsorption isotherm in the presence of NOM and the adsorption isotherm obtained from the IAST model. The EBC isotherm determined through this procedure can then be used to successfully predict the adsorption isotherms of the target compound in the same NOM solution at a range of initial concentrations [58, 61].

A reduction of adsorption of microcontaminants can also be caused by full or partial blockage of the pores by NOM, resulting in hindered kinetics, and reduction of available adsorption sites [44, 55, 58, 62]. Clearly, if the pores that are blocked by NOM are small, similar in size to the target microcontaminant, the distinction between the mechanisms of direct competition and pore blockage becomes unclear. The pore blockage mechanism is generally considered to be of lesser importance in the application of activated carbon for drinking water treatment [55, 58, 62].

As well as the direct competitive effect, which occurs on both PAC and GAC, NOM is responsible for "fouling" of GAC [1, 55]. When activated carbon is used in this form it is continuously exposed to water containing NOM, whereas the target compounds are often only present for particular periods; for example water sources are more prone to taste and odour problems during the warmer months. During the fouling process the NOM adsorbs onto the carbon surface, a process called preadsorption, and occupies adsorption sites that are no longer available when the filter is challenged by the target compound. The NOM compounds can also block or restrict activated carbon pores, resulting in hindered diffusion through the pore structure [55]. In addition, when NOM adsorbs onto the activated carbon surface it significantly changes the surface properties, in particular the charge [51].

4.2. Adsorption of taste and odour compounds

4.2.1. Tastes and odours of interest to the water industry

In general, tastes and odours of a natural origin are produced by algae and bacteria. Some of the descriptors of naturally occurring odours in drinking water are given in Table 2.

Table 2 Some common descriptors of naturally occurring odours in drinking water, and possible source organisms. Adapted from the AWQC internal document "Problem Phytoplankton, including Cyanobacteria, in Australian Freshwaters"

Odour descriptor	Potential causative microorganisms
Earthy/ Musty	*Anabaena sp. Aphanizomenon, actinomycetes Oscillatoria, Planktothrix, Pseudanabaena*
Grassy	*Actinastrum Closterium Ankistrodesmus Cosmarium Monoraphidium Pediastrum Scenedesmus Spirogyra Staurastrum Ulothrix*
Fishy	*Chlamydomonas Dictyosphaerium Eudorina Gonium Pandorina Volvox*
Cucumber	*Synura Uroglenopsis*
Geranium	*Asterionella Aulacoseira Cyclotella Fragilaria Tabellaria*
Septic	*Chlamydomonas Cladophora Gloeocystis Hydrodictyon*
Sulphurous	*Microcystis*
Garlic	*Chara*
Violet	*Cryptomonas Dinobyron Mallamonas*

Geosmin and 2-methylisoborneol (MIB) are the most common of the naturally-occurring taste and odour compounds. The compounds are produced by cyanobacteria and actinomycetes bacteria, and are an issue for water suppliers world-wide. Geosmin and MIB are small (0.6-0.8 nm) [64] moderately polar alicyclic tertiary alcohols; their molecular structures are shown in Fig. 6. The musty-earthy taste and odour the compounds impart to drinking water are detectable by consumers at approximately 10 ng L^{-1}. As the compounds are not removed during conventional treatment, or by chlorination, activated carbon is often used to reduce the concentration of the compounds to acceptable levels.

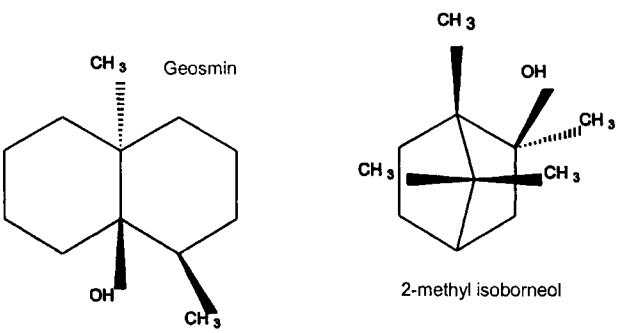


Fig.6. Molecular structures of the musty-earthy odour compounds geosmin and 2-methyl isoborneol (MIB)

4.2.2. Influences on the adsorption of MIB and geosmin

From the size of the MIB and geosmin molecules it could be predicted that they would adsorb in pores in the primary or smaller secondary micropore range, (approximately <1.0 nm) or pores slightly larger than the size of the molecules. The removal of MIB and geosmin from potable water has been the focus of many studies. Most studies have shown large differences between activated carbons [65-67]. In particular, wood-based chemically activated carbons performed consistently worse than coal- or wood-based carbons [58, 63]. While this was often attributed to surface chemistry of the wood-based carbons, more recent rigorous pore volume data indicate it is most likely due the lower volume of primary micropores present in these carbons [61]. For example, Pendleton *et al.* [63] related the low adsorption of MIB by chemically-activated wood-based carbons to the oxygen content of the surface. They hypothesised that the more hydrophilic surface would display less affinity for the MIB molecule than for water molecules, and thus the competition between the two molecules would result in lower adsorption of MIB. A later study by Pelekani [68], on the same carbon as that used by Pendleton et al. [63], showed that the wood-based carbons may not have had pores in the size range suitable for MIB adsorption. Fig. 7 illustrates the importance of high-quality micropore volume distributions for the interpretation of adsorption onto activated carbon. Fig. 7a shows the relationship between the volume of pores of width 1.0 nm and the equilibrium adsorption of MIB for the 7 activated carbons shown in Table 1 (this pore width gave the highest linear regression value). There is a direct linear relationship, with no apparent effect of the very different surface chemistries of the carbons. As mentioned earlier, pore volume distributions are obtained using nitrogen adsorption at very low relative pressures. For comparison, Figure 7b shows the relationship obtained when less rigorous pore volume distribution analysis is used; an effect of surface chemistry could easily be (erroneously) assumed from this data.

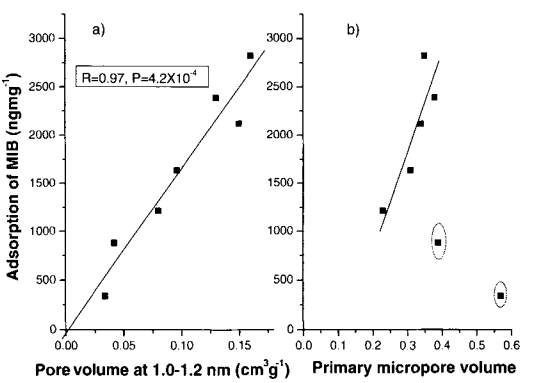


Fig.7. The relationship between a) MIB adsorption and pore volume distribution for a rigorous pore volume analysis and b) pore volumes determined at relative pressures above that necessary for the accurate determination of the volume of pores in the relevant size range

Another interesting aspect of the adsorption of these compounds is the difference in their ease of removal. Fig. 8 shows the percent removal of MIB and geosmin using a range of activated carbons. As expected, the different activated carbons display a range of removal

efficiencies for the two compounds. It is also clear that geosmin is more readily removed than MIB by all activated carbons. This could be due to several factors, such as the shape of the geosmin molecule – flatter in shape and possibly more amenable to adsorption in the slit-shaped pores in the activated carbon, resulting in higher adsorption energy – its slightly lower solubility, lower competition with NOM, or a combination of these factors. The figure shows that the trends are very similar for removal of the two compounds, although the actual percent removals are different. This suggests that the removal of geosmin is also to a large extent controlled by the pore size distribution of the carbons.

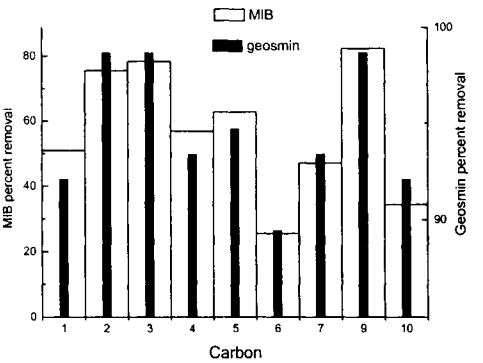


Fig. 8. Percent geosmin and MIB removal for ten activated carbons, Initial concentration, 100 ng L^{-1}, carbon dose, 30 mg L^{-1}. Results obtained in Hope Valley reservoir water, dissolved organic carbon concentration, 5.6 mg L^{-1}

4.2.3. Effect of NOM on the adsorption of MIB and geosmin

The application of activated carbon for the removal of natural contaminants invariably takes place in the presence of NOM. In all cases the NOM is present at concentrations much higher than the contaminant of interest, and has a significant detrimental effect on the adsorption of microcontaminants. Fig. 9 illustrates the effect of NOM on adsorption of MIB onto six PACs; a 90-98% reduction in surface concentration is seen at a solution concentration of 50 ng L^{-1} of MIB. The effect of NOM is slightly different for each carbon, although the linear relationship between pore volume distribution and MIB adsorption still holds, with a coefficient of determination (r^2) of 0.96 and a p value < 0.002 for pores of 1-1.2 nm width. The high level of adsorption competition seen in Fig. 9 is almost entirely due to the presence of very low molecular weight compounds (around 200 g mol^{-1}) in the NOM directly competing for adsorption sites [57, 58, 62]. Fig. 10 shows the molecular weight distribution of the NOM used in the experiments shown in Fig. 9. The peak representing the main competing species is indicated. As it is such a low percentage of the total NOM, bulk parameters such as DOC or UV absorbance cannot give an indication of the degree of competition to expect. In two studies the EBC model was used to estimate the concentration of NOM competing with MIB for adsorption sites in a series of molecular weight fractions isolated by different methods [57, 61]. The authors estimated that the concentration of NOM directly competing with MIB was between <1 and 10 percent of the total in the different NOM fractions.

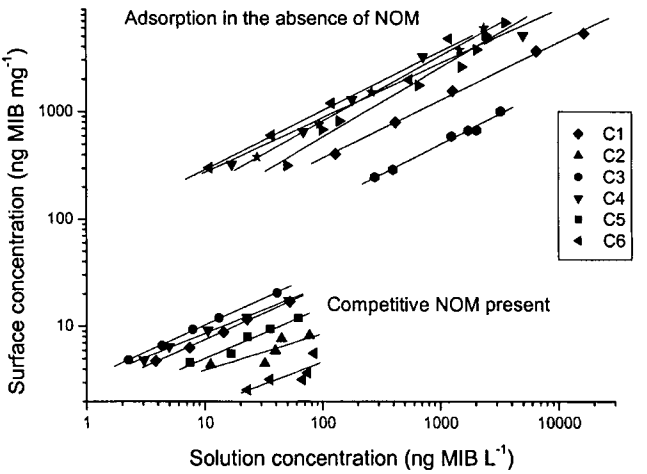


Fig. 9. Adsorption isotherms, MIB on six activated carbons in the presence and absence of NOM

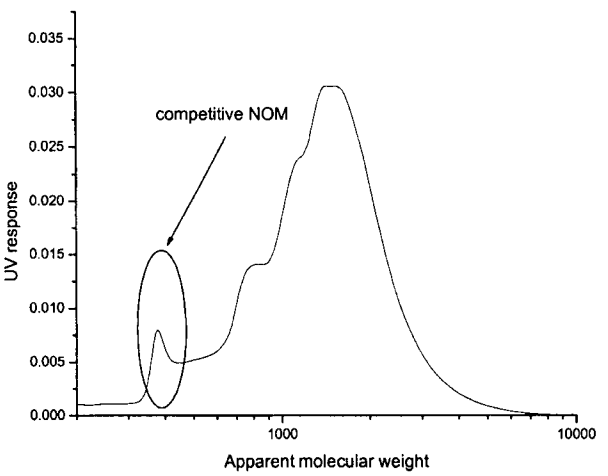


Fig 10. Molecular weight distribution of NOM showing the peak responsible the majority of the competitive adsorption effect with MIB.

4.3. Adsorption of algal toxins

4.3.1. Algal toxins of interest to the water industry

Although a range of algal toxins has been identified, the most common worldwide, and therefore the most widely studied, is the microcystin group of compounds [68]. The microcystins are of importance to the international water industry as they are hepatotoxins and have also been implicated as tumour promoters [69]. They are cyclic molecules, consisting of seven amino acid groups, and have molecular weights of around 1000 g mol^{-1}. More than sixty variants of the microcystin toxins have been identified, differing from one another mainly in variations of two of the amino acid groups indicated by X and Y in Fig.11, although minor variations to the other amino acids are also seen in some variants [68]. The microcystins are named according to the amino acids in the positions indicated, ie microcystin XY. The most common of variant of the toxin, microcystin LR (mLR), incorporates **L**eucine and a**R**ginine in the variable positions. All microcystins contain the Adda side chain, the structural unit largely responsible for the toxicity of the compounds through protein phosphatase inhibition. The structure of mLR is shown in Fig. 11. Although mLR is the most commonly reported of the microcystins, it is seldom the only microcystin found in a bloom situation, and is often not present at all, with the other variants predominating. The structures of three of the other most common variants, mLA (**L**eucine/**A**lanine), mYR (t**Y**rosine/a**R**ginine) and mRR (a**R**ginine/a**R**ginine) are shown in Fig. 12.

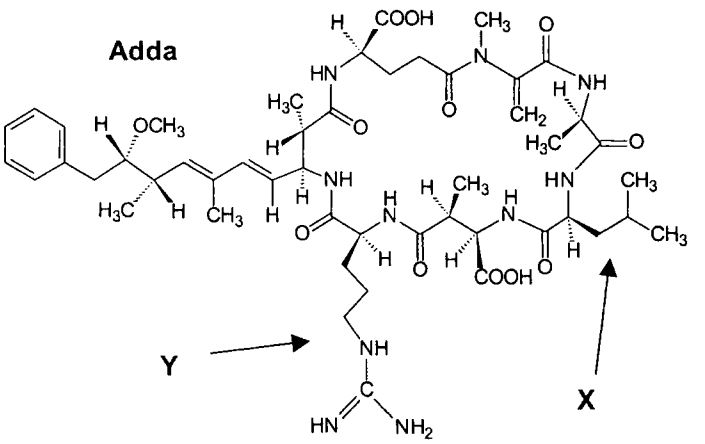


Fig. 11. Molecular structure of microcystin LR

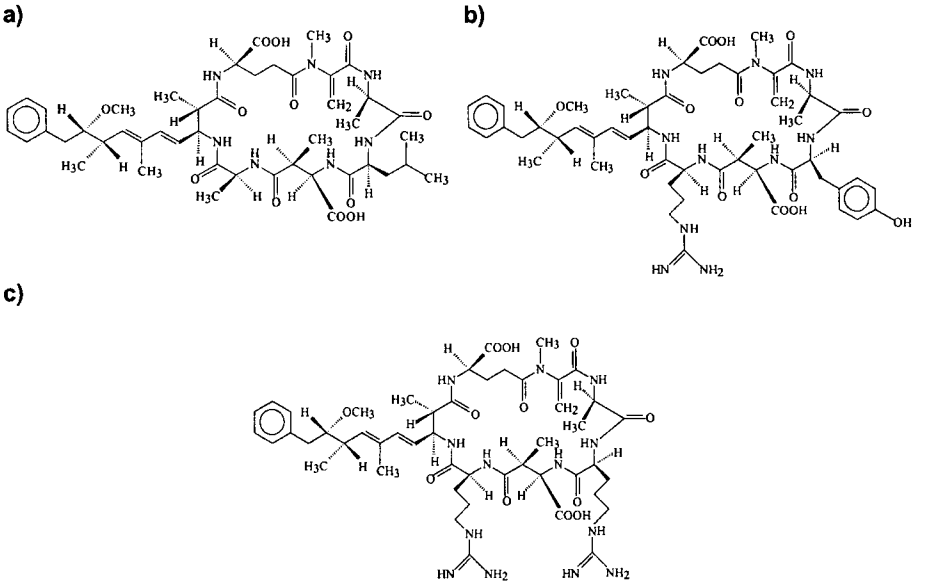


Fig. 12. Molecular structures of the microcystin toxins a) mLA, b) mYR and c) mRR

4.3.2. Influences on the adsorption of microcystins

The information available in the literature on the adsorption of mLR onto activated carbon indicates that, as with the adsorption of most microcontaminants, the removal efficiency is dependent on the type of activated carbon and the water quality conditions [28, 70-72].

Newcombe and Nicholson [72] have reported a direct linear relationship between the adsorption of the toxin and the volume of pores between 2 and 50 nm, with a linear regression giving parameters R^2= 0.97, P<0.0001 and N=9. Microcystin-LR is a hydrophobic molecule with an estimated size of 1.2 -2.6 nm, and the results suggest that the adsorption is mainly controlled by dispersion forces and the pore size distribution of the activated carbon, with surface chemistry probably only playing a minor role.

Another very interesting relationship is shown in Fig. 13. The maximum adsorption value (taken from the adsorption isotherm) of an NOM fraction with molecular weight 500-3000 g mol^{-1}, at pH 3, was directly related to the maximum adsorption of mLR under similar conditions, at pH 7.5. Clearly the NOM fraction must be both of similar size and display a similar adsorption mechanism (in the absence of electrostatic interactions) to the algal toxin mLR.

Microcystin LR is seldom the only microcystin present in a toxic algal bloom, and in some regions mLR is not the most commonly-occurring variant [73]. Very little information is available in the literature on the effect of water treatment processes on other variants. One investigation of the adsorption of microcystin variants other than mLR used relatively impure toxin extracts [74].

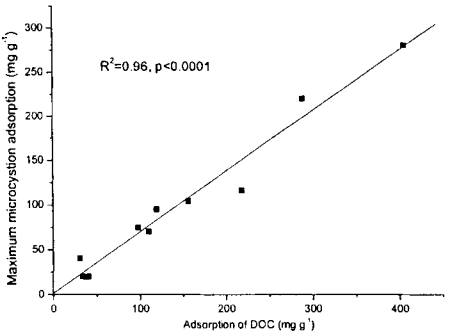


Fig. 13. Maximum microcystin-LR adsorption vs. NOM adsorption at pH 3

The authors suggested differences seen in the adsorption of the microcystin variants could have been due to different contaminant levels in the spiking material. The UKWIR undertook a computer modelling study to compare the octanol/water partition coefficients of nine microcystin variants [75]. With this information, and molecular size data, the authors concluded that the variants should respond similarly to water treatment processes, and, in particular, that the variants would adsorb onto activated carbon to the same, or greater, extent as the commonly-studied variant microcystin LR. Studies since, on four microcystin variants, have shown large differences in the adsorption of the variants [70, 76]. Fig. 14 shows the adsorption of four microcystin variants. The variants show removals between 10 and 80% under identical conditions of water quality and activated carbon dose and type. The cause of this difference is currently under investigation as the trend is unexpected. The physical properties of the variants (Table 3) indicate the expected trend would be opposite to the experimental trend, as it would be expected that the adsorption of these compounds would increase with increasing hydrophobicity and decreasing molecular weight.

Table 3
Properties of microcystins variants. The trend in hydrophobicity is based on the relative hydrophobicities of the substituent amino groups on the microcystins

	m-RR	m-YR	m-LR	m-LA
Variable amino acid (X)	arginine	tyrosine	leucine	leucine
Variable amino acid (Y)	arginine	arginine	arginine	alanine
Toxicity (LD 50 μg kg^{-1}) [1]	600	70	50	50
Net charge, pH 7	0	-1	-1	-2
Molecular weight	1037	1044	994	909
	→ Increasing hydrophobicity			

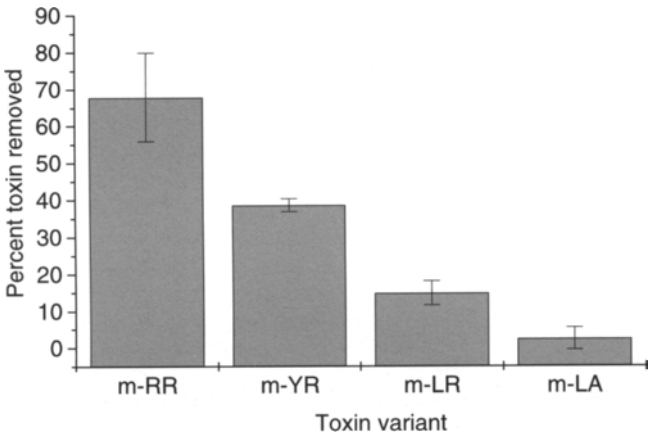


Fig.14. Adsorption of four microcystin variants by 2 mg L^{-1} powdered activated carbon C6

4.3.3. Effect of NOM on the adsorption of microcystins

The adsorption of microcystins has been shown to be strongly affected by NOM [76, 77]. Cook and Newcombe [76] reported that the adsorption of four microcystin variants was dependent on the DOC concentration, when other conditions of the system were held constant. This is probably a result of the direct competitive effect, where the competitive NOM, those compounds approximately the same molecular weight as the microcystins, makes up the bulk of the NOM. For example, the molecular weight distribution shown in Fig.10 is typical of a source water where PAC may be applied. The NOM of a similar size to the microcystin toxins, around 1000 g mol^{-1}, represents a main molecular weight fraction of the NOM. Therefore the bulk characterisation parameter, DOC, gives an indication of the concentration of competing compounds for microcystins. In contrast, the NOM providing the most competition with MIB and geosmin is low in concentration compared with the bulk DOC, and simple water quality parameters such as DOC or UV254 could not be used to give an indication of the potential for competition.

5. SUMMARY

Activated carbon adsorption is a very complex process. The adsorbents available for use in water treatment display a wide range of physical and chemical properties due the variations in raw materials and activation extent and method. This variation in properties leads to a very large variation in the processes involved in the application of activated carbon for a particular water treatment goal. The examples given in this chapter are the removal of natural organic material, the taste and odour compounds MIB and geosmin, and some variants of the microcystin class of toxins. The effect of NOM on the adsorption of the other compounds was also discussed.

It is generally recognised that both surface chemistry and pore volume distributions will play a role in adsorption onto activated carbon; however, the relative importance of these two factors is not immediately apparent, and will vary from adsorbate to adsorbate and similarly

between adsorbents. Solution adsorption data, and the relative importance of physical and chemical properties of the adsorbent, can be easily misinterpreted if pore volume distribution data are not of the highest quality and accuracy.

Despite the fact that the carbons discussed here showed a very large range of surface properties, the major influence on the adsorption of some important natural contaminants (also with a wide range of properties) was found to be the pore volume distributions of the carbons. In specific cases the removal of adsorbate could be interpreted in terms of the direct relationship between the pore size distribution of the carbon and the adsorbate size. Only the adsorption of charged natural organic material showed the influence of electrostatic effects related to the charge of the adsorbate and of the carbon surface. This does not necessarily suggest that there was no influence of surface chemistry, as the next chapter will explain; however, in most of the cases described in this chapter any influences of surface chemistry were overshadowed by the much stronger impact of the size of the pores.

In practical application of activated carbon for the removal of microcontaminants another important impact is the competitive effect provided by NOM present in the water. This effect is significant in terms of the additional cost associated with additional activated carbon required to overcome the competition. In general it has been found that the degree of competition depends on the concentration of organic compounds present in the NOM of approximately the same molecular weight as the target microcontaminant. Identifying and controlling this competitive factor is a difficult issue for the cost-effective application of activated carbon.

REFERENCES

[1] H. Sontheimer, J. Crittenden and S. Summers, Activated Carbon for Water Treatment; DVGW-Forschungsstelle, Engler-Bunte-Institute, Germany, 1988.
[2] R. Bansal, J. Donnet and F. Stoekli, Active Carbon, Marcel Dekker, New York, 1988.
[3] H.F. Stoeckli, Carbon , 28 (1990) 1.
[4] IUPAC Manual of Symbols and Terminology, Appendix 2, Pt 1, Colloid and Surface Chemistry. D. Pure Appl. Chem. 31(1972) 578.
[5] M. Daley, Y. Tandon, J. Economy and E. Hippo, Carbon, 34 (1996) 1191.
[6] V.M. Gunko and S.V. Mikhalovsky, Carbon, 42 (2004) 843.
[7] C. Brasquet, B. Rousseau, H. Estrade-Szwarckopf and P. Le Cloirec, Carbon, 38 (2000) 407.
[8] K. Gurudatt and V. Tripathi, Carbon, 36 (1998) 1371.
[9] T. Furuta, F. Minoura, C. Kim, K. Oshida, G. Dresselhaus and M. Dresselhaus, Supramolecular Sci., 5 (1998) 261.
[10] D. Lozano-Castello, D. Cazorla-Amoros and A. Linares-Solano, Chem. Eng. Technol., 26 (2003) 852.
[11] M. El-Merraoui, M Aoshima and K. Kaneko, Langmuir, 16 (2000) 4300.
[12] G. Newcombe, R. Hayes and M. Drikas, Colloids Surf., 78 (1993) 65.
[13] M. Studebaker, E. Huffman, A. Wolfe and L. Nabors, Ind. Eng. Chem., 48 (1956)175.
[14] H.P. Boehm, Advances in Catalysis, Vol 16. Academic Press, New York, 1966.
[15] M.O. Corapcioglu and C.P. Huang, Carbon, 25 (1987) 569.
[16] R.W. Coughlin, F.S. Ezra and R.N. Tan, J. Colloid Interface Sci., 28 (1968) 386.
[17] C. Ishizaki and I. Marti, Carbon, 19 (1981) 409.
[18] H. Boehm, Carbon, 32 (1994) 759.
[19] J. Menendez, J. Phillips, B. Xia and L. Radovic, Langmuir 12 (1996) 4404.
[20] B. Puri, Chemistry and Physics of Carbon, P. Walker, Jr, Ed. Marcel Dekker, New York, Vol 6 1970.
[21] C. Monterra, M. Low and A. Severdia, Carbon, 22 (1984) 5.
[22] C. Leon y Leon, J. Solar, V. Calemma and L. Radovic, Carbon, 30 (1992) 797.

[23] D. Leon y Leon and L. Radovic, Interfacial chemistry and electrochemistry of carbon surfaces. Chemistry and Physics of Carbon, P. Thrower (ed.), Marcel Dekker: New York, 1992.
[24] J. Menendez, M.J. Illian-Gomez, C. Leon y Leon and L. Radovic, Carbon, 33 (1995)1655.
[25] L. Radovic, J. Ume and A. Scaroni, On tailoring the surface chemistry of activated carbons for their use in purification of aqueous effluents. Proceedings of the 6th International Conference on Fundamentals of Adsorption, M. Le Van, M. Norwell (eds.), Kluwer, The Netherlands, 1996.
[26] J. Menendez, B. Xia, J. Phillips and L. Radovic, Langmuir, 13 (1997) 3414.
[27] R.S. Summers and P.V. Roberts. J. Colloid Interface Sci., 122 (1988) 382-397.
[28] C. Donati, M. Drikas, R. Hayes and G. Newcombe, Water Res., 28 (1994) 1735.
[29] R. Cini, F. Pantani and G. Sorace, Physicochemical aspects of the use of activated carbon in drinking water treatment. In I.H. Suffet and M.J. McGuire (eds), Activated Carbon Adsorption of Organics from the Aqueous Phase, Vol.1. Ann Arbor Science, Michigan, 1980.
[30] S.D. Faust and O.M. Aly, Chemistry of Water Treatment. Ann Arbor Science, Ann Arbor, Michigan, 1983.
[31] L.K. Koopal, Zeitschrift fur Wasser und Abwasser Forschung, 16 (1983) S91.
[32] G. Muller, C.J. Radke and J.M. Prausnitz, J. Colloid Interface Sci. 103 (1985) 466.
[33] R.W. Coughlin, F. S. Ezra and R. N. Tan, J. Colloid Interface Sci. 28 (1968) 386.
[34] W. Adamson, Physical Chemistry of Surfaces. Wiley-Interscience, London, 1967.
[35] M. Greenbank and M. Manes, J. Phys. Chem., 85 (1981) 3050.
[36] K. Urano, Y. Koichi and E. Yamamoto, J. Colloid Interface Sci., 861 (1982) 43.
[37] M.J. McGuire, I.H. Suffet and J.V. Radziul, J. - Am. Water Works Assoc., 70 (1978) 565.
[38] J. Jain and V.L. Snoeyink, J. Water Pollution Control Federation, 4512 (1973) 2463.
[39] F.J. Stevenson, Geochemistry of soil humic substances, G. Aitken, D. McKnight, R. Wershaw, P. MacCarthy (eds.) Humic Substances in Soil, Sediment and Water: Geochemistry, Isolation and Characterisation, John Wiley and Sons, New York, 1985.
[40] G. Choudry, Humic Substances - Structural, Photophysical, Photochemical and Free Radical Aspects of Interactions with Environmental Chemicals, Gordon and Breach, New York, 1984.
[41] F. Frimmel and R. Christman, Humic Substances and their Role in the Environment, John Wiley and Sons, New York, 1988.
[42] J.-P. Croue, G. Korshin and M. Benjamin, Characterisation of Natural Organic Matter in Drinking Water, AwwaRF Report No. 90780, AWWA Research Foundation, Denver, Co., USA, 2000.
[43] K. Ebie, F.,Li and T. Hagishita, Water Supply, 13 (1995) 65-70.
[44] S. Kasoaka, Y. Sakata, E. Tanaka and R. Naitoh, Int. Chem. Eng., 29(4) (1989) 734-742.
[45] J.E. Kilduff, T. Karanfil, Y. Chin and W.J. Weber, Environ. Sci. Technol., 30 (1996)1336-1343.
[46] R.W. Coughlin, F.S. Ezra and R.N. Tan, J. Colloid Interface Sci., 28 (1968) 386-396.
[47] P. LaFrance and M. Mazet, J. - Am. Water Works Assoc., 81(4) (1989) 155-162.
[48] G. Newcombe, M. Drikas, S. Assemi and R. Beckett, Water Res., 31(5) (1997) 963-972.
[49] G. Newcombe, A Study of Natural Organic Material and its Adsorption onto Activated Carbon. PhD thesis, University of South Australia, Adelaide, 1999.
[50] G. Newcombe, J. Colloid Interface Sci., 164 (1994) 452.
[51] G. Newcombe and M. Drikas, Carbon, 35(9) (1997) 1239-1250.
[52] M. Bjelopavlic, G. Newcombe and R. Hayes J. Colloid Interface Sci., 210 (1998) 271-280.
[53] A. Derylo-Marczewska, Langmuir, 9 (1993) 2344.
[54] J.E. Kilduff and A. Wigton, Environ. Sci. Technol., 33 (1999) 250.
[55] J.E. Kilduff, T. Karanfil and W.J. Weber Jr., J. Colloid Interface Sci., 205 (1998) 280.
[56] G. Newcombe, M. Drikas and R. Hayes, Water Res, 31 (1997) 1065.
[57] G. Newcombe, J. Morrison, C. Hepplewhite and D.R.U. Knappe, Carbon, 40 (2002) 2147.
[58] C.J. Radke and J.M. Prausnitz, Amer. Inst. Chem. Eng. J., 18 (1972) 761.
[59] J.C. Crittenden, P.J. Luft, D.W. Hand, J.L. Oravitz, S. Loper and M. Ari, Environ. Sci. Technol., 19 (1985) 1037.
[60] D. Knappe, Predicting the Removal of Atrazine by Powdered and Granular Activated Carbon. PhD Thesis, University of Illinois, Urbana-Champaign, 1996.

[61] C. Hepplewhite, G. Newcombe and D.R..U. Knappe, Water Sci. Technol., 49 (2004) 257.
[62] C. Pelekani and V.L. Snoeyink, Carbon, 39 (2001) 25.
[63] P. Pendleton, S. Wong, R. Schumann, G. Levay, R. Denoyel and J. Rouquerol, Carbon, 35 (1997) 1141.
[64] S. Lalezary, M. Pirbazari and M. McGuire, J. - Am. Water Works Assoc., 78 (1986)76.
[65] D. Hertzing, V. Snoeyink and N. Wood, J. - Am. Water Works Assoc., 69 (1977) 223.
[66] E. Vik, R. Storhaug, H. Naes, H. Utkilen, Water Sci. Technol., 20 (1988) 229.
[67] C. Pelekani, The Role of Pore Size Distribution in Competitive Adsorption on Activated Carbon; PhD Dissertation, Universtiy of Illinois at Urbana-Champaign, 1999.
[68] R.S. Yoo, W.W. Carmichael, R.C. Hoehn and S.E. Hrudey, Cyanobacterial (Blue-Green Algal) Toxins: A Resource Guide. American Water Works Association Research Foundation report no.90693 , 1995.
[69] I. Falconer, Health implications of cyanobacterial (blue-green algal) toxins, D.A. Steffensen and B.C. Nicholson (eds), Proceedings of an International Workshop, Toxic Cyanobacteria Current Status of Research and Management Australian Water Quality Centre, Adelaide, Australia, 1994.
[70] D. Cook, and G. Newcombe, Water Sci.Technol. Water Supply, 2 (2002) 201.
[71] J. Hart and P. Stott, Microcystin-LR Removal from Water (Report FR0367), Marlow, Buckinghamshire: Foundation for Water Research, 1993.
[72] G. Newcombe and B. Nicholson (2004) J. Water Supply: Res. Technol. – Aqua, 53(4) 227-239.
[73] Falconer, J. Bartram, I. Chorus, T. Kuiper-Goodman, H. Utkilen, M. Burch, and G. Codd Safe levels and practices, I. Chorus and J. Bartam (eds.), Toxic Cyanobacteria in Water, A Guide to Their Public Health Consequences, Monitoring and Management , E & FN Spon, World Health Organization, 1999.
[74] Z. Mohamed, W. Carmichael, J. An, and H. El-Sharouny, Environ. Toxicology, 14 (1999) 197.
[75] UKWIR Algal Toxins, Occurrence and Treatability of Anatoxin a and Microcystins. Report number 97/DW/07/05, 1997.
[76] D. Cook and G. Newcombe, Effect of natural organic matter concentration and character on the adsorption of microcystin analogues onto PAC. Proceedings of the American Water Works Association Water Quality Technology Conference, November 11-13, 2002, Seattle, USA. CD ROM
[77] C. Donati, M. Drikas, R. Hayes, and G. Newcombe (1993). Water, J. Australian Water Wastewater Assoc., 20(3), 25.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 9: Surface chemistry effects in activated carbon adsorption of industrial pollutants

D.R.U. Knappe

Department of Civil, Construction, and Environmental Engineering,
North Carolina State University, Raleigh, NC 27695-7908, USA

1. INTRODUCTION

Activated carbon adsorption processes are frequently identified as the best available technology for the removal of industrial pollutants such as pesticides, chlorinated organic solvents, and fuel hydrocarbons from drinking water [1]. Although commonly employed in water treatment since the 1930s [2], activated carbon adsorption processes are implemented on the basis of design procedures that remain largely empirical; i.e., water treatment professionals rely on the output of "black-box" models that are calibrated with experimental data specific to an individual treatment scenario. The lack of more mechanistic and thus more broadly applicable models is primarily related to (1) the heterogeneous pore structure and surface chemistry of activated carbons, and (2) the complex nature of the solutions from which industrial pollutants need to be removed; apart from the targeted industrial pollutant(s), potable water sources contain other known and unknown industrial pollutants, naturally occurring pollutants (e.g., taste and odor compounds or cyanotoxins), and natural organic matter (NOM). Aquatic NOM, which is ubiquitous in drinking water sources, refers to a mixture of organic compounds derived from decaying terrestrial vegetation and microbial material [3].

The principal goal of this chapter is to describe the effects of activated carbon surface chemistry on the adsorption of industrial pollutants from water. To facilitate this rather complex undertaking, fundamental interactions controlling the adsorption of aqueous pollutants on activated carbon will initially be reviewed in Section 2. Subsequently, the characteristics of activated carbons (Section 3) and representative industrial pollutants (Section 4) will be described. Finally, the effects of activated carbon surface chemistry on the adsorption of industrial pollutants from ultrapure water and from solutions containing NOM will be discussed (Section 5). Throughout this chapter, emphasis will be given to solution conditions that apply to the treatment of drinking water. In potable water sources, industrial pollutants are typically present at trace concentrations that are well below the pollutant's aqueous solubility limit. When adsorbing industrial pollutants from very dilute solutions, the effects of activated carbon surface chemistry on adsorption capacity are more pronounced because surface coverages are low [e.g., 4-6]. In contrast, the total micropore volume or the commonly measured BET surface area primarily determine the adsorption capacity of activated carbon for a pollutant at higher surface coverages, which may be achieved in the treatment of highly contaminated waters or in solvent recovery operations [e.g., 6, 7].

2. FUNDAMENTALS OF ADSORPTION

To clearly illustrate the effects of activated carbon surface chemistry on adsorption, fundamental interactions controlling the adsorption of aqueous pollutants need to be considered first. The adsorptive removal of industrial pollutants from aqueous solution is affected by the following three interactions [8]:

- pollutant (adsorbate) – activated carbon (adsorbent) interactions,
- water (solvent) – activated carbon (adsorbent) interactions, and
- pollutant (adsorbate) – water (solvent) interactions.

Interactions between industrial pollutants and activated carbon pore surfaces are primarily controlled by (1) non-specific dispersive interactions (i.e., van der Waals interactions) between pollutant molecules or ions and the condensed polyaromatic sheets (basal planes) that constitute the building blocks of activated carbons and (2) electrostatic interactions (i.e., Coulombic interactions) between ionic industrial pollutants and charged activated carbon pore surfaces [5, 9, 10]. Non-specific dispersive interactions are always attractive, and their strength tends to increase with increasing adsorbate size, polarizability (i.e., the ease with which uneven electron distributions can be induced in a molecule), planarity, and, for aromatic pollutants, electron-donating (or decreased electron-withdrawing) strength of ring substituents [7, 8, 10-12]. In addition, the strength of dispersive interactions between aromatic adsorbates and activated carbon pore surfaces may be affected by the density of delocalized π-electrons on the activated carbon basal planes [10]. Section 3 describes in more detail the effects of activated carbon surface chemistry on the density of delocalized π-electrons while Section 4 summarizes physicochemical characteristics of representative organic pollutants. It should be pointed out that other intermolecular interactions may control the adsorption of some aqueous industrial pollutants on some activated carbons (e.g., electron donor-acceptor complexes may develop between (1) carbonyl groups on the activated carbon surface and the aromatic ring of nitrophenols [4] or (2) unsaturated carbon atoms located at the basal plane edges and aromatic adsorbates [13]). Whether or not a specific adsorption mechanism is feasible in dilute aqueous solutions depends on the strength of the interaction between the pollutant and a specific adsorption site relative to that between water and the same adsorption site.

Coulombic adsorbate-adsorbent interactions are strongly influenced by solution pH and ionic strength. Solution pH controls the (de)protonation of surface functional groups, which, in turn, determines the surface charge of activated carbon pores. Furthermore, solution pH controls the (de)protonation of ionizable industrial pollutants. Therefore, solution pH determines whether electrostatic interactions between the pollutant and the activated carbon surface are attractive (e.g., adsorption of anilinium cations on a negatively charged activated carbon surface) or repulsive (e.g., adsorption of phenolate anions on a negatively charged activated carbon surface). With increasing ionic strength, the strength of Coulombic interactions decreases at a given distance of separation between the adsorbate and the adsorbent [5]. Thus, at low surface coverages, an increase in ionic strength negatively affects the adsorption of counter-ions while it positively affects the adsorption of co-ions [14, 15].

Interactions between the solvent (i.e., water) and the activated carbon surface are primarily controlled by the availability of surface functional groups to which water can adsorb via hydrogen bonding. In general, these surface functional groups contain oxygen, but nitrogen-containing functional groups and mineral impurities in activated carbons may also serve as specific sites for water adsorption. For activated carbons with similar physical

characteristics, a greater availability of specific adsorption sites for water results in a smaller adsorption capacity for organic micropollutants [e.g., 6, 16-20].

Finally, interactions between the adsorbate and the solvent (i.e., water) play an important role in the adsorption of aqueous pollutants. Generally, the more hydrophobic the pollutant, the greater is its tendency to escape from aqueous solution and to adsorb on the activated carbon surface, and this tendency is known as solvent-motivated adsorption [21]. To eliminate effects related to pollutant – water interactions, aqueous-phase pollutant concentrations in an adsorption isotherm plot should be normalized by the aqueous solubility of the pollutant at the temperature of the experiment [4]. Aqueous solubilities of representative industrial pollutants are presented in Section 4.

3. ACTIVATED CARBON CHARACTERISTICS

Typically, water treatment plants employ activated carbons in powdered or granular form, and these activated carbons are manufactured from relatively heterogeneous base materials such as bituminous or sub-bituminous coal, lignite, peat, coconut shells, or wood. Activated carbons (including activated carbon fibers) can also be prepared from homogeneous polymeric base materials such as polyacrylonitrile, cellulose or phenolic resin [22], but such activated carbons are not currently applied in drinking water treatment. The manufacture of activated carbons involves two important steps: (1) carbonization and (2) activation. During carbonization, the base material is heated in an inert atmosphere to remove volatile matter and to produce condensed, polyaromatic sheets that are the fundamental building blocks of activated carbons [23]. The chars produced during the carbonization step exhibit a poorly developed pore structure, and the interstices between the condensed, polyaromatic sheets may be partially filled with tarry substances [23]. Further development of the internal pore structure is most commonly achieved by thermal activation with steam and CO_2 as oxidants at temperatures between 800 and 1100°C. Most activated carbons for water treatment are produced by sequential carbonization/thermal activation processes. Alternatively, carbonization and activation can be accomplished concurrently when a chemical activating agent, such as phosphoric acid, is combined with the base material, and the mixture is subsequently heated in the absence of oxygen. The latter process, which is commonly referred to as 'chemical activation,' is used in the manufacture of some wood-based activated carbons that are employed in water treatment. Both the base material and manufacturing conditions affect activated carbon pore structure and surface chemistry [23, 24], and the following sections give a brief overview of activated carbon pore structure and a more detailed description of activated carbon surface chemistry.

3.1. Pore structure

As discussed in more detail in the previous chapter, activated carbon is a highly porous material with an internal surface area that typically ranges from about 800-1500 $m^2 g^{-1}$ [23]. The pores that give rise to this large surface area can be envisioned as spaces between irregularly arranged graphite-like platelets [25] or condensed, polyaromatic sheets [23]. To classify pores according to size, the International Union of Pure and Applied Chemistry differentiates between (1) micropores (<2 nm width), (2) mesopores (2 - 50 nm width), and (3) macropores (> 50 nm width) [26]. Powdered and granular activated carbons for water treatment typically exhibit a heterogeneous pore structure, in which micropores, mesopores and macropores are all present [22].

The size of adsorbent pores affects the adsorption of organic contaminants in two important ways. First, adsorption strength increases with decreasing pore size because (1) contact points between the adsorbate and the adsorbent surface increase [27] and (2) adsorption potentials between opposing pore walls begin to overlap once the micropore width is less than twice the adsorbate diameter [e.g., 28, 29]. Second, size exclusion limits the adsorption of contaminants of a given size and shape if pores are too small. These observations suggest that adsorption takes place in the smallest pores that are accessible to a given pollutant. Because many organic contaminants are small (i.e., total surface areas of about 100-300 $Å^2$ [30], which corresponds to spherical diameters of about 5.5-10 Å), the above observations suggest that the presence of small micropores is important for their removal from aqueous solution. For example the results of Li et al. [19] suggest that the adsorption capacity of trichloroethene (TCE) is proportional to the volume of micropores with widths of about 7-10 Å while that of methyl tertiary-butyl ether (MTBE) is proportional to the volume of micropores with widths of about 8-11 Å. These pore sizes are approximately 1.3 to 1.8 times the kinetic diameter of the adsorbate molecules. Using a newer software package to calculate micropore size distributions, a recent study [20] indicated that TCE adsorbs primarily in pores with widths of about 5-8 Å, suggesting that pores with dimensions very close to those of the adsorbate are required to effectively remove industrial pollutants from dilute aqueous solutions.

3.2. Surface chemistry

At the edges of the condensed, polyaromatic sheets that constitute the building blocks of activated carbons, heteroatoms, i.e. atoms other than carbon, are present, and these heteroatoms define the chemical characteristics of activated carbon pore surfaces. A typical elemental composition of activated carbon is approximately 88% C, 6-7% O, 1% S, 0.5% N, and 0.5% H with the remainder being mineral matter (i.e., ash) [23]. However, the elemental composition of activated carbons can vary substantially from these average values; e.g., the oxygen content can range from as low as 1% to as high as 25% [23] while the ash content can range from about 1 to 20% [31]. Because of its abundance and its profound effects on activated carbon hydrophilicity and surface charge, oxygen is generally the most important heteroatom from a standpoint of activated carbon surface chemistry. As discussed in more detail below, oxygen-containing functionalities (1) represent specific sites for water adsorption, (2) affect the density of delocalized π-electrons on the basal plane, and (3) control the pH-dependent surface charge of activated carbons.

3.2.1. Acidic functional groups

Activated carbons exhibit an acidic character (i.e., the ability to neutralize bases) when activated at relatively low temperatures. Also, activated carbons can assume an acidic character when exposed to (1) oxygen between 200 and 700°C following high-temperature treatment in an inert atmosphere or (2) aqueous solutions containing oxidants such as hydrogen peroxide, ammonium persulfate, nitric acid, or nitric and sulfuric acid mixtures [e.g., 32-34]. The acidic character of activated carbons is primarily attributed to surface oxygen groups such as carboxylic acid, lactone, phenol, and lactol groups (Fig. 1) that are located at the edges of the polyaromatic basal plane sheets [e.g., 34, 35]. The dissociation constants of acidic functional groups (K_a values) on activated carbon surfaces vary over several orders of magnitude. E.g, carboxylic acid groups exhibit pK_a values in the range of 3 to 6, lactone groups in the range of 7 to 9, and phenol groups in the range of 8 to 11 [36]. Results of potentiometric titrations suggest that some activated carbons exhibit just one or two

readily distinguishable pK_a values [e.g., 37] while others have a broad distribution of pK_a values [e.g., 38].

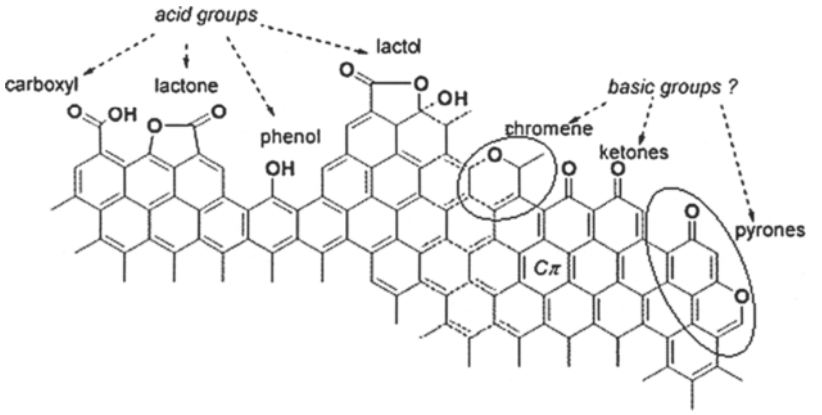


Fig. 1. Proposed acidic and basic oxygen-containing functional groups on activated carbon surfaces; Cπ indicates that basicity may also be associated with delocalized π electrons on the polyaromatic basal plane (Reprinted with permission from Fuente et al. [39]. Copyright 2003 American Chemical Society)

3.2.2. Basic functional groups

The ability of activated carbons to bind protons (H^+) is commonly associated with basic oxygen-containing functional groups, delocalized π-electrons of the polyaromatic sheets (basal planes), nitrogen-containing functional groups, and/or inorganic impurities (metal oxides). Which factor or combination of factors primarily gives activated carbons their basic character is still a matter of debate [see 35]. As shown in Fig. 1, oxygen-containing functionalities such as chromene- and pyrone-type structures as well as diketone or quinone groups may be responsible for the basicity of activated carbons [34, 35, 40-42]. Pyrone-type structures; i.e., structures containing carbonyl and ether oxygen atoms in non-neighboring rings, have been the focus of recent investigations, and the results of quantum chemical calculations and infrared spectroscopy suggest that pyrone-type structures may indeed contribute to the basicity of activated carbons [39, 43]. Depending on the positions of the carbonyl and ether oxygen atoms, protonated pyrone-type structures are predicted to exhibit pK_a values between 4 and 13 [39].

Other researchers have argued that the basic character of activated carbons arises primarily from a high density of delocalized π-electrons on the condensed polyaromatic basal plane sheets (see Cπ in Fig. 1) [25, 44]. These electron-rich Lewis base sites are capable of forming cation-π complexes (i.e., H_3O^+-π complexes), but the proton binding strength associated with these sites appears to be unknown to date [35]. The density of the delocalized π-electrons is affected by the nature of the functional groups attached to the basal plane edges; e.g., carboxylic acid groups withdraw electrons while phenol groups donate electrons [15, 45]. Generally, the basic character of activated carbons increases during heat-treatment with such gases as hydrogen, nitrogen, or helium because these treatments not only remove inherently acidic oxygen-containing functional groups, but also oxygen-containing functionalities that withdraw π-electrons from the basal planes. Apart from increasing the

basic character of activated carbons, an increasing density of delocalized π-electrons appears to increase the strength of dispersive interactions between aromatic adsorbates and activated carbon basal planes [10, 15, 16, 46].

Apart from oxygen-containing functional groups and delocalized π-electrons, nitrogen-containing functional groups may contribute to the basicity of activated carbons. When treated with ammonia at high temperatures, N is incorporated into the polyaromatic basal planes primarily in the form of pyridine-type (or acridine-type) structures [47-50] that are capable of binding protons in the pH 4-6 range [35]. Although a likely contributor to the basicity of activated carbons tailored by heat-treatment with ammonia, pyridine-type functional groups are not expected to play a key role in the basicity of typical activated carbons used in water treatment applications. Amines could also contribute to the basicity of activated carbons [36], but amine groups do not appear to be stable at temperatures >400-500°C [35].

Finally, metal oxides (i.e. ash constituents) may represent a little studied contributor to activated carbon basicity. As shown by Montes-Morán et al. [35], acid-washing of an activated carbon caused a decrease in ash content and a concurrent decrease in basicity, as measured by the equilibrium pH that resulted when contacting activated carbons with water.

3.2.3. Surface charge of activated carbons

Because both acidic and basic surface functional groups are typically present at the edges of the polyaromatic basal plane sheets, the surface charge of activated carbons changes as a function of solution pH. Fig. 2 depicts acid/base reactions of common activated carbon surface functional groups and their associated pK_a values. Fig. 2 further illustrates that the protonation of basic groups yields surface sites that are positively charged while the deprotonation of acidic groups yields surface sites that are negatively charged. The pK_a values associated with the protonated functional groups suggest that both positive and negative surface sites can co-exist on the activated carbon surface at a given solution pH. For example, in the pH 4 to 6 range, some pyrone-type and pyridine-type structures may exist in the protonated cationic form while some carboxyl groups may exist in deprotonated anionic form. As a result, activated carbon surfaces are amphoteric [10]. Although positively and negatively charged sites can both be present on activated carbon surfaces, currently available analytical techniques only permit the determination of the net surface charge.

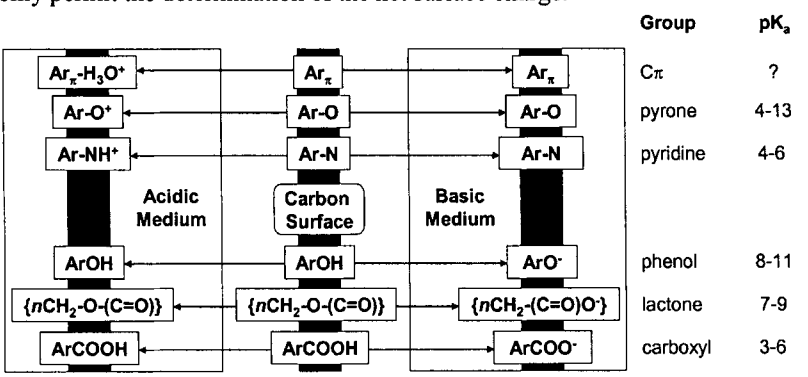


Fig. 2. Effect of surface functional groups and solution pH on activated carbon surface charge (after Radovic et al. [10], amended with information from [36] and [39])

Fig. 2 suggests that the net surface charge of activated carbons is positive in acidic solutions and negative in basic solutions. However, the magnitude of the surface charge at a given solution pH strongly depends on the surface chemistry of individual activated carbons (see Fig. 5 in the previous chapter as well as Newcombe et al. [51]). Fig. 5 in the previous chapter also illustrates that the dissociation of protonated functional groups with increasing solution pH leads to an almost continuous decrease in net positive surface charge or increase in net negative surface charge over the tested pH range (~4-10). As pointed out in section 2, the surface charge of activated carbons determines whether Coulombic interactions between ionic pollutants and the activated carbon pore surfaces are attractive or repulsive.

4. CHARACTERISTICS OF REPRESENTATIVE INDUSTRIAL POLLUTANTS

For treated drinking water, the US EPA has currently set maximum contaminant levels (MCLs) for 53 organic compounds or compound groups (e.g., polychlorinated biphenyls, haloacetic acids, trihalomethanes) [1]. In addition, the contaminant candidate list (CCL) of the US EPA contains 43 organic compounds or groups of compounds (e.g., triazines) that may require regulation in the future [52]. Apart from these pollutants, water treatment professionals have to be concerned about emerging anthropogenic contaminants such as endocrine disruptors and pharmaceutically active compounds as well as chemical agents that may be used in a potential terrorist attack.

Regulated and non-regulated contaminants exhibit a broad range of physicochemical characteristics that affect their adsorption behavior. The adsorbability of aqueous organic compounds is largely determined by the following factors [e.g. 7, 8, 11, 15]:

- a lower **aqueous solubility** of the adsorbate implies an increasing preference of the adsorbate to leave the aqueous phase and adsorb onto the activated carbon surface,
- the adsorbate's **molecular size and shape** determine which adsorption sites are physically accessible and which pore sizes produce the highest sorption energies,
- an increasing **polarizability** of the adsorbate implies an increased strength of dispersion interactions between the adsorbate and the adsorbent, and
- the adsorbate's **acid dissociation constant** determines its charge and degree of (de)protonation at a given solution pH and thus affects Coulombic interactions between the adsorbate and charged activated carbon pore surfaces.

In addition, adsorbate characteristics such as planarity, hydrogen-bond donor and/or acceptor strengths, and electron donating or withdrawing strengths of aromatic ring substituents may affect the adsorbability of organic contaminants from aqueous solutions [10, 12, 53, 54].

Table 1 summarizes aqueous solubilities (C_s), molecular weights (MW), molar volumes (V_m), polarizabilities (α), and pK_a values for organic contaminants that are currently regulated by the US EPA or are on the CCL of the US EPA. As shown in Table 1, aqueous solubilities of industrial pollutants vary over more than nine orders of magnitude (0.0002 mg L^{-1} for dioxin to about $1x10^6$ mg L^{-1} for dichloroacetic acid). The broad range of aqueous solubilities indicates that the adsorbability of organic contaminants, based on their "motivation" to leave the aqueous phase, varies greatly.

The molecular weights of the pollutants summarized in Table 1 vary from about 60 to 600 g mol^{-1}, and their molar volumes range from about 60 to 400 cm^3 mol^{-1}. Assuming spherical molecules, the latter values suggest that organic pollutants in Table 1 exhibit diameters in the range of about 6 to 11 Å. Assuming further that pollutants preferentially

adsorb in pores of similar dimensions, adsorption capacities for the pollutants summarized in Table 1 should be largest on activated carbons with relatively large volumes of small micropores. However, which micropore size distribution will produce the maximum adsorption capacity for a given contaminant depends on the contaminant's size and shape, [e.g., 19, 20].

Polarizabilities of the pollutants shown in Table 1 range from about 6 to $45x10^{-24}$ cm^3. Generally, polarizabilities increase with increasing molecular weight and aromaticity. Thus, strong van der Waals interactions can be expected between activated carbon basal planes and compounds such as di(2-ethylhexyl)phthalate, an aromatic compound of relatively high molecular weight, or benzo[a]pyrene, a polyaromatic hydrocarbon (PAH).

The pK_a values in Table 1 illustrate that several industrial pollutants are present in ionic form at solution pHs typically encountered in water treatment. For example, 2,4-D and Dalapon are predominantly present as singly charged anions at neutral pH. Thus, Coulombic interactions in addition to van der Waals interactions will be important in the adsorption of these pesticides from aqueous solution. While the majority of the contaminants summarized in Table 1 are neutral molecules at solution pHs typically encountered in water treatment, a number of emerging contaminants, such as common antimicrobial compounds, occur as anions, cations, or zwitterions at typical water treatment pHs [eg. 63].

5. EFFECTS OF SURFACE CHEMISTRY ON ADSORPTION

Many researchers have investigated the effects of activated carbon surface chemistry on the adsorption of organic contaminants from water, and an extensive review on the subject was recently published [10]. To explain the effects of activated carbon surface chemistry on the adsorption of aqueous organic contaminants, both solvent-adsorbent (Section 5.1) and adsorbate-adsorbent interactions (Sections 5.2 and 5.3) need to be considered.

5.1. Water adsorption

In aqueous-phase applications, activated carbon pores are initially filled with water. Therefore, a volume of water equal to that required by the pollutant needs to be displaced from the activated carbon pores for adsorption of the pollutant to take place [7]. Consequently, the strength with which water interacts with the activated carbon surface affects the availability of adsorption sites for aqueous pollutants.

The affinity of water for the hydrophobic basal planes of activated carbons is low. Because the comparatively strong hydrogen bonds between water molecules need to be disrupted for water to adsorb on a surface, water adsorption is therefore not energetically favorable on the hydrophobic basal planes. However, when the edges of activated carbon basal planes are decorated with oxygen-containing functional groups, water is able to form hydrogen bonds with these "primary active centers" [64, 65]. Subsequently, additional water molecules can form hydrogen bonds with the water molecules adsorbed at the primary active sites, and the resulting water clusters begin to fill the activated carbon pores [64-66]. As the concentration of primary active centers increases, the affinity of water for the activated carbon surface increases, and the displacement of adsorbed water by pollutant molecules becomes energetically more difficult [67]. Furthermore, the results of a recent study suggest that water molecules adsorbed at the primary active centers are not displaced by phenol in either dilute or concentrated aqueous solution [6]. As a result of the strong, specific interaction between water and oxygen-containing surface functional groups, the available surface area for phenol adsorption decreased by 71 m^2 for each mmol O on the activated carbon surface [6].

Table 1
Characteristics of organic pollutants currently regulated by the US EPA or on the contaminant candidate list of the US EPA

Chemical name	Formula	CAS #	C_s (mg/L)	MW (g/mol)	V_m (cm^3/mol)	$\alpha * 10^{24}$ (cm^3)	pK_a #
Acetochlor*	$C_{14}H_{20}ClNO_2$	34256-82-1	223	269.77	240.9	29.91	
Alachlor@	$C_{14}H_{20}ClNO_2$	15972-60-8	240	269.77	240.9	29.95	
Aldrin*	$C_{12}H_8Cl_6$	309-00-2	0.017	364.91	210.7	30.81	
Atrazine	$C_8H_{14}ClN_5$	1912-24-9	34.7 (26°C)	215.68	169.8	23.19	1.7 (+/0) [55]
Benzene	C_6H_6	71-43-2	1790	78.11	89.4	10.40	
Benzo[a]pyrene	$C_{20}H_{12}$	50-32-8	0.00162	252.31	196.0	35.80	
Bromobenzene*	C_6H_5Br	108-86-1	446 (30°C)	157.01	105.6	13.45	
Carbofuran	$C_{12}H_{15}NO_3$	1563-66-2	320	221.25	194.4	23.65	
Carbon tetrachloride	CCl_4	56-23-5	793	153.82	90.6	10.32	
Chlordane	$C_{10}H_6Cl_8$	57-74-9	0.056	409.78	226.4	31.82	
Chlorobenzene	C_6H_5Cl	108-90-7	498	112.56	101.3	12.34	
2,4-D	$C_8H_6Cl_2O_3$	94-75-7	677	221.04	148.4	19.39	2.73 (0/-) [55]
Dalapon	$C_3H_4Cl_2O_2$	75-99-0	502000	142.97	95.4	10.77	1.74-1.84 (0/-) [55]
DCPA mono-acid degradate*	$C_9H_4Cl_4O_4$	887-54-7	18.3&	317.94	187.6	25.58	acid
DCPA di-acid degradate (chlorthal)*	$C_8H_2Cl_4O_4$	2136-79-0	175&	303.91	162.2	23.66	acid
DDE*	$C_{14}H_8Cl_4$	72-55-9	0.04	318.02	226.8	31.66	
Diazinon*	$C_{12}H_{21}N_2O_3PS$	333-41-5	40	304.35	260.4	31.57	
1,2-Dibromo-3-Chloropropane (DBCP)	$C_3H_5Br_2Cl$	96-12-8	1230 (20°C)	236.33	116.1	14.35	
o-Dichlorobenzene	$C_6H_4Cl_2$	95-50-1	156	147.00	113.3	14.28	
p-Dichlorobenzene	$C_6H_4Cl_2$	106-46-7	81.3	147.00	113.3	14.28	
1,1-Dichlorethane*	$C_2H_4Cl_2$	75-34-3	5040	98.96	84.6	8.31	
1,2-Dichlorethane	$C_2H_4Cl_2$	107-06-2	8600	98.96	84.3	8.33	
1,1-Dichloroethylene	$C_2H_2Cl_2$	75-35-4	2420	96.94	79.2	8.15	
cis-1,2-Dichloroethylene	$C_2H_2Cl_2$	156-59-2	6410	96.94	77.9	8.35	
trans-1,2-Dichloroethylene	$C_2H_2Cl_2$	156-60-5	4520	96.94	77.9	8.35	
Dichloromethane	CH_2Cl_2	75-09-2	13000	84.93	67.8	6.49	
2,4-Dichlorophenol*	$C_6H_4Cl_2O$	120-83-2	4500 (20°C)	163.00	111.7	15.03	7.85 (0/-) [11]
1,2-Dichloropropane	$C_3H_6Cl_2$	78-87-5	2800	112.99	101.2	10.15	
1,3-Dichloropropane*	$C_3H_6Cl_2$	142-28-9	2750	112.99	100.8	10.16	
2,2-Dichloropropane*	$C_3H_6Cl_2$	594-20-7	344&	112.99	100.8	10.15	
1,1-Dichloropropene*	$C_3H_4Cl_2$	563-58-6	749&	110.97	94.2	10.12	
1,3-Dichloropropene*	$C_3H_4Cl_2$	542-75-6	2800 (20°C)	110.97	94.4	10.18	
Dieldrin*	$C_{12}H_8Cl_6O$	60-57-1	0.195	380.91	205.9	30.71	

Table 1
Characteristics of organic pollutants currently regulated by the US EPA or on the contaminant candidate list of the US EPA

Chemical name	Formula	CAS #	C_s (mg/L)	MW (g/mol)	V_m (cm^3/mol)	$\alpha * 10^{24}$ (cm^3)	pK_a #
Di(2-ethylhexyl) adipate	$C_{22}H_{42}O_4$	103-23-1	0.78 (22°C)	370.57	399.4	42.59	
Di(2-ethylhexyl) phthalate	$C_{24}H_{38}O_4$	117-81-7	0.27	390.56	397.0	45.42	
2,4-Dinitrophenol*	$C_6H_4N_2O_5$	51-28-5	2790 (20°C)	184.11	111.5	16.34	4.01 (0/-) [11]
2,4-Dinitrotoluene*	$C_7H_6N_2O_4$	121-14-2	200	182.13	129.3	17.50	
2,6-Dinitrotoluene*	$C_7H_6N_2O_4$	606-20-2	352[&]	182.13	129.3	17.50	
Dinoseb	$C_{10}H_{12}N_2O_5$	88-85-7	52	240.21	178.1	23.80	4.62 (0/-) [56]
Dioxin (2,3,7,8-TCDD)	$C_{12}H_4Cl_4O_2$	1746-01-6	0.0002	321.98	195.9	28.53	
1,2-Diphenylhydrazine*	$C_{12}H_{12}N_2$	122-66-7	221	184.24	156.2	23.94	
Diquat	$C_{12}H_{12}Br_2N_2$	85-00-7	708000 (20°C)	344.05	N/C	N/C	++ (permanent)
Disulfoton*	$C_8H_{19}O_2PS_3$	298-04-4	16.3 (20°C)	274.41	233.2	28.83	
Diuron*	$C_9H_{10}Cl_2N_2O$	330-54-1	42	233.09	170.1	23.25	
Endothall	$C_8H_{10}O_5$	145-73-3	100000 (20°C)	186.16	120.7	15.67	3.4 (0/-), 6.7 (-/--) [55]
Endrin	$C_{12}H_8Cl_6O$	72-20-8	0.25	380.91	205.9	30.71	
EPTC (s-ethyl-dipropylthiocarbamate)*	$C_9H_{19}NOS$	759-94-4	375	189.32	195.4	21.99	
Ethylbenzene	C_8H_{10}	100-41-4	169	106.17	122.2	14.19	
Ethylene dibromide	$C_2H_4Br_2$	106-93-4	3910	187.86	87.9	10.61	
Fonofos*	$C_{10}H_{15}OPS_2$	944-22-9	15.7 (20°C)	246.33	210.3	27.26	
Glyphosate	$C_3H_8NO_5P$	1071-83-6	12000	169.07	100.6	12.30	<2 (+/+-), 2.6 (+-/+--) 5.6 (+--/+---) 10.6 (+---/---) [57]
Haloacetic acids (HAA5)							
Monochloroacetic acid	$C_2H_3ClO_2$	79-11-8	858000	94.50	67.5	7.02	2.87 (0/-) [58]
Dichloroacetic acid	$C_2H_2Cl_2O_2$	79-43-6	10^6 (20°C)	128.94	79.3	8.93	1.26 (0/-) [58]
Trichloroacetic acid	$C_2HCl_3O_2$	76-03-9	44000	163.39	90.3	10.85	0.52 (0/-) [58]
Monobromoacetic acid	$C_2H_3BrO_2$	79-08-3	93790 [&]	138.95	69.3	8.17	2.90 (0/-) [58]
Dibromoaceric acid	$C_2H_2Br_2O_2$	631-64-1	22100 [&]	217.84	82.9	11.21	1.39 (0/-) [58]
Heptachlor	$C_{10}H_5Cl_7$	76-44-8	0.18	373.32	207.7	29.88	
Heptachlor epoxide	$C_{10}H_5Cl_7O$	1024-57-3	0.2	389.32	203.0	29.78	
Hexachlorobenzene	C_6Cl_6	118-74-1	0.0062	284.78	161.1	22.05	
Hexachlorobutadiene*	C_4Cl_6	87-68-3	3.2	260.76	149.3	19.48	
Hexachlorocyclopentadiene	C_5Cl_6	77-47-4	1.8	272.77	148.8	20.24	
p-Isopropyltoluene (p-cymene)*	$C_{10}H_{14}$	99-87-6	23.4	134.22	155.7	17.94	
Lindane	$C_6H_6Cl_6$	58-89-9	7.3	290.83	182.5	22.45	

Table 1
Characteristics of organic pollutants currently regulated by the US EPA or on the contaminant candidate list of the US EPA

Chemical name	Formula	CAS #	C_s (mg/L)	MW (g/mol)	V_m (cm^3/mol)	$\alpha * 10^{24}$ (cm^3)	pK_a #
Linuron*	$C_9H_{10}Cl_2N_2O_2$	330-55-2	75	249.09	176.5	23.95	
Methoxychlor	$C_{16}H_{15}Cl_3O_2$	72-43-5	0.1	345.65	268.2	34.91	
Methyl bromide*	CH_3Br	74-83-9	15200	94.94	58.2	5.71	
2-Methylphenol (o-cresol)*	C_7H_8O	95-48-7	25900	108.14	104.1	13.06	10.28 (0/-) [11]
Methyl-t-butyl ether (MTBE)*	$C_5H_{12}O$	1634-04-4	51000	88.15	117.4	10.67	
Metolachlor*	$C_{15}H_{22}ClNO_2$	51218-45-2	530 (20°C)	283.79	257.7	31.73	
Metribuzin*	$C_8H_{14}N_4OS$	21087-64-9	1050 (20°C)	214.29	163.5	22.71	1.0 (+/0) [59]
Molinate*	$C_9H_{17}NOS$	2212-67-1	970	187.30	176.9	21.03	
Naphthalene*	$C_{10}H_8$	91-20-3	31	128.17	123.5	17.48	
Nitrobenzene*	$C_6H_5NO_2$	98-95-3	2090	123.11	101.2	13.00	
Organotins*							
Bis(tri-n-butyltin) oxide	$C_{24}H_{54}OSn_2$	56-35-9	19.5	596.12	N/C	N/C	
Oxamyl (Vydate)	$C_7H_{13}N_3O_3S$	23135-22-0	280000	219.26	177.6	21.81	
Pentachlorophenol	C_6HCl_5O	87-86-5	14	266.34	147.6	20.85	4.83 (0/-) [11]
Picloram	$C_6H_3Cl_3N_2O_2$	1918-02-1	430	241.46	133.3	19.89	2.3 (0/-) [55]
Polychlorinated biphenyls (PCBs)§							
2,2',5-Trichlorobiphenyl	$C_{12}H_7Cl_3$	37680-65-2	0.4	257.55	190.5	25.97	
Prometon*	$C_{10}H_{19}N_5O$	1610-18-0	750	225.29	198.7	25.71	4.3 (+/0) [55]
RDX*	$C_3H_6N_6O_6$	121-82-4	59.7	222.12	117.0	17.30	
Simazine	$C_7H_{12}ClN_5$	122-34-9	6.2 (22°C)	201.66	152.9	21.37	1.62 (+/0) [55]
Styrene	C_8H_8	100-42-5	310	104.15	115.3	14.73	
Terbacil*	$C_9H_{13}ClN_2O_2$	5902-51-2	710	216.66	172.4	21.29	9.5 (0/-) [55]
Terbufos*	$C_9H_{21}O_2PS_3$	13071-79-9	5.07	288.43	249.8	30.65	
1,1,2,2-Tetrachloroethane*	$C_2H_2Cl_4$	79-34-5	2830	167.85	107.8	12.14	
Tetrachloroethylene	C_2Cl_4	127-18-4	206	165.83	100.3	12.07	
Triazines*							
Atrazine-desethyl*	$C_6H_{10}ClN_5$	6190-65-4	3200 (22°C)	187.63	136.1	19.22	
Cyanazine*	$C_9H_{13}ClN_6$	21725-46-2	170	240.69	179.4	24.99	0.63 (+/0) [55]
Trihalomethanes							
Chloroform	$CHCl_3$	67-66-3	7950	119.38	79.5	8.39	
Bromodichloromethane	$CHBrCl_2$	75-27-4	3030 (30°C)	163.83	81.3	9.54	
Dibromochloromethane	$CHBr_2Cl$	124-48-1	2700 (20°C)	208.28	83.1	10.68	
Bromoform	$CHBr_3$	75-25-2	3100	252.73	84.9	11.82	

Table 1
Characteristics of organic pollutants currently regulated by the US EPA or on the contaminant candidate list of the US EPA

Chemical name	Formula	CAS #	C_s (mg/L)	MW (g/mol)	V_m (cm^3/mol)	$\alpha * 10^{24}$ (cm^3)	pK_a #
Toluene	C_7H_8	108-88-3	526	92.14	105.7	12.32	
Toxaphene+		8001-35-2					
Octachlorobornane	$C_{10}H_{10}Cl_8$		0.55 (20°C)	413.81	253.6	32.82	
2,4,5-TP (Silvex)	$C_9H_7Cl_3O_3$	93-72-1	71	269.51	177.3	23.15	2.84 (0/-) [60]
1,2,4-Trichlorobenzene	$C_6H_3Cl_3$	120-82-1	49	181.45	125.2	16.22	
1,1,1-Trichlorethane	$C_2H_3Cl_3$	71-55-6	1290	133.40	95.7	10.23	
1,1,2-Trichloroethane	$C_2H_3Cl_3$	79-00-5	4590	133.40	96.0	10.23	
Trichloroethylene	C_2HCl_3	79-01-6	1280	131.39	89.1	10.21	
2,4,6-Trichlorophenol*	$C_6H_3Cl_3O$	88-06-2	800	197.45	123.7	16.97	6.19 (0/-) [11]
1,2,4-Trimethylbenzene*	C_9H_{12}	95-63-6	57	120.19	138.2	16.14	
Vinyl chloride	C_2H_3Cl	75-01-4	8800	62.50	68.0	6.29	
Xylenes							
o-Xylene	C_8H_{10}	95-47-6	178	106.17	121.9	14.23	
p-Xylene	C_8H_{10}	106-42-3	162	106.17	121.9	14.23	
m-Xylene	C_8H_{10}	108-38-3	161	106.17	121.9	14.23	

C_s experimentally determined aqueous solubility at 25°C (unless otherwise noted); from EPI Suite v. 3.12 database [61]
V_m molar volume at 20°C; estimated with ChemSketch v. 5.0 [62]
α polarizability; estimated with ChemSketch v. 5.0 [62]
a blank field indicates that a pollutant is either not ionizable at aqueous solution pH values (0-14) or that the pK_a value of the pollutant is not known; information in parentheses represents the charges of the protonated/deprotonated species for a given pK_a; references from which the pK_a information was obtained are shown in square brackets
* currently on the contaminant candidate list (CCL) of the US EPA
@ the CCL also lists alachlor ethanesulfonic acid (alachlor ESA) as well as other acetanilide pesticide degradation products
& aqueous solubility at 25°C estimated from octanol-water partition coefficient using EPI Suite v. 3.12 [61]
N/C could not be calculated
§ mixture of polychlorinated biphenyl congeners
+ mixture of polychlorinated bicyclic terpenes

5.2. Adsorption of nonionic industrial pollutants

The adsorption of nonionic industrial pollutants on activated carbon surfaces is primarily controlled by non-specific dispersive interactions. Whether or not specific adsorbate-adsorbent interactions are important depends in part on the characteristics of the adsorbate; i.e., whether it is a H-bond donor or a H-bond acceptor or whether it is an aromatic compound with electron-donating or electron-withdrawing substituents. In general, mechanisms controlling the adsorption of aliphatic pollutants may be more easily explained than those controlling the adsorption of aromatic pollutants.

5.2.1. Aliphatic industrial pollutants

Relatively few studies have evaluated the effects of activated carbon surface chemistry on the adsorption of aliphatic, rather than aromatic, molecules from aqueous solution [18-20]. Using 15 characterized activated carbons, Li et al. [19] studied the effects of surface chemistry on the adsorption of (1) trichloroethene (TCE), which is relatively hydrophobic and does not form H-bonds, and (2) methyl tertiary-butyl ether (MTBE), which is relatively hydrophilic and contains an ether oxygen atom that can serve as a H-bond acceptor. Because TCE cannot form H-bonds, Li et al. [19] hypothesized that TCE interacts primarily with activated carbon basal planes via van der Waals interactions that are not strongly affected by changes in surface chemistry. To test this, TCE adsorption experiments were conducted in the non-polar solvent cyclohexane. For four activated carbon fibers (ACFs) with similar pore structures and different surface chemistries, TCE adsorption capacities were similar (Fig. 3A), a result that agrees with the above hypothesis. In the presence of water, however, Fig. 3B illustrates that the TCE adsorption capacity decreased with increasing adsorbent hydrophilicity, which was expressed as the sum of the oxygen and nitrogen contents. The latter result was interpreted to be a consequence of enhanced water adsorption on oxygen- and nitrogen-containing surface groups. As suggested by previous researchers, water clusters may prevent pollutant access to hydrophobic regions on the activated carbon surface, reduce the interaction energy between the pollutant and the adsorbent surface, and/or effectively block pollutant access to micropores [16, 17, 46, 67-69].

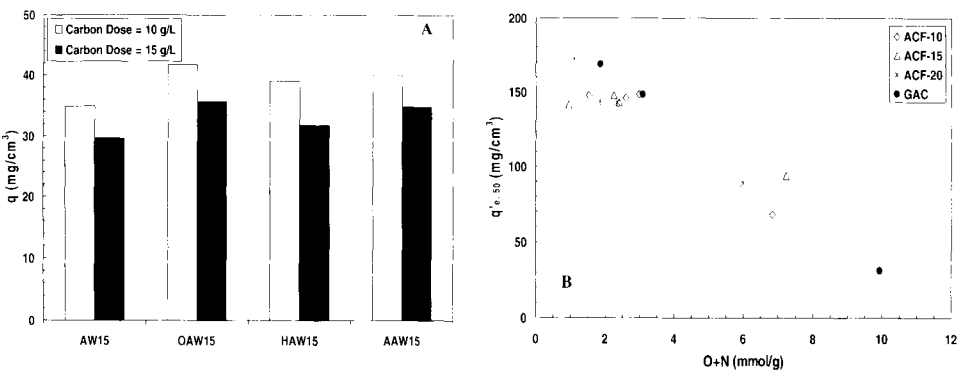


Fig. 3. Effect of adsorbent hydrophobicity on TCE adsorption from (A) cyclohexane and (B) water. TCE adsorption capacities were normalized by the pore volume in the size range of 7 to 10 Å, and aqueous TCE adsorption capacities were determined at an equilibrium liquid phase concentration of 50 μg L^{-1}. AW – acid-washed, OAW – H_2O_2 oxidized, HAW – heat-treated in hydrogen, AAW – heat-treated in ammonia (from Li et al. [19])

For the hydrogen-bond acceptor MTBE, it was expected that MTBE would interact specifically with oxygen-containing groups on activated carbon surfaces that can serve as hydrogen bond donors. As shown in Figure 4A, MTBE adsorption experiments in cyclohexane substantiated that enhanced MTBE adsorption occurred on an oxidized, acidic carbon (OAW15). Therefore, preferential adsorption of MTBE on oxygen-containing functional groups such as carboxylic acid and phenolic hydroxyl groups must have occurred because hydrogen bonds can form between the hydrogen atoms of these surface groups and the ether-oxygen of MTBE. Nonetheless, in aqueous solution, MTBE adsorption capacities decreased with increasing oxygen- and nitrogen-group contents (Fig. 4B). In the presence of water, hydrogen bonds preferentially form between water and these surface groups. Consequently, the availability of specific adsorption sites; i.e., hydrogen-bond donor sites, for MTBE decreased greatly in aqueous solution. However, the loss of such sites to water adsorption alone cannot explain why the MTBE adsorption capacity of ACFs with higher oxygen- and nitrogen-group contents in aqueous solution is lower than that of more hydrophobic ACFs with a smaller surface concentration of hydrogen-bond donor sites. It is plausible that hydrogen-bond donor sites become essentially unavailable to MTBE in aqueous solution and that MTBE adsorption therefore shifted from a site-specific hydrogen-bonding mechanism in cyclohexane to a non-specific dispersive mechanism in aqueous solution, where MTBE adsorbs instead on the graphitic basal planes of the adsorbent. Hence, the lower MTBE adsorption capacity of oxidized adsorbents with a larger concentration of surface-oxygen and nitrogen groups most likely resulted from the increased formation of water clusters that reduce access to the graphitic basal planes, reduce the interaction energy between MTBE and the adsorbent surface, and/or block pore entrances [16, 17, 46, 67-69].

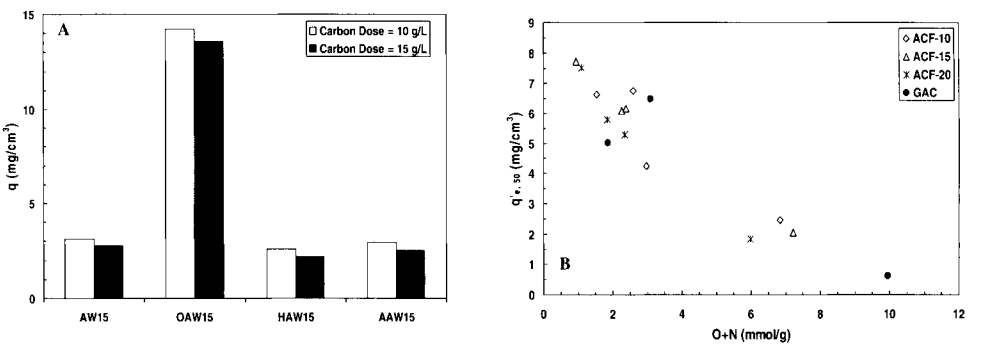


Fig. 4. Effect of adsorbent hydrophobicity on MTBE adsorption from (A) cyclohexane and (B) water. MTBE adsorption capacities were normalized by the pore volume in the size range of 8 to 11 Å, and aqueous MTBE adsorption capacities were determined at an equilibrium liquid phase concentration of 50 μg L^{-1}. AW – acid-washed, OAW – H2O2 oxidized, HAW – heat-treated in hydrogen, AAW – heat-treated in ammonia (from Li et al. [19])

5.2.2. Aromatic industrial pollutants

With respect to interactions between aromatic molecules and activated carbon pore surfaces, three important phenomena are discussed in the literature: (1) π-π dispersion interactions that are enhanced as density of π-electrons associated with the basal plane of the adsorbent and with the aromatic ring of the adsorbate increases [10, 46], (2) hydrogen-bonding interactions between the adsorbate and surface functional groups [68, 70], and (3) interactions due to the

formation of electron donor-acceptor complexes between the activated carbons surface (e.g. carbonyl groups [4] or unsaturated carbons at basal plane edges [13]) and the ring of an aromatic adsorbate. The importance of one or more of these mechanisms on the adsorption of aromatic molecules from aqueous solutions needs to be contrasted, however, with the importance of water adsorption on the activated carbon surface. The phenol adsorption data depicted in Fig. 5 nicely illustrate the implications of this statement.

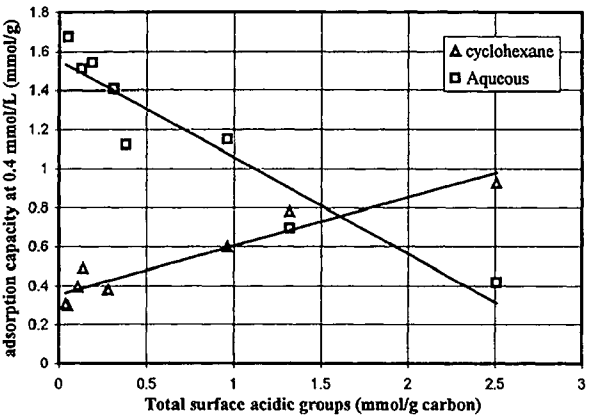


Fig. 5. Effect of activated carbon surface acidity on the adsorption of phenol from cyclohexane and from pH neutral water (Copyright (2004) from Arafat et al. [70]. Reproduced by permission of Taylor & Francis Group, LLC., http://www.taylorandfrancis.com)

In the absence of water, the phenol adsorption capacity increased with increasing surface acidity (or surface oxygen content), and this result is consistent with the formation of hydrogen bonds between the OH groups of phenol and oxygen-containing functionalities on the activated carbon surface [70]. In the presence of water, however, the phenol adsorption capacity of activated carbons decreased with increasing surface acidity (Figure 5) because of preferential water adsorption on oxygen-containing surface groups. This result as well as that of Li et al. [19] for MTBE shows that enhanced water adsorption on hydrophilic adsorbents negates the benefits of providing specific adsorption sites capable of forming hydrogen-bonds with aliphatic and aromatic adsorbates.

With respect to the π-π dispersion argument, adsorption data for nitrobenzene, which does not form hydrogen bonds, are illustrative [17, 68, 71]. In aqueous solution, the adsorption capacity of activated carbons for nitrobenzene decreased with increasing surface oxygen content [17, 68, 71]. This trend is consistent with the argument that a lower density of delocalized π-electrons on the basal planes of activated carbons with higher oxygen contents negatively affected π-π dispersion interactions. However, this trend is also consistent with the enhanced adsorption of water on activated carbons with higher oxygen contents. Thus, it is unclear to what extent enhanced water adsorption and to what extent a decrease in delocalized π-electron density lowered the adsorption capacity of oxidized activated carbons for nitrobenzene. To complicate matters, Franz et al. [68] implied that a higher π-electron density on the basal planes enhances the adsorption of nitrobenzene via the formation of π-π donor-acceptor complexes [72] between the electron-rich activated carbon basal plane and the

electron-poor aromatic ring of nitrobenzene, an alternative explanation that is consistent with the result of a higher nitrobenzene adsorption capacity on non-oxidized activated carbons.

While the results of immersion calorimetry suggest that the mechanism of benzene adsorption is independent of activated carbon surface chemistry [73, 74], adsorption isotherms for benzene [70, 75], nitrobenzene [68], and xylene [75] from non-polar solvents do show that lower adsorption capacities are obtained with increasing activated carbon acidity (or surface oxygen content). Pinto and co-workers [70, 75] attributed this trend to a lower affinity of hydrophobic aromatic adsorbates for polar surfaces. While, these results are also consistent with the π-π dispersion argument, they appear to be *in*consistent with the π-π donor-acceptor argument. For the latter mechanism, one would expect that the interaction of electron-rich aromatic adsorbates such as xylene would be enhanced as electron-withdrawing functionalities such as carboxylic acid groups are incorporated into the activated carbon surface, a trend that is opposite to what was observed [75].

With respect to the formation of electron donor-acceptor complexes, many recent results do not appear to support the argument that such charge-transfer complexes form between carbonyl groups on the activated carbon surface and the aromatic ring of an adsorbate [see 10, 15]. However, very recent evidence suggests that electron donor-acceptor complexes may form between unsaturated carbon atoms at the basal plane edges and the aromatic ring of an adsorbate [13] or between electron-rich and electron-poor aromatic ring structures [76].

5.3. Adsorption of ionic industrial pollutants

In addition to dispersive interactions, Coulombic interactions control the adsorption of ionic industrial pollutants on charged activated carbon pore surfaces. If the pore surface and the pollutant are oppositely charged, Coulombic interactions are attractive and enhance adsorption. However, if the pore surface and the pollutant are similarly charged, Coulombic interactions are repulsive, and adsorption of the pollutant can only occur if the attractive dispersive interaction at a given adsorption site is greater than the repulsive Coulombic interaction [5, 9]. Müller et al. [5, 9] formulated a model that takes into account (1) the distribution of adsorption potentials describing the dispersive interactions between the molecular or ionic pollutant and the heterogenerous activated carbon surface and (2) Coulombic interactions between the ionic pollutant and the charged activated carbon surface. The latter interactions depend on solution pH, which controls the charge and extent of (de)protonation of the pollutant as well as the net charge of activated carbon pore surfaces (Fig. 6A). In addition, an increase in solution ionic strength shields Coulombic interactions and thus negatively affects the adsorption of counter-ions while it positively affects the adsorption of co-ions in dilute solutions [14, 15].

Applying the model of Müller et al. [5], Radovic et al. [71] simulated the adsorption of aniline on three model activated carbons with different surface chemistries, and the simulation results nicely illustrate the implications of Coulombic interactions on the adsorption of ionizable contaminants. In their example, Radovic et al. [71] considered an acidic carbon (C1), an amphoteric carbon (C2), and a basic carbon (C3), where C1 represents the product following oxidative treatment of C2, and C3 represents the product following heat treatment of C1. Fig. 6A depicts the effect of solution pH on the net surface charge of carbons C1 (negative above pH 3), C2 (positive below pH 6.5, negative above pH 6.5, and C3 (positive below pH 10). For carbon C2, Figure 6B illustrates that the adsorption of aniline (pK_a = 4.6) begins to drop rapidly as the solution pH becomes more acidic. This effect can be explained by repulsive Coulombic interactions between anilinium cations and the net positive surface charge of C2; with decreasing pH, both the magnitude of the positive surface charge of C2

and the concentration of anilinium cations increase. In contrast, the acidic carbon C1, which exhibits a net negative surface charge above pH 3, exhibits an adsorption maximum for aniline/anilinium at a solution pH of 4.6, the pK_a of the anilinium cation. This result can be explained by attractive Coulombic interactions between anilinium cations and the net negative surface charge of C1. At solution pHs below pH 3, repulsive Coulombic interactions begin to dominate, and the adsorption capacity of C1 for anilinium cations decreases rapidly.

Fig. 6B also illustrates the effects of acidic surface functional groups on dispersion interactions between aniline molecules and activated carbon basal planes. These interactions may become weaker as a result of a decreased density of delocalized π-electrons on the basal planes of the oxidized activated carbon C1. Reducing the strength of dispersion interactions in the model by factors of two (D/2), five (D/5), and ten (D/10), Figure 6B illustrates that the aniline adsorption capacity decreases greatly at higher pH values, at which adsorption of the neutral aniline molecule is not affected by Coulombic interactions.

Fig. 6C depicts differences in anilinium/aniline adsorption that may be expected as the acidic activated carbon C1 is converted to a basic carbon (C3) by heat treatment. For the adsorption of anilinium cations, C3 exhibits a lower adsorption capacity than C1 because of repulsive Coulombic interactions between anilinium cations and the positively charged surface of C3. For the heat-treated carbon C3, Fig. 6C also illustrates the effect of enhanced dispersive interactions, which may result from a higher density of delocalized π-electrons on the basal planes of C3, on the adsorption of aniline molecules; apart from the baseline case (C3E), the strength of dispersion interactions was increased in the model by factors of two (Dx2), five (Dx5), and ten (Dx10).

While Fig. 6 nicely illustrates how changes in Coulombic and dispersive interactions affect the adsorption capacity of an ionizable pollutant, the modeling results do not directly address the effects of solvent-adsorbent interactions. The affinity of water for the activated carbon surface is expected to increase with increasing acidity or surface oxygen content; i.e., from C3 to C1. Thus, enhanced water adsorption will likely have a greater negative effect on the aniline/anilinium adsorption capacity of C1 than C3. However, by adjusting the strength of dispersive interactions, Radovic et al. [71] indirectly simulated the importance of enhanced water adsorption on acidic functional groups even though changes in the strength of dispersive interactions were discussed in light of changes in the density of delocalized π-electrons on activated carbon basal planes.

5.4. Effects of natural organic matter on the adsorption of industrial pollutants

In drinking water treatment, natural organic matter (NOM) adsorption can (1) precede the adsorption of micropollutants ("NOM preloading") [e.g., 77-81] as is typical for packed bed adsorption processes involving granular activated carbon (GAC), or (2) occur concurrently with the adsorption of micropollutants [e.g., 27, 82-85] as is typical in adsorption processes involving powdered activated carbon (PAC). In either case, NOM adsorption lowers the micropollutant adsorption capacity of activated carbon relative to that measured in the absence of NOM.

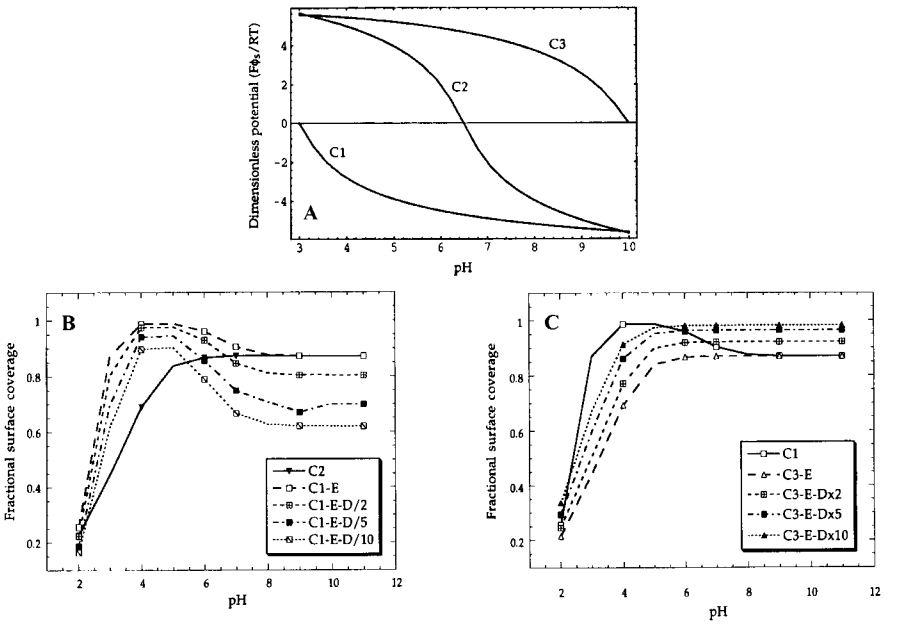


Fig. 6. Effect of solution pH on the surface charge of three activated carbon samples (A), and effects of solution pH and activated carbon surface charge/chemistry on aniline adsorption (B – effect of oxidative treatment of C2) and (C – effect of heat treatment of C1). C2 represents as-received carbon, C1 represents product following oxidative treatment of C2, C3 represents product following heat-treatment of C1 (from Radovic et al. [71]).

Two studies have systematically evaluated adsorbent surface chemistry effects on the adsorption of micropollutants from waters containing NOM. Kilduff et al. [86] evaluated the effects of adsorbent surface chemistry on TCE adsorption capacity for activated carbons that were preloaded with NOM. The percent decrease in TCE adsorption capacity resulting from NOM preloading became smaller with increasing surface acidity of coal-based activated carbons. However, while the effect of NOM preloading became smaller, the single-solute TCE adsorption capacity also decreased with increasing surface acidity, and the latter effect dominated over the former. Therefore, the net result was that the most hydrophobic adsorbent, i.e. the activated carbon with the lowest surface acidity, exhibited the largest TCE adsorption capacity following NOM preloading [86]. In contrast to the results obtained for coal-based activated carbon, decreasing the surface acidity of wood-based activated carbons did not increase the effect of NOM preloading on the percent reduction in TCE adsorption capacity [86]. The adsorbent with the lowest surface acidity among the wood-based activated carbons was again the most effective for TCE adsorption both in the presence and absence of preloaded NOM [86].

Quinlivan et al. [87] and Knappe et al. [88] studied the effects of adsorbent pore structure and surface chemistry on the adsorption of micropollutants in the presence of co-adsorbing NOM. The adsorption of both TCE and MTBE was evaluated. As expected, co-adsorbing NOM constituents competed with micropollutants for adsorption sites, and micropollutant adsorption capacities in natural water were lower than in ultrapure water as a

result. However, the percent reduction in TCE or MTBE adsorption capacity resulting from the presence of competing NOM components was not strongly affected by the surface chemistry of 15 adsorbents, whose oxygen contents ranged from about 1.4 to 16 % (w/w).

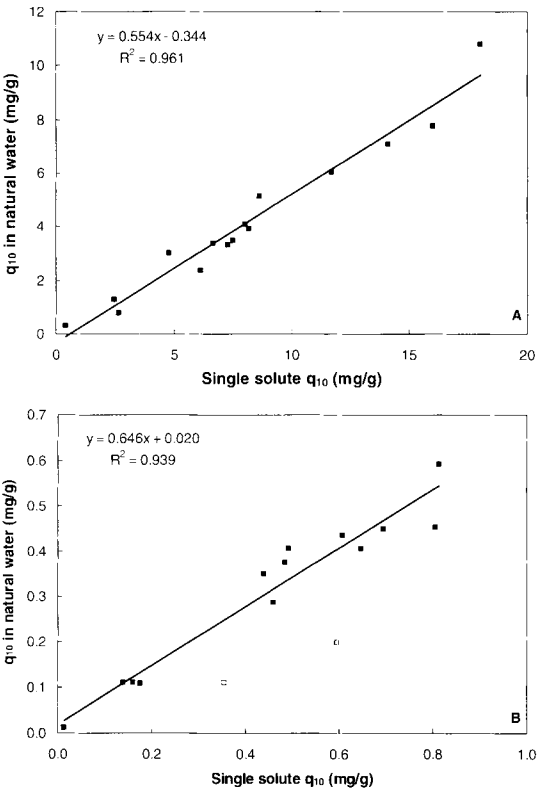


Fig. 7. Correlations between TCE (A) and MTBE (B) adsorption capacities in natural water and ultrapure water. Adsorption capacities are compared at an equilibrium liquid phase concentration of 10 μg L^{-1}, and the initial micropollutant concentration in the natural water experiments was 100 μg L^{-1}. The two open symbols in panel B were not included in the correlation; the lower adsorption capacities obtained in the presence of NOM with these two adsorbents was interpreted to be a result of pore blockage effects caused by adsorbed NOM (see refs. [87] and [88] for additional details)

As shown in Fig. 7, the average TCE and MTBE adsorption capacities in natural water were about 55% and 65%, respectively, of those in ultrapure water. Consequently, the relationships between pollutant adsorption capacity in natural water and adsorbent hydrophilicity were very similar to those shown in Figs. 3B and 4B for TCE and MTBE, respectively [87, 88]. These results as well as those of Kilduff et al. [86] suggest that enhanced water adsorption on hydrophilic carbons detrimentally affects the adsorption of aqueous organic contaminants both in the presence and absence of NOM.

6. CONCLUSIONS

Both the physical and chemical characteristics of activated carbons determine their adsorption capacity for organic micropollutants from aqueous solution. In terms of pore structure, activated carbons exhibit the highest adsorption capacity when the volume of pores with widths approximating the size of the targeted pollutant is maximized. With respect to surface chemistry, hydrophobic adsorbents exhibit larger adsorption capacities for organic micropollutants than hydrophilic adsorbents with similar physical characteristics, and this observation is principally attributable to enhanced water adsorption on hydrophilic activated carbons. Hydrogen bonding between water and oxygen-containing functional groups on the activated carbon surface is sufficiently strong that it negates the benefits of providing specific adsorption sites capable of forming hydrogen bonds with aliphatic or aromatic adsorbates. Furthermore, the negative effects of enhanced water adsorption were dominant in two studies that evaluated the effects of preloaded and co-adsorbing NOM on the adsorption capacity of organic micropollutants.

For the removal of ionic pollutants, Coulombic interactions are important in addition to dispersive interactions. For anionic pollutants, enhanced adsorption would therefore be expected on basic activated carbons that carry a net positive surface at the pH of the water being treated. Similarly, attractive Coulombic interactions between cationic pollutants and negatively charged surfaces of acidic activated carbons can be expected.

Overall, this overview suggests that the selection of an activated carbon with a large adsorption capacity for a targeted organic micropollutant requires knowledge about the characteristics of the adsorbate (molecular size, charge dependence on solution pH) as well as the adsorbent (pore size distribution, oxygen content, surface charge dependence on solution pH). From a standpoint of activated carbon characterization techniques, current characterization tools such as iodine number or BET surface area do not provide the information required to select effective carbons for the removal of micropollutants from drinking water sources. Instead, activated carbons should be characterized in terms of micropore size distribution, elemental composition (C, H, N, O, ash), and surface charge dependence on solution pH.

REFERENCES

[1] F.W. Pontius, J. - Am. Water Works Assoc., 95 (2003) 57.
[2] M.N. Baker, The Quest for Pure Water, 2nd ed., American Water Works Association, Denver, CO, 1981, p. 454.
[3] D.M. McKnight and G.R. Aiken, Sources and Age of Aquatic Humus, D.O. Hessen and L.J. Tranvik, (eds.), Aquatic Humic Substances, Springer-Verlag, Berlin, Germany, 1998, pp. 9-39.
[4] J.S. Mattson and H.B. Mark, Jr., Activated Carbon - Surface Chemistry and Adsorption from Solution. Marcel Dekker, New York, 1971.
[5] G. Müller, C.J. Radke and J.M. Prausnitz, J. Colloid Interface Sci., 103 (1985) 466.
[6] E. Fernandez, D. Hugi-Cleary, M.V. López-Ramón and F. Stoeckli, Langmuir 19 (2003) 9719.
[7] M. Manes, Activated Carbon Adsorption Fundamentals. In: R.A. Meyers, (ed.), Encyclopedia of Environmental Analysis and Remediation, John Wiley & Sons, New York, NY, 1998, pp. 26-68.
[8] G. Belfort, Environ. Sci. Technol., 13 (1979) 939.
[9] G. Müller, C.J. Radke and J.M. Prausnitz, J. Phys. Chem., 84 (1980) 369.
[10] L.R. Radovic, C. Moreno-Castilla and J. Rivera-Utrilla, Carbon Materials as Adsorbents in Aqueous Solution, L.R. Radovic (ed.), Chemistry and Physics of Carbon, vol. 27, Marcel Dekker, Inc., New York, NY, 2001, pp. 227-416.

[11] R.P. Schwarzenbach, P.M. Gschwend and D.M. Imboden, Environmental Organic Chemistry, 2nd ed., John Wiley & Sons, Hoboken, NJ, 2003.
[12] G. Cornelissen, M. Elmquist, I. Groth and Ö. Gustafsson, Environ. Sci. Tech., 38 (2004) 3574.
[13] E. Castillejos-Lopéz, D.M. Nevskaia, V. Muñoz, I. Rodríguez-Ramos and A. Guerrero-Ruiz, Langmuir, 20 (2004) 1013.
[14] M. Bjelopavlic, G. Newcombe and R. Hayes, J. Colloid Interface Sci., 210 (1999) 271.
[15] C. Moreno-Castilla, Carbon, 42 (2004) 83.
[16] O.P. Mahajan, C. Moreno-Castilla and P.L. Walker, Jr., Sep. Sci. Technol., 15 (1980) 1733.
[17] Y. Kaneko, M. Abe and K. Ogino, Colloids Surf., 37 (1989) 211.
[18] T. Karanfil and J.E. Kilduff, Environ. Sci. Technol., 33 (1999) 3217.
[19] L. Li, P.A. Quinlivan and D.R.U. Knappe, Carbon, 40 (2002) 2085.
[20] T. Karanfil and S.A. Dastgheib, Environ. Sci. Tech., 38 (2004) 5834.
[21] W.J. Weber, Jr. and F.A. DiGiano, Process Dynamics in Environmental Systems, John Wiley & Sons, New York, NY, 1996, pp. 345-356.
[22] S. Kasaoka, Y. Sakata, E. Tanaka and R. Naitoh, Int. Chem. Eng., 29 (1989) 101.
[23] R.C. Bansal, J.B. Donnet and F. Stoeckli, Active Carbon, Marcel Dekker, Inc., New York, NY, 1988.
[24] M. Smisek and S. Cerny, Active Carbon. Manufacture, Properties and Applications, Elsevier Publishing Company, New York, NY, 1970.
[25] C.A. Leon y Leon, J.M. Solar, V. Calemma and L.R. Radovic, Carbon, 30 (1992) 797.
[26] K.S.W. Sing, D.H. Everett, R.A.W. Haul, L. Moscou, R.A. Pierotti, J. Rouquerol and T. Siemieniewska, Pure Appl. Chem., 57 (1985) 603.
[27] G. Newcombe, M. Drikas and R. Hayes, Water Res., 31 (1997) 1065.
[28] M.M. Dubinin, Chem. Rev., 60 (1960) 235.
[29] K.S.W. Sing, J. Porous Mat., 2 (1995) 5.
[30] S. Okouchi, H. Saegusa and O. Nojima, Environ. Int., 18 (1992) 249.
[31] H. Jankowska, A. Swiatkowski, and J. Choma, Active Carbon, Ellis Horwood, New York, NY, 1991.
[32] B.R. Puri, Surface Complexes on Carbons, P.L. Walker Jr., (ed.), Chemistry and Physics of Carbon, Vol. 6, Marcel Dekker, New York, 1970, pp. 191-282.
[33] B.R. Puri, Physicochemical aspects of carbon affecting adsorption from the aqueous phase, M.J. McGuire and I.H. Suffet (eds.), Advances in Chemistry Series, Vol. 202, American Chemical Society, Washington, DC, 1983, pp. 77-93.
[34] H.P. Boehm, Carbon, 32 (1994) 759.
[35] M.A. Montes-Morán, D. Suárez, J.A. Menéndez and E. Fuente, Carbon, 42 (2004) 1219.
[36] C.A. Leon y Leon and L.R. Radovic, Interfacial chemistry and electrochemistry of carbon surfaces, P.A. Thrower (ed.), Chemistry and Physics of Carbon, Vol. 24, Marcel Dekker, New York, 1992, pp. 213-310.
[37] T.J. Bandosz, J. Jagiello, C. Contescu and J.A. Schwarz, Carbon, 31 (1993) 1193.
[38] Y. El-Sayed and T.J. Bandosz, J. Colloid Interface Sci., 273 (2004) 64.
[39] E. Fuente, J.A. Menéndez, D. Suárez and M.A. Montes-Morán, Langmuir 19 (2003) 3505.
[40] V.A. Garten and D.E. Weiss, Austr. J. Chem., 10 (1957) 309.
[41] H.P. Boehm and M. Voll, Carbon, 8 (1970) 227.
[42] A. Contescu, M. Vass, C. Contescu, K. Putyera and J.A. Schwarz, Carbon, 36 (1998) 247.
[43] E. Fuente, J.A. Menéndez, M.A. Diez, D. Suárez, and M.A. Montes-Morán, J. Phys. Chem. B, 107 (2003) 6350.
[44] S.S. Barton, M.J.B. Evans, E. Halliop and J.A.F. MacDonald, Carbon, 35 (1997) 1361.
[45] H. Tamon and M. Okazaki, J. Colloid Interface Sci., 179 (1996) 181.
[46] R.W. Coughlin and F.S. Ezra, Environ. Sci. Technol., 2 (1968) 291.
[47] B. Stöhr, H.P. Boehm and R. Schlögl, Carbon, 29 (1991) 707.
[48] R.J.J. Janssen and H. van Bekkum, Carbon, 33 (1995) 1021.
[49] S. Biniak, G. Szymanski, J. Siedlewski and A. Swiatkowski, Carbon, 35 (1997) 1799.
[50] C.L. Mangun, K.R. Benak, J. Economy and K.L. Foster, Carbon, 39 (2001) 1809.

[51] G. Newcombe, J. Morrison and C. Hepplewhite, Carbon, 40 (2002) 2135.
[52] US EPA, Unregulated Drinking Water Contaminants, http://www.epa.gov/safewater/dw_unregcontaminants.html, accessed December 2004.
[53] D.C. Luehrs, J.P. Hickey, P.E. Nilsen, K.A. Godbole and T.N. Rogers, Environ. Sci. Technol., 30 (1996) 143.
[54] J.C. Crittenden, S. Sanongraj, J.L. Bulloch, D.W. Hand, T.N. Rogers, T.F. Speth and M. Ulmer, Environ. Sci. Technol., 33 (1999) 2926.
[55] C.D.S. Tomlin, (ed.), The Pesticide Manual, 13th ed., British Crop Protection Council, Hampshire, UK, 2003.
[56] C.R. Worthing (ed.), The Pesticide Manual, 7th ed., British Crop Protection Council, Croydon, UK, 1987.
[57] P. Sprankle, W.F. Meggitt and D. Penner, Weed Sci., 23 (1975) 229.
[58] J.A. Dean, Lange's Handbook of Chemistry, 15th ed., McGraw-Hill, New York, NY, 1999.
[59] R.D. Wauchope, T.M. Buttler, A.G. Hornsby, P.W.M. Augustijn-Beckers and J.P. Burt, Rev. Environ. Contam. Toxicol. 123 (1992) 1.
[60] C.R. Worthing (ed.), The Pesticide Manual, 8th ed., British Crop Protection Council, Thornton Heath, UK, 1987.
[61] US EPA, EPI Suite v 3.12, http://www.epa.gov/opptintr/exposure/docs/episuitedl.htm, accessed December 2004.
[62] Advanced Chemistry Development, ACD/ChemSketch 5.0 Freeware, http://www.acdlabs.com/download/chemsk.html, accessed December 2004.
[63] Z. Qiang and C. Adams, Water Res., 38 (2004) 2874.
[64] M.M. Dubinin, E.D. Zaverina and V.V. Serpinsky, J. Am. Chem. Soc., 77 (1955) 1760.
[65] D. Mowla, D.D. Do and K. Kaneko, Adsorption of Water Vapor on Activated Carbon: A brief overview, L.R. Radovic (ed.), Chemistry and Physics of Carbon, vol. 28, Marcel Dekker, Inc., New York, NY, 2003, pp. 229-262.
[66] J.K. Brennan, T.J. Bandosz, K.T. Thomson and K.E. Gubbins, Colloids Surf. A, 187-188 (2001) 539.
[67] P. Pendleton, S.H. Wong, R. Schumann, G. Levay, R. Denoyel and J. Rouquerol, Carbon, 35 (1997) 1141.
[68] M. Franz, H.A. Arafat and N.G. Pinto, Carbon 38 (2000) 1807.
[69] E.A. Müller, F.R. Hung and K.E. Gubbins, Langmuir, 16 (2000) 5418.
[70] H.A. Arafat, F. Ahnert and N.G. Pinto, Sep. Sci. Technol., 39 (2004) 43.
[71] L.R. Radovic, I.F. Silva, J.I. Ume, J.A. Menéndez, C.A. Leon y Leon and A.W. Scaroni, Carbon, 35 (1997) 1339.
[72] C.A. Hunter, K.R. Lawson, J. Perkins and C.J. Urch, J. Chem. Soc., Perkin Trans., 2 (2001) 651.
[73] F. Stoeckli and A. Lavanchy, Carbon, 38 (2000) 475.
[74] M.V. López-Ramón, F. Stoeckli, C. Moreno-Castilla and F. Carrasco-Marín, Langmuir, 16 (2000) 5967.
[75] F. Ahnert, H.A. Arafat and N.G. Pinto, Adsorption, 9 (2003) 311.
[76] D. Zhu, S. Hyun, J.J. Pignatello and L.S. Lee, Environ. Sci. Technol., 38 (2004) 4361.
[77] R.S. Summers, B. Haist, J. Koehler, J. Ritz, G. Zimmer and H. Sontheimer, J. - Am. Water Works Assoc., 81 (1989) 66.
[78] T.F. Speth, J. Environ. Eng., 117 (1991) 66.
[79] M.C. Carter and W.J. Weber Jr., Environ. Sci. Technol., 28 (1994) 614.
[80] J.E. Kilduff, T. Karanfil, and W.J. Weber, Jr., J. - Am. Water Works Assoc., 90 (1998) 76.
[81] D.R.U. Knappe, V.L. Snoeyink, P. Roche, M.J. Prados and M.M. Bourbigot, J. - Am. Water Works Assoc., 91 (1999) 97.
[82] I.N. Najm, V.L. Snoeyink, and Y. Richard, J. - Am. Water Works Assoc., 83 (1991) 57.
[83] T.E.T. Gillogly, V.L. Snoeyink, J.R. Elarde, C.M. Wilson and E.P. Royal, J. - Am. Water Works Assoc., 90 (1998) 98.
[84] D.R.U. Knappe, Y. Matsui, V.L. Snoeyink, P. Roche, M.J. Prados and M.M. Bourbigot, Environ. Sci. Technol., 32 (1998) 1694.

[85] G. Newcombe, J. Morrison, C. Hepplewhite and D.R.U. Knappe, Carbon, 40 (2002) 2147.
[86] J.E. Kilduff, R. Srivastava and T. Karanfil, Stud. Surf. Sci. Catal., 144 (2002) 553.
[87] P.A. Quinlivan, L. Li and D.R.U. Knappe, Water Res., 39 (2005) 1663.
[88] D.R.U. Knappe, L. Li, P.A. Quinlivan and T.B. Wagner, Effects of Activated Carbon Characteristics on Organic Contaminant Removal, American Water Works Association Research Foundation, Denver, CO, 2003.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 10: Manganese removal

B. Chiswell and S.-H. D. Huang

National Research Centre for Environmental Toxicology,
University of Queensland,
39 Kessels Road, Coopers Plains,
Queensland, 4108 Australia

1. INTRODUCTION

A survey of end-user perception of problems associated with their reticulated water identified taste, odour and colour as being their prime concerns [1]. It is thus in this context, somewhat ironic that many taste/odour complaints are often associated with the presence of chlorine, whose introduction to drinking water has been described as the greatest advance in public health in the last 150 years [2], while colour is often associated with the presence of manganese oxides at levels far below those which threaten human health. Thus, although the WHO and USEPA guideline limits for manganese in drinking water are currently 100 and 50 $\mu g\ L^{-1}$ respectively [3, 4], this is set on the basis of aesthetic rather than health considerations Indeed a recent comprehensive review of the toxicity of manganese in drinking water cites little evidence of problems until concentrations reach some hundred times of this value, and most evidence points to manganese toxicity being associated with inhalation of manganese oxide dust rather than with ingestion of drinking water [4].

Nevertheless, the water treatment plant operator who chooses to ignore "dirty" water complaints from the consumers is certain to rapidly be made aware of the fact that black stains in clothing, which are very difficult to remove, bath and swimming pool waters which are black/brown in colour, and drinking water that looks very unpalatable (particularly when added to whisky) are politically unacceptable to authorities providing water. Indeed work [5] done on manganese-containing deposits in reticulation systems over a decade ago indicates that the WHO limit is far too high for "aesthetic" purposes (at least in some situations), and that deposition in the reticulation system occurs down to manganese concentrations of less than 20 $\mu g\ L^{-1}$ in potable water. This work recommended a manganese limit of 10 $\mu g\ L^{-1}$; a recommendation whose force is only now becoming clear. However, the problems of manganese-associated depositions are not confined to the consumers alone; unless manganese input into the reticulation system is controlled, solid manganese-containing encrustations can substantially reduce pipe volume and thus water throughput.

2. ENVIRONMENTAL MANGANESE CHEMISTRY

2.1. Raw water manganese sources

Manganese is the twelfth most abundant metal in the earth's crust and occurs in over 300 different minerals [6]. Its atomic number is 25, which places it next to iron in the periodic table of the elements; a feature of great practical significance in water treatment.

Nearly all iron ores and even purified iron compounds contain manganese as a "contaminant", a fact that cannot be overlooked in plants in which iron flocculation is employed. Manganese is an essential element occurring in a very large number of plants, and a prominent United States Food and Nutrition Board recommends an "adequate" adult intake of approximately 2mg per day [4].

However, not only does manganese appear widely in the earth's crust and in organic matter, it has widespread industrial use, particularly in steel manufacture. Thus, the metal's natural environmental occurrence may be enhanced by mining activities and by decomposition of manufactured manganese-containing products, such as steel, fertilizer, fungicide, livestock feed and even unleaded gasoline.

Thus it follows that virtually all raw waters entering a treatment plant will contain some manganese; the extent of the manganese concentration in the water will depend in particular on the geochemical makeup of the environment from which the water (surface or sub-surface) is drawn, and (usually) to lesser extents on the fauna composition of any catchment, and on the possibility of anthropogenic inputs associated with agricultural processes or waste dumping.

Although there is substantial evidence of neurological damage to humans from inhalation of manganese (di)oxide dust, there is a relevant paucity of data indicating human manganese toxicity from drinking water [7]; for further information the reader should refer to the USEPA report dealing with the health effects of manganese in drinking water [4].

2.2. Relevant oxidation states

Although manganese can be prepared in a wide range of oxidation states in the laboratory, there are only three oxidation states of environmental importance, viz. (II), (III) and (IV); further higher oxidation states, except in rare cases, do not occur in the environment, although Mn(VII), as permanganate, is widely used in removal of manganese from water.

In waters of pH below 6, the favoured form of manganese is the very soluble hydrated $Mn_{(aq)}^{2+}$ cation. Above pH 8.5, this soluble cation is converted to relatively insoluble (hydr)oxides (or depending on carbon dioxide concentration, carbonates), which can be present as Mn(II), (III) or (IV) species, and indeed even mixed oxidation state species (Fig. 1) [8]. However, one cannot assume that insoluble manganese oxide species in low pH waters will dissolve rapidly. The solubility of such oxides is generally low, and thus dissolution is kinetically slow. Thus insoluble Mn(II), (III) and (IV) oxides are often found in large concentrations in dam sediments, even when such sediments are reducing in nature [9].

One further species which bedevils manganese removal is colloidal manganese (di)oxide. Although theoretically speaking an insoluble form of Mn(IV), for natural freshwaters there is no reason to believe that colloidal manganese oxide species are not present across the range of oxidation states (II), (III), and (IV), with mixed oxidation states being very likely; these species will not be removed by filtration processes and may be difficult to flocculate. Colloidal manganese dioxide is a very stable product (the authors have kept stable solutions in the laboratory for years with no sign of decomposition), formed by treating soluble Mn(II) salt solutions with permanganate in basic solution [11]; a reaction which can mirror the permanganate method of manganese removal. As will be shown, attempts to remove this species of manganese, which will be detected analytically as soluble (filterable) manganese, by such an oxidation process may create problems downstream in the water treatment process.

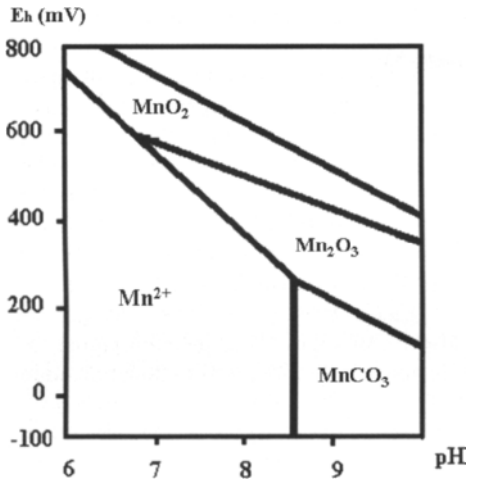


Fig. 1. Eh versus pH diagram of manganese (adapted from Degrémont [10])

The fact that most incoming raw waters to a treatment plant have pH values in the 6-8.5 range means that there is a delicate balance between Eh and pH which determines the manganese species present. Furthermore, the fact that both oxidising/reducing capacity and the pH of the water are adjusted during water treatment and subsequent handling in the reticulation system means that the actual manganese species present can readily change their nature, and thus their solubility, during treatment and distribution.

A further complication in understanding manganese speciational change is the very important role played by microorganisms in facilitating $Mn_{(aq)}^{2+}$ oxidation to insoluble oxide species [12]; indeed, such microorganisms can also bind colloidal manganese dioxide [13]. Thus, although there is little evidence that $Mn_{(aq)}^{2+}$ forms either soluble or insoluble manganese complexes in surface freshwaters [14], manganese deposits in dams/lakes and pipes may be associated with extensive deposits of both organic and inorganic material. However, for the water treatment plant operator using raw water from a dam/lake that undergoes seasonal stratification, the speciational changes taking place in the sediments below the thermocline are those that will yield most of the manganese removal problems. Sedimental manganese (and iron) release has been well-documented over many years [15]. Reducing, anoxic conditions in the sediment below a thermocline yield a situation in which pH falls favouring reduction of Mn(III) and (IV) species, and oxygen is abstracted from manganese oxides for both chemical and microbial processes. The result is the production of soluble $Mn_{(aq)}^{2+}$ cations which migrate through the hypolimnic waters leading to high concentrations of soluble manganese.

2.3. Chemical speciation and analysis

Speciation of manganese in freshwaters has commonly been undertaken by filtration through a 0.45μm (or 0.22μm) membrane; filterable manganese (often termed "soluble") is thus operationally separated from particulate ("insoluble") manganese. [16] There is a wealth of evidence [17] that soluble manganese in surface waters is present as the simple aqueous Mn^{2+} cation, and is normally not complexed by either inorganic or organic species.

Conversely, insoluble manganese is present in such waters as manganese oxides, often of mixed oxidation states [18].

The above approach to speciation totally ignores the possibility that filterable manganese may contain colloidal manganese oxide species, usually of oxidation state greater than (II). Where this is the case, errors can be made by over-calculating the amount of oxidising agent required to remove soluble manganese in the treatment plant. That colloidal manganese (di)oxide is not speciated at the standard treatment plant is hardly surprising, as such speciation is both expensive and very problematic. For those with large amounts of time and money, electron paramagnetic resonance spectroscopy (EPRS) appears to offer a speciation approach: $Mn^{2+}_{(aq)}$ cations give excellent EPR spectra, which cannot be confused with those obtained from higher oxidation state manganese species. However, manganese oxides, particularly colloidal species, are likely to adsorb Mn^{2+} ions. Thus the EPR signal for Mn^{2+} may be quenched in the presence of colloidal manganese dioxide, leading to a reduction in the calculated concentration of Mn^{2+}.

Measurement of oxidising equivalent [19] may be used to determine the amount of oxidising agent required to convert all lower oxidation states of manganese to the +4 state. Such an approach does not actually speciate the manganese present, but offers an operational approach to assessing how much oxidising agent is required for manganese removal. The fact that the method will also assess other entities present that may also use oxidising agent is not a disadvantage, as there is much evidence [20] that organic species in water may be oxidised prior to manganese species and thus it will be essential during manganese removal using added oxidant to use enough oxidant to oxidise such species as well as any Mn(II). Problems that may arise through the production of oxidised organics will be discussed in later sections.
A colorimetric reagent, leucoberbelin blue, has been described as a method of determining the concentration of Mn(III) and Mn(IV) oxides in the presence of Mn(II) and other metals in microbiological culture media [21]. The method appears to work well for fine particulates in such media, but as the reaction is essentially heterogeneous, there is a delay time for accumulated oxide deposits. We have found that there are interferences, which appear to be variable in extent, when the method is applied to either raw or treated waters in a treatment plant. Certainly, the presence of any oxidising agent such as chlorine (not usually present in media cultures!) or permanganate in water rapidly reacts with the dyestuff.

A recent simplistic attempt by us to separate $Mn^{2+}_{(aq)}$ from colloidal manganese dioxide by passage of relevant solutions through a column containing cation exchange resin was of little use. Whereas for separate solutions, $Mn^{2+}_{(aq)}$ was totally removed by the resin, but colloidal MnO_2 was not removed at all, mixtures of the two species had variable amounts of manganese removed; the atomic absorption spectroscopy (AAS) detection system that was used to determine manganese concentration in the column effluent could not speciate Mn^{2+} from colloidal MnO_2. Again, it seems likely that the Mn^{2+} adsorbs to the colloidal MnO_2; such adsorption may facilitate the passage of the Mn^{2+} through the column, but may also help retention of the MnO_2.

Although the speciation of colloidal manganese is difficult, and not usually attempted, the determination of filterable and particulate manganese is straightforward. Acidification of filtrates ensures their stability as Mn(II), while acidic digestion of particulates, assumed to be oxides, reduces them to a soluble form (presumably also Mn(II)). Thus analysis of the filtrate and particulate species can be undertaken on manganese solutions. Although flame AAS is a suitable tool for manganese determination of soluble manganese, the level of detection of the element is realistically only approximately $20\mu\ L^{-1}$, which is not low enough to confirm that an adequate removal process has taken place. However, carbon furnace AAS or inductively

coupled plasma analysis are both able to yield accurate results for manganese at concentrations of less than 5μg L^{-1}. For laboratories, in which such equipment is not available, there are literally hundreds of spectrophotometric methods of analysing for manganese in aqueous solution [22, 23]; however, many of them will not be sensitive enough for detection of the low levels required to assess successful manganese removal.

Speciation of manganese oxides found as particulates in freshwaters can be undertaken, with varying amounts of success, using X-ray photon spectroscopy (XPS), and by surface analysis techniques. Although pure mineralogical forms of the various possible oxides, e.g. MnO, MnO(OH), Mn_2O_3, MnO_2, may be identified by X-ray diffraction methods, the species present in such waters are likely to be mixed and strongly hydrated. In the XPS technique the oxidation state of the manganese is determined by a binding energy associated with an electron transition in the innermost electronic levels of the atom. Unfortunately for manganese, the energies (given in electron volts) of the peaks obtained for binding energies for the various manganese oxides are close together (often within experimental error), and distinction of oxidation state of the manganese atom owes as much to art as to science [9]. Add to this the fact that to obtain XPS data, one concentrates the study on a very limited area of the sample (usually when studying manganese oxides on the darkest area), and it can be seen that there is a major problem associated with how accurately the results represent a homogeneous sample.

3. MANGANESE ASSOCIATED PROBLEMS

3.1. Raw water source problems

There seems little doubt that neutral pH water drawn from a source which is well-oxygenated will contain manganese oxides as the main manganese species, while water drawn from a source of low to zero dissolved oxygen can be expected to have manganese present as soluble $Mn^{2+}_{(aq)}$ [24]. Thus, where water is drawn from a summer stratified dam or lake, the epilimnic water will contain mainly oxidised manganese species, while the hypolimnic waters will contain almost exclusively the soluble Mn(II) species.

3.2. Treatment plant problems

Treatment plant problems with manganese are associated with failures in processes used to oxidise $Mn^{2+}_{(aq)}$ during treatment. Detailed discussion of this topic will be delayed until Section 4.2 of this chapter, however it can be noted here that the wish of most treatment plant operators to waste as little water as possible can result in recycling of backwash (and even sludge dewatering) waters into the plant; such waters invariably possess high manganese concentrations [25]. Also, as noted previously, the pH conditions and manganese oxidation processes are favourable to in-plant colloidal manganese dioxide formation; virtually nothing is known about such an occurrence.

In recent years there has been a movement away from the use of alum as a flocculent in water treatment in favour of the use of iron salts. Such a change has been driven, to some extent, by the toxicity of aluminium residuals, and the claimed connection with Alzheimer's disease [26]. The need to control flocculation pH more closely when iron salts are used is a problem which is not directly connected with manganese, but the fact that all iron compounds/solutions contain substantial concentrations (up to 0.05% in even analytical reagent grade iron salts) is a matter of concern to the treatment operator. Replacement of alum with iron flocculent will almost certainly lead to the necessity for manganese removal later in the treatment process.

3.3. Reticulation problems

There is a growing body of evidence [21] that regular input of soluble manganese at a level greater than 5 μg L^{-1} into a reticulation system may lead to consumers' "dirty water" complaints. The problem of dirty black deposits in pipes carrying manganese-bearing water has been well-documented over a long period in many regions of the world. As long ago as 1967, a publication [27], on manganese-containing deposition in pipelines carrying water to hydroelectric generating plants in Tasmania, initiated an interest by one of the authors (B.C.) in this problem.

In sections of the distribution system in which residual chlorination of the potable water is still present, water discolouration will be produced by chemical oxidation of the soluble manganese, while in sections in which residual chlorination is no longer effective, microbially mediated pipe deposits will occur [5]. Such deposits will usually contain manganese as well as a number of other possible elements, e.g. aluminium (if used in flocculation), iron and calcium (particularly in harder waters). These deposits will often be quite stable on the pipe surfaces leading gradually to a reduction in pipe diameter and delivery volumes, but they can be also removed by processes such as water surges, and will yield highly discoloured water to the consumer [28].

Pedomicrobium manganicum has been shown to often be associated with manganese oxide deposits in reticulation systems with little or no residual disinfection. The microorganism is very common in surface waters and is an exceedingly capable manganese oxidiser; so much so that it has been used in approaches to in-plant microbial manganese removal.

4. REMOVAL APPROACHES

4.1. In source water

Using the generalisation noted in Section 3.1, countless reservoirs/dams have had their hypolimnic waters oxygenated by a number of different approaches. Undoubtedly the most commonly used approach to such oxygenation has been the technique of dam destratification by the introduction of air via a diffuser close to the dam base [29]. The process does not aim to pump enough air into the dam to fully oxygenate the overlying water, rather the oxygen flow destabilises the dam water structure, breaks down the thermocline and sets up current patterns so that aerial oxygen is introduced into all the waters of the dam. Although there is good evidence that well-structured destratification, whether achieved by pumping air into the dam or by such techniques as use of impellers [30], will markedly reduce soluble manganese levels in the storage waters [31], there are some negatives to the process that should be kept in mind. Firstly, destratification raises the overall temperature of the water body to approximate that of the surface waters [32]. Such a result may negate attempts to "blend" treatment plant intake waters to achieve the most advantageous water temperature for subsequent treatment. Secondly, and as a consequence of this increase in waterbody temperature, the balance of algal species in the water may change markedly, with possibly large increases in unwanted toxic species. It should also be noted that water usage may change dam water levels, so that destratification is only partially successful; as the aeration unit is usually situated close to the raw water intake, this may not be important. Thirdly, destratification may mix the dam water, increasing actual raw water manganese, albeit in an insoluble rather than soluble form. This may cause problems in a poorly-run treatment plant.

It should also be kept in mind by water treatment authorities that, where the treatment plant and the storage lake/reservoir are some distance apart and the raw water must pass through a lengthy delivery main, dam destratification may help reduce the input of manganese into the main with concomitant reduction in in-pipe deposition and reduced requirement for pipeline cleaning.

Techniques other than destratification have been used to seal the sediment from the hypolimnic water without breaking the lake thermocline. These include hypolimnetic oxygenation, and sealing of the sediment with clay-type materials. These approaches will only be successful if the source of manganese to the hypolimnium is from the dam sediment rather from inflow waters; although the authors' experience [33] is that sedimental release is the main source of soluble manganese in storages, there is a need to ensure this is the case by close study of both input waters and sedimental release before embarking on costly sediment sealing.

The possibility that destratification, or indeed a combination of natural factors, may yield colloidal manganese oxide species in the dam raw water, has not been addressed at this stage for reasons associated with the difficulty of analytical determination of such species.

Finally, it should be kept in mind that in-storage treatment for manganese may not be the best approach to its removal. In a well-run treatment plant it may be a cheaper approach to react quickly to increased levels of raw water manganese, typically occurring with dam "turn-over", with in-plant removal approaches. In this context, it is the authors' experience that most treatment plants using waters from a destratified storage have in-plant manganese removal systems also present.

4.2. In-plant

In-plant removal of manganese can be achieved by use of chemical oxidants, including air, or by microbial oxidation; both Knocke et. al. [34] and Degrémont [10] have extensively reviewed the various approaches. The choice of point of oxidation will depend upon the actual physical makeup of equipment of the plant and upon the nature of input water.

Chemical oxidation may be undertaken, often with pH raising, upon raw water entering the plant, or upon water that has undergone flocculation prior to, or during, filtration. Aeration of incoming raw water, if physically possible, is a useful initial approach to manganese removal; it can be seen as an extension/replacement for reservoir destratification. Both chlorine and permanganate can be added to the raw water, usually with pH elevation above 8, to remove manganese in flocculation, although it should be noted that pre-chlorination of this type is very likely to yield chlorinated organics.

For plants with flocculation, sedimentation and filtration, and with raw water with even a moderate organic content, the choice of oxidation site will usually be after sedimentation, when substantial removal of organics has been achieved,; such an approach reduces the amounts of chemical oxidants required for manganese removal, and reduces the likelihood of formation of chlorinated organic products such as trihalomethanes (THMs). Such products are known toxicants, and most potable water treatment plants will have specified limits upon the allowable levels of THMs in their finished water; a WHO website states that a number of countries have set guidelines or standards for THMs in drinking water from 0.025 to 0.25 mg L^{-1}.

For direct filtration plants, in which there is no provision for sedimentation, it may be preferable to add oxidant prior to flocculant, although it has been established that not only do organics consume chemical oxidants, often prior to manganese oxidation [35], but that both humics [36] and fulvics [37] can reduce manganese oxides. It also should be kept in mind that

the ideal pH for formation of oxidised manganese species is at least 8, and preferably 9 [36]; attempts to add oxidants together with low pH flocculants such as alum are likely to be self-negating.

Microbial manganese removal is traditionally undertaken using a biologically active filter system [38], although Sly *et. al.* [13, 39] have studied the possibility of by-passing the use of sedimentation and chemical oxidation totally by achieving biological removal from input raw water.

The three most commonly used oxidants for manganese removal are air, chlorine and potassium permanganate. Although permanganate will react rapidly with Mn(II) at pH 8, chlorine and air only oxidise Mn(II) slowly at this pH, and for these latter oxidants the success of manganese removal will depend upon catalytic behaviour of the manganese oxide formed [40-42].

Although all three commonly used chemicals can be used alone, there is much evidence that they are most successful in manganese removal when used in conjunction with an active catalytic manganese (di)oxide presence; such a presence can be prepared upon filtration media such as sand or coal. This approach will be discussed fully below.

The use of permanganate as an oxidant for manganese removal can be very successful, but there are problems that should be kept in mind. Although the theoretical oxidation reaction stoichiometry is two moles of permanganate will oxidise three moles of Mn(II), producing Mn(IV) oxide in each case, this stoichiometry is not met. Stumm and Morgan [43] have pointed out that autocatalysis by manganese dioxide will promote the reaction, and there is also the fact that some of the oxide formed will have an oxidation state of less than +4. An estimate of the permanganate "demand" of the water being treated is undoubtedly preferable to adding the reagent until the water "just turns pink"; addition of excess manganese to the water being treated, by over-use of permanganate, is not recommended, particularly if there is a possibility that oxidised manganese deposits can be resolubilized if treatment goes awry, e.g. anoxic sedimentation chambers. Also a pinkish tinge in the potable water is not generally considered advisable. Although the chances of consumers receiving pink water are very limited, there is no point in introducing more deposited manganese into the reticulation system.

Less commonly used oxidants include chlorine dioxide and ozone. Other oxidants such as hydrogen peroxide do not oxidise $Mn^{2+}_{(aq)}$; this is not surprising as peroxide has been known since the 19^{th} century to be able to reduce higher oxidation state manganese compounds in acid solution liberating oxygen [44].

Removal of manganese by addition of either chlorine dioxide or ozone to the water entering the filtration media is highly successful, although a more costly approach to those using the more common oxidants. Both reagents rapidly oxidise $Mn^{2+}_{(aq)}$ to insoluble manganese oxides, but both have their serious drawbacks.

Chlorine dioxide, produced on site by reacting concentrated hydrochloric acid and sodium chlorite is not an easy gas to handle. Although not likely to explode in aqueous solution, the material is very corrosive on seals, and is highly toxic as a gas. Furthermore, its reaction by-products include chlorite and chlorate whose presence in reticulated water is problematic. Having said this, chlorine dioxide very successfully removes manganese over the pH range of natural waters without producing the chlorinated by-products of the type obtained when using chlorine. Masschelein's comprehensive work on chlorine dioxide [45] should be studied by those contemplating its use.

Ozone is also a most successful oxidant [46]. However overdosing may lead to solubilization of manganese oxides by oxidation to permanganate [47], with subsequent

manganese oxide deposits in the treatment plant or the distribution system. Studies [48] also indicate that waters with a high organic content will consume large amounts of ozone thereby either negating its manganese removal potential, or markedly increasing the cost of water treatment.

To this stage of the discussion, manganese removal in-plant has been treated as a simple oxidation to insoluble oxide step followed by filtration of the deposited material, which can then be removed from the filter medium by backwashing. However, this approach fails to take into account the very important role played by the deposited manganese oxides in removing further manganese from the water being filtered.

The high adsorption ability of manganese dioxide for cationic species is well-documented [43], and its use in removing manganese from potable water has been commercially available as the "manganese greensand" process for many years [49]. In this process manganese(II) ions are adsorbed to a zeolite substrate which is then treated with oxidant in the presence of base to yield a highly active greenish coloured manganese oxide layer; the process very successfully removes manganese from solution, but can prove to be costly on a large scale. The need for zeolite in this process has been shown not to be essential, and a number of water treatment plants have achieved successful manganese removal by treating typical filter media, viz. anthracite or sand, with permanganate solution, and then adding hydroxide to provide a manganese oxide coated media.

Much more recent work by Knocke and co-workers [50], building on the observation that in-plant filter media appeared to "cure" over time, and build up a layer of manganese-adsorbing material [51], has opened up a new approach to manganese removal based on what has been termed a "modified greensand process".

In a seminal paper published in 1988, Knocke et al. [52] described work which showed efficient manganese removal, with the addition of free chlorine, by filtration through naturally oxide-coated anthracite and sand. The work done by Knocke and co-workers over the last decade and a half has been both extensive and comprehensive. An excellent starting point to develop an understanding of how important it has been to manganese removal is the report prepared in 1990 for the American Water Works Association Research Foundation [50]; the report also contains an excellent reference survey, some items of which are included in this chapter.

To summarize the conclusions to this report, we reiterate the five made by Knocke et al. [50]. Firstly, the ability of oxide-coated media to adsorb Mn(II) is a function of both pH and surface oxide concentration. Secondly, Mn(II) uptake by the oxide surface is first-order in rate. Thirdly, and very importantly, in the absence of free chlorine, below pH 9.0, there is no auto-oxidative reaction between surface oxide and adsorbed Mn(II). Fourthly, in the presence of free chlorine, oxidation of the Mn(II) does not occur in the pre-filter water, but at the oxide surface, where adsorbed Mn(II) is readily oxidised yielding a fresh oxide surface to continuously promote manganese removal. Fifthly, the presence of surface manganese oxide deposits does not appear to change the filter characteristics of the filter media for a number of water treatment plants studied.

Of interest in the many observations made by Knocke et al. [50], is that continuous chlorine addition to the filter may not be essential in all cases. In cases where water pH > 7, soluble Mn(II) < 0.2 mg L^{-1}, and there is sufficient filter bed depth, intermittent chlorine treatment may suffice to ensure significant surface oxide layer on the media. Where the plant operation aims to reduce possible chlorinated organic production, this is obviously desirable. Also worthy of note is the fact that the work reports that major manganese removal occurs in the very top layers of the filter bed. This result is paralleled by work by Eley and Nicholson

[53], who also found that at elevated pH (9.0), chlorination of the water entering the filter led to deposition of the MnO_x in the very top layers of the filter bed. We have also observed in our work using a large scale pilot plant at the Gold Coast, Australia, that sodium hypochlorite, rather than chlorine dosed onto the filter, removes the manganese completely in the very top layers of the anthracite filter bed; we suspect that the elevated pH of the hypochlorite solution yields instantaneous nucleation of the MnO_x, which is then removed by filtration/adsorption onto the media, even though the measured pH of the water entering the filter does not appear to change from its usual operational value of 7.0 – 7.5. Prolonged running of the pilot plant's four filter columns gives no indication that there is any MnO_x build-up down the columns.

Hargette and Knocke [40] also note that at pH > 7, chlorine oxidation led to rapid MnO_x deposition and "particle filtration" on the media. However, Knocke's work, using free chlorine rather than sodium hypochlorite, and usually carried out on water of lower pH (5-7) [5-7, 50, 52], indicates clearly that at these pH values little Mn(II) oxidation is achieved by chlorine alone, and that MnO_x adsorption and catalyzation of Mn(II) oxidation plays an important role [40].

For those desirous of obtaining fuller information on Knocke's methods for coating anthracite and sand filter media with MnO_x, reference to work published in 1997 [54] is recommended. Those wishing to develop a deeper understanding of rates and mechanisms of the process are referred to Knocke et al. 's work elsewhere [55-58].

Of interest in elucidating the processes taking place upon the filter media have been attempts to characterize the manganese oxide containing coatings present upon them. Eley and Nicholson [53] showed that the coatings on sand media, from a Scottish water treatment plant using iron coagulation, were amorphous in composition and, in terms of metals, consisted primarily of manganese and iron oxides. Later work by Merkle et al. [59] found somewhat similar results, although as alum coagulant had been used, this element rather than iron was found to be present. These workers suggest $Al_5Mn_{13}O_{28}{\cdot}8H_2O$ (giving the Mn a positive 3.15 charge) as the only (very tentatively) identified mineral presence in the coating, but in general the work demonstrates the difficulty, noted in earlier sections, of obtaining definitive data from what are very heterogeneous/amorphous deposits. On the basis of the traditionally used Mn_{3s} multiplet peak splitting of XPS, the authors suggest that the manganese is present in the +4 state. We also have studied the manganese oxide deposits from the very top layers of our filtration pilot plant, and our XPS data also tend to indicate that Mn(IV) is present, although the previously noted problems of interpretation of XPS manganese oxide spectra cannot rule out the presence of Mn(III).

Other chemical approaches to manganese removal in-plant can be generally described as those connected with sequestration. Of these methods, the best known is the use of polyphosphate. This entity can coordinate well with Mn(II) [60], and indeed there have been suggestions that addition of polyphosphate to reticulated water could prevent manganese deposition in the reticulation system. Nevertheless, there is evidence that the use of polyphosphate prior to filtration and oxidation promotes manganese oxide deposition and removal.

The above study [60] also looked at the use of sodium silicate as a sequestering agent; the approach does not look inviting. Humic acid will bind Mn(II) at a low pH(< 6.5) [61], but other work [17] has shown that for freshwater levels of manganese at neutral pH values there is little complex formation.

Use of biological in-plant manganese removal became a commercially viable approach during the last two decades of the 20th century. The approach, which either removes or substantially reduces the in-plant use of chemicals, undesirable by-products and sludge waste,

is intuitively attractive. Its use in Europe is now some two decades old [62, 63], and although not applicable to all raw waters [50], it has great potential for waters of higher pH, with dissolved oxygen grreater than 5 mg L^{-1} and a relatively high (300 – 400 mV) Eh value.
Sommerfield [64] notes some five bacteria, commonly present in freshwaters, which can oxidise manganese(II), viz. *Leptothrix, Crenothrix, Siderocapasa, Siderocytes* and *Metallogenium*. Oxidation mechanisms appear to be twofold: intracellular enzymatic, and extracellular catalytic action. Of particular interest for practical manganese removal is the ability of these species to promote extracellular oxidation by excreted polymer action [65]. Sly *et. al.* [13] have utilized *Pedomicrobium manganicum*, isolated from a water distribution system, supported on magnetite particles in a continuous flow recycling fluidised bed bioreactor at pH approx. 7 – 8, to remove up to 100% of input water manganese (0.25 – 8.5 mg L^{-1}) with a residence time of 21 hours. An up-scaled pilot plant designed to study the viability of this approach is currently in operation at Hinze Dam on the Gold Coast near Brisbane.

Watanabe et al. [66] have successfully utilized powdered activated carbon as a support for *Leptothrix* and *Hypomicrobium* in a bioreactor designed to oxidise Mn(II) and ammonia and enhance humic removal, while the wide applicability of the approach is seen in the use of *Leptothrix discophora* to remove Mn(II) in mine drainage water [67].

Over the last two decades, commercially large treatment plants employing membrane filtration have become more common. Reverse osmosis membranes will exclude Mn^{2+} ions, but their use is a wasteful process in terms of water; more commonly used membranes have a pore size which requires that successful manganese removal will be achieved by oxide production prior to filtration.

4.3. In reticulation system

Failure to reduce levels of manganese entering the reticulation system to the order of 10 μg L^{-1} or less, will almost certainly lead to manganese–associated build up in the system; the piping hardware that constitutes a water distribution network acts as a superb chromotographic system. However, any removal method applied at this later stage, whether chemical, flushing or pigging (with a slug forced through the lines by water pressure) will yield a highly coloured (usually black) sludge. Where high levels of chlorine or chlorine dioxide are added to the reticulated water to destroy bacterial growth in the lines, one can expect a high level of consumer complaints. The public perception of the sludge obtained when flushing water mains is almost as damaging in its perception of discoloured tap water. Nevertheless, flushing of mains is still widely performed by water authorities; one can only observe that better control of the contaminants in the water produced at the treatment plant is a much better way to proceed politically and almost certainly economically.

There is some evidence that chloramination of potable water gives a longer disinfectant residual in the reticulation system, thus cutting down on bacterial growth; whether this reduces the need for mains cleaning is not clear.

4.4. At point-of-use

From the point of view of water authorities, point-of-use water treatment is evidence of poor water treatment operations. However, there is a substantial body of the public who are convinced that the water from their household tap requires treatment – particularly if it is to be used for drinking purposes. Much of this treatment is to remove taste and odour associated with chlorine, or chlorinated products, or organic contaminants such as methylisoborneol (MIB) or geosmin.

There are literally hundreds of proprietary point-of-use home water purifiers on the market, all of which will remove particulate manganese, and some dissolved manganese, from potable water. Small scale reverse osmosis units can remove virtually anything, soluble or not from the water [68], while filter cartridges of various porosities can be used depending upon the "purity" of water required [69]. Ion exchange systems can also be used to remove unwanted (usually) cationic species such as manganese from water. These systems all have their drawback; reverse osmosis is wasteful of water [68], as appreciable volumes of input water are run over the membrane to remove species trapped on it; filter cartridge systems must be effectively maintained and cartridges regularly replaced [69].

One approach that is probably worthy of consideration by the householder is the provision of filters in the water inlets of the household washing machine. Automatic machines in particular use valves which turn on and off with substantial vibration giving physical shock to the plumbing and releasing any pipe deposits into the machine. Black stains, which are difficult to remove, on otherwise clean laundry are not a popular item; although such stains can be removed by the use of a solution of sodium dithionate, the associated sulfur dioxide can attack both fabrics and their dyes.

Industrial users of reticulated water, who have needs to ensure purity, are likely to use reverse osmosis or distillation for their needs.

4.5. Manganese removal – summary

Like all contaminants in potable water, the procedures used to remove manganese should aim at a fixed value in the finished water, preferably less than 10 $\mu g\ L^{-1}$. It has been surprising to us that so much of the R&D work reviewed for this chapter dealt with percentage removal, and we point out that on the above absolute figure, 95% removal from water with more than 110 $\mu g\ L^{-1}$ of manganese will not meet the criterion.

Dam destratification or various oxygenation techniques will convert soluble manganese into insoluble forms, although it may not reduce the overall total manganese entering the treatment plant. Changes in other parameters, e.g. algal growth and variety, associated with in-dam processes should be considered before proceeding down this line of approach.

Procedures for in-plant manganese removal will be virtually essential if manganese is present in the dam water, even if dam manipulation is used. The best and probably most economical approaches for successful removal appear to lie in adding oxidant to the water entering the filtration beds. However, there may be sound reasons for adding oxidant earlier in the process. Obviously, where any form of membrane filtration is used, the addition of oxidant may be unnecessary and even counter-productive.

REFERENCES

[1] D. Comnie, S. Evans, R. Gale and P. Kitney, Water, 29 (2002) 20.
[2] G.C. White, Handbook of Chlorination, Van Nostrand Reinhold, New York, 1972, p.278.
[3] WHO, Guidelines for Drinking Water Quality, 2nd ed. Vol.1 Recommendations, WHO, Geneva, 1993.
[4] United States Environmental Protection Agency, Health Effect Support Document for Manganese, Alexandria, Virginia, US, 2002, pp.1-1.
[5] L.I. Sly, M.C. Hodgkinson and V. Arunpairojana, Appl. Environ. Microbiol., 56 (1990) 628.
[6] N.N. Greenwood and A. Earnshaw, Chemistry of the Elements, Pergamon, Oxford, 1984, pp.1211 - 1213.

[7] B. Chiswell and D. Johnson, Manganese, in Handbook on Metals in Clinical and Analytical Chemistry, H.G. Seiler, A. Sigel, and H. Sigel (eds.), Marcel Decker, New York, 1994, pp.467 - 478.
[8] B. Chiswell, E.D. McKenzie and L. F. Lindoy, Manganese, in Comprehensive Coordination Chemistry, G. Wilkinson (ed.), Pergamon Press, London, 1988, pp.1 - 122.
[9] M. Zaw and B. Chiswell, Talanta, 42 (1994) 27.
[10] Degremont, Water Treatment Handbook, Degremont, Paris, 1991, pp.1211 - 1216.
[11] J.J. Morgan and W. Stumm, J. Colloid Sci., 19 (1964) 347.
[12] W.C. Ghiorse and P. Hirsch, Arch. Microbiol., 44 (1979) 213.
[13] L.I. Sly, V. Arunpairojana and D.R. Dixon, Appl. Environ. Microbiol., 56 (1990) 2791.
[14] B. Chiswell and M. Zaw, Hydrol. Procceses, 3 (1989) 277.
[15] B. Chiswell and G. Rauchle, Proc. Roy. Soc. Qld., 97 (1986) 53.
[16] B. Chiswell, D.R. Dixon, G. Hamilton, L.I. Sly and T.D. Waite, Chemistry in Australia, 59 (1992) 400.
[17] B. Chiswell and M.B. Mokhtar, Talanta, 33 (1986) 53.
[18] M. Zaw and B. Chiswell, Talanta, 42 (1995) 27.
[19] D. Johnson and B. Chiswell, Talanta, 40 (1993) 533.
[20] W. Stumm and J.J. Morgan, Aquatic Chemistry, John Wiley, New York, 1996, pp.464-9.
[21] L.I. Sly, M.C. Hodgkinson and V. Arunpairojana, FEMS Microbiol. Ecol., 53 (1988) 175.
[22] B. Chiswell, G. Rauchle and M. Pascoe, Talanta, 37 (1990) 237.
[23] B. Chiswell and K. R. O'Halloran, Talanta, 38 (1991) 64.
[24] B. Chiswell and M. B. Mokhtar, Talanta, 33 (1986) 669.
[25] G. Hamilton, Water, 29 (2002) 34.
[26] P. Zatta, R. Lucchini, S.J. van Rensburg and A. Taylor, Brain Res. Bull., 62 (2003) 15.
[27] P.A. Tyler and K.C. Marshall, Arch. fur Mikrobiologie, 56 (1967) 344.
[28] R.A. Prince, I. Goulter and G. Ryan, Water, 30 (2003) 62.
[29] T.F. McAuliffe and R.S. Rosich, Research Report No 9, Urban Water Research Assocn., (1989), pp.1 - 233.
[30] P. Morgan and K. Bugden, Water, 29 (2002) 30.
[31] T. Asaeda and J. Imberger, Dynamics Report ED-88-250, Centre Water Research, Univ. Western Australia, (1988).
[32] B. Chiswell and M. Zaw, Environ. Monitor Assess., 19 (1991) 433.
[33] B. Chiswell and S-H. Huang, Report to SEQWB, National Research Centre for Environmental Toxicology, Univ. of Queensland, (2001).
[34] W.R. Knocke, J.E. Van Benschoten, M. Kearney, M. Soborski and D.A. Reckhow, Alternative Oxidants for the Removal of Soluble Iron and Manganese, AWWA Research Foundation Report, AWWA, Denver (1990).
[35] D. Giannissis, P. LeCloirec, C. Dorange and G. Martin, J. Water Supply: Research Technol. - Aqua, 4 (1985) 197.
[36] D. Bannerjee and N.H. Nesbitt, Geochim. Cosmochim. Acta, 65 (2001) 1703.
[37] T.D. Waite, I.C. Wrigley and R. Szymczak, Environ. Sci. Technol., 22 (1988) 778.
[38] H. Shorney, E. Lee, D. Smith, W.R. Knocke and B.W. Long, Proc. Water Qual. Technol. Conf., 1998.
[39] L.I. Sly, V. Arunpairojana and D.R. Dixon, Australia Patent No. WO 92-AU579 Univ. Q'land/CSIRO, (1993).
[40] A.C. Hargette and W.R. Knocke, J. Environ. Eng., 127 (2001) 1132.
[41] S. Aydin, N. Tufekei and S. Arayici, Fresenius Environ. Bull., (2001) 386.
[42] O.J. Hao, A.P. Davis and P.H. Chang, J. Environ. Eng., 117 (1991) 359.
[43] W. Stumm and J.J. Morgan, Aquatic Chemistry, John Wiley, New York, 1996, pp.683-686.
[44] S. Taylor, Inorganic and Theoretical Chemistry, W. Heineman, London, 1947, p.p.7.
[45] W.J. Masschelein, Chlorine Dioxide: Chemistry and Environmental Impact of Oxychlorine Compounds, R.P. Rice (ed.), Ann Arbor, Michigan, 1979.

[46] L. Wang, B. Wang, D. Wang, W. Zhang, Y. Yang, Y. Du and Q. Kong, J. Water Supply: Research Technol. - Aqua, 51(2002) 209.
[47] A. Wilczak, W.R. Knocke, R.E. Hubel and E.M. Aieta, J. - Am. Water Works Assoc., 81(1993) 98.
[48] D. Gregory and K.H Carlson, Ozone: Sci. Eng., 23 (2001) 149.
[49] E. Nordell, Water Treatment, Van Nostrand Reinhold Co., New York, 1961.
[50] W.R. Knocke, S. Occiano and R. Hungate, Removal of Soluble Manganese from Water by Oxide Coated Filter Media, AWWA Research Foundation Report, AWWA, Denver (1990).
[51] C. Weng, D.L. Hoven and B.J. Schwartz, J. - Am. Water Works Assoc., 78 (1986) 83.
[52] W.R. Knocke, J.R. Haman and C.P. Thomson, J. - Am. Water Works Assoc., 80 (1988) 65.
[53] M. Eley and K. Nicholson, Environ. Geochem. Health, 15 (1993) 85.
[54] P.B. Merkle, W.R. Knocke and D. Gallagher, J. Environ. Eng., 123 (1997) 542.
[55] W.R. Knocke, S.C. Occiano and R. Hungate, Res. Technol., 83 (1991) 64.
[56] K.H. Carlson and W.R. Knocke, J. Environ. Eng., 125 (1999) 892.
[57] P.B. Merkle, W.R. Knocke, D.L. Gallagher and J.C. Little, J. Environ. Eng., 123 (1997) 650.
[58] B.M. Coffey, D.L. Gallagher and W.R. Knocke, J. Environ. Eng., 119 (1993) 679.
[59] P.B. Merkle, W.R. Knocke, D. Gallagher, J. Junta-Rosso and T. Solberg, J. - Am. Water Works Assoc., 88 (1996) 62.
[60] R.B. Robinson, G.D. Reed, D. Christodos, B. Frazier and V. Chidambariah, Sequestering Methods of Iron and Manganese Treatment, AWWA Research Foundation Report, AWWA, Denver (1990).
[61] M. Nakayama, R. Fujiyoshi and S. Sawamura, J. Radioanal. Nucl. Chem., 250 (2001) 433.
[62] J.P. Boudou, P. Kaiser and J.M. Philipot, Water Supply, 3 (1985) 151.
[63] I. Cameron, Water, 23(1996) 25.
[64] E.D. Sommerfield, Iron and Manganese Removal Handbook, Amer. Water Works Assoc., Denver, (1999).
[65] L.I. Sly, M.C. Hodgkinson and V. Arunpairojana, Appl. Environ. Microbiol., 56 (1990) 2791
[66] Y. Watanbe, K. Kimura and T. Suzuki, Water Sci. Technol., 22 (2000) 9.
[67] K. Sasaki, M. Endo, K. Kurosawa and H. Konno, Mater. Trans., 43 (2002) 2773.
[68] F. Bergsrud, B. Seelig and R. Derickson, Treatment Systems for Household Water Supplies - Reverse Osmosis, http://www.ext.nodak.edu/extpubs/h2oqual/watsys/ae1047w.htm, in NDSU Extension Service. 1992, North Dakota State University.
[69] F. Bergsrud, B. Seelig and R. Derickson, Treatment Systems for Household Water Supplies-Activated Carbon Filtration, http://www.ext.nodak.edu/extpubs/h2oqual/watsys/ae1029w.htm, in NDSU Extension Service. 1992, North Dakota State University.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 11: Arsenic removal during drinking water treatment

M. Jekel[a] and G. L. Amy[b]

[a]Technical University of Berlin, Dept. of Water Quality Control, Secr. KF 4, Strasse des 17. Juni 135, D-10623 Berlin, Germany

[b]UNESCO-IHE Institute for Water Education, Westvest 7, 2601 DA Delft, The Netherlands

1. BACKGROUND

1.1. Arsenic chemistry

The common oxidation states of arsenic in natural water sources are +3 or As(III) (arsenous acid, arsenite) and +5 or As(V) (arsenic acid, arsenate) as shown in the inorganic hydrolysis species: H_3AsO_3, $H_2AsO_3^-$, $HAsO_3^{2-}$, AsO_3^{3-} and H_3AsO_4, $H_2AsO_4^-$, $HAsO_4^{2-}$, AsO_4^{3-}, respectively. The dissociation constants of arsenic(III) are $pK_{A1} = 9.22$; $pK_{A2} = 12.10$; $pK_{A3} = 13.40$ [1], implying that at pH 9.2 the As(III) is 50% dissociated to $H_2AsO_3^-$. At lower pH values typical of natural waters, most As(III) exists as the neutral molecule H_3AsO_3. The dissociation constants of As(V) are $pK_{A1} = 2.22$; $pK_{A2} = 6.96$; $pK_{A3} = 11.5$. Thus, at pH 6.96 about 50% of As(V) exists as a monovalent anion $H_2AsO_4^-$ and 50% as a divalent anion $HAsO_4^{2-}$. In the typical pH-ranges of natural waters, only three out of eight species ($H_2AsO_4^-$, $HAsO_4^{2-}$, H_3AsO_3) are dominant and therefore influence the removal techniques. The redox reaction of the As(III)/As(V) system can be described by the following equation:

$$H_3AsO_4 + 2H^+ + 2e^- \rightarrow H_3AsO_3 + H_2O \quad E_0 = +0.56\ V \tag{1}$$

If the Nernst equation is applied for the redox potential at pH 7, a value of E_{pH7} of close to 0 V is calculated for equal concentrations of both species under otherwise standard conditions. To oxidize 99% of the trivalent arsenic to pentavalent species, a redox potential of about +60 mV is necessary, indicating a relatively easy oxidation, even by dissolved oxygen. However, the kinetics of homogeneous oxidation are very slow and may only lead to conversion rates in the range of a few percent per week [2]. Catalytic effects, either by homogeneous or heterogeneous catalysts or by bacterial enzymes, increase the rate of oxidation by orders of magnitude to become interesting techniques for arsenic oxidation ahead of removal processes (see Section 5).

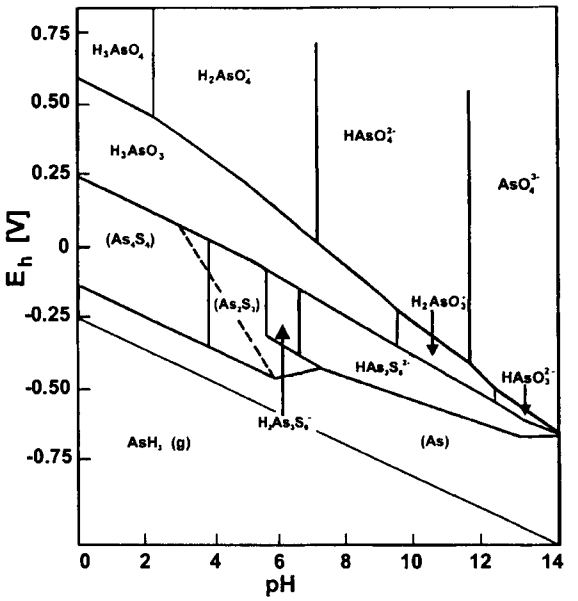


Fig. 1. E_h-pH diagram for arsenic [3]

The overall stability diagram of arsenic is shown in Fig. 1. showing species dominance as a function of E_H and pH.

The pentavalent arsenic system displays a high similarity with the phosphoric acid/ortho-phosphate system in its general dissociation and surface related reactions. The As(V)-species exhibit generally a higher affinity to surfaces and are thus more effectively removed in respective treatment processes, caused by their anionic charges. It can be predicted that a strong competition between arsenate and orthophosphate occurs in these reactions and must be taken into account when considering removal processes.

1.2. Arsenic occurrence

Arsenic occurs worldwide. Small amounts can be detected in rain, rivers, lakes, groundwater and seawater. However, it is occasionally present at much higher concentrations, especially in groundwater, where arsenic concentrations vary regionally due to the geochemical conditions, over a range from <0.5 up to 5000 μg L^{-1}. This is of concern when groundwater is used as a source for drinking water. Arsenic poisoning from natural arsenic in drinking water has been reported from Taiwan, Chile, Argentina, Mexico, China, West Bengal (India) and Bangladesh. In the Bengal Basin, perhaps 40 million people drink water containing "excessive" arsenic [4]. In Germany, arsenic concentrations below 10 μg L^{-1} are common in groundwater, but concentrations up to 250 μg L^{-1} occur in sandstone aquifers [5]. In contrast to Germany, where arsenic in ground water is a minor issue, the problem is generally more widespread in the U.S., with arsenic concentrations exceeding 10 μg L^{-1} more frequently observed in the Southwest of the USA [6]. In the USA, As(V) is the dominant species while, in Bangladesh, As(III) is the dominant species.

1.3. Arsenic regulations

The insights into arsenic health effects (especially its carcinogenicity) as well as exposure and risk assessments were the reasons that the World Health Organization (WHO) re-evaluated its guideline value and adopted 0.01 mg L^{-1} (10 μg L^{-1}) as a new provisional guideline value in 1993. However, based on health criteria, the guideline value for arsenic in drinking water should be less than 0.01 mg L^{-1} [7]. In Germany, this same value was set as a national standard for drinking water in 1996. The U.S. EPA also established 0.01 mg L^{-1} as the enforceable Maximum Contaminant Level (MCL) for arsenic in 2001. In addition, a health-based, non-enforceable Maximum Contaminant Level Goal (MCLG) for arsenic of zero was established. Public water systems must comply with the 0.01 mg L^{-1} EPA standard beginning January 23, 2006.

1.4. Overview of arsenic removal options

To remove arsenic from surface and ground water supplies, various treatment techniques can be considered, depending on water quality, the overall treatment goals, treatment costs and complexity. Fig. 2 provides an overview of the conventional and advanced treatments and removal options.

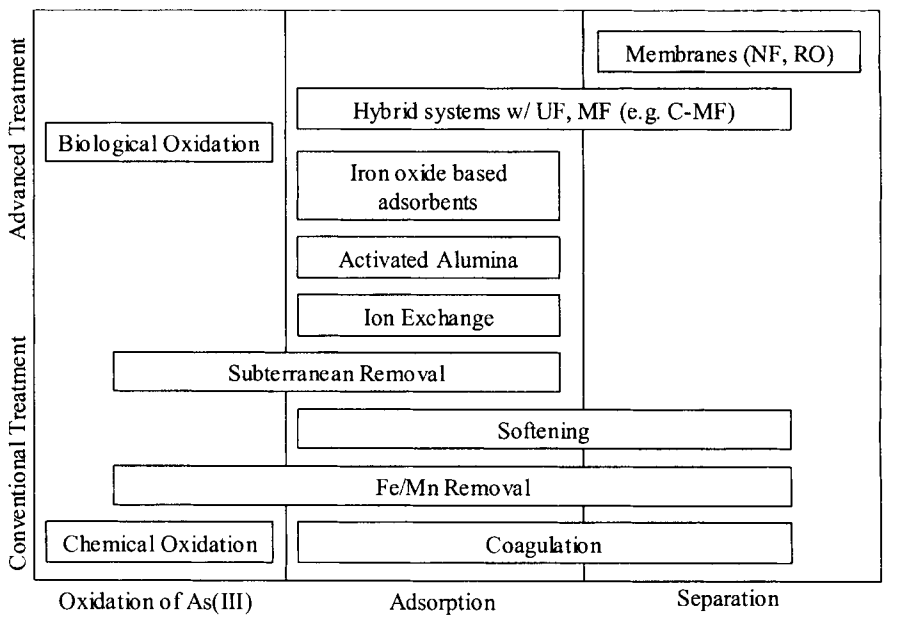


Fig. 2. Overview of arsenic removal options

2. OPTIMIZED CONVENTIONAL TREATMENT

2.1. Coagulation

The coagulation process has traditionally been used to remove particles and colloids, however, it is also applicable to the removal of arsenic. Due to its proven reliability,

coagulation followed by clarification (optional) and filtration is probably the most widely used treatment for arsenic removal. After the dosing of a metal coagulant, a hydroxide precipitate is formed which takes up the arsenic by adsorption. To remove the arsenic-loaded particles from the water, a separation technique, such as filtration over granular media or microfiltration has to be combined with the coagulation process.

The bi- and tri-valent salts of iron or alum can be used as coagulants. Both undergo hydrolysis to various products [8], as summarized in the following equation:

$$Me(H_2O)_6^{3+} \longleftrightarrow Me(OH)_3 + 3H_3O^{3+} \quad (2)$$

In the case of a ferrous salt, iron(II) needs to be oxidized to iron(III), before becoming involved in the hydrolysis reaction shown in Eq. (3).

$$4Fe^{2+} + O_2 + 4H_3O^{+} \longleftrightarrow 4Fe^{3+} + 6H_2O \quad (3)$$

Ferric hydroxides can be formed over a relatively wide pH-range between pH 5 and pH 11 while the use of alum coagulants is restricted to pH 6 to 7.2. The added coagulants can be removed to a very low residual, if the poorly soluble hydroxides are formed at a proper pH and if they are filtered effectively. The use of a polymeric flocculant enhances the performance of the following separation process and may be necessary to remove smaller flocs.

Best removals are achieved with As(V) and the ferric salts, if the pH-value is below 7.2 to 7.5. Under optimized conditions in terms of Fe^{3+}-dosage and pH, efficiencies approach 99 % if the As-content of the raw water is in the range of 0.1 to 1 mg L^{-1}. The residual As-content after solids removal is then below 0.01 mg L^{-1} [9- 13].

In contrast to As-V precipitation by Fe^{3+}-salts, As(III) removals are generally much lower with the same coagulant. Efficiencies of only 50 - 60 % have been found [12]. Figure 3 presents percentage removals for four combinations. Aluminium salts are generally less effective for both As-species, but removal rates of 80 – 90 % are possible for As(V) in the pH-range 6 - 7, where aluminium itself is rather insoluble.

The dosage of the coagulant depends on the specific raw water ion composition, particularly on the pH and the concentration of arsenic and phosphate present. As a rough approximation, the molar concentration of iron should be at least 10 – 20 times higher than the molar arsenic concentration [14]. Due to strong competition, phosphate should be accounted for as arsenic equivalent when using this estimate. Typical doses are between 0.5 and 10 mg Fe L^{-1} while dosage of alum requires higher molar concentrations.

Dosing of ferrous salts (Fe(II)) is followed by oxidation to ferric ions (Fe(III)) and thus leads to the same coagulant system. Table 1 shows the advantages and disadvantages of both coagulants from an operational point of view [5]. In the cited previous studies, the need for the oxidation of As(III) is pointed out and was performed with chlorine. If the removal rates are not sufficient, even though larger dosages of ferric ions (5 – 10 mg L^{-1} of Fe^{3+}) are applied at pH-values below 7.2, it can be concluded that As(III) is present and should be oxidized.

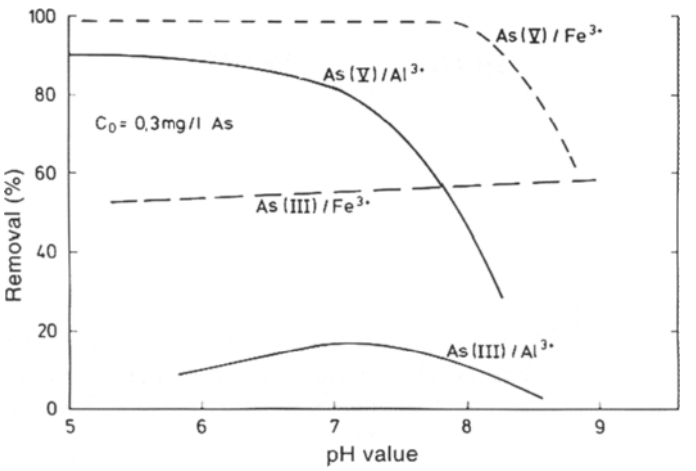


Fig. 3. Influence of pH on the removal of As(III) and As(V) from groundwater by ferric and aluminium ions. Initial arsenic concentration: 0.3 mg L^{-1}; dosage of metal ions: 0.09 mmol L^{-1} (modified from Sorg and Logsdon [12])

Table 1
Advantages (+) and disadvantages (-) of iron(II) and iron (III) coagulation [5]

Coagulation using Fe(III)		Coagulation using Fe(II)	
+	Long storage life of chemicals	+	Good filterability of formed flocs, single media filter sufficient up to 2 mg Fe/L
+	Low price of chemicals	+	Longer operation time, lower water loss
+	Liquid chemicals can be dosed directly	+	Lower sludge volume
+	Simultaneous removal of particles / colloids possible	+	Simple control of the filtration process by pressure loss
-	Lower filterability of the flocs, use of dual media filters or polymeric coagulant aids may be necessary	+	Simpler dosage (turbulent flow in a pipe sufficient)
-	Smaller operation time	+	Lower consumption of buffer capacity
-	Higher sludge volume	-	Restricted storage life of chemicals
-	Control of the filtration process is more complex	-	Preparation of dosing solution necessary
-	Complex dosing, quick and complete mixing necessary	-	Particle removal by additional coagulant dose not possible, poorer aggregation
-	Higher consumption of buffer capacity	-	Aeration may become necessary for oxidation
-	Potential risk of impure dosed chemicals	-	Additional compressed-air machine necessary to break the loaded filter bed

2.2. Softening

Lime treatment to remove or reduce carbonate hardness (softening) is an efficient process for As(V) removal. Sorg and Longsdon [12] presented data, indicating more than 90

%-removal (C_0 = 0.4 ppm) if the pH is above 10.5. As-III could be removed to about 75 % at pH-values of above 11.0. Oxidation of As(III) to As(V) prior to lime softening treatment will increase removal efficiencies, if As(III) is the predominant form. The mechanism of removal may be adsorption onto the calcium carbonate and $Mg(OH)_2$ formed (at high pH) or it may be a direct precipitation of calcium arsenates, similar to the phosphate precipitation under similar conditions. The importance of high pH and concomitant $Mg(OH)_2$ formation is stated in Amy et al. [15].

A survey of softening facilities by AwwaRF [16] indicated that the percentage of arsenate removal is constant regardless of the initial concentration. These findings are contradictory to the above mentioned study [12]. Arsenate removal could be further improved by addition of iron although a lower softening pH is required commensurate with more optimum conditions for $Fe(OH)_3(s)$ precipitation. Consistently, lower As(V) removal was observed in facilities precipitating only calcium carbonate when compared to facilities precipitating calcium carbonate and magnesium and ferric hydroxide. The presence of sulfate or carbonate anions showed no significant interference with As(V) removal at pH 11. However, arsenate removal is reduced in the presence of carbonate at pH 10 to 10.5 and the presence of orthophosphate at pH less than 12.0.

Lime softening is a complex treatment process which produces a considerable amount of sludge. Due to the elevated arsenic concentrations, sludge disposal might be difficult and expensive. Lime softening would only be considered as a reasonable treatment alternative if the treatment goals include the removal of hardness and/or heavy metals. However, raw waters containing both high hardness and high arsenic concentrations are relatively rare [14].

2.3. Fe/Mn Removal

Iron/manganese removal is commonly applied in facilities treating ground water. Arsenic can also be removed, if the iron concentrations are high and the arsenic concentrations are low enough. Dissolved Fe(II) and Mn(II) are oxidized in the filter bed in a catalytic or biological process. Hydroxides/oxides are formed and retained together with soluble arsenate, which is removed by precipitation and adsorption reactions. In addition, As(III) is oxidized to As(V) during the process.

Borho [17] states that arsenic can be efficiently removed by Fe/Mn removal up to an influent concentration of 35 $\mu g\ L^{-1}$. Even at molar Fe(II) to As(III) ratios as high as 25:1, As(III) will sufficiently be oxidized, if the raw water contains more than 100 $\mu g\ L^{-1}$ manganese. An empty-bed contact time of 8 to 10 minutes is recommended. Baldauf [3] reports arsenic removal down to 3 $\mu g\ L^{-1}$ from a reduced raw water containing 1.9 $mg\ L^{-1}$ Fe(II), 0.34 $mg\ L^{-1}$ Mn(II), and 33 $\mu g\ L^{-1}$ arsenic. This is consistent with the results from field studies [18], where arsenic concentrations could be reduced down to 3 µg/L (2.3 $mg\ L^{-1}$ Fe and 20.3 $\mu g\ L^{-1}$ As in the source water) and 11.3 $\mu g\ L^{-1}$ (1.2 $mg\ L^{-1}$ Fe and 48.5 $\mu g\ L^{-1}$ As in the source water), respectively. The latter example shows that arsenic removal is heavily dependent on the amount of iron present in the influent. Thus, an additional dosing of Fe(II) can enhance the removal efficiency. On the other hand, high concentrations of Fe(II) may inhibit oxidation processes by covering the reactive surfaces with oxidation products and As(III) may pass the fixed-bed filter. A breakthrough of Mn(II) can occur at the same time. Sufficient aeration, additional dosing of oxidants or multistage processes can prevent these problems.

3. ADSORPTION PROCESSES

3.1. Metal Oxides

Filtration over granular adsorbents like activated alumina or granular ferric (hydr)oxides is a promising method for arsenic removal. Dosing of additional chemicals is usually not required, although pre-oxidation of As(III) to As(V) generally enhances the performance. Either none or only backwash sludge is produced, and the only residual is loaded (spent) adsorbent, which can either be regenerated or disposed of in a landfill. There has been recent interest in considering high capacity adsorbents as throwaway materials, thus avoiding the brine residual problem upon regeneration. The low requirements of operating and maintaining such treatment systems makes these techniques suitable for wellhead or point of entry (POE) ground water treatment systems, which are frequently used in rural areas. Activated alumina has been used for more than 20 years. The development of iron-based adsorbents in recent years such as granular ferric hydroxide (GFH) has provided an attractive alternative for these systems, since less chemical pre-treatment is required and the operational life is longer compared to anion-exchange resins or activated alumina.

3.2. Activated alumina

Activated alumina is a commercial porous oxide with specific surfaces of 200-300 $m^2 g^{-1}$ and can be used for phosphate, fluoride and arsenic adsorption, especially in a fixed-bed mode. It is rather specific for these anions, and can be regenerated with dilute NaOH and sulfuric acid.

The adsorption of phosphate and arsenate onto activated alumina is quite selective. The sorption equilibria of As(V) on activated alumina were measured by Rosenblum and Clifford [19] as a function of the major parameters. The best pH value for As(V) was 5.5 - 6, where the alumina is protonated, but the anions of the acid added (to lower the pH) are not competitors in adsorption. The alumina capacities were about 5 - 15 g kg^{-1} at equilibrium concentrations of 0.05 – 0.2 ppm. Sulfate anions especially and to a lesser extent, chloride anions, can reduce the capacity by up to 50 % (at 720 ppm sulfate). Frank and Clifford [20] demonstrated with pilot filters that As(III) is not removable by activated alumina, but As(V) is adsorbed and breakthrough may occur at 10,000-20,000 bed volumes (C_0 = 100 ppb, pH = 6).

The adsorption capacity is strongly pH dependent and significantly reduced at pH > 7.3 [21]. Optimal capacities are reached at pH-values between 5.5 and 6.5. In experiments with ground waters (20 µg L^{-1} As(V), pH 7.7 – 8.2), an effluent concentration of 10 µg L^{-1} (MCL) was reached after 4,000 bed volumes. By adjusting the pH to 5.5 – 6.5, capacities of 35,000 bed volumes could be reached [22]. However, most of the ground waters need a pH-adjustment by adding mineral acids or CO_2. Moreover, subsequent neutralization and/or regeneration of the activated alumina is necessary to achieve adequate operation times and minimize disposal costs. Depending on the buffer capacity of the water, considerable amounts of acid or base, respectively, have to be used. In the case of regeneration, the remaining concentrate has to be neutralized. The produced liquid waste stream is highly enriched with arsenic and has to be treated before disposal. With regard to the reliability of the process, there is the possibility that desorption of arsenate can occur during outages of the acid dosage system [21, 23]. Thus, the complexity of the treatment process is significantly increased, outweighing many of the advantages of an adsorption process.

3.3. Iron oxide based adsorbents

The high capacities of amorphous ferric hydroxide for arsenate led to the development of granular ferric hydroxides (GEH® or GFH™) in 1990 – 1994 at the Technical University of Berlin, Department of Water Quality Control. The technique provides a simple filtration process over granular adsorbent media, commonly without any dose of chemicals and without pH adjustment.

Currently more than 60 plants for arsenic removal from drinking water are in operation in Germany, other European countries, and in the USA and Japan. There are about 7000 hand pump plants in West Bengal, India that utilize GFH.

Compared to activated alumina, GFH shows approximately 5 - 10 times higher adsorption capacity for As(V). Competition by phosphate is significant, although lower compared to activated alumina. GFH adsorbs As(V) over a broad pH range, although capacities are diminished at pH > 8. It has been found that As(III) also adsorbs onto GFH to a significant, but lesser extent than As(V) [24].

Experience from 19 full-scale treatment plants using GFH in packed-bed systems has been published [25]. At arsenic concentrations between 10 and 40 $\mu g\ L^{-1}$ and pH values between 6 and 8, from 50,000 up to 250,000 bed volumes until breakthrough at 10 $\mu g\ L^{-1}$ are reached. Usually, the filters are backwashed every 2 to 6 weeks. The exhausted material is changed after approximately 2 to 3 years and can be disposed of in a landfill. Operation and maintenance of the adsorbers is simple, and no chemicals need to be dosed. Demonstration-scale investigations on competitive adsorption, effluent water quality, and backwash behaviour have proven the operational reliability of GFH adsorbers [26].

Ruhland and Jekel [27] evaluated different arsenic removal technologies with regard to technical, ecological, and economical aspects. Based on data obtained from a case study of direct filtration with $FeCl_3$, adsorptive filtration with $FeSO_4$, and adsorption onto GFH, adsorption onto GFH was rated as the preferred method of arsenic removal.

3.4. Ion exchange

Due to the different dissociation equilibria, only the pentavalent As is present as an anion, either mono- or di-valent, in the near-neutral pH-range and can be exchanged in strong base ion exchange resins. The process and operation are similar to activated alumina adsorption. In contrast to iron oxide based adsorbents, however, arsenic is not selectively removed. Other commonly encountered ions in the water, especially sulfate, show a high affinity to ion exchange resins and are usually present in significantly higher concentrations than arsenic. Therefore, the number of adsorption sites available for arsenic is reduced, and capacities of only a few hundred bed volumes are reached [28] in the presence of significant sulfate ion concentrations. Thus frequent regeneration is necessary. Competition by sulfate imposes the additional risk of arsenic release from the resin and accumulation in the effluent water (chromatographic peaking effect). Due to these restrictions, ion exchange has been considered less viable for drinking water treatment.

However, ion exchange is becoming more applied and accepted in the U.S. for general potable water treatment. Some research has been conducted to overcome the shortcomings of ion exchange by recycling the used regenerant (brine recycling) [29]. Also, recent work has focused on MIEX (a magnetically impregnated resin) in slurry reactors as an alternative to fixed bed contactors [30].

3.5. Redox reactive adsorbents

There has been limited work on redox-reactive adsorbents such as, for example, manganese dioxide, MnO_2, which can function as both an oxidant, oxidizing As(III) to As(V), and an adsorbent [31].

3.6. Alternative adsorbents

A recent literature survey [24] found that a large number of adsorbents have been tested for both As(V) and As(III) removals, including both commercially available and developmental (experimental) materials. Many of these media fall within the general category of metal (hydro)xides, with iron oxides being very common. Another common group of materials are surface-modified natural materials, often in the form of a coating layer on the surface of an inert media with an iron oxide coating, or impregnation. The major interferants have been shown to be phosphate, silica, carbonate, sulfate, and fluoride. pH has been shown in many cases to exhibit a major effect on media performance.

Among the commercially available media are MIEX resins and sulfur modified iron (SMI), while developmental media include iron oxide coated sand (IOCS). MIEX has been studied by Sinha et al. [30], and shown to have As(V) removal efficiencies similar to traditional strong base anion (SBA) resins, with sulfate having the most influential impact on adsorption capacity. Sulfur modified iron (SMI), with a metallic iron and elemental sulfur composition, has been studied by Sinha et al. [30] with effective removal of both As(V) and As(III) observed but with some operational problems including iron dissolution, suggesting that it is also a redox-reactive medium. Other commercially available media include iron-modified (impregnated) activated alumina (AA) and granular activated carbon (GAC) and zeolites. IOCS, sand coated with ferric hydroxide, has been successfully employed in fixed bed reactors to treat wastes containing various dissolved metal species, including arsenic [31]; Benjamin et al. [32] reported that arsenite binds extremely strongly to IOCS.

3.7. Regeneration

Media regeneration often involves caustic soda (NaOH) for metal oxide based adsorbents and salt (NaCl) for ion exchange media. In both cases, a liquid residual waste stream is produced, requiring treatment and/or disposal (see subsequent section). In the case of metal oxide adsorbents, there is interest in disposal of throwaway high-capacity media as spent adsorbents. In ion exchange, there is interest in brine recycle as a means of reducing the volume of waste brine. There has been some work on coagulation of ion exchange brines.

4. ARSENIC REMOVAL BY MEMBRANE PROCESSES

4.1. High pressure membranes

A study by Hering and Elimelech [33] concluded that both RO membranes and "tight" NF membranes could effectively remove As(III) and As(V) from synthetic and natural waters. The authors noticed no difference in the rejection rates of As(III) and As(V). It was also shown that co-occurring inorganic solutes affected arsenic rejections slightly. Finally, the authors concluded that the size of arsenic species controlled their separation behavior in RO membranes and tight NF membranes.

Recent results from bench and pilot studies [15] indicate that arsenic speciation, membrane surface charge, and source water quality are important variables in controlling the ability of a high-pressure membrane to reject arsenic. Within the range of conditions tested in

this study, flux, recovery, and initial As(V) concentration had little effect on arsenic rejection. Based on the significant number of RO and NF membranes tested, rejection of As(III) was consistently lower than As(V). Summarizing the arsenic rejections observed, As(V) rejections ranged between 85% and 99% for the NF and RO, while As(III) rejections ranged between 5% to 87%. Considering the RO elements only, As(V) rejections ranged from 95% to 99% while As(III) rejections ranged from 61% to 87%. The membrane with the highest As(III) rejection of all the membranes tested was the 'tightest' RO membrane, having the highest operating pressure and lowest specific flux of any membrane tested. Membrane charge was found to be a more important factor than membrane class (RO or NF) or molecular weight cutoff (MWCO) in determining the ability of a membrane to reject As(V). All of the membranes tested exhibited a negative surface charge over the pH conditions tested. The negatively charged NF membranes, all possessing a MWCO greater than the molecular weight of arsenate, consistently demonstrated substantial As(V) rejection rates. An important mechanism of As(V) rejection was postulated to be electrostatic repulsion between the membrane and anionic arsenate molecules. The study concluded that NF was the membrane of choice for As(V) removal while As(III) removal requires RO unless a pre-oxidation step is implemented.

4.2. Low pressure membranes

Coagulation/microfiltration technology (C-MF) has been recognized by the USEPA as an emerging technology for the removal of arsenic from drinking water. Therefore, pilot tests have been conducted at several locations in the U.S. [15, 34, 35], confirming that iron coagulation – direct microfiltration is a feasible method for arsenic removal from ground waters possessing natural arsenic concentrations up to 100 $\mu g\ L^{-1}$. Unlike conventional coagulation/filtration, this process does not require a flocculation step to create a larger particle, since the size of the formed $Fe(OH)_3$ particles needs only to be larger than the pore size of the microfilter, i.e. 0.1-0.2 μm. As a consequence, only rapid mixing is required, generating particles in the size range of 2-10 μm. A mixing intensity of around 1000 s^{-1} and a minimum detention time of 20 s is reported [34]. Regarding feed flux, it was observed that C-MF can be operated at high flux rates (> 153 $L\ m^{-2}\ h^{-1}$). To reduce treatment costs, membrane flux can be optimized by reducing the coagulant dose because lower solids loading on the membrane results in higher membrane flux rates. Thus, optimization may include pH pretreatment to lower the required coagulant dose and the resulting solid formation.

Microfiltration, coupled with low dose ferric chloride pretreatment, was demonstrated as a technically feasible method for arsenic removal from a ground water possessing a natural arsenic concentration of approximately 18 $\mu g\ L^{-1}$ during a pilot test [15]. Coagulant doses from 2 to 10 mg l^{-1} as $FeCl_3$, coupled with the 0.2 micron microfilter, resulted in arsenic rejections ranging from 39% to 89%. The system demonstrated the capability to meet a target 5 $\mu g\ L^{-1}$ arsenic (1/2 of the U.S. MCL) over a two week period at a system recovery of > 90%.

Similarly, another study with feedwater containing 40 $\mu g\ L^{-1}$ As showed reliable removal down to 2 $\mu g\ L^{-1}$ As [35]. Arsenic removal was strongly controlled by pH, with lower pH associated with better removals and lower coagulant doses required. The authors distinguish a low iron-low pH approach (Fe dose = 2.8 mg L^{-1}) and a high iron-high pH approach (Fe dose = 7.0 mg L^{-1}, pH = 7.3) to achieve low filtrate As concentrations. The choice of which options to use will depend on the relative cost and complexity of additional pH pretreatment for the low iron-low pH approach or the increased cost of ferric coagulant use and higher sludge production for the high iron-high pH approach. Residuals from the C-

MF process can be classified as nonhazardous and have been shown to pass the toxicity characteristic leaching procedure (TCLP) [35].

Decreasing costs for microfiltration membranes and low operation and maintenance requirements make C-MF more competitive with other technologies for arsenic removal, especially with granular iron media favored for small arsenic removal systems [34].

Sinha et al. [30] studied hydrous iron oxide particles (HIOPs) as a potential adsorbent for use in a hybrid (coupled) HIOPs-UF membrane system. The HIOPs adsorbent, a semi-crystaline iron oxide identified as predominantly hematite, achieved adsorption capacities similar to granular iron oxide media (see previous sections).

5. PRE-OXIDATION

5.1. Chemicals

The effective removal of arsenic from water usually requires the complete oxidation of As(III) to achieve the required drinking water standard. There are various means of oxidation available, but in drinking water treatment, additional aspects must be considered, such as the limited list of chemicals, residual oxidants and oxidation by-products or the oxidation of other inorganic and organic water constituents. Thus, the national drinking water and treatment regulations will therefore be important in the selection of the most feasible oxidant.

The need for a special oxidation step is caused by the very slow kinetics of oxidation by dissolved oxygen [2]. Oxygen would be a preferred oxidant as some problems with other chemicals can be avoided. There are also technical and operational problems whenever chemicals are added continuously in small-scale water treatment plants, where operators are not available all the time. Thus, a pre-oxidation step adds operational complexity.

Effective oxidants include free chlorine, hypochlorite, ozone, permanganate and hydrogen peroxide/Fe^{2+} (Fenton's reagent), but not chloramines [20] and peroxosulfate. Chlorine is widely used for this purpose, but may lead to chlorination by-products, namely trihalomethanes (THM's), from reactions with natural organic matter (NOM). In Germany, the application of chlorine is allowed only for disinfection purposes, not for general oxidation.

Ozone, widely used in surface water treatment for oxidation and disinfection, is very effective, but not feasible for As(III) oxidation. It must be produced on-site from air or oxygen and is a strong oxidant, which would also partly oxidize the NOM, an unwanted reaction in cases of geogenic arsenic in groundwater.

The most feasible chemical oxidants found to date are potassium permanganate and Fenton's reagent (H_2O_2/Fe^{2+}), if the removal of the As(V) is to be accomplished by precipitation/coagulation and rapid filtration. Permanganate as a weak oxidant oxidizes As(III), ferrous and manganese ions specifically and rapidly. The Mn(IV)-hydroxide formed must be filtered afterwards, together with the precipitated As(V). The presence of Fe(II) and Mn(II) leads to a higher consumption of $KMnO_4$. A dose of approximately 1 – 2 mol $KMnO_4$ per mol As(III) is usually needed. It was found that the DOC concentration has no significant influence on the dose necessary. Thus, problems regarding the formation of biodegradable organic carbon should not be expected [17].

The singular application of H_2O_2 may be successful, if the raw water contains already sufficient ferrous ions to induce OH-radical formation. The mass ratio of H_2O_2 to Fe^{2+} should be varied to find the optimum for a maximum oxidation. The dosages required are usually in the range of 0.3 - 2 mg L^{-1}. The ferric ions formed will precipitate As(V). The residual H_2O_2 may not exceed 0.1 mg L^{-1} (German standard), which can be achieved by careful dosage

adjustment. High concentrations of bicarbonate reduce the oxidation of As(III) by scavenging of the OH-radicals. The use of H_2O_2 may also lead to a partial oxidation of NOM.

In comparing these chemical oxidation processes, preference may be given to $KMnO_4$ in the case where there is a subsequent precipitation/filtration step and to chlorine or hypochlorite, if regulations permit.

5.2. Biological oxidation

The biological oxidation of As(III) in fixed-bed reactors represents a simple treatment process which does not require additional chemicals. Investigation of the feasibility of this process as a pre-treatment for arsenic removal showed a sustainable and mainly biological oxidation using common filter material such as sand or pumice. After adaptation periods of a few days to a few weeks, fixed-bed filters operated at filter rates of 25 m h^{-1} and fed with As(III) concentrations of over 100 µg L^{-1} showed high oxidation rates and reliably oxidized As(III) down to effluent concentrations of 2 µg L^{-1} [36]. Such a biological filter is generally insensitive to changes in the operational conditions. However, backwash events will reduce the oxidation rate for a short period of time. This technique is particularly suited for the treatment of ground waters where arsenic removal is the only treatment goal. These investigations suggest that biological oxidation also occurs in other filter media, e.g. in iron removal plants. The possibility of biological arsenic oxidation in GFH / activated alumina fixed-bed reactors may represent an interesting focus of research aiming at the oxidation and removal of As(III) in one reactor only [36].

6. RESIDUALS MANAGEMENT

There are several potential methods of disposal for residuals derived from arsenic treatment technologies. Depending on the form of the waste (solid or liquid), options include: disposal in a sanitary landfill, disposal in a secure hazardous waste landfill, discharge to a sanitary sewer (conveying the liquid waste to a wastewater treatment plant (WWTP)), land application/disposal (e.g., evaporation ponds), and deep well injection into a non-potable aquifer [15].

Optimization of conventional treatment technologies leads to an arsenic-enriched sludge which, after dewatering, represents a solid residual. There are two general approaches for dealing with residuals from arsenic adsorptive treatment: (i) disposal of *throwaway* high-capacity adsorbents (e.g., GFH) or (ii) treatment and disposal of process (e.g., membranes) or regeneration brines (e.g., NaCl from ion exchange). In the former case, the residual is a solid, while in the latter case, the residual is a liquid. The most likely disposal method for solid residuals, dewatered sludges or spent adsorbents, is landfilling. If the solid waste material passes a hazardous waste test (e.g., TCLP (toxicity characteristics leaching procedure) or WET (waste extraction procedure)), a sanitary landfill can be used, otherwise, a more costly secure hazardous waste landfill with liners is necessary. An attribute of regenerable adsorbents (e.g., ion exchange) and high pressure membranes (nanofiltration or reverse osmosis) is that they are effective volume reduction/contaminant concentration processes; for example, RO treatment results in a brine of 20 % of the feed water containing virtually all of the arsenic. Depending on local regulations, brine disposal may be possible (e.g., sanitary sewer discharge or deep well injection) while, in other cases, brine treatment may be necessary in which the liquid form of the arsenic residuals is converted to a solid form utilizing precipitation techniques. For example, iron coagulation of an ion exchange brine regeneration effluent will promote adsorption of arsenic on ferric or aluminium hydroxide

flocs, leading to a sludge enriched in arsenic which, upon dewatering, can be disposed of as a solid waste. While coagulation of lower volume, higher concentration waste streams are generally more cost-effective, the effects of ionic strength on coagulation chemistry must be considered. Since the disposal of a residual in a landfill requires that the residual does not have any free water, conventional-treatment sludge must first be dewatered before disposal as solid residuals. The nanofiltration/reverse osmosis process, a technology that produces a reject stream with a more dilute arsenic concentration compared to ion exchange brines, is the only technology that has a possibility of being discharged to a sanitary sewer. However, the liquid residuals generated from an arsenic treatment facility may be classified as an industrial waste since it contains contaminants which may impact the WWTP; pretreatment regulations may constrain this option. In considering evaporation ponds, significantly high evaporation rates will be necessary and regulations may require that evaporation ponds be lined to protect underlying groundwater; moreover, periodic physical removal of salts may be necessary. Deep-well injection must ensure that the aquifer is non-potable and that well construction protects overlying potable aquifers.

A key question pertaining to solid residuals is their potential classification as hazardous wastes. In North America, there are two tests commonly applied as a definition of hazardous wastes: the toxicity leaching characteristics procedure (TCLP) and the waste extraction test (WET) [15]. Both tests attempt to simulate landfill leaching conditions, with the latter test being more aggressive. The specification for a hazardous waste is that the leachate arsenic concentration is 100 times the drinking water standard.

7. APPROPRIATE TECHNOLOGIES FOR DEVELOPING COUNTRIES

Many of the technologies discussed within this chapter are oriented toward developed countries. However, there are many problems with high levels of arsenic in developing countries and countries in transition (e.g., Bangladesh and India). The criteria for selection of appropriate technologies include low-cost, ease of operation, possibility of decentralized application, and sustainability. For developing countries, membranes are not economically feasible, pre-oxidation adds operational complexity, and many problems exist at the community (small systems) level. In contrast, high-capacity iron-oxide based adsorbents, capable of adsorbing As(V) and As(III), are attractive. Within this context, such adsorbents would be used as throwaway adsorbents if they passed the appropriate leaching (e.g., TCLP) test.

REFERENCES

[1] J.F. Ferguson and J. Gavis, Water Res., 6 (1972) 1259.
[2] D.A. Clifford, L. Ceber and S. Chow, Arsenic(III)/Arsenic(V) separation by chloride-form ion exchange resins, XI, Proc Am. Waterworks Assoc. Water Qual. Technol. Conf, Norfolk, VA, 1983.
[3] G. Baldauf, gwf Wasser Special, 136 (1995) 99.
[4] WHO, United Nations Synthesis Report on Arsenic in Drinking Water, World Health Organization, 2001.
[5] A. Ruhland, K. Karschunke and M. Jekel, bbr Fachmagazin für Brunnen- und Leitungsbau, 3 (2003) 53.
[6] A.H. Welch, S.A. Watkins, D.R. Helsel and M.F. Focazio, U.S. Geological Survey Fact Sheet, 063-00, 2000.
[7] WHO, Guidelines for drinking-water quality, World Health Organization, 1993.

[8] C.F. Baes and R.E. Mesmer, The Hydrolysis of Cations, John Wiley & Sons, Inc., New York, 1976.
[9] E. Bellack, J. - Am. Water Works Assoc., 63 (1971) 454.
[10] J.H. Gulledge and J.T. O'Connor, J. - Am. Water Works Assoc., 65 (1973) 548.
[11] Y.S. Shen, J. - Am. Water Works Assoc., 65 (1973) 543.
[12] T.J. Sorg and G.S. Logsdon, J. - Am. Water Works Assoc., 70 (1978) 379.
[13] M. Jekel, Removal of trace elements during drinking water treatment (in German), DVGW-Schriftenreihe Wasser, (1986) 259-272.
[14] M. Jekel, Coagulation processes for arsenic removal, in M. Jekel (Ed.), Arsenic in Drinking Water Treatment (in German), DVGW-Schriftenreihe Wasser, Eschborn, 1993.
[15] G. Amy, M. Edwards, M. Benjamin, K. Carlson, J. Chwirka, P. Brandhuber, L. McNeill and F. Vagliasindi, Arsenic Treatability Options and Evaluation of Residuals Management Issues, AWWA Research Foundation, Denver, 2000.
[16] L.S. McNeill and M. Edwards, J. - Am. Water Works Assoc., 89 (1997) 75.
[17] M. Borho, Arsenic Removal in Groundwater Treatment by an Optimized Combination of Oxidation and Coagulation Processes (in German). Institute of Water Quality Control and Waste Management. Garching, Technical University of Munich, 1996.
[18] US-EPA, Technologies and Costs for Removal of Arsenic from Drinking Water, prepared by International Consultants, Inc. and Malcolm Pirnie, Inc. under contract 68-C6-0039 for EPA OGWDW, 2000.
[19] E.R. Rosenblum and D.A. Clifford, The Equilibrium Arsenic Capacity of Activated Alumina, Summary rep. EPA-600/52-83-107, 1984.
[20] P. Frank and D.A. Clifford, Arsenic (III) Oxidation and Removal from Drinking Water, Summary report EPA/600/52-86/021, 1986.
[21] U. Hildebrandt, Investigations of Adsorption of Arsenate(V) on Activated Alumina (in German), Technical University of Berlin, 2000.
[22] U. Hildebrandt and G. Hölzel, Arsenic Removal during Drinking Water Treatment (in German), Final report for BMBF, 1998.
[23] J.D. Chwirka, B.M. Thomson and J.M. Stomp, J. - Am. Water Works Assoc., 92 (2000) 79.
[24] G. Amy, H. Chen, P. Brandhuber, N. Graziano, U. von Gunten, Z. Chowdhury, S. Kommineni, K. Banerjee and M. Jekel, Impact of Water Quality Parameters on Adsorbent Treatment Technologies for Arsenic Removal, AWWA Research Foundation, Denver, 2004.
[25] W. Driehaus, Water Sci. Technol.: Water Supply, 2 (2002) 275.
[26] M. Jekel and R. Seith, Water Sci. Technol.: Water Supply, 18 (2000) 628.
[27] A. Ruhland and M. Jekel, Water Sci. Technol.: Water Supply, 2 (2002) 267.
[28] U. Hildebrandt and G. Hölzel, Wasser Abwasser, 138 (1997) 190.
[29] J. Kim, M.M. Benjamin, P. Kwan and Y.J. Chang, J. - Am. Water Works Assoc., 95 (2003) 77.
[30] S. Sinha, N. Lee, Y. Yoon, J. Yoon, G. Amy, Y.J. Chang and S. Reiber, Evaluating Emerging and Innovative Technologies for Arsenic Removal - A Bench Scale Study. Proc. Am. Waterworks Assoc. Water Qual. Technol. Conf., Seattle, USA. 2002
[31] K. Subramanian, T. Viraraghavan, S. Tanjore and T. Phommavong, Water Qual. Res. J. Can., 32 (1997) 551.
[32] M.M. Benjamin, R.S. Sletten, R.P. Bailey and T. Bennett, Water Res., 30 (1996) 2609.
[33] J. Hering and M. Elimelech, Arsenic Removal by Enhanced Coagulation and Membrane Processes, AWWA Research Foundation, Denver, 1996.
[34] J.D. Chwirka, C. Colvin, J.D. Gomez and P.A. Mueller, J. - Am. Water Works Assoc., 96 (2004) 106.
[35] G. Ghurye, D. Clifford and A. Tripp, J. - Am. Water Works Assoc., 96 (2004) 143.
[36] R. Seith and M. Jekel, Vom Wasser, 89 (1997) 283.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 12: *Cryptosporidium*/sand interactions during filtration

Robert F. Considine[a], David R. Dixon[b] and Calum J. Drummond[c]

[a]Cooperative Research Centre for Water Quality and Treatment, Bag 3, Salisbury, South Australia 5108, Australia

[b]Particulate Fluids Processing Centre, Dept of Chemical and Biomolecular Engineering, University of Melbourne, Parkville, Victoria 3010, Australia

[c]CSIRO Molecular and Health Technologies, Bag 10, Clayton South, Victoria 3169, Australia

1. INTRODUCTION

1.1. *Cryptosporidium* and drinking water

Surveillance data for *Cryptosporidea* indicate the prevalence of the parasite throughout the environment, and sources are probably present in every surface water catchment [1-3]. *Cryptosporidea* have been found in rivers and streams, lakes and reservoirs, raw and treated sewage, and treated surface waters. The parasite is increasingly regarded as an important cause of enteric disease [4] and several waterborne outbreaks of cryptosporidiosis have been documented. Outbreaks have occurred in the United Kingdom [5] as well as the United States [6-7] most notable of which was the outbreak in Milwaukee [8] where up to 400,000 people were infected and 100 deaths of immune-compromised patients were attributed to the disease. Occasional outbreaks have occurred with no measured changes in source water quality or treatment processes [9]. Such episodes, combined with the serious consequences of cryptosporidiosis outbreaks, have necessitated research on reliable processes to defend drinking water supplies from contamination with *Cryptosporidea.*

Transmission of cryptosporidiosis occurs via the faecal-oral route, and involves the ingestion of oocysts [10, 11]. Oocysts are biological globules 3-7 μm in diameter, and consist of a biologically active core surrounded by a thick, environmentally stable coating. Contact with oocysts can occur from any faecally contaminated medium, including surfaces, and contaminated food or water supplies [12]. A significant consequence of the thick coating is the resistance of the oocysts to conventional disinfection by chlorination processes [13-14].

The application of multiple barriers in the provision of safe drinking water is advocated by the World Health Organization's Guidelines for Drinking Water Quality [15]. This concept involves the establishment of several barriers to water quality hazards. Protecting surface waters from faecal contamination, understanding and managing the fate of contaminants through waterways and storages, removing or inactivating *Cryptosporidea* by water treatment and protecting the distribution system from contamination are all important stages in the provision of drinking water safe from the *Cryptosporidium* risk. The resistance to chlorination, coupled with the ubiquity of the oocysts in the environment however, reinforces the importance of physical removal of oocysts by sand-bed filtration to ensure safe drinking water at the point of treatment [16, 17].

1.2. DLVO theory of surface forces in liquids

The ubiquity of systems that involve surface interactions in liquids, combined with the need to understand the corresponding forces of interaction, has driven theoretical developments in the field. A common approach has been to build on the significant contribution of Derjaguin, Landau, Verwey, and Overbeek who reported the force of interaction between surfaces in liquids as a sum of the van der Waals and electrical double layer interactions in the 1940s [18, 19]. This fundamental approach, known as DLVO theory, has been widely applied in order to understand the stability and rheological properties of a vast range of colloidal systems. The original DLVO calculations were based on the interaction between incompressible and uniformly charged surfaces of ideal geometry (typically infinite flat plates) interacting in electrolytes consisting of point charges dispersed in a continuum phase (such as water). Validation of the theoretical approach has been obtained experimentally for the interaction between atomically smooth mica sheets in dilute monovalent electrolyte solutions [20-25].

DLVO Theory accounts for the interaction between two surfaces that are dispersed in a continuum phase, as they approach one another. The theory assumes that the total energy of interaction can be computed as the sum of two additive interactions, ie, the van der Waals and the electrical double layer interactions. The DLVO force of interaction ($F_{sphere\text{-}sphere}$) between two particles (of inter-surface separation H) is related to the interaction energy between two flat plates ($W_{flat\text{-}flat}$) by the Derjaguin approximation:

$$\frac{F_{sphere-sphere}}{R} = 2\pi W_{flat-flat} \tag{1}$$

Where R is the harmonic mean of the radii of the two spheres. Extensive reviews of DLVO theory is provided in various colloid and surface chemistry textbooks [26, 27] and will not be reviewed further here.

1.3. Chapter scope

In our work, an atomic force microscope (AFM) has been used to measure the surface forces of interaction between *Cryptosporidea* and sand. Originally developed as a surface-imaging tool, the AFM has been used to measure the surface properties of polymers, semiconductors, and biomolecules [28]. The capacity of the AFM to measure surface forces as well as topography is also well established [29]. With minor modification, the AFM can be used to measure the surface forces between dissimilar colloidal particles in solution, using the 'colloid-probe' approach [30, 31]. The surface forces of interaction between individual *Cryptosporidea* and sand particles in aqueous solutions of dissolved calcium and dissolved organic carbon is reported. The surface force measurements have been compared and contrasted with the corresponding surface charges obtained from electrophoretic measurements. The measurements are interpreted in terms of a model of the *Cryptosporidium* surface, which includes both physical and electrical components. The investigations demonstrate the importance of surface forces in *Cryptosporidium*/sand interactions.

The model sand surfaces used were silica colloids manufactured from soda lime glass with a reported density of 2.48 $g.cm^{-3}$ (cf 2.32 $g.cm^{-3}$ for cristobalite SiO_2) [32]. The isoelectric point (iep, defined as the pH at which the surface has an averaged zero zeta potential) and ζ-potentials as a function of pH (see below) are consistent with earlier reports of silica surfaces [33]. The oocysts were sourced from Waterborne Inc (Louisiana, USA), isolated from experimentally infected calves' faeces (originally 'Iowa' isolate source) by

sucrose and Percoll density gradient centrifugation, identified by direct immunofluorescence microscopy with genus-specific monoclonal antibodies, and supplied without surface treatment at a concentration of around 1 x 10^6 oocysts / mL in phosphate buffered saline (PBS) containing penicillin, streptomycin, and gentamicin. The source of oocysts has been compared to others by electrophoresis and found to have surface charge properties typical of un-altered oocysts [34].

The procedure for preparing the oocysts for force versus separation measurements involved anchoring the underside of oocysts to coated silicon wafers, using a localised fixative process. The smooth silicon wafer (Silica Source Technology Corp) was exposed to a plasma deposition (power = 20 W, frequency = 200 kHz) of 3-amino-propylene (time = 25 seconds, pressure = 0.125 mm Hg). The prepared surface was then incubated with a 0.5% w/w solution of glutaric dialdehyde (Aldrich Chemical Company, Inc., obtained as 25% w/w glutaric dialdehyde in water) for approximately 1 hour. Excess glutaric dialdehyde was removed by rinsing with copious amounts of Milli-Q water. A drop of the oocyst dispersion was then placed on the glutaric dialdehyde surface and left to dry in an evacuated environment (25 mmHg), and the dried oocysts were then washed with Milli-Q water. The minimum drying time (12 hrs) was used, and comparison of the oocyst-silica force of interaction showed no significant change over forty hours, suggesting that the surface is stable over the course of an experiment. Comparison of glutaric dialdehyde/silica and oocyst/silica force curves confirmed that the glutaric dialdehyde film did not contaminate the scanned region of the oocyst.

2. SURFACE MEASURMENT METHODS

2.1. Atomic force microscopy

The general operation of the AFM has been reviewed in detail elsewhere [28-29] and will be only briefly described here. The essential components of the AFM are a piezo-electric ceramic, force sensor (cantilever spring), laser/photodiode assembly, and a feedback system. AFM measurements can be performed either in various liquids (such as water) or gases (such as air) and all measurement reported here have been made in liquids. The silicon wafer, on which the *Cryptosporidium* oocyst was mounted, rests on a piezo-electric ceramic, the position of which has been controlled to sub-Angstrom resolution by the application of a voltage bias across the piezo. The force sensor, which when used for topographical measurements typically consists of a cantilever spring with integrated pyramidal tip, is held immediately above the sample.

In the case of topography measurements, the AFM has been used in *constant deflection* mode, and the imaging process can be summarised as:

- Voltage bias applied across piezo-electric ceramic and ceramic expands;
- Cantilever spring senses sample (by repulsive electrical double layer interactions [35]) and the cantilever undergoes positive vertical deflection, ie repulsion;
- Feedback system adjusts piezo-electric ceramic to position the sample in order to maintain constant deflection.

In this way, the movement of the piezo-electric ceramic can be used to obtain sample height information. In the case of force versus separation measurements, the force sensor and laser/photodiode assembly is used to detect the position of the second interacting surface, which has been a colloidal silica particle attached to the cantilever spring. The measurement process can be summarised as follows:

- Voltage bias applied across piezo-electric ceramic and ceramic expands;
- Sample moves towards the force sensor; and
- Force sensor deflects as separation is reduced, and deflection can be equated to force via Hooke's law ($F = -K_sH$, where K_s is the cantilever spring constant).

Once the surfaces are in contact, the bias on the piezo can be reversed, and the force as the two surfaces are separated (ie adhesion) can be measured. A schematic diagram of the AFM set-up to measure forces is presented in Fig. 1.

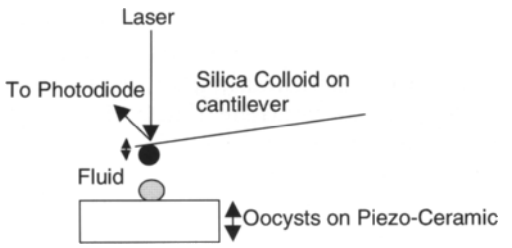


Fig. 1. Schematic drawing of the AFM operation for force measurements between *Cryptosporidium* oocysts and a model sand surface

2.2. Particle micro-electrophoresis

A RANK Bros Mark II particle micro-electrophoresis apparatus was used to obtain the micro-electrophoretic mobility data. The basic operation of the RANK Bros Mark II apparatus is summarised here and more detailed information can be obtained in standard colloid texts [36]. A schematic diagram of the apparatus is presented in Fig. 2. The measurement process involves a rectangular glass cell in which the colloids are dispersed in the electrolyte of interest. Platinum black electrodes are located at either end of the cell, through which the electrical field is applied. A microscope objective is positioned such that the movement of individual particles within the cell can be monitored in the applied field. In this way, the velocity of the particles in an applied field can be measured. The location of the microscope focal plane within the electrophoresis cell is critical due to the hydrodynamic flow that develops within the cell when the electrical field is applied. The hydrodynamic flows are a result of counter-ion association with the cell walls (streaming) and reverse flow at hydrostatic equilibrium that obeys Poiseuille's law. The combination of these flows in the rectangular cell used in this chapter means that the liquid itself is only stationary at two distinct stationary planes located within the cell.

Mobilities were converted to the electrophoretic (ζ) potential by the Helmholtz-Smolouchowski equation which is valid for large values of κR (where κ^{-1} is the Debye length and R is the particle radius). In the measurements reported here, the velocity of particles in applied electric fields ranging from 50 - 100 V was measured. At each pH, the particle velocity was measured at both stationary planes and with the electric field reversed, involving in excess of 40 measurements, from which the times were averaged.

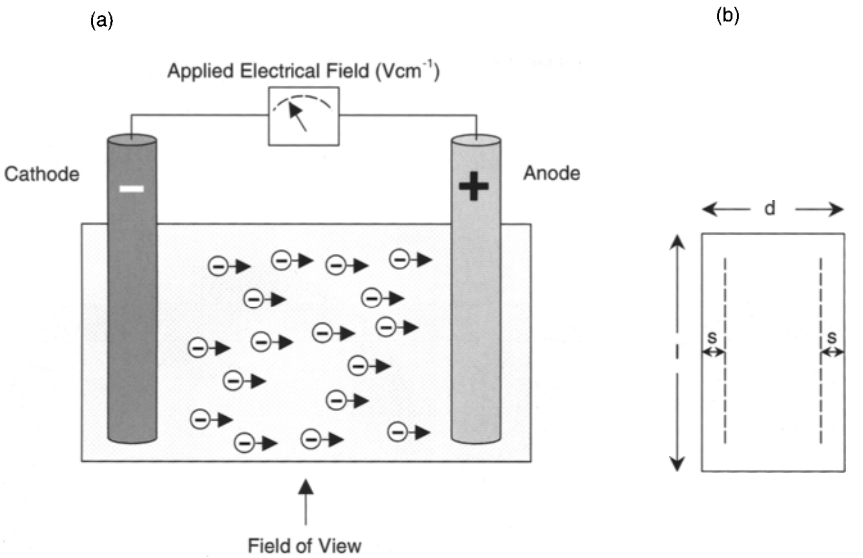


Fig. 2. Schematic representation of particle micro-electrophoresis (a) as viewed through the optical microscope for the case of anionic colloids and (b) rectangular cell geometry (side view) with *d* being the internal cell width, *l* the internal cell height and *s* being the location of the stationary planes within the cell

3. SURFACE DATA

The surface chemistry of silica particles in water is dominated by the presence of silanol groups which are weakly acidic and largely responsible for the very low iep values. In contrast the oocyst surface is believed to be composed of a complex protein mixture in which there is sufficient charge repulsion between ionisable surface groups to stretch anchored surface proteins into solutions, thereby creating a brush-like conformation.

3.1.Topography

The topographical features of both the silica spheres and oocysts were measured in aqueous electrolyte using the AFM. Images of the silica spheres and oocysts have been presented in Fig. 3.

The silica spheres have a root-mean-square (RMS) roughness of around 1-2 nanometres (nm) for a 2 x 2 micron (μm), *xy* plane fitted scan. The roughness is distributed as gentle undulations of a few hundred nm in the lateral dimensions and around 10 to 20 nm in the vertical dimension.

In contrast, the surface topography of the oocysts is substantially rougher with an RMS roughness of a 2 x 2 μm, *xy* plane fitted scan of around 5 nm with a peak to valley height of around 50 nm across the scan. Rather than gentle undulations, the oocyst roughness appears as discrete asperities, ranging in size from around 50 nm to 250 nm in the lateral dimensions. The finding that the oocysts are relatively rough is in accord with earlier observations [37] where oocysts were found to possess an RMS of 17nm across a 2 μm, *xy* plane fitted scan.

The variation in precise roughness values may be indicative of the variation amongst oocysts. The presence of a 'suture' (or fold) in the oocyst surface, and its purpose in the protozoan lifecycle, has received some attention in earlier transmission electron and optical microscopy studies. Occasionally the oocyst sutures were clearly visible although this was not always the case. This may indicate that the oocyst deposition is random, and not oocyst surface site specific (ie oocysts attach to the glutaric dialdehyde suture-side-up as readily as suture-side-down).

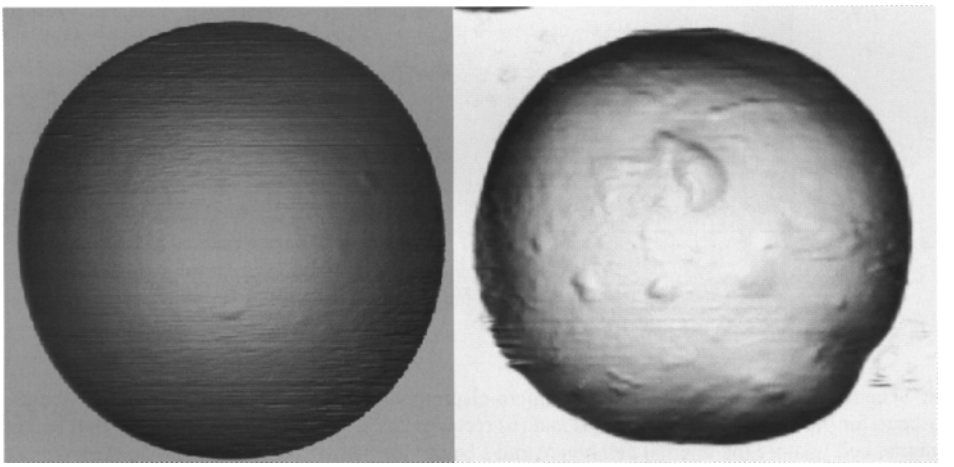


Fig. 3. Surface topography of the interacting surfaces. The model sand surface (left) and the *Cryptosporidium* oocyst surface (right) were measured in aqueous electrolyte and on horizontal scans of around 3-4 μm

3.2. Particle micro-electrophoresis

3.2.1. Monovalent electrolyte

The particle micro-electrophoretic mobility of the silica spheres and oocysts were measured as a function of pH at ionic strengths of 1, 2 and 5 mM KNO_3 (Fig. 4, where the mobility has been converted to a ζ-potential). The silica spheres exhibit an isoelectric point (*iep*) lower than pH 3, and a plateau ζ-potential of around -(80-90) mV at high pH, which is typical of colloidal silica [38]. It has been shown elsewhere [36] that the influence of increasing ionic strength on the ζ-potential for most inorganic oxides is to decrease the absolute magnitude of the ζ-potential. Further, for the mono-valent electrolyte used here (KNO_3) little specific absorption (ie shift in the *iep*) is expected to occur. The large standard deviation among the total of forty measurements for each pH (error bars in Fig. 4) make trends in the ζ-potential due to the influence of ionic strength difficult to distinguish, other than that the *iep* appears stationary as the ionic strength is increased. The oocysts also exhibit an *iep* less than pH 3 and the potential at high pH is only around –(15-18) mV. The origin of the surface charge is likely to be in the form of ionisable carboxylic acid groups incorporated in oocyst surface proteins as described by Ranucci *et al.* [39]. Increasing the ionic strength has been found to decrease the ζ-potential without altering the *iep*, consistent with neither the potassium or nitrate ions being potential determining. It should be noted, however, that the surface force measurements

(see below) indicate that the surface of the oocysts cannot be described purely in terms of electrical double layer theory, which may have implications for the conversion of mobility to ζ-potential. Nevertheless, the measured ζ-potential maintains the consistent definition as the potential at the plain of shear.

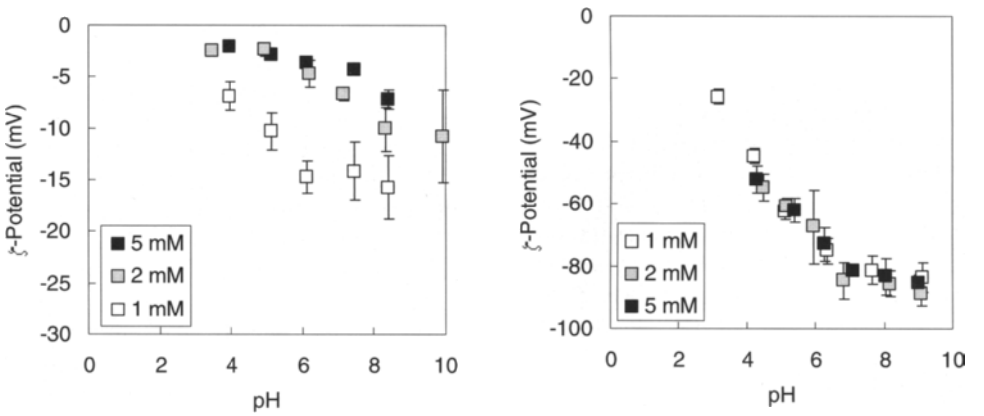


Fig. 4. ζ-potential, derived from mobility data, for the silica particles (left) and oocysts (right) as a function of pH in 1, 2 and 5 mM KNO_3 solutions (open, grey and filled symbols, respectively). The error bars correspond to the standard deviation for the total of forty velocities, measured under an applied field, at each solution condition

3.2.2. Divalent electrolyte

The electrophoretic mobility, in terms of the ζ-potential, was measured in waters of varying hardness (covering the range of values commonly encountered in raw water supplies), and is presented in Fig. 5. The effect of increasing ζ-potential with increasing ionic strength is certainly apparent at high pH in Fig. 5. Also evident in Fig. 5 is a trend towards an increasing *iep* with increasing calcium concentration. This observation is consistent with specific surface adsorption of Ca^{2+} at the oxide-water interface (ie Ca^{2+} is a potential determining ion), and similar conclusions for various other colloidal dispersions in the presence of calcium-based electrolytes have been reported [36,39, 40]. It is also apparent that the ζ-potential of oocysts becomes less negative with increasing calcium concentration (Fig. 5) and that the *iep* also increases with increasing calcium ion concentration. The increase of the *iep* is consistent with specific surface adsorption, probably via calcium ion / surface carboxylate complexation, as has been observed on carboxylic acid terminated latex colloids [36].

3.2.3. Natural organic matter

The microelectrophoretic mobility data in electrolyte solutions containing natural organic matter (NOM) were also measured and are presented in Fig. 6. For the silica colloids, the ζ-potential are spread around values of -(70-75) mV at high pH, with no systematic correlation with increasing NOM being evident (the amount of NOM has been quantified at total organic carbon, TOC). At low pH, the influence of NOM on the electrokinetic data is hardly distinguishable. The slight decrease in absolute magnitude of the ζ-potential at high pH is consistent with the adsorption of NOM at the silica-solution interface leading to a decrease in

the net charge at the shear plane. Buleva and Petkanchin [41] concluded, based on an electro-optical study, that a specific interaction promoted the adsorption of humic substances onto silica surfaces. Certainly the adsorption of NOM onto other oxides such as iron and alumina has been well established [42-45]. It should also be noted that the force measurements (see below) are consistent with the NOM solutions being of higher net ionic strength compared to the 1 mM KNO_3 solution. Therefore, the contribution of NOM to the net ionic strength may contribute to the slight reduction in ζ-potential at high pH, as observed for other oxides with increasing ionic strength [36]. In any event, the influence of the NOM is only very slight, and given the absence of a systematic trend with increasing TOC, it is difficult to make conclusive statements based on the electrokinetic data alone. Further discussion of the ζ-potential of the silica colloids in the NOM solutions is provided below in the context of the AFM data.

On the other hand, the ζ-potential of oocysts has been found to become more negative in the presence of NOM, with an increase in the *iep* also being observed (Fig. 6). The increase in the *iep* and the absolute magnitude of the ζ-potential are consistent with the adsorption of NOM at the oocyst surface. It may be that the adsorption of NOM molecules possessing a sufficiently high negative charge to increase the surface charge density at the plane of shear proceeds due to a non-electrical attraction and despite the anticipated electrostatic repulsion. This hypothesis is supported by similar interpretations drawn from adsorption measurements of humic substances on bacteria [46] and other negatively charged colloids [41].

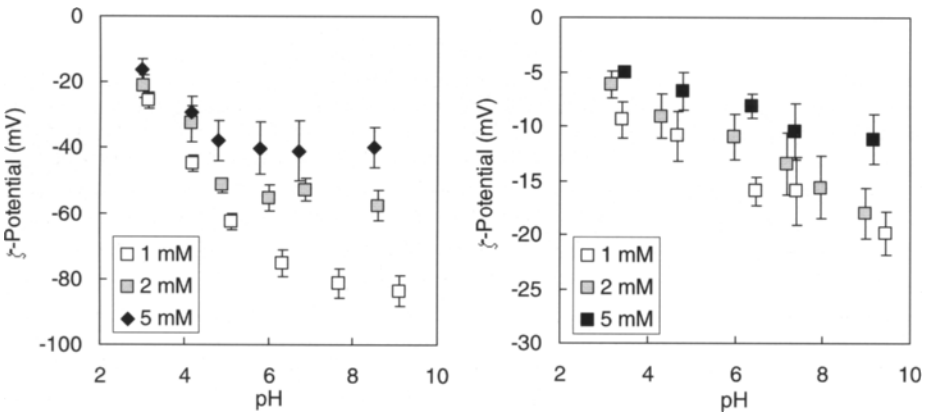


Fig. 5. Electrostatic (ζ-) potential of silica colloids (left) and oocysts (right) in potassium and calcium nitrate solutions of net ionic strengths 1, 2, and 5 mM, prepared as the following molarities: 1 mM KNO_3 (1 mM), 1 mM KNO_3 and 0.25 mM $Ca(NO_3)_2$ (2 mM), as well as 1 mM KNO_3 and 1.25 mM $Ca(NO_3)_2$ (5 mM). Error bars correspond to the standard deviation of values derived from at least forty mobility measurements

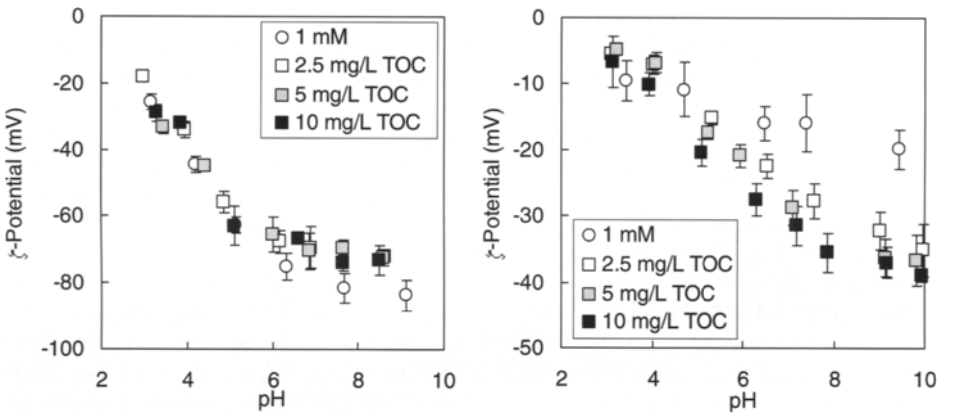


Fig. 6. Electrostatic (ζ-) potential of silica colloids (left) and oocysts (right) in 1 mM KNO_3, 1 mM KNO_3 and 2.5 mg L^{-1} TOC, 1 mM KNO_3 and 5 mg L^{-1} TOC, and 1 mM KNO_3 and 10 mg L^{-1} TOC. Error bars correspond to the standard deviation of values derived from at least forty mobility measurements

3.3. Surface forces on approach

The surface forces were measured as the separation between individual oocysts and silica surfaces were reduced. The force of interaction is presented in milli-Newtons (mN), and is normalised by the harmonic mean of the radii of the surfaces (m^{-1}), to facilitate comparison between particles of different radii. The normalised force, as a function of separation (nm) for the surface approach of an oocyst and a silica particle is presented in Fig. 7. The data consists of three distinct regions; at very long range the cantilever experiences zero force and so does not deflect with piezo movement, at intermediate range the cantilever experiences repulsive surface forces and deflects non-linearly with piezo movement, and after contact the cantilever deflects linearly with piezo movement. The calibration and interpretation of the force-separation data reported in this chapter has been described in detail elsewhere [47].

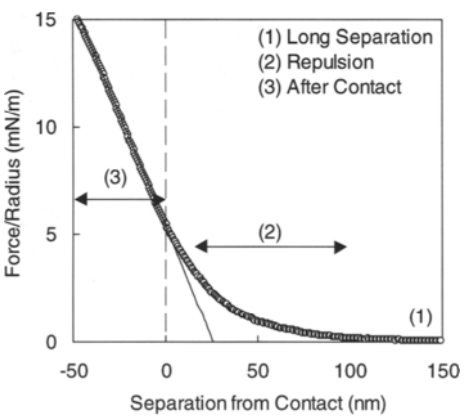


Fig. 7. Typical force-separation data for the approach of an oocyst and a silica particle in 1 mM KNO_3 at pH = 8.9

3.3.1. Origin of the repulsive surface force

All force-separation curves between oocysts and the model sand surface (silica) exhibited a repulsive force at separations prior to contact, on surface approach. In order to investigate the origin of the repulsive force, a conventional approach has been used to evoke theoretical models and obtain fits with experimental data [26]. In this chapter it will be shown that the oocyst surface is not well approximated by DLVO theory alone and that a steric force coexists with the electrical double layer force. Therefore, the repulsive force has been modelled using DLVO theory combined with a model of a polylectrolyte brush based principally on the developments of Pincus [48, 49].

Many theoretical models of the force of interaction between polymer-bearing surfaces have been proposed [48-55]. Calculations for charged polymers adsorbed at interfaces reveal a brush-like conformation that extends further with increasing charge density or decreasing salt concentration [54-55]. Neglecting chain stiffening and excluded volume effects, Pincus [48, 49] quantified the interaction forces in terms of the counter-ion osmotic pressure as a function of separation distance H to be:

$$\Pi \approx \frac{2fN_B kT}{d^2 H} \tag{2}$$

where f is the fraction of monomers carrying an ionic charge, N_B is the number of monomers within the hydrophilic block, d is the grafted interchain distance, k is the Boltzmann constant, and T is the temperature. Abraham et al. [56] showed that the energy of interaction between two polyelectrolyte brushes of individual width L_B can be given by:

$$F/R = \frac{4\pi fkTN_B}{d^2} \ln\left[\frac{2L_B}{H}\right] \tag{3}$$

The above equation corresponds to the interaction of two polymer bearing surfaces. However, the experimental situation is that the silica surface is well described by electrical double layer theory in 1 mM KNO_3 [57] and the brush is logically located on the oocyst surface. Therefore, the parameter $2L_B$ has been taken to be the brush width on the oocyst surface. Although the grafted inter-chain distance d is unknown it has been fixed at 2.31 nm for all calculations. This value was chosen since the oocyst wall has been shown to consist of predominantly proline, cystine, and histidine [39]; the largest of these amino acids is histidine, measuring 2.31 nm in the longest linear dimension predicted from the van der Waals radius of the constituent atoms. Although the value of d has been arbitrarily fixed, it is unlikely to vary as a function of pH, and does permit the investigation of other parameters that describe the steric force of interaction. The value of the temperature T has also been fixed in these calculations at 298 K. The remaining terms correspond to the fraction of ionicity, f, and the number of monomers within the hydrophilic block, N_B. Both of these dimensionless terms describe the chemical nature of the extended brush, and are expected to be related, that is, as f increases N_B will probably also increase, given the increase in charge on the polypeptide backbone. Therefore N_B and f have been combined to form $N_B.f$, where large values correspond to a brush that is composed of molecules that are highly charged and predominantly hydrophilic.

The measured normalised force versus separation data for the interaction between an oocyst and a silica sphere, in 1 mM KNO_3 at pH = 8.9, has been presented in Fig. 8. The DLVO fit has been presented as thin solid lines with the upper and lower limit both being

shown. The calculation has been made for the interaction of two surfaces governed by a Hamaker constant of 1 x 10^{-20} J and diffuse layer potentials corresponding to the measured ζ-potentials, immersed in the bulk electrolyte concentration of 1 mM KNO_3. In order to obtain an order of magnitude fit, the origin of the plane of charge in the DLVO calculation has had to be shifted around 35 nm from the point of contact. The data is seen to be well described by DLVO theory at separations > 35 nm with close agreement with the predicted decay length (9.5 nm). At separations < 35 nm but > 10 nm, the force of interaction is well described by the theory of Pincus (thick line). The calculation has been based on a brush width ($2L_B$) of 50 nm with a $N_B f$ of around 0.3, and all other parameters fixed. For comparison, the fraction of ionicity and number of monomers in the hydrophilic block has been estimated from the amino acid sequence provided by Ranucci et al. [39]. The constituent amino acids have been designated as ionisable (acidic or basic) and non-ionisable (polar or non-polar) according to conventional biochemical definitions [58]. Grouping the total number of characterised amino acids (1,252) in this way it was found that 273 were ionisable ($f = 0.2$) and that the molecule was predominantly hydrophilic (58%).

The overlay of the experimental data and theoretical curves presented in Fig. 8 permit the diagnosis of the origin of the exponential repulsive force. At separations greater than 35 nm, the interaction is well described by electrical double layer theory. At separations between 10 and 35 nm the interaction is well described as the collapse of a polyelectrolyte brush of width 50 nm. The implication from the theoretical modelling is that protein 'hairs' extend into solution from the surface of *Cryptosporidium*, and act as a significant repulsive barrier.

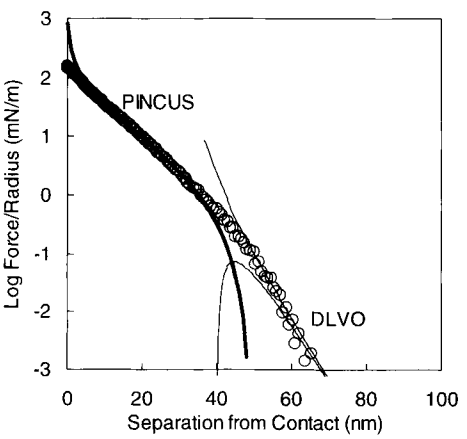


Fig. 8. Normalised force of interaction between an oocyst and a silica sphere (open symbols) in 1 mM KNO_3 at pH 8.9. The thin solid lines correspond to the result of the DLVO calculation based on the corresponding ζ-potentials and the bulk electrolyte concentration, with the DLVO plane of charge origin being shifted around 35 nm to obtain an order of magnitude fit. The solid line corresponds to the result of the force predicted from a Pincus model of a polyelectrolyte brush of width 50 nm

Not all oocysts exhibited force-separation curves that could be fit with DLVO and Pincus predictions in the same manor as for Fig. 8. The inability to obtain agreement with DLVO theory is probably a combined result of the significant surface roughness and extension of the oocyst surface proteins into the electrical double layer. Indeed, it is somewhat

surprising that frequent agreement with DLVO theory can be obtained at all, given the complexity of the surface. In order to make comparisons between all oocysts, the range of the repulsive force has been analysed in terms of the decay length (Fig. 9) and magnitude (Fig. 10).

3.3.2. Characterisation of the repulsive force in various water qualities

The characterisation of the repulsive force on surface approach by descriptive parameters aids in the assessment of the influence of experimental variables, in this case varying ionic strength (including hardness) and natural organic matter. To this end, the repulsive force on surface approach has been characterised by its decay length, or more precisely the log linear (natural log) slope of the force at separations between 20 and 50 nm from contact. The decay length values have been presented in Fig. 9, as a function of pH, calcium concentration and NOM.

It is apparent that a systematic decrease in the measured decay length occurs with an increase in ionic strength in both the KNO_3 and $Ca(NO_3)_2$ solutions. For comparison with the DLVO and Pincus theories, the power law dependence of the experimentally measured decay length ($1/\kappa_{exp}$) on increasing electrolyte concentration (C_S) can be obtained from the slope of a log-log plot of the decay length versus electrolyte concentration. Although only three electrolyte concentrations were assessed the quality of the power law fit was found to be quite high (well within the error associated with the decay length) with an r^2 value of 0.998 and gave the following power law dependence:

$$1/\kappa_{exp} \propto C_S^{-0.23} \tag{4}$$

The measured exponent is significantly smaller than that anticipated from DLVO theory according to the non-linearised Poisson-Boltzmann equation, which predicts a dependence of the Debye length (defined as $1/\kappa_{Debye}$ and a measure of the width of the electrical double layer) on the electrolyte concentration around $1/\kappa_{Debye} \propto C_S^{-0.5}$. Pincus [49] predicted the force between adsorbed polyelectrolyte layers to be less sensitive to Debye screening. The measured exponent is closer to that predicted by Pincus, who calculated a dependence of the brush width on the electrolyte concentration of $L_B \propto C_S^{-0.33}$. The slight discrepancy in exponents may have arisen due to the complexity of the surface (eg heterogeneity and roughness).

The behaviour of the repulsive force decay length with pH and added salt is in agreement with generalised picture of the oocyst surface. That is, oocyst surface proteins extend into solution due to charge repulsion between charged groups distributed along surface molecules. As the pH is reduced the ionisable groups approach their *iep*, therefore reducing the range of the charge interactions and the surface molecules collapse. Also, as the electrolyte concentration is increased, the charge interactions are screened, also resulting in a collapse of the surface molecules.

Natural organic matter is generally regarded as a complex mixture of biological molecules, containing hydrophobic acidic moieties, mostly of carboxylic acid and phenolic functionalities. Therefore, electrostatic association of NOM and the oocyst surface molecules would generally not be expected at neutral to high pH, since both the oocyst and NOM would be predominantly negatively charged. It is somewhat surprising then that the decay lengths in NOM solutions are generally lower than the corresponding decay lengths in 1 mM KNO_3. However, the addition of NOM also involves the addition of NOM counter-ions, and a net

increase in the total ionic strength. The increase in total ionic strength would be anticipated to lead to increased charge screening and counter ion complexation, and hence shorter experimental decay lengths. Characterisation of the NOM by size exclusion chromatography (SEC) with UV detection suggests that the predominant fraction of the NOM is composed of hydrophobic molecules, accounting for more than 67% of the total composition [59]. Consequently it is possible that some components of the NOM adsorb by a non-electrostatic mechanism involving hydrophobic and/or hydrogen bonding.

At this stage, it is worth re-considering the electrophoretic data in calcium and NOM solutions in light of the decay length results. The influence of a 'surface hairy layer' on the electrophoretic potential has been a subject of discussion in the colloid and surface chemistry field for some time [60]. Essentially, compression of the hairy layer is anticipated to lead to a decrease in the hydrodynamic radius and an increase in the electrophoretic mobility. Comparison of the effective changes in the decay length for calcium and NOM, however, reveal that the decay length is more substantially affected by calcium than NOM, indicating that charge screening, rather than the extension of the hairy layer, may be dominating the ζ-potential of oocysts in the calcium nitrate solutions.

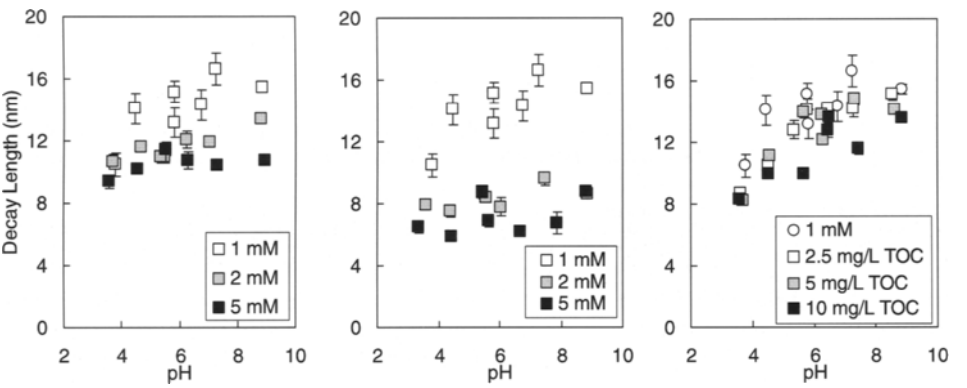


Fig. 9. Repulsive surface force decay length as a function of pH for three ionic strengths in monovalent electrolyte KNO_3 (left) divalent electrolyte (centre) prepared as the following molarities: 1 mM KNO_3 (1 mM), 1 mM KNO_3 and 0.25 mM $Ca(NO_3)_2$ (2 mM), as well as 1 mM KNO_3 and 1.25 mM $Ca(NO_3)_2$ (5 mM) and natural organic matter (right) prepared as 1 mM KNO_3, 1 mM KNO_3 and 2.5 mgL^{-1} TOC, 1 mM KNO_3 and 5 mgL^{-1} TOC, and 1 mM KNO_3 and 10 mgL^{-1} TOC

The repulsive force on approach has also been analysed in terms of the log linear (natural log) intercept of the force at zero separation, extrapolated from separations between 20 – 50 nm. The corresponding values are presented in Fig. 10. The relatively large variation in values of the magnitude presented in Fig. 10 immediately suggests the derived values are not significantly different. The variation in magnitude may be attributed to surface roughness, that is, the compression of discrete asperities on the oocyst surface during a force measurement may give rise to non-systematic variation in the magnitude. Despite the large variation, however, it is can be stated that the values of the magnitude are large compared to the predictions based on the particle micro-electrophoretic mobilities and DLVO theory (which equate to magnitudes of the order of 1-2 mNm^{-1}). Given that it has been shown that the region separating the DLVO force and the experimental total force can be described by a

Pincus model of polyelectrolyte brush compression, the large values of the magnitude are probably a result of brushlike compression of the oocyst surface molecules. The influence of electrolyte composition to the magnitude of interaction can not considered to be significant, especially in light of the large variation over multiple ionic strengths.

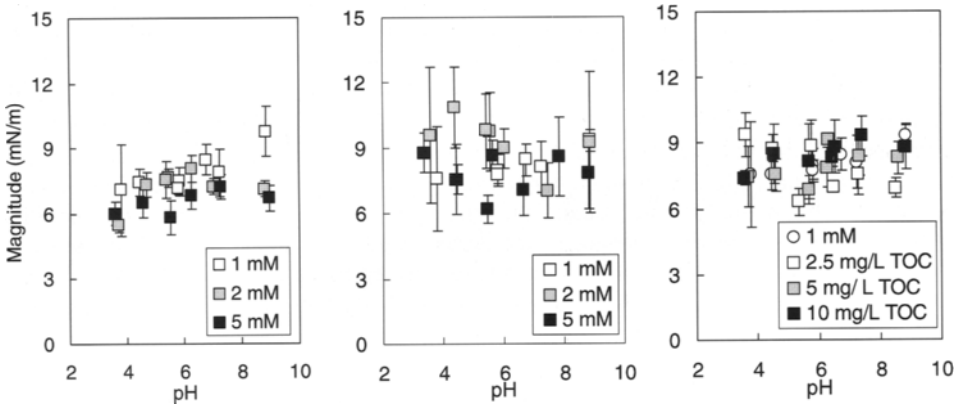


Fig. 10. Repulsive surface force magnitude as a function of pH for three ionic strengths in monovalent electrolyte KNO_3 (left) divalent electrolyte (centre) prepared as the following molarities: 1 mM KNO_3 (1 mM), 1 mM KNO_3 and 0.25 mM $Ca(NO_3)_2$ (2 mM), as well as 1 mM KNO_3 and 1.25 mM $Ca(NO_3)_2$ (5 mM) and natural organic matter (right) prepared as 1 mM KNO_3, 1 mM KNO_3 and 2.5 mgL^{-1} TOC, 1 mM KNO_3 and 5 mg L^{-1} TOC, and 1 mM KNO_3 and 10 mg L^{-1} TOC

3.4. Surface forces on separation

Despite the repulsive surface force on approach, most oocysts were found to adhere to the model sand surface (silica) to varying extents. A representative force curve exhibiting adhesion upon surface separation is shown in Fig. 11. The force curve upon retraction consists of two different regions: one where the measured force almost mirrors the approach (initial repulsive force at close separation); and the other where pull-off 'spikes' occur at various distances ranging from 0 to over 100 nm from contact. In comparison, the interaction between opposing silica surfaces is entirely repulsive on separation due to the predominance of an electrical double layer repulsion and the absence of an adhesion phenomena [56]. In the case of oocyst adhesion to the silica surface, the origin of the adhesion is likely to be quite different from the origin of the repulsion observed on surface approach. It has been shown that surface specific interactions are largely responsible for the adsorption of proteins from solution on to inorganic oxides [61]. Therefore, once the two surfaces are in contact specific surface adsorption of oocyst surface proteins may occur. On separation, the bulk compression of surface proteins leads to a repulsive force as observed on surface approach, but discrete adsorption of surface proteins (either in the molecular form or incorporated as macroscopic asperities) leads to occasional adhesion of the oocyst surface to the glass bead, despite the significant repulsive force.

The location of the snap-back to zero force ($H_{Jump\text{-}out}$, indicated by the vertical arrow in Fig. 11) corresponds to the maximum separation for attachment between the two surfaces. The normalised adhesion force ($F/R_{Adhesion}$) can be computed as a function of the spring constant of the cantilever (K_s), the harmonic radius (R) of the two surfaces and the maximum separation at which $H_{Jump\text{-}out}$ occurs:

$$\left(\frac{F}{R}\right)_{Adhesion} = \frac{-K_S H_{Jump-out}}{R} \tag{5}$$

The various values of the adhesion force ($F/R_{Adhesion}$) are plotted as a function of pH, potassium and calcium nitrate concentration and NOM in Fig. 12. Inspection of Fig. 12 reveals little correlation of the measured adhesion with mono-valent ionic strength, while the adhesion in divalent electrolyte and natural organic matter is slightly lower. Reduction in the adhesion is consistent with collapse of the hairy layer reducing the maximum extension of the surface proteins, leading to a tighter surface molecule confirmation.

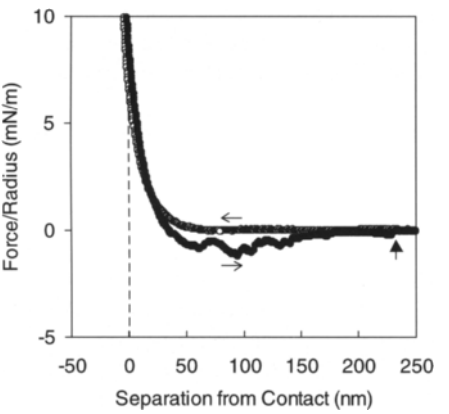


Fig. 11. Typical force-separation data for the separation of an oocyst and a silica particle in 1 mM KNO_3 at pH = 5.8. The horizontal arrows denote the direction of interparticle approach (left pointing arrow is for data collected as separation is narrowed) and the vertical arrow indicates the point of detachment between the surfaces

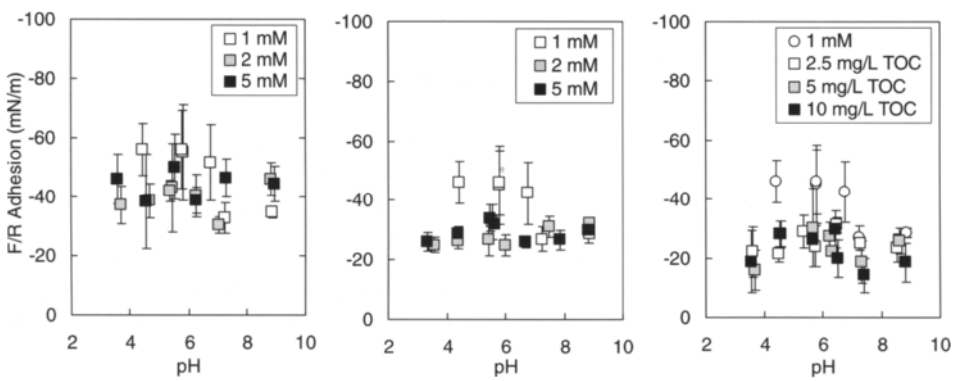


Fig. 12. Normalised force of adhesion as a function of pH for three ionic strengths in monovalent electrolyte KNO_3 (left) divalent electrolyte (centre) prepared as the following molarities: 1 mM KNO_3 (1 mM), 1 mM KNO_3 and 0.25 mM $Ca(NO_3)_2$ (2 mM), as well as 1 mM KNO_3 and 1.25 mM $Ca(NO_3)_2$ (5 mM) and natural organic matter (right) prepared as 1 mM KNO_3, 1 mM KNO_3 and 2.5 mg L^{-1} TOC, 1 mM KNO_3 and 5 mg L^{-1} TOC, and 1 mM KNO_3 and 10 mg L^{-1} TOC

4. CONCLUSIONS

Surface charge (ζ-potential) and surface forces (atomic force microscopy) measurements have provided fundamental insights into the sand-filtration process for removal of *Cryptosporidium* from water. The ζ-potential data prove the oocysts to be negatively charged above pH 3, and the ζ-potential decreases with increasing ionic strength and increases (ie. becomes more negative) with increasing NOM. The force of interaction between a single oocyst and a model sand surface has been measured using an AFM. The measurements revealed that a repulsive force always exists between the surfaces under the range of conditions measured, and that the magnitude of the force is under-predicted by DLVO calculations using the corresponding ζ-potential. The range of the force is reduced upon addition of calcium and NOM. The force of adhesion between the oocyst and model sand surface has also been examined. The findings provide an enhanced understanding of the surface of *Cryptosporidium* and how it interacts with other surfaces.

Oocyst surfaces possess anchored proteins that are stretched into solution by charge repulsion of surface ionisable groups distributed along the polypeptide backbone, leading to the formation of a charged brush-like structure. The plane of shear frequently lies somewhere toward the edge of the brush, as evidenced by comparison with DLVO theory, principally based on the measured ζ-potential of both surfaces. Upon addition of dissolved calcium ions, there is increased charge screening and an effective decrease in charge density at the plane of shear, as well as compression of the hairy layer. Upon addition of NOM, adsorption leads to an increase in charge at the plane of shear accompanied by a compression of the hairy layer. It has been found that while the hairy layer also imposes a repulsive force on surface separation, discrete adsorption of oocyst surface proteins can occur.

The measurements indicate that surface forces dominate the interaction of *Cryptosporidium* with sand in waters of varying hardness and natural organic matter. Of course, the use of coagulants in conventional water treatment plants alter the surface properties of charged particles in water, and it is likely this practice which ensures the reliability of water filtration. Nonetheless, the inherent repulsion between unmodified *Cryptosporidium* and sand reinforces the need for a multi-barrier approach in supplying drinking water.

REFERENCES

[1] W. Quintero-Betancourt and L. Botero De Ledesma, Int. J. Environ. Health Res., 10 (2000) 51.
[2] G. Ionas, J.J. Learmonth, E.A. Keys and T.J. Brown, Water Sci. Technol., 38 (1998) 57.
[3] J.T. Lisle and J.B. Rose, J. Water Supply Res. Technol.-Aqua, 44 (1995) 103.
[4] P.C. Okhuysen, C.L. Chappell, C.R. Sterling, W. Jakubowski and H.L. DuPont, Infect. Immun., 66 (1998) 441.
[5] The National Cryptosporidium Survey Group, J. Inst. Water Environ. Manage., 6 (1992) 697.
[6] H. Solo-Gabriele and S.J. Neumeister, J. - Am. Water Works Assoc., 88 (1996) 76.
[7] M.H. Kramer, B.L. Herwaldt, G.F. Craun, R.L. Calderon and D.D. Juranek, Morbidity and Mortality Weekly Report, 45 (1996) 1.
[8] J.N.S. Eisenber, E.Y.W. Seto, J.M. Colford, A. Olivieri and R. Spear, Epidemiology, 9 (1998) 255.
[9] P.A. Roefer, J.T. Monscvitz and D.J.J. Rexing, J. - Am. Water Works Assoc., 88 (1996) 95.
[10] H.L. DuPont, C.L. Chappell, C.R. Sterling, P.C. Okhuysen, J.B. Rose and W. Jakubowski, N. Engl. J. Med., 332 (1995) 855.

[11] P.C. Okhuysen, C.L. Chappell, C.R. Sterling, W. Jakubowski, and H.L. DuPont, Infect. Immun., 66 (1998) 441.
[12] J.K. Donnelly and E.I. Stentifold, Lebensmittel-Wissenscharft und Technologie, 30 (1997) 111.
[13] C. Campbell, S. Tzipori, G. Hutchison and K.W. Angus, Vet. Rec., 111 (1982) 414.
[14] D.G. Korich, J.R. Mead, M.S. Madore, N.A. Sinclair and C.R. Sterling, Appl. Environ. Microbiol., 56 (1990) 1423.
[15] Guidelines for Drinking Water Quality, World Health Organization, Geneva, 2004.
[16] J. Clancy and C. Fricker, Water Qual. Int., July/August (1998) 37.
[17] D. Fogel, J. Isaac-Renton, R. Guasparini, W. Moorehead and J.J. Ongerth, J. - Am. Water Works Assoc., 85 (1993) 77.
[18] B.V. Deryaguin and L. Landau, Acta Physicocochim., URSS, 14 (1941) 633.
[19] E.J.W. Verwey and J.Th.G. Overbeek, Theory of the Stability of Lyophobic Colloids, Elsevier, 1948.
[20] R.G. Horn, D.R. Clarke and M.T. Clarkson, J. Mater. Res., 3 (1988) 413.
[21] J.N. Israelachvili and G.E. Adams, J. Chem. Soc. Faraday Trans. I, 74 (1978) 975.
[22] R.M. Pashley, J. Colloid Interface Sci., 80, (1981) 153.
[23] R.M. Pashley, J. Colloid Interface Sci., 80 (1981) 531.
[24] J.N. Israelachvili, Adv. Colloid Interface Sci., 16 (1982) 31.
[25] R.M. Pashley and J.N. Israelachvili, J. Colloid Interface Sci., 97 (1984) 446.
[26] J.N. Israelachvili, Intermolecular and Surface Forces, Academic Press, London, 1992.
[27] P.C. Hiemenz, Principles of Colloid and Surface Chemistry, Marcel Decker, New York, 1986.
[28] L.A. Bottomley, Anal. Chem., 70 (1998) 425.
[29] B. Cappella and G. Dietler, Surf. Sci. Rep., 34 (1999) 1.
[30] H.J. Butt, Biophys. J. 60 (1991) 1438.
[31] W.A. Ducker, T.J. Senden and R.M. Pashley, Nature, 353 (1991) 239.
[32] R.C. Weast, (ed.) CRC Handbook of Chemistry and Physics, 60th ed., CRC Press, USA, 1981.
[33] P.G. Hartley, I. Larson, and P.J. Scales, Langmuir, 13 (1997) 2207.
[34] R.F. Considine, D.R. Dixon and C.J. Drummond, Water Res., (2002) 3421.
[35] T.J. Senden, C.J. Drummond and P.Kekicheff, Langmuir, 10 (1994) 358.
[36] R.J. Hunter, Zeta Potential in Colloid Science, Academic Press, New York, 1981.
[37] R.F. Considine, D.R. Dixon and C.J. Drummond, Langmuir, 16 (2000) 1323.
[38] P.G. Hartley, I. Larson and P.J. Scales, Langmuir, 13 (1997) 2207.
[39] L. Ranucci, H.-M. Muller, G.L. Rosa, I. Reckmann, M.A.G. Morales, F. Spano and E. Pozio, Infect. Immun., 61 (1993) 2347.
[40] J. Callejas Fernandez, F.J. de las Nieves, R. Martinez Garcia and R. Hidalgo-Alvarez, Colloids Surf., A, 61 (1991) 123.
[41] M. Buleva and I. Petkanchin, Colloids Surf., A, 151 (1999) 225.
[42] I. Sondi, O. Milat and V.J. Pravdic, J. Colloid Interface Sci., 189 (1997) 66.
[43] J.L. Zhou, S. Rowland, F.C. Mantoura and J. Braven, Water Res., 28 (1994) 571.
[44] M.G. Day, B.T. Hart, I.D. McKelvie and R. Beckett, Colloids Surf. A, 89 (1994) 1.
[45] B. Gu, J. Schmitt, Z. Chen, L. Liang and J.F. McCarthy, Environ. Sci. Technol., 28 (1994) 38.
[46] J.B. Fein, J.F. Boily, K. Guclu and E. Kaulbach, Chem. Geol., 162 (1999) 33.
[47] R.F. Considine, C.J. Drummond and D.R. Dixon, Langmuir, 17 (2001) 6325.
[48] P. Pincus and T. Witten, Europhys. Lett., 3 (1987) 315.
[49] P. Pincus, Macromolecules, 24 (1991) 2912.
[50] S.J. Alexander, J. Phys. Paris, 38 (1977) 983.
[51] P.G. De Gennes, Macromolecules, 15 (1982) 492.
[52] S.J. Miklavic and S. Marcelja, J. Phys. Chem., 93 (1988) 6718.
[53] S.T. Milner, T.A. Witten and M.E. Cates, Macromolecules, 22 (1989) 853.
[54] E.B. Zhulina, O.V. Borisov and T.M. Birshtein, J. Phys. II Fr., 2 (1992) 63.
[55] F. von Goeler and M. Muthukumar, Macromolecules, 28 (1995) 6608.
[56] T. Abraham, S. Giasson, J.F. Gohy and R. Jérôme, Langmuir, 16 (2000) 4286.
[57] R. F. Considine and C. J. Drummond, Langmuir, 17 (2001) 7777.

[58] T.D. Brock and M.T. Madigan, J.M. Martinko and J. Parker (eds.), Biology of Microorganisms, 8th ed., London, 1997.
[59] B. Bolto, A. Abbt-Braun, D.R. Dixon, R. Eldridge, F. Frimmel, S. Hesse, S. King, and M. Toifl, Water Sci. Technol., 40 (1999) 71.
[60] J.E. Seebergh and J.C. Berg, Colloids Surf. A, 100 (1995) 139.
[61] J.K. Stuart and V. Hlady, Langmuir, 11 (1995) 1368.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 13: Dewatering of water treatment plant sludges

Peter Scales

Particulate Fluids Processing Centre, Department of Chemical and Biomolecular Engineering University of Melbourne, Victoria, 3010, Australia. peterjs@unimelb.edu.au

1. INTRODUCTION

The dewatering of suspensions in the water industry is becoming a critical issue as the trend towards reuse of solids from the treatment of water gathers momentum. Whilst it has been commonplace in many places in the world to dispose of solids from potable water coagulation and filter backwash to sewer, the real cost of this activity is now being questioned. This is particularly the case where the influent to waste water processing sites constitutes significant quantities of these materials. Alternative utilisation for the solids fraction from potable water processing has long been sought and uses such as soil conditioning, cement production and brick-making are replacing traditional options such as landfill and discharge to sewer [1].

Hand in hand with the development of alternative disposal options is the need to dewater suspensions to high solids, usually to a dry cake, prior to transport. Dewatering techniques such as thickening followed by either filtration or centrifugation are the methods of choice. The difficulty is often the range of suspensions produced in the water industry and the number of options available. It is not always clear which is the best or lowest risk option to pursue. To aid in this assessment, this chapter presents a first principles approach to the understanding of the role of the type of solids produced and their dewaterability as a way forward in the selection and operation of dewatering equipment.

1.1. Dewatering principles

The solid-liquid separation process as it applies to water rich particulate slurries (dewatering), is commonly carried out in one of the unit operations of clarification, filtration or centrifugation, either alone or in combination. An example is where a low solids suspension from water coagulation and/or filter backwash activities is fed to a clarifier to produce an initial upgrade of the solids in the clarifier output that is subsequently fed to a mechanical filter or centrifuge [2].

There are a number of theories that describe the fundamentals of dewatering. In addition, there are a large number of model descriptions of unit operations such as clarifiers and filters in the literature. The difficulty for the practitioner is often to decide which theory or model is applicable in a particular situation. Nominally, this is not easy since different principles apply to different types of materials and it is not always clear where sludges from potable water treatment fit into the matrix.

To help to define a basis on which to move forward, it is appropriate to consider the type of materials that need to be dewatered in the water industry. A typical water treatment plant takes in water from a river or reservoir source and exposes the water to a series of treatments. The first of these is usually coagulation, where the addition of a coagulant causes

particulate and natural organic matter (NOM) to be enmeshed in a coagulant rich floc. The water then either goes through a sedimentation basin to allow these flocs to settle and then to filtration, or directly to filtration. Collection of the settled flocs from the sedimentation basin and from the filter through a backwash operation produces a low solids suspension (sludge) of precipitated coagulant, rich in particulates and NOM. On this basis, the sludge produced through potable water treatment plant operations will depend on the type of coagulant chosen but will typically be the result of the use of a metal salt coagulant (i.e. aluminium sulphate (alum), ferric chloride (ferric) or poly-aluminium chloride (PACl)) that will precipitate as a metal hydroxide at neutral to slightly acidic pH's. Other additives such as polymeric organic coagulants (i.e. poly dodeceyldimethyl ammonium chloride) or long chain polymeric organic flocculant style molecules used alone or in combination with the metal salt coagulants will change the nature of the sludge produced. Whatever the case, it is usual that the sludge will contain a metal hydroxide matrix that is the dominant component of the sludge on a mass basis.

As a starting point for analysis, it is important to recognise that a metal hydroxide rich water treatment sludge, when subjected to a pressure filtration process, will show a strong dependence of the output solids of the device on the applied pressure, assuming of course that adequate time has been allowed for filtration. This observation is a strong indication that such sludges may be classed as compressible materials. This is significant in that it excludes many of the simplified textbook methodologies typically used to produce an accurate and quantitative description of dewatering (especially filtration) since many are limited to non-compressible materials [3]. Despite this fact, these simplified methodologies are often applied and without being too critical, provide an order of magnitude description of the expected outcomes from the dewatering unit operation of interest and use average, parameter inputs [3].

Given that the sludges from water treatment plants are compressible materials, the theoretical description of the relevant dewatering operations often requires numerical computation approaches and is not straight forward. Scaling of the expected behaviour of similar sludges is less complex and one can use the fundamental inputs to the numerical descriptions as scalar arguments without compromise. The caveat is that the input parameters to the models are sludge material properties that are both non-linear functions of the solids content of the sludge and are highly dependent on the physical and chemical makeup of the sludge, the conditions under which the sludge was formed (i.e. mixing conditions) and the presence of additives such as high molecular weight polymeric flocculants.

To describe the solid - liquid separation process most theories start from fundamental momentum balances on the solid and fluid phases of an element of the solid-fluid mixture [4]. These are as follows for the solid and liquid phases respectively (ignoring diffusional motion):

$$F_d + F_g + F_{ps} = 0 \tag{1}$$

$$F_d + F_g + F_{pl} = 0 \tag{2}$$

where F_d is the drag force due to relative movement of solid and liquid, F_g is the gravitational body force and F_{pl} and F_{ps} are terms relating to the liquid and solid pressure gradients respectively.

The drag and solids pressure terms require two basic physical quantities for their definition. The first is a measure of the inter-phase drag due to relative movement of solid and

liquid. For filtration, this is often quantified through measurement of the permeability of the filter cake. In sedimentation, the measurement is through hindered settling. The second is a measure of the relationship between the local solids concentration and the local particle pressure. For filtration this can be thought of as the compressive behaviour of the filter cake and for thickening, the resistance of a networked bed of particles to collapse. Indeed, some measure of permeability and compressive behaviour of a filter cake or particulate slurry is a fundamental component of all the dewatering theories. Further to the above, these parameters need to be measured under conditions consistent with the industrial operation if dewatering is to be predicted and understood. Unfortunately, there is a large volume of parallel work in the literature with little cross reference. Rather than add to this literature, this chapter looks to take one approach and where possible, simply refer to other equivalent methodologies.

In the filtration arena, most work employs a specific cake resistance, α and solids pressure or stress, p_s, to describe the permeability and compressive behaviour of a solids suspension respectively. The description is based on the initial work by Darcy, extended to the case of filtration [5] and the subsequent work of Tiller, Shirato and co-workers [6-8].

In another approach (the one that will be utilised here), Buscall and White [9] developed a description of dewatering from a rheological perspective. They utilised a hindered settling function, $R(\phi)$, to describe the permeability (it is actually an inverse permeability parameter). This parameter is equivalent to a flow resistance when considering flow within a networked sediment or filter cake and employs a correction to the Stokesian settling rate due to inter-particle interactions in the case of un-networked sedimentation. The parameter ϕ refers to the volume fraction of particles in the suspension and implies in the case of $R(\phi)$ that it is a parameter that changes with the volume fraction of particles. Indeed, experimental data shows it to be a highly non-linear parameter that can vary by up to eight orders of magnitude between low and high solids concentration. Further to the above, a compressive yield stress, $P_y(\phi)$ is used to describe the compressive behaviour. This parameter describes the integral strength that a solids network possesses and its ability to support load as a function of its solids concentration.

The Buscall and White approach treats the $R(\phi)$ and $P_y(\phi)$ parameters as material properties. This implies that they will be particular to a system in a given state (eg. particle size, surface chemistry, temperature, state of flocculation, aggregate formation conditions, etc.). This approach has been applied to filtration [10-19] as well as sedimentation [20], thickening [21, 22] and centrifugation processes [23]. A parallel and closely related approach to that of Buscall-White has been developed by Bürger and co-workers [24].

Other work on dewatering comes from the geotechnical engineering arena and this body of work is also acknowledged as important, especially in understanding the compression of solids. However, application of the theory is typically to consolidation at very high solids and its relevance to a broad range of dewatering devices is limited. In later work Stickland et al. [25] have compared the rheological work discussed above with the geotechnical approach.

In summary, determination of both a solids dependent inter-phase drag, hindered settling function or specific resistance to filtration due to relative movement of solid and liquid and a solids compressibility or solids stress is mandatory in the quantitative characterisation of dewatering. Many tests are available to measure these parameters although some of the tests and parameters are not comprehensive in that they do not cover the full range of solids concentrations applicable to the dewatering operation to be considered. Very few approaches consider the applicability of the parameters to a broad range of dewatering devices and applications, indeed, these appear to be an exception rather than a rule [7, 26]. Rather than critique all of these approaches, a single comprehensive approach will be

discussed. The relationship to particular tests that are common to the industry will be considered as part of the development. Only the description of Buscall and White [9] will be considered here although the relationship to the other theoretical descriptions is available [27].

1.2. Characterisation of dewatering

Characterisation of the dewatering of sludges from the treatment of potable water has traditionally but not exclusively been performed using techniques such as capillary suction time (CST) and specific resistance to filtration (SRF). In particular, optimisation of the dose of polymeric flocculants has used CST routinely. However, it is largely accepted that CST is dependent on a number of non-fundamental parameters [28-30], which render it unsuitable for the purposes of comparing different sludge types or sludges from different plants [30]. This technique is also highly dependent on the initial solids concentration used in testing and it is unsuitable for predictive or scalar modelling of dewatering. SRF is more fundamental measure of dewatering and may give useful information regarding the permeability of non-compressible sludges but it is highly dependent on the applied pressure and the initial solids fraction for compressible sludges such as those from potable water treatment.

To fully characterize compressible sludges requires testing at a range of pressures, up to the required operating pressure of the device in question, starting at the initial solids concentration. For thickening operations, knowledge of sedimentation behaviour is also required. Comprehensive characterisation to allow quantitative prediction of dewatering will therefore require a combination of characterisation techniques along with appropriate phenomenological models of the device in question. The time scale required to achieve a comprehensive characterisation (approximately one day per sample) is sometimes not operationally desirable but it is essential in providing an understanding of the differences between samples and in providing a quantitative as well as distinct from a qualitative comparison between the dewaterability of sludges.

The dewatering characteristics of water treatment sludges can vary for a number of reasons including changes to:

- The quality of the raw water being treated (turbidity and colour)
- The dose and type of coagulant
- The dose and type of high molecular weight polymeric flocculants.

Four common primary coagulants used in water treatment are alum (aluminium sulphate), ferric salts (sulphate or chloride), poly (aluminium chloride) (PACl) and cationic polyelectrolytes. Inorganic coagulants remain the workhorse reagents for removing particulate and natural organic matter (NOM) impurities from raw water for potable utilisation, as detailed in Chapter 2, despite a number of disadvantages. The most important of these are the low solids content and difficult dewaterability of the resulting sludge [31-33]. It is generally accepted that sludge dewaterability is worse at higher coagulant dose, but this effect has apparently not been quantified, at least not in terms of the comprehensive methodologies suggested herein. In the case of the aluminium and ferric salts, the dewatering characteristics of the sludge may also be dependent on the formation conditions of the metal hydroxide precipitate. In addition, each plant may operate with different coagulation mechanisms that can vary from charge neutralisation to precipitation, sweep floc or enhanced coagulation [34].

It was noted previously that the solid-liquid separation theory developed by Buscall and White [9] requires as inputs to the theory, the hindered settling and compressibility behaviour of the sludge. These two fundamental parameters are non-linear functions of the solids

concentration and can be combined when solving the theoretical governing equations to give the diffusivity, $D(\phi)$. $D(\phi)$ is a material characteristic which is independent of both the initial solids concentration of the characterisation test and the type of process used to extract the data (i.e. vacuum or pressure filtration). The diffusivity describes the interplay between the compressibility, the solids content and the hindered settling behaviour of a sludge and represents a true dewaterability parameter in that along with either the hindered settling behaviour or compressibility, it is an input to models for the prediction of dewaterability. It should be noted that as a measure of dewaterability, SRF describes a particular set of circumstances and is a finite subset of $D(\phi)$. The diffusivity is given as [12];

$$D(\phi) = \left[\frac{\delta P_y(\phi)/\delta(\phi)}{R(\phi)}\right](1-\phi)^2 \qquad (3)$$

To help visualize the form of $D(\phi)$, it has been shown mathematically to be inversely proportional to the time it takes to dewater a sludge [12]. On this basis, simple comparisons of dewaterability can be made by comparing $D(\phi)$ on an equal volume fraction basis (i.e. by plotting $D(\phi)$ versus ϕ). Harbour et al. [33, 35] used this parameter to describe and compare the dewatering properties of mineral slurries, selected inorganic sludges and water treatment plant sludges. Methods for the calculation of $D(\phi)$ from compressibility and hindered settling function data are detailed in the literature [36]. Comparisons of different materials as for example, the predicted throughput of a filter utilising these materials as a feed, can also be useful but this approach relies implicitly on the accuracy and suitability of the associated phenomenological device model.

1.3. Dewatering equipment and methods

The rigorous modelling of dewatering processes (i.e. thickening, filtration, centrifugation) requires the measurement of the dewatering parameters that are the key inputs to these same models, namely $P_y(\phi)$ and $R(\phi)$. A key barrier to the widespread use of rigorous modelling based on first principles phenomenological approaches is the ability to determine the compressibility and hindered settling function parameters quickly. This is an important problem since with desktop computational facilities being available to the masses, the arguments against the use of computationally intensive numerical as distinct from analytical, or even worse, empirical approaches to the modelling of dewatering devices, is lowered. Therefore, the need is for robust experimental procedures for the extraction of compressibility and hindered settling function across a range of materials and additives (i.e. flocculants).

It was discussed earlier that it is likely for any one processing situation, that a range of techniques is required to comprehensively characterise a sludge sample in terms of dewaterability. Although the ultimate aim is no doubt to produce a single measurement, this is not currently possible. Using the extraction of the compressibility as an example, centrifuge techniques were originally proposed and refined however, they were found to be slow and not able to simultaneously provide a measure of $R(\phi)$ [37, 38]. In response to this problem, Green et al. [14] developed a constant pressure filtration technique that allowed simultaneous extraction of $R(\phi)$. As with other filtration techniques described in the literature [39], this was still slow. Advances using a pressure stepping technique have overcome many of these time issues and it is now possible to characterise the compressibility and hindered settling behaviour of a sample in an automated fashion in a few hours [18, 19]. This being the case, some of the opposition to the use of rigorous dewatering models can be dismissed; albeit, the

data from filtration testing is often in itself inadequate since data is required from the feed solids to the output solids of the device. In the case of filtration, laboratory filter tests can quite often achieve this goal but more often than not, low solids data is also required. This data is below the gel point, ϕ_g, or null stress solids of the material and filtration and centrifugation techniques are unsuitable.

The gel point of a suspension is a useful parameter. It represents the concentration at which a coagulated or flocculated suspension first forms a particulate network and is the point of onset of the compressibility of the material, since it is assumed that there is no means by which the particles can transmit stress below this transition. Typical curves representing the compressibility and hindered settling behaviour of a material are shown in Fig. 1 and 2 respectively. The gel point is shown as a line on the plots.

It is interesting to note that in the hindered settling plot (Fig. 2), that there is a significant change in the data below the gel point concentration (at least two orders of magnitude). It should be understood that the examples given in Figs. 1 and 2 are theoretical and represent materials with a gel point that is slightly higher than that observed for many potable water sludges. Despite this, the trend in the data relative to the gel point is typical of a large variety of materials, including alum and ferric rich sludges. The region below the gel point is representative of un-networked settling behaviour and data in this region is of relevance to the prediction, design and operation of clarification and thickening devices as well as being important to the prediction and operation of centrifuges and filters, especially if the feed concentration of the device is less than the gel point. This is often the case since the flocculation of suspensions is more efficient at concentrations below the gel point, a direct consequence of the observation that the gel point also represents a transition to non-Newtonian flow behaviour. This is characterised by the presence, in most flocculated suspensions, of yielding behaviour [40, 41].

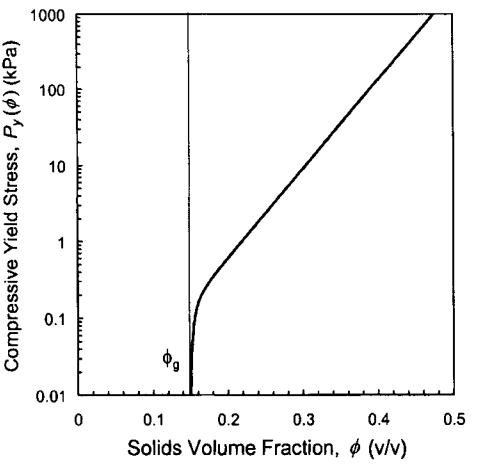


Fig. 1. Example of the compressive yield stress, $P_y(\phi)$, of a material, in this case a theoretical functional form that has shown to be representative of a large range of materials, including sludges from potable water treatment

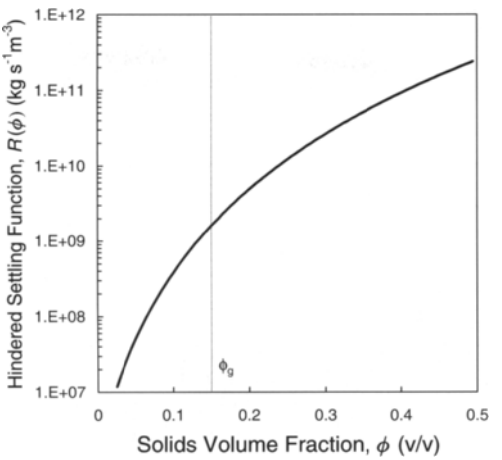


Fig. 2. Example of hindered settling function, $R(\phi)$, of a material, in this case a theoretical functional form that has shown to be representative of a large range of materials, including sludges from potable water treatment

The extraction of hindered settling data below the suspension gel point is usually obtained from simple batch settling data in a straight walled cylinder. This involves recording the height of the suspension interface with time, a relatively straightforward experiment requiring a stop watch and a cylinder.

The mechanics of the process are described by numerous authors [9, 20, 42-45]. Initially, the solids settle to form a clear layer of supernatant ($\phi = 0$) at the top of the column, and simultaneously, a consolidating bed ($\phi \geq \phi_g$) forms at the base. Depending upon the sedimentation type, a transition zone of un-networked suspension ($\phi_0 \leq \phi \leq \phi_g$) can also form between the initial suspension zone ($\phi = \phi_0$) and the consolidating bed. From the start of the test, the initial suspension zone settles (linear settling region) until it vanishes, a point in time characterized by non-linear behaviour in a height versus time plot. Settling then continues until only the bed and supernatant remain (non-linear settling region characterized by a rarefaction wave). The bed may then settle further toward equilibrium, when all of the compressive stress is borne by the solids bed (bed settling or compression region). A range of different types of settling behaviour are possible depending on the initial solids concentration of the material relative to the gel point and the density of the flocs in the suspension. These include a shock between the linear settling region and the bed settling region, with or without a rarefaction region, or a smooth transition between all three regions.

Typically, for potable water treatment sludges, settling is characterized by behaviour in which there is a linear settling region with a smooth transition to a rarefaction region and then into bed compression. This type of behaviour is demonstrated in Fig. 3.

The extraction of the hindered settling and gel point information from the test is not simple. A number of workers have proposed methods to characterize the gel point, ϕ_g and compressive yield stress, $P_y(\phi)$ from batch settling tests [6, 18] with estimation of $R(\phi)$ classically limited to the initial solids concentration, ϕ_0, of the test. This utilizes the slope of the linear settling region. Kynch [46] proposed a graphical method for estimating $R(\phi)$ for *concentrations* less than the gel point that has been used extensively despite the fact that physically correct solutions could not be guaranteed. A new analytic method to estimate $R(\phi)$

from batch settling data has now been proposed and is useful for the extraction of data up to concentrations just below the gel point [47]. It is only useful in the case where a rarefaction wave exists, which happens to be the most usual case for sludge from potable water treatment.

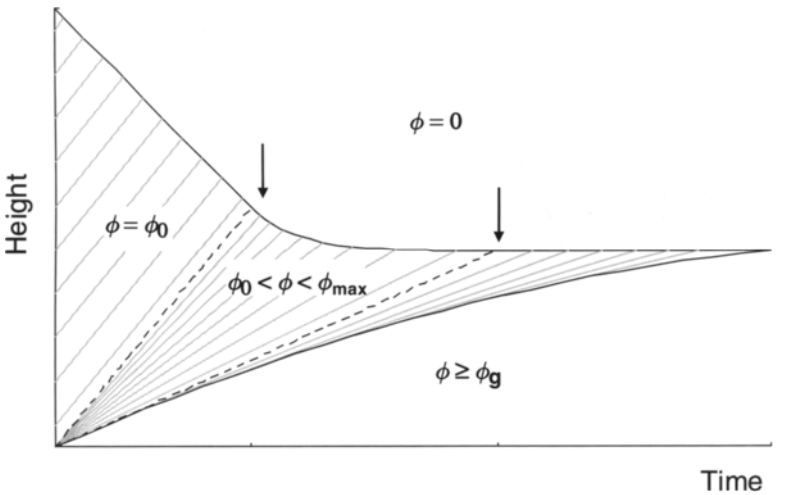


Fig. 3. Expected batch settling behaviour for a sludge from a potable water operation. The arrows represent the transition (smooth in this case) between the linear settling to rarefaction and rarefaction to bed compression regions of the plot

Therefore, using a combination of batch settling and stepped pressure filtration tests, it is now possible to determine the material parameters of interest to the rigorous prediction of dewatering behaviour in a relatively short (a few hours) space of time.

2. SLUDGE DEWATERING

2.1. Inorganic coagulants

2.1.1. *Alum*

Alum is one of the most common coagulants used in water treatment, especially for waters with low colour. It has quite a complex chemistry and many different species, including polymeric aluminium species are known to form under various conditions with factors such as dose, pH, rate of addition of hydroxide and mixing conditions all known to affect the type of products formed [48]. This complexity means alum is also very versatile and conditions can be manipulated to attain different coagulation mechanisms [49]. These subjects are covered in detail in Chapters 2 and 3.

Using the comprehensive approach to the comparison of dewatering described earlier, Figs. 4 and 5 show the compressibility and hindered settling function behaviour respectively of alum sludge from the treatment of low colour water. The hindered settling data in the region less than the gel point (Fig. 5) and the low solids compressibility data, shown as a fitted line in Fig. 4, was extracted from a settling test using the method described by Lester et al. [47]. The compressibility and hindered settling information at higher solids was extracted from a stepped pressure filtration test in the range 20 – 300 kPa, which covers a large number

of filtration processes. Higher pressure data would be required for application to higher pressure filtration devices, such as plate and frame membrane filters operating above 1000 kPa. From the $R(\phi)$ and $P_y(\phi)$ as measured at each pressure, Eq. (3) was then used to calculate $D(\phi)$.

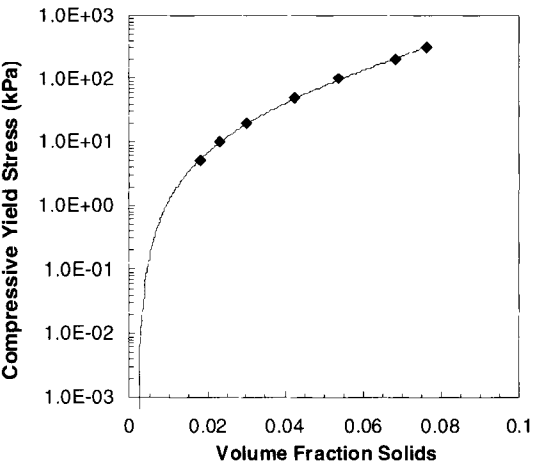


Fig. 4. Compressibility as a function of volume fraction of solids for a sludge generated from the treatment of a low colour water with alum

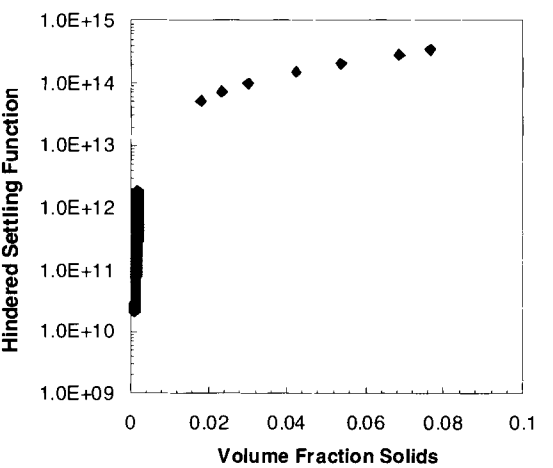


Fig. 5. Hindered settling function as a function of volume fraction of solids for a sludge generated from the treatment of a low colour water with alum

The data in Fig. 4 show the material to have a gel point at a volume fraction of approximately 0.003, a very low number. This is indicated by the extrapolation of the data to the horizontal axis. The low gel point is not unexpected since examination of the inter-relationship between alum dose and the percentage of solids in the sludge shows that the gel point is affected by the presence of alum once the percentage of alum in the sludge exceeds a value of approximately five percent. The compressibility behaviour is totally dominated by alum beyond a value of approximately fifty percent. This is a useful guide to the level of alum

above which addition of more alum is unlikely to be of any further detriment to sludge compressibility. It is perhaps unfortunate that modern utilisation practices for alum whereby enhanced or excess doses are used to aid NOM removal results in sludges that are well beyond the fifty percent alum condition. An indication of the expected trend in compressibility with increasing amounts of alum to solids is shown in Fig. 6. The data here was taken from the work of Dixon et al. [32]. Close examination of the data shows that the compressibility data from Fig. 4 is quite close to that expected for sludge with 100% alum. Therefore, the sludge is very difficult to dewater (i.e. curves shifted to the left are increasingly difficult to compress to high solids).

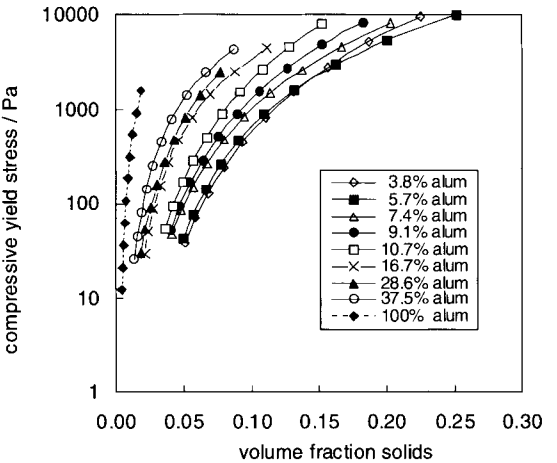


Fig. 6. Low pressure compressibility of a clay suspension flocculated with alum where the ratio of clay to alum is varied systematically. The data is from the work of Dixon et al. [32]

The hindered settling data shows many of the trends expected from the discussion associated with Fig. 2. There is a gap in the data between the very low solids points extracted from hindered settling tests and the filtration data at higher (albeit not very high) solids. Current practice is to interpolate between these two regions although it would be preferable if a technique were available that could produce data in this region to avoid this procedure.

In order to compare the dewatering behaviour using a single parameter, the solids diffusivity is a useful parameter and as intimated earlier, once hindered settling and compressibility data are available, calculation of $D(\phi)$ is quite direct. A typical example of the diffusivity across a range of solids concentrations for a wide variety of alum sludges taken from both operational treatment works and pilot plant studies is shown in Fig. 7. The data in this instance is from the work of Harbour et al. [33].

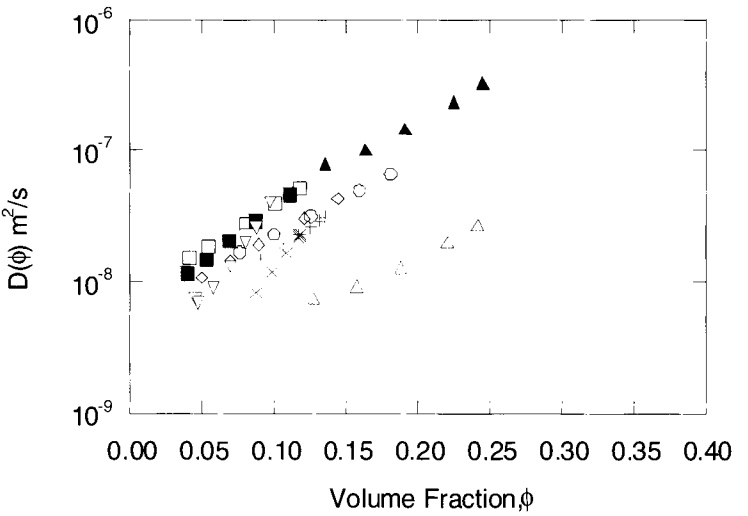


Fig. 7. Diffusivity as a function of volume fraction of solids for a range of sludges using aluminium sulphate as the coagulant. The data is from the work of Harbour et al. [33]

The conditions covered in the work are representative of a wide range of operational scenarios and except for one obvious outlier, the data is grouped reasonably closely. The outlier (open upright triangles on the bottom right) represents a case where a combination of alum and polyelectrolyte was utilised as the coagulant. This was a low colour, high turbidity water and the resultant sludge was found to be substantially more compressible than alum only sludges. The hindered settling behaviour showed it to be less permeable than comparable alum rich sludges. The consequence is a lower diffusivity.

As a guide, since $D(\phi)$ is inversely proportional to the time to filter or dewater a sample, an overall improvement in dewaterability will be associated with both improved compression and permeability behaviour. This will result in a diffusivity curve that moves both up and to the right in Fig. 7. It is generally observed that most efforts to improve dewaterability (i.e. through the addition of polymeric flocculants), cause the diffusivity to move up but not to the right. In summary, procedures that improve the compressibility will move the diffusivity to the right and those that improve the permeability (inversely related to the hindered settling behaviour) will move the diffusivity up. In this context, the practice of replacing alum with polyelectrolyte will produce a more compressible albeit less permeable sludge. The addition of polyelectrolyte after the coagulation process is likely to result in changes in dewaterability but only as an improvement in the permeability characteristic of the sludge.

2.1.2. Ferric

Ferric coagulants are becoming the coagulant of choice for waters with low turbidity and high colour. The high colour requires high doses of coagulant and sludges produced generally have a high coagulant to solids ratio. As a consequence, the expected variability in dewatering as a result of subtleties of dose, are less likely than in the alum case since it is likely that the resultant sludge will always be rich in ferric. The diffusivity, calculated from the compressibility and hindered settling data of a range of sludges taken from operational potable water treatment plants in both Australia and the UK, is shown in Fig. 8. As with the alum data, this data is also taken from the work of Harbour et al. [33].

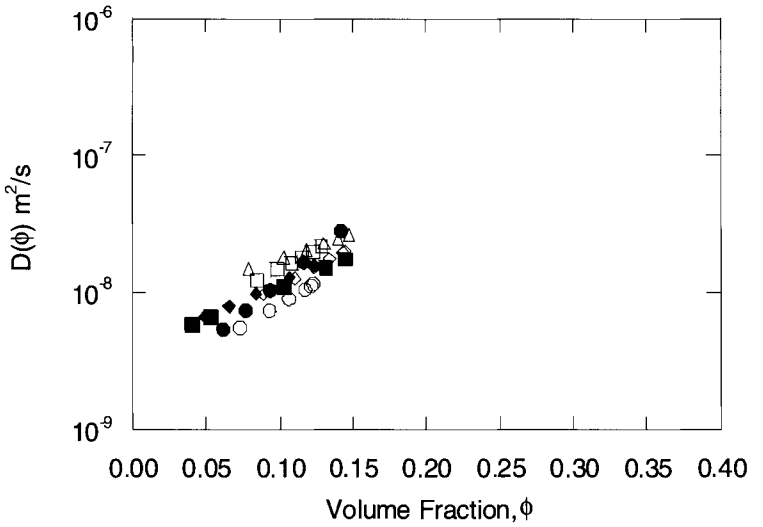


Fig. 8. Diffusivity as a function of volume fraction of solids for a range of sludges using ferric chloride as the coagulant. The data is from the work of Harbour et al. [33]

The data is very closely grouped indicating that these sludges have very similar properties even though they are formed under different conditions in terms of raw water quality and clarification conditions. In general, they have a lower compressibility than the alum sludges in Fig. 7, failing to reach 15% solids by volume at the highest pressures. The diffusivity is similar in magnitude to the alum sludges, indicating a similar permeability. It is expected on this basis that under equivalent conditions, ferric sludges will filter at a similar rate but to lower solids than most coagulant rich alum sludges.

Overall, the differences in the behaviour of the ferric sludges are small and they effectively have similar dewatering characteristics. On this basis, differences in dewatering performance observed with, for example, filtration equipment between different plants are more likely to be caused by changes in operational parameters such as initial solids concentration or solids loading.

2.1.3. Poly-aluminium chloride

Poly-aluminium chloride is often used in water treatment to control the aluminium species present and to produce a lower volume of sludge, presumably because of the greater control over the species present during precipitation. Better performance in terms of colour and turbidity removal is sometimes noted for specific waters [50]. It is interesting to note however that in pilot trials of the dewatering behaviour of sludges produced using poly-aluminium chloride, that there appears to be little or no benefit over the use of aluminium sulphate. Given the discussion associated with the rate of use and dewatering behaviour in the case of aluminium sulphate, the expected dewaterability should indeed be close to that for aluminium sulphate. The diffusivity data of Fig. 9 shows this to be the case.

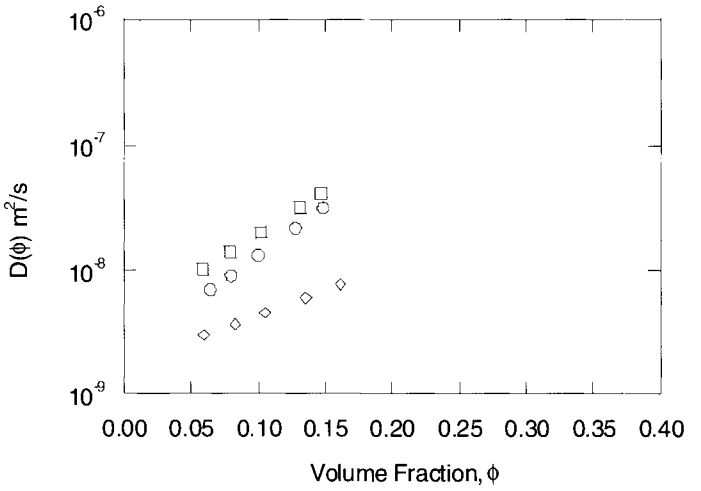


Fig. 9. Diffusivity as a function of volume fraction of solids for a range of sludges produced using poly-aluminium chloride as the coagulant. The data is from the work of Harbour et al. [33]

2.1.4. Polymer only sludges

For high turbidity, low colour waters, there are a modest number of operational plants around the world that use polyelectrolyte as the primary coagulant. These are typically cationic and of high molecular weight. Sludges produced from these activities are expected to be dominated by the dewaterability characteristics of the particulate material causing the turbidity rather than that of a metal ion coagulant. The expected trend on this basis is that sludges from these operations should be easier to compress to high solids although the expected trend in the hindered settling behaviour is harder to visualise. Fig. 10 shows diffusivity data for an operational plant in Australia. The data are clearly of greater compressibility (moved to the right but lower permeability, moved downwards) than any of the metal ion coagulant sludges. The data corresponding to the open and closed circles in Fig. 10 refers to the same sludge, except that the one corresponding to the closed circles has received an extra dose of polyelectrolyte (post formation) in order to aid the dewatering rate. The observed trend is that the data is shifted upwards (improved permeability) and slightly to the left (decreased compressibility). This is further confirmation that the role of polyelectrolyte addition to sludges for dewatering enhancement only acts to improve the permeability and not the compressibility of the sludge.

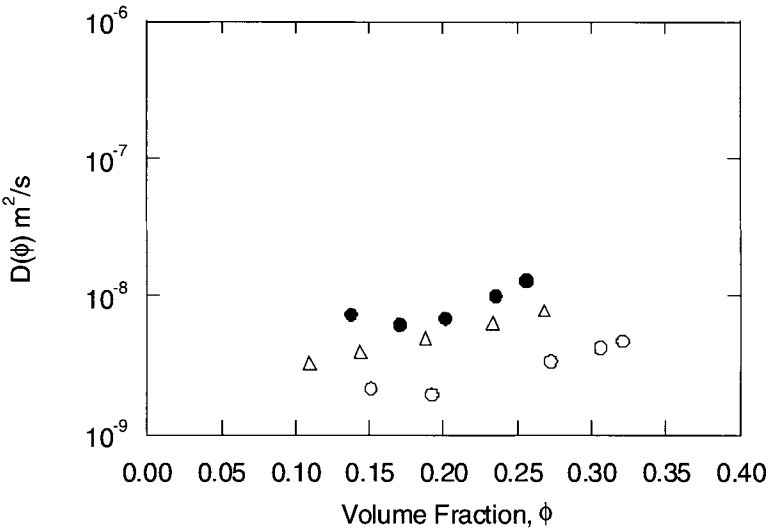


Fig. 10. Diffusivity as a function of volume fraction of solids for a range of sludges produced using poly-electrolyte as the coagulant. The data is from the work of Harbour et al. [33]

3. DEWATERING MODELS

There are a number of dewatering devices common to the potable water treatment industry and although it is beyond the scope of this chapter to provide an in-depth analysis of each of these devices, the fundamental parameters affecting performance of a number of devices in the context of the above discussion will be considered. The devices of interest are clarifiers, filters and centrifuges.

3.1. Clarification

Clarification is a process that occurs in any type of clarifier, thickener or settler that concentrates solids via settling and formation of a network structure or bed in the bottom of the device. A continuous flow gravity clarifier takes a slurry input and produces an underflow at a higher solids concentration and an overflow of clarified liquid. The primary mechanism for dewatering is gravity acting on the density difference between the solid particles and the carrier liquid. Effective operation requires that solid particles have a settling rate greater than the liquid rise such that particles are not carried to the overflow. The distinction between a clarifier, settler or thickener is not always clear but the definition that will be used here is one in which the solids concentration of the underflow of the device in the case of thickening is the result of compressional processes operating in a network of particles or bed. In simple terms, a clarifier is a thickener without a bed of solids.

There is limited design information for clarifiers and thickeners in the literature based on phenomenological modelling of the device using material property inputs such as the compressibility and hindered settling behaviour, as measured in the laboratory. The simplest design scenario comes from a simple flux balance such that the particle free settling rate is matched to approximately a third to a fifth of the upward fluid velocity. Using this

methodology, or a graphical Kynch approach [46], an operational clarifier can be designed with limited experimental information.

Many incremental improvements in the performance of industrial clarifiers have been based on selecting conditions that produce desired properties in settling test behaviour. Examples include selecting conditions that produce the desired settling rate, final sediment solids concentration, viscosity or shear yield stress. Though well entrenched and loosely based on material properties, such as permeability, compressibility and shear rheology, this type of empirical method does not enable quantitative prediction of clarifier performance, particularly if there is a solids bed in the device.

A more sophisticated approach is to predict steady state or transient clarifier performance based on fundamentally based equations and floc material properties. Numerous authors [51, 52] have presented fundamentally based equations and computational algorithms for predicting transient performance. These calculations enable prediction of how long it takes for process variations to have an effect on process performance. The problem is that the numerical methodologies are complex and computationally time consuming. Prediction of steady state clarifier or thickener performance is easier and computationally fast, allowing a range of scenarios to be investigated. Again, numerous authors [21, 22, 53] have presented algorithms to predict steady state clarifier and thickener performance.

The main output of steady state predictive models is the steady state solids flux (usually in tonnes of solid $hour^{-1}$ m^{-2} of device cross sectional area) as a function of underflow solids concentration. The calculation assumes input of parameters such as the feed solids concentration, hindered settling function and compressibility for the material, solid and liquid densities, device dimensions and operational solids bed height. An example of such a prediction for a water treatment sludge (ferric rich) is shown in Fig. 11. As indicated, the data output is in terms of the expected underflow solids concentration as a function of the solids flux.

The shape and trends in the data in Fig. 11 requires some explanation. The data is for a range of operational bed heights, ranging from 1 to 10 meters. The model assumes that no solids can exit the device through the overflow such that at high solids flux, the device will act like a pipe (suspension will be pumped out at close to the rate at which it enters) and the feed and underflow solids concentrations will be equivalent. As the solids flux decreases, water recovery starts and the solids concentration in the underflow increases. Up until the underflow reaches the gel point concentration of the sludge, the device will be permeability limited. There is no scope for compression up to this point because a percolating network of particles is not available to transmit stress. At lower flux, the operation of the device becomes compressibility limited and the underflow solids output will depend on the height of the solids bed. At very low solids fluxes, Fig. 11 can be used to predict settling pond or very long term settling behaviour. The upper line in the plot (only distinguishable at low solids flux) represents a permeability limit. It is clear that as a function of solids flux, the device operation is permeability limited at high flux but moves to compressibility limited behaviour at low flux. Taking this concept further suggests that settling rate and the position of the gel point for a sludge will determine the output underflow solids at high flux and the rate of escape of liquid from the networked bed of particles will determine underflow solids at low flux.

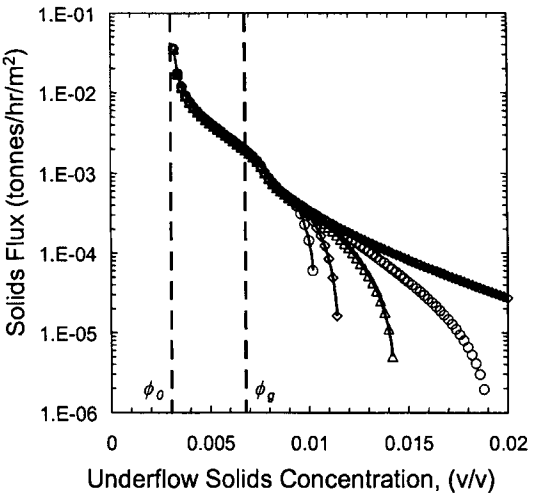


Fig. 11. Steady state thickener prediction of the expected throughput as a function of the expected underflow solids for a ferric rich water treatment sludge. The data set, from left to right represents prediction of operation with 1, 2, 5, 10 meters of bed height with the data set to the upper right representing permeability limited operation

The data from Fig. 11 presents a comprehensive overview of steady state clarifier or thickener performance based on relatively simple modeling using material property inputs. The addition of flocculants to improve permeability represents a new case for modeling and the expected improvement can be quantified as a predicted increase in solids flux at a given solids underflow concentration or as improved solids output at fixed flux.

3.2. Filtration

Commercial application of a comprehensive approach to the prediction of dewatering incorporating measured sludge properties such as hindered settling and compressibility behaviour to filtration is limited. The large number of filter types on the market, including gravity belt, belt press, ceramic disk, drum, plate and frame and tube presses means utilization of a first principles phenomenological approach to the prediction of filtration is a significant task given that each device requires development of a specific model. The task is reduced somewhat in the case of potable water treatment sludges since their attributes of being difficult to compress and poorly permeable makes the use of some filter types inefficient.

The need to filter sludge is dependent upon a number of economic and technical drivers, not the least of these being the disposal route [1]. Often, there is a need to produce a cake with appropriate rheological properties for transport (i.e. a high yield stress) as well as a need to reduce moisture to lower transport costs. In this context, low pressure filtration devices are often inappropriate for potable water sludge. The low permeability of water treatment plant sludge also means that batch processing is favoured due to an ability to filter longer. Therefore, whether by accident or design, gravity belt filters for upgrading solids as a feed to further processing, belt press filters and plate and frame filters are the preferred unit operations for potable water sludge. The latter is the highest pressure device and has the flexibility, through batch operation, to deal with the low permeability of the sludge.

Traditionally, plate-and-frame presses are designed from scale-up of pilot-scale tests or laboratory tests of the specific resistance to filtration (SRF). As noted in the introduction,

these scalar approaches are adequate for non-compressible materials but quickly loose accuracy for compressible sludges [54]. In addition, neither approach accounts for changes to operating conditions such as initial concentration, thus a more general approach applicable to compressible materials is needed. A model to describe the volume fraction distribution during plate and frame filtration based on the piston-driven filtration model of Landman et al.[11] has now been developed [55]. Utilization of such a model allows predictions of filter performance across a broad range of operational scenarios and allows optimization of filter throughput [12] and output. This type of approach is significantly more complex but has a host of operational advantages over traditional methodologies [3]. First principles phenomenological models for gravity belt filtration have also been developed [56] but not for belt press filtration.

3.3. Centrifugation

Centrifugation is becoming an increasingly important option for potable water sludge dewatering. It provides some advantages in terms of footprint and continuity of throughput over filtration although in terms of prediction of performance, the industry still relies heavily on scale up factors from pilot plant studies for design and operation [3]. In addition, as with filtration devices, there are a large number of centrifuge types from which to choose, namely sedimenting, filtering, solid bowl, and disk centrifuges, many having the option to operate in either batch or continuous mode. It is really only solid bowl centrifuges (sometimes referred to as decanter centrifuges) that are of great interest to future operations. This is a consequence of the large volumes of sludge and the low permeability and poor compressive characteristics of the sludge.

Despite the fact that there appear to be no first principles models of decanter centrifuges available in the literature, an analysis of the important parameters in centrifuge operation can be initiated by drawing an analogy between simple centrifugation processes and one dimensional continuous thickening [21]. This is a useful starting point for understanding the parameters that are important to the operation of a centrifuge. It can be argued that just as with thickening processes, centrifugation conditions will operationally fall somewhere between compressibility limited (where the output solids concentration will be consistent with the compressive limit of that material at a given applied pressure or centrifugal force) and permeability limited operation (where the output solids concentration will be less than the compressive limit of that material at a given applied pressure or centrifugal force) as the throughput increases. As with thickening, this situation can be offset by changing operational parameters such as the amount of polymeric flocculant added to the sludge. This will improve sludge permeability. There is an additional benefit in the case of centrifugation of being able to increase the applied pressure through increasing the rotation speed of the centrifuge and in the case of scrolling centrifuges, being able to maximize the torque on the scroll, which by implication means that the rheology of the solids being removed from the centrifuge will be at a maximum.

4. SUMMARY

The dewatering of sludges from the treatment of potable water is complex and sludge behaviour in dewatering shows a strong dependence of the coagulant and/or flocculant type and amount used to remove both particulate and natural organic matter from solution. The characterisation of the dewaterability of these sludges is not easy but if performed comprehensively, allows quantitative comparison of the expected dewatering behaviour. The

methodologies to characterise dewatering in a comprehensive manner and the subsequent outcomes of dewatering device modelling using this approach has been outlined.

REFERENCES

[1] R.F. van Nieuwenhuyze, N.E. McTigue and R.G. Lee, Beneficial applications and innovative sludge disposal methods, D.A. Cornwell and H.M.M. Koppers (eds.) Slib, Schlamm, Sludge, American Water Works Association Research Foundation and KIWA Ltd., 1990.

[2] H.M.M. Koppers, T.A. Rolan, C. Vandermeyden, N.E. McTigue, H.A. Henke, H. Martin and R.F. van Nieuwenhuyze, Advanced treatment technologies, D.A. Cornwell and H.M.M. Koppers (eds.), Slib, Schlamm, Sludge American Water Works Association Research Foundation and KIWA Ltd. 1990.

[3] D.A. Dahlstrom, R.C. Bennett, R.C. Emmett, P. Harriot, T. Laros, W. Leung, C. McCleary, B. Morey, J.Y. Oldshue, G. Priday, C.E. Silverblatt, J.S. Slottee, J.C. Smith and D.B. Todd, Chapter 18: Liquid-solid operations and equipment, in Perry's Chemical Engineering Handbook, Seventh Edition, R.H. Perry, D.W. Green, and J.O. Maloney, (eds). McGraw Hill. 1998.

[4] R.G. de Kretser, P.J. Scales and D.V. Boger, Compressive: An overview, Annu. Rheol. Rev.: 2002 (Bri. Soc. Rheol.), 2003.

[5] B.F. Ruth, G.H. Montillon and R.E. Montonna, Ind. Eng. Chem., 25 (1993) 76.

[6] F. Tiller and M. Shirato, AIChE J., 10(1) (1964) 61.

[7] F. Tiller and N.B. Hsyung, Water Sci. Technol., 28 (1993) 1.

[8] M. Shirato, T. Murase, E. Iritani and N. Hayashi, Filtration Separation, 20 (1983) 404.

[9] R. Buscall and L.R. White, J. Chem. Soc. Faraday Trans. I, 83 (1987) 873.

[10] K.A. Landman, L.R. White and M. Eberl, AIChE J., 41 (1995) 1687.

[11] K.A. Landman, C. Sirakoff and L.R. White, Phys. Fluids A, 6 (1991) 1495.

[12] K.A. Landman and L.R. White, AIChE J., 43 (1997) 3147.

[13] C. Shen, W.B. Russel and F.M. Auzerais, AIChE J., 40 (1994)1876.

[14] M.D. Green, K.A. Landman, R. De Kretser and D. V. Boger, Ind. Eng. Chem. Res., 37(10) (1998) 4152.

[15] M. Eberl, K.A. Landman and P.J. Scales, Colloids Surf., A, 103 (1995) 1.

[16] P. Scales, R. de Kretser, A. Stickland and S. Usher, Filtration process modelling and optimisation using fundamental theory, Proc.Filtech Europa, Dusseldorf, Germany, 2001.

[17] R. de Kretser, S. Usher, D. Lester and P. Scales, Compressional dewatering of suspensions - From fundamentals to application, D. Zhang (ed.), Proc. Chemeca 2000, Opportunities and Challenges for the Resource and Processing Industries, 2000.

[18] R. de Kretser, S. Usher, P. Scales, K. Landman and D. Boger, AIChE J., 47 (2001) 1758.

[19] S. Usher, R. de Kretser and P. Scales, AIChE J., 47 (2001) 1561.

[20] I. Howells, K.A. Landman, A. Panjkov, C. Sirakoff and L.R. White, Appl. Math. Model., 14 (1990) 77.

[21] K.A. Landman, L.R. White and R. Buscall, AIChE J., 34 (1988) 239.

[22] S. Usher and P.J. Scales, Chem. Eng. J., 111(2-3) (2005) 253.

[23] W.F. Eckert, J.H. Masliyah, M.R. Gray, and P.M. Fedorak,, AIChE J., 42 (1996) 960.

[24] R. Bürger, F. Concha and K.H. Karlsen, Chem. Eng. Sci., 56 (2001) 4537.

[25] A.D. Stickland, P.J. Scales and J.R. Styles, Geotech. Testing J., 28(6) (2005) 596.

[26] F. Tiller, C. S. Yeh, C. D. Tsai, and W. Chen, Filtration and Separation, 24 (1987) 121.

[27] R.G. de Kretser and P.J. Scales, Trans. Filtration Soc., accepted 2006.

[28] S.K. Dentel, Water Sci.Technol., 36(11) (1997) 1.

[29] J.T. Novak, M.L. Agerbaek, B.L. Sorensen and J.A. Hansen, J. Environ. Eng., 125(9) (1999) 816.

[30] P.A Vesiland, J.Water Pollut. Control Fed., 60(2) (1988) 215.

[31] J. Bratby, Coagulation and Flocculation. Uplands Press, Croydon, UK. 1980.

[32] D.R. Dixon, R. Eldridge, N.P. Lee and P.J. Scales, J. Water Supply Res. Technol. - Aqua, 53

(2004) 545.
[33] P. Harbour, N. Anderson, A. Aziz, D. Dixon, P. Hillis, P. Scales, A. Stickland and M. Tillotson, J. Water Supply Res. Technol. - Aqua, 53 (2004) 29.
[34] P.H King, H.M.M. Koppers, C. Vandermeyden, M.W.M. van Eekeren and N.C. Wortel, Optimizing sludge characteristics and minimizing generation, D.A. Cornwall and H.M.M. Koppers (eds.), Slib, Schlamm, Sludge, American Water Works Association Research Foundation and KIWA Ltd. 1990.
[35] P. Harbour, A. Aziz, P. Scales and D. Dixon, Prediction of the dewatering of selected inorganic sludges, in Proceedings of the IWA Conference 'Sludge Management Entering the 3rd Millennium - Industrial, Combined, Water and Wastewater Residuals,' Taipei, Taiwan, 2001.
[36] K.A. Landman, J.M. Stankovitch, and L.R. White, AIChE J., 45 (1999) 1875.
[37] L. Bergstrom, C.H. Schilling and I.A. Aksay, J. Am. Ceram. Soc., 75 (1992) 3305.
[38] M.D. Green, M. Eberl and K.L. Landman, AIChE J., 42 (1996) 2308.
[39] R.J. Wakeman, M.N. Sabri and E.S. Tarleton, Powder Technol., 65 (1991) 283.
[40] P.J. Scales, P.C. Kapur, S.B. Johnson and T.W. Healy, AIChE J., 44 (1998) 538.
[41] P.C. Kapur, P.J. Scales, D.V. Boger, and T.W. Healy, AIChE J., 43 (1997) 1171.
[42] K.A. Landman, L.R. White and M. Eberl, Adv. Colloid Interface Sci., 51 (1994) 175.
[43] P. Diplas and A.N. Papanicolaou, J. Environ. Eng., 123 (1997) 659.
[44] R. Bürger and W.L. Wedland, Math. Methods Appl. Sci., 21 (1998) 865.
[45] R. Bürger and E.M. Tory, Powder Technol., 108 (2000) 74.
[46] G.J. Kynch, Trans. Faraday Soc., 48 (1952) 166.
[47] D. Lester, S.P. Usher and P.J. Scales, AIChE J., 51(4) (2005) 1158.
[48] J.M. Duan and J. Gregory, Adv. Colloid Interface Sci., 100 (2003) 475 .
[49] J.M. Duan and J. Gregory, Pure Appl.Chem., 73 (2001) 2017.
[50] S. Ndiongue, R. Desjardins and M. Prevost, Environ. Technol., 21 (2000) 67.
[51] A.D. Martin, Water Res., 38 (2004)1568.
[52] R. Bürger, S. Evje, K. Hvistendahl Karlsen and K.A. Lie, Chem. Eng. J., 80 (2000) 91.
[53] P. Garrido, R. Burgos, F. Concha and R. Bürger, Miner. Eng., 16 (2003) 85.
[54] E.S. Tarleton and S.A. Willmer, Chem. Eng. Res. Des., 75 (1997) 497.
[55] A.D Stickland, R.G. de Kretser, P.J. Scales, S.P. Usher, P. Hillis and M. Tillotson, Chem. Eng. Sci., 61 (2006) 3818.
[56] A.D. Martin, AIChE J., 50 (2004) 1418.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chaper 14: Biocorrosion in drinking water distribution systems

I.B. Beech[a] and J.A. Sunner[b]

[a]School of Pharmacy and Biological Sciences, University of Portsmouth, St. Michael's Building, White Swan Road, Portsmouth PO1 2DT, UK

[b]Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana 59717, USA

1. INTRODUCTION

Biofilms, i.e. microbial growth at interfaces, are found on almost every submerged solid surface in aqueous habitats, including piping in drinking water treatment and distribution systems. A recent review by Hall-Stoodley et al. [1], as well as references therein, discusses the role of biofilms in natural environments and in human health.

In any given city in the world, there are hundreds of kilometres of piping made of different materials, such as copper, galvanized steels, stainless steels, plastics and even lead. All of these materials are susceptible to microbial colonization under both laminar and dynamic flow regimes, as well as under stagnant conditions (e.g. in "dead legs" in the system). It is important to realize that even the high flow rates typically encountered in fully pressurized water mains will not inhibit the development of a biofilm (Fig. 1).

Although the oligotrophic, i.e., low nutrient, environment of drinking water distribution systems is generally non-supportive of planktonic microbial growth, this is not necessarily the case for biofilms. Indeed, many species of bacteria that are able to use inorganic compounds, such as metal oxides and hydroxides, as electron acceptors, can thrive as biofilms on metallic surfaces of tubing material used in the distribution network. However, they are not readily detectable in the bulk liquid phase of the system.

Biofilms can cause a wide range of water quality and operational problems in drinking water distribution pipe networks. They contribute to loss of distribution system disinfectant residuals, increased bacterial levels, reduction of dissolved oxygen, taste and odour changes, red, black or blue water problems and hydraulic roughness [2]. Biofilms can also act as "safe havens" for pathogenic strains of bacteria, such as *E. coli*, *Pseudomonas* and *Legionella*, enabling their survival and proliferation and facilitating genetic transfer of antibiotic resistance, as well as virulence factors within bacterial populations [3]. Apart from fouling of the material surfaces, biofilms can adversely affect properties of material on which they are formed by promoting interfacial physico-chemical reactions that would not occur in the absence of microbes [4].

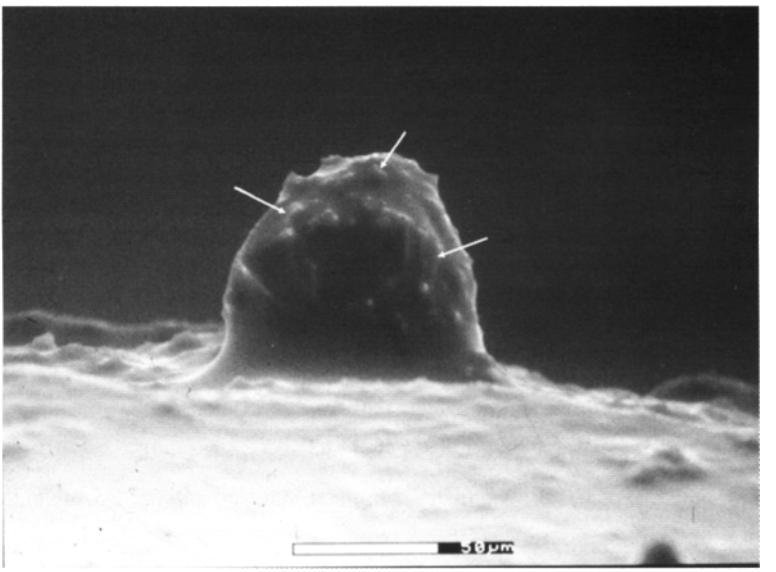


Fig.1. Environmental scanning electron micrograph (ESEM) of a 3 week old bacterial biofilm on the surface of AISI 316 stainless steel tubing removed from newly installed drinking water distribution system operating under high flow regime. White arrows indicate bacterial cells embedded in EPS matrix

For metallic materials, microbially-mediated, undesirable changes in material properties leading to their deterioration is referred to as biocorrosion or microbially-influenced corrosion (MIC). In water distribution systems, deterioration of piping material is of considerable concern. Typically, water departments expect up to 30%-40% losses due to leaks caused by both chemical corrosion and MIC. However, the contribution of biocorrosion to the overall corrosion process is not easy to assess. Therefore the role of MIC is often ignored, unless the corrosion rates in a given system cannot be explained as solely due to abiotic factors. Indeed, MIC, as a significant phenomenon, has been disregarded, particularly among engineers. The slow progress in establishing the importance of biocorrosion to material damage has is also been due to limitations of electrochemical, surface and analytical techniques to identify, localize and prevent or control biofilm-induced corrosion reactions.

Despite the problems, progress in disseminating understanding of deleterious effects of biofilms on metallic materials has been achieved. An excellent review on prevention and monitoring, as well as on failure analysis and controlling of MIC in industrial systems, has recently been published by Scott [5]. Corrosion of iron pipes in distribution systems has been reviewed extensively by McNeill and Edwards [6].

2. BIOFILMS AND CORROSION

The main constituents of biofilm are the microbial cells, their extracellular polymeric substances (EPS), and inorganic precipitates. The latter originate from the bulk aqueous phase or are formed as corrosion products of the substratum. EPS, comprising macromolecules such as polysaccharides, proteins, nucleic acids and lipids, constitute the biofilm matrix and are

often referred to as glycocalix or slime [7]. Microorganisms and/or their metabolites, *e.g.* enzymes active within the EPS matrix, organic and inorganic acids, as well as volatile compounds, such as ammonia or hydrogen sulphide, can alter electrochemical processes at the biofilm/metal interface [8].

2.1. Electrochemical corrosion

The physico-chemical interaction between a metallic material, and its environment that leads to changes in the material properties is termed corrosion [9]. The latter is an electrochemical process in which electrons are transferred from the metal, through a series of redox reactions, to an electron acceptor, often molecular oxygen, in contact with the metallic surface. As a result the metal is oxidized, which leads to its dissolution, and the electron acceptor is reduced. The most familiar form of corrosion is rusting of ferrous materials, which occurs when such materials are exposed to oxygen and water. The net chemical reaction of iron rusting can be expressed as follows:

$$Fe\ (0) + 1/2O_2 + H_2O \rightarrow Fe^{2+} + 2\ OH^- \qquad (1)$$

The Fe (II) product of this reaction is further oxidized to Fe (III) and, under neutral conditions in the presence of oxygen, typically forms the amorphous solid $Fe(OH)_3$. The latter can be converted to other iron oxides e.g. hematite (Fe_2O_3) and oxyhydroxides, e.g. goethite (α-FeOOH).

In most instances, the oxidation reaction slows to a low rate after a period of time because the oxidation products adhere to the metal surface and form an oxide/hydroxide layer that serves as a diffusion barrier to other reactants. These layers, referred to as passive layers, form a protective barrier to further oxidation of the underlying metal.

Environmental conditions can influence the equilibrium concentrations and diffusion rates of reactants and products of the oxidation reaction, metallic materials used for equipment fabrication are matched to the physical/chemical conditions under which the equipment is intended to operate. Changes in the environmental conditions can affect the stability of the passive layers and hence the susceptibility of the metal to corrosion. In general terms, any conditions that promote or retard metal oxidation will accelerate or inhibit corrosion, respectively.

2.2. Microbially-influenced corrosion

Due to the morphology of biofilms, which in most cases are present as a heterogeneous, non-continuous deposit or coating, on the metallic surface, MIC occurs as localized attack in form of pitting (Fig. 2). In water distribution systems, this type of corrosion often leads to extensive perforation of piping material.

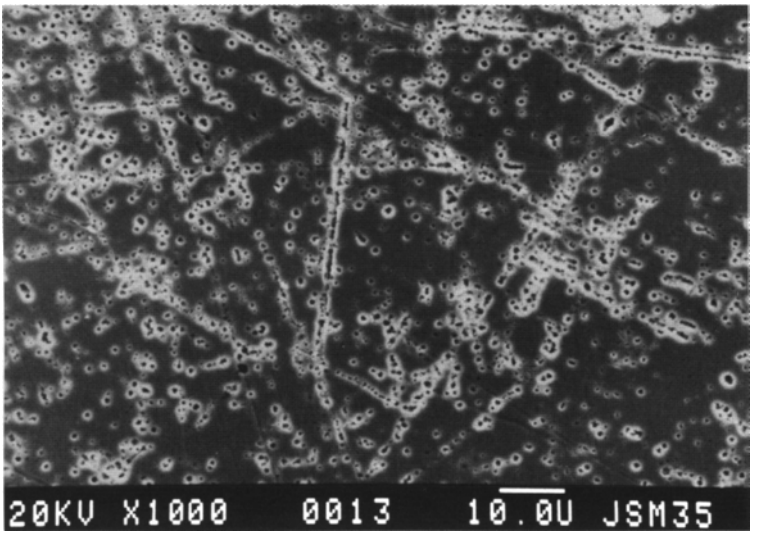


Fig.2. Scanning electron microscopy (SEM) micrograph demonstrating pitting corrosion of AISI 316 stainless steel tubing after the removal of 3 weeks old biofilm formed in the presence of bacterial consortium isolated from corrosion failure in cast iron potable water mains. The biofilm was grown in the laboratory under continuous flow conditions mimicking the flow regime, temperature and chemistry of water in the UK drinking water distribution system from which the bacterial consortium originated

Although microorganisms detected within biofilms on corroded metallic materials include bacteria, fungi and algae, the majority of research efforts have focused on the role that bacteria play in corrosion processes. The main types of bacteria found on such materials are sulphate reducing bacteria (SRB), sulphur-oxidising bacteria (SOB), iron-oxidising/reducing bacteria (IOB/IRB), manganese-oxidising bacteria (MOB), as well as bacteria secreting organic acids and slime [4,8 and references therein]. Several recent reviews discuss in detail mechanisms by which SRB, SOB and MRB/IRB can promote corrosion of iron and ferrous alloys [10-12].

Bacteria representing all the groups listed above have been found in biofilms recovered from different types of piping material in water treatment and distribution systems [13, 14] and the deteriorating effect of these consortia on e.g. stainless steels, cast iron or copper has been demonstrated [15]. However, the contribution of any particular genera or species to the corrosion process has not been established. It is now acknowledged that bacteria form synergistic communities within biofilms that are able to catalyse electrochemical processes through co-operative metabolism in ways that a single species has difficulty to initiate and/or maintain [16]. It is, therefore, important to realize that in any given system, including drinking water distribution networks, biocorrosion is seldom linked to a single bacterial species or to a unique mechanism. Typically, both the aggressive and the inhibitory effects of bacterial populations on corrosion reactions are due to complex interactions, involving both biofilm and corrosion products on the material surface.

It is equally important to consider the fact that documenting the presence of microorganisms on a corroded metallic surface, even if they are species known to produce metabolic by-products aggressive toward metals, is not sufficient evidence for their contribution to the corrosion process [17]. Similarly, the number of biofilm microorganisms detected at a corroded site does not necessarily correlate with the extent of corrosion [18]. It has been argued that the active metabolic capability of the microbes is most likely the key contributing parameter. To date, no direct evidence of a relationship between specific microbial metabolic rates and observed corrosion rates has been demonstrated, although emerging data strongly indicate that such a relationship exists [19-20].

2.2.1. Latest hypothesis of MIC

Studies of bacterial interaction with metallic materials led to the formulation of a unifying electron-transfer hypothesis of biocorrosion, using MIC of ferrous metals as a model system [11]. According to this hypothesis, biocorrosion is a process in which metabolic activities of microorganisms supply insoluble products, which are able to accept electrons from the base metal. This sequence of biotic and abiotic reactions produces a kinetically favored pathway of electron flow from the metal anode to the universal electron acceptor, oxygen. Although convincing and based on sound scientific evidence, this theory does not take into account (i) the part the EPS matrix plays in electron transfer nor (ii) the involvement of ultimate electron acceptors other than oxygen. Indeed, the theory has recently been challenged based on a study of marine biocorrosion of carbon steel under anoxic conditions [21]. A model demonstrating the likely involvement of biofilm EPS in electron transfer has also been proposed [8].

2.3. Biocorrosion viewed as a biomineralization process

A current trend in biocorrosion research focuses on biomineralization processes. Investigations of cell interactions with mineral surfaces is of obvious importance to MIC, as passive layers of oxides and oxyhydroxides formed on metallic surfaces, as well as abiotically and biologically produced corrosion products, including metal sulphides formed in the presence of SRB, are prime examples of such minerals. A number of recent studies report the use of state-of-the-art analytical techniques to elucidate the importance of biogenically formed nanocrystals in electron transfer processes, as well as in sequestering metal pollutants such as e.g. Pb (II) [22-24 and references therein].

Of particular interest to corrosion are processes leading to the dissolution of protective metal oxide/hydroxide films on metallic surfaces due to microbial dissimilatory reduction reactions i.e. using metal reduction as a mechanism for generating energy. The impact that bacterial metal reduction has on corrosion of iron and its alloys has been extensively reviewed [10, 25]. In the case of stainless steels, passive layers can either be lost or replaced by less stable reduced metal films that allow further corrosion to occur. Numerous bacteria are known to promote corrosion of iron and its alloys through dissimilatory reduction reactions [26]. In contrast to assimilatory iron reduction, which allows incorporation (assimilation) of iron into proteins, dissimilatory iron reduction involves electron transfer to iron as part of the fermentative or respiratory pathways. Iron reducing bacteria are readily recovered from cast iron pipes in water mains, for example. One of the best examples of a microorganism causing biocorrosion due to iron reduction is *Shewanella oneidensis,* formerly classified as *S. putrefaciens*. This is an extensively studied, Gram-negative, facultatively anaerobic bacterium. It oxidizes various carbon substrates by reductively dissolving Fe(III)-containing minerals, such as ferrihydrite, goethite and hematite. The biocorrosion of steel in the presence of *S. oneidensis* has been documented [10]. The corrosion rate was first measured by Little et

al. [27], who also showed that the rate depended on the type of oxide film under attack. In a later study it was confirmed that iron oxides, such as hematite, accumulated a higher density of *S. oneidensis* cells and showed a greater accumulation of Fe(II) than did magnetite [20].

2.4. Biocorrosion and bacterial species specificity

The study of Neal et al. [20], as well as other investigations, clearly demonstrated that in the presence of identical bacterial populations under the same growth conditions, different materials including metallic substrata, are colonised to a varied extent [4] (Fig. 3).

Furthermore, it is now accepted that the composition of the primary biofilms that develop on material surfaces depends on physico-chemical properties of the material and that different bacterial species preferentially colonize different materials [28 and references therein]. Studies have shown that biofilms developed more quickly on iron pipe surfaces than on plastic polyvinyl chloride (PVC) pipes. Fast colonization on iron occurred, despite the fact that adequate corrosion control was applied, i.e. the water was biologically treated to reduce assimilable organic carbon (AOC) levels and chlorine residuals were consistently maintained [29, 30]. It has also been reported that in water distribution systems, iron pipes supported a more diverse microbial population than did PVC pipes [31].

It is now recognized that the pipe surface itself can influence the composition and activity of biofilm population. Therefore, the choice of materials in water distribution systems is of paramount importance when one aims to reduce fouling and corrosion problems. A variety of materials in contact with water may leach materials that support bacterial growth. For example, pipe gaskets and elastic sealants (containing polyamide and silicone) can be a source of nutrients for bacterial proliferation. Pump lubricants should be non-nutritive to avoid bacterial growth in treated water [32]. Coating compounds for storage reservoirs and standpipes can contribute organic polymers and solvents that may support growth of heterotrophic bacteria. Liner materials may contain bitumen, chlorinated rubber, epoxy resin or tar-epoxy resin combinations that can promote bacterial growth [33].

3. THE EFFECT OF BIOFILMS ON COPPER AND LEAD CONTAINING MATERIALS

The choice of pipe material and the accumulation of corrosion products can dramatically impact the ability to control the effects of biofilms in drinking-water systems. Biofilms are known to significantly influence the accumulation, transport and transformation of heavy metals. In water distribution systems containing copper and lead tubing, the presence of biological deposits can lead to the increased release of Pb (II) or Cu (II) into drinking water. In addition to corrosion problems, this causes considerable health concerns [34]. The use of chemical treatments, for example the application of free chlorine or chloramine to control biofilms in drinking water systems, can also produce adverse effects, such as increasing the corrosion of lead-containing materials, resulting in Pb (II) contamination [35].

In the UK the link between the water main and the building service pipe is typically made of plastic. However, the majority of buildings have copper pipe distribution networks. Indeed, copper is widely accepted as a durable, corrosion resistant conduit for potable water. Although lead piping is comparatively rare in the UK, as there was a very effective programme of grant aided replacement, it is still extensively used in some other countries, such as the US.

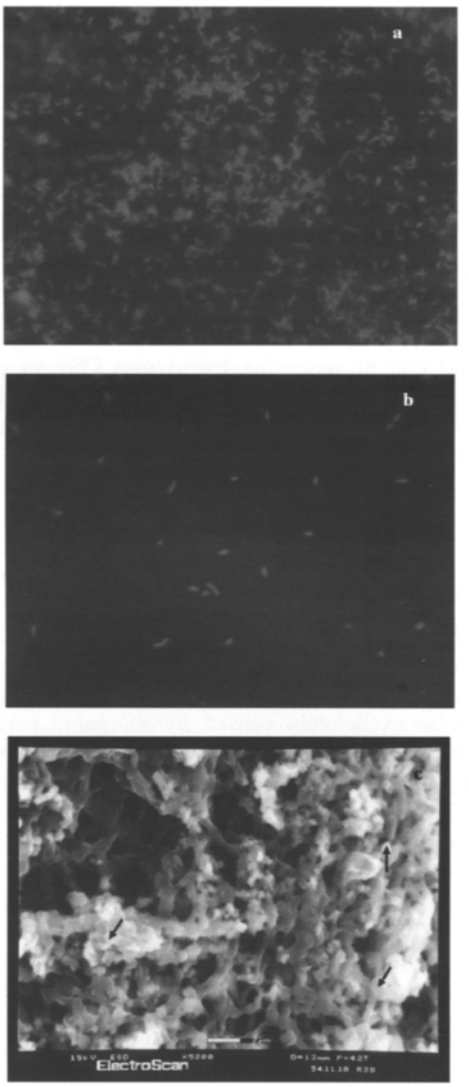


Fig.3. A light epifluorescence micrograph of DAPI stained bacterial cells in a 6-day-old biofilm formed under dynamic flow conditions in laboratory experiments on surfaces of (a) copper tubing and (b) AISI 316 stainless steel tubing. The biofilm was formed in a mixed culture of *Pseudomonas spp.* isolated from a surface of corroding copper tubing in potable water distribution system in the UK. Bacterial resistance to copper was confirmed by growing cells under different Cu(II) ion concentration regimes. ESEM micrograph (c) shows bacterial cells with associated copper precipitates (black arrows)

3.1. Lead

Generally, reports on corrosion of lead pipes in drinking water systems do not address the contribution of MIC to the overall process. Since corrosion of lead-containing service lines and of solder or brass plumbing materials is responsible for significant contamination of drinking water, the partitioning and mobility of Pb (II) ions in drinking water supplies is of main concern [36, 37]. It is accepted that bacterial-mineral composites are important for the retention of heavy metals, including Pb (II), due to their large sorption capacity. Numerous studies demonstrated that the ability of biofilms to retain Pb (II) is pH and metal concentration dependent and that biofilm can compete with abiotic and biotically generated metal oxide-hydroxide surfaces for Pb binding [38]. As already stated, these types of organic and inorganic coatings are commonplace on surfaces of metallic pipes in water distribution systems. In particular, biogenically formed Mn oxides have a very high capacity for Pb (II) uptake, exceeding that of e.g. Fe oxides/oxyhydroxides [39]. In some cases, biofilms and associated deposits can have a protective effect on lead pipe surfaces, thereby reducing metal ion release. In contrast, disruption of these deposits due to e.g. chemical treatment, flow increase or naturally occurring detachment of biofilms could cause the release of associated trace metal, as well as acceleration of corrosion.

3.2. Copper

A prime example of excessive metal corrosion by-product release is the "blue water" phenomenon observed in copper tubing. "Blue water" is caused by copper-containing compounds and occurrences have been reported from Europe, USA, New Zealand, Australia, and Japan. Copper contamination has been most prevalent in dead-legs or infrequently used cold-water services of institutional buildings. Traditionally, copper corrosion by-product release was thought to be exclusively caused by inorganic reactions and controlled by chemical properties of water. However, models based strictly on inorganic chemical reactions failed to predict the high concentrations of copper detected in some potable water, particularly at low alkalinity (pH 7.8 to 9.5). Studies have demonstrated that MIC can potentially contribute excessive copper concentrations in bulk water even when the water composition meets the regulations for potable use [40]. Investigations of excessive copper by-products release in chlorinated water supplies in New Zealand, Australia, and in the US have shown that the problem is restricted to the extremities of the recirculation systems, where it is difficult to maintain chlorine residual and control of microbial growth is lost. The topic of blue water corrosion has been reviewed in detail elsewhere [15]. In summary, the biological mechanisms most likely responsible for increased copper by-product release from copper tubing are (i) the accumulation of acidic microbial metabolites on the metal surface (ii) the binding of copper by microbial cells and their EPS and (iii) alteration of the physico-chemical properties of the copper oxide film due to biofilm growth within the oxide layers. While blue water corrosion does not necessarily compromise the integrity of the copper tubing, the process does result in the contamination of the water supply and wastewater sediments.

Copper-based metals are not immune to MIC and there are many examples of biofouling and biocorrosion failures of copper pipes in potable water systems [41]. Perforated copper pipes are frequently retrieved from domestic cold water systems. Typically, such failures are associated with stagnation. However, it can be difficult to discriminate between biological and purely electrochemical pitting. A review discussing phenomenology, water characteristics and corrosion morphology of copper tubing is available and these issues are therefore not addressed here [42].

Several mechanisms of Cu biocorrosion have been proposed and these are discussed by Geeesey et al. [15]. Recent work by Busalmen et al. [43] demonstrated that the type of Cu oxide film could play an important part in the MIC of copper-containing materials. A study of the effect of the extracellular enzyme catalase, produced by *Pseudomonas* species, on the corrosion behaviour of aluminium brass revealed that the presence of catalase strongly influenced the reduction kinetics of H_2O_2 produced during oxygen reduction. An increase in cathodic currents (up to 60%) relative to controls was measured in the presence of catalase, when the surface film was composed mainly of CuO. With only a submonolayer of Cu_2O covering the surface, there was no change in the cathodic current. When Cu_2O and CuO were present simultaneously, a 35% increment in the cathodic current was recorded. Enzymatic activities are readily detected in biofilms. Nonetheless, the possible importance of reactions mediated by these enzymes has only recently been considered as relevant to biocorrosion [4 and references therein]. Extracellular enzymes active within the biofilm matrix, such as catalases, peroxidases and superoxide dismutases are involved in reactions of oxygen reduction, therefore in principle they might facilitate corrosion by accelerating the overall cathodic reaction. However, it is important to realize that the ability of such enzymes to accelerate oxygen reduction depends on the chemistry of surface films.

4. CONCLUSIONS

The ecology of microbial communities and their metabolic output are of great importance for the effects of biofilms on the corrosion behaviour of metallic materials in any natural or man-made environment, including potable water treatment and distribution systems. A vast amount of work has been carried out in the areas of bacterial genomics and molecular environmental ecology. Reviewing work in this field is outside the scope of this chapter. Briefly, already existing data indicate an enormous physiological versatility of bacterial species and suggest that species specificity could explain why, despite identical environmental conditions, biofilms composed of bacteria from different genera, or even biofilms composed of different bacterial strains belonging to the same genus, differ in their ability to deteriorate colonised materials. Conversely, the type of material used in a given system is likely to influence the composition of the microbial population that will settle on its surface. The intrinsic relationship between microorganisms and materials ought to be seriously considered when designing piping networks.

Undisputedly, biofilm-influenced corrosion within drinking water treatment and distribution system is dependent upon a complex interaction of chemical, physical, operational and engineering parameters. No single factor could account for all the occurrences of biofilm development. Hence all of the above parameters ought to be considered in devising a solution to the MIC problem. Even for systems that do not currently experience biocorrosion, operators may want to more closely examine biofilm control strategies as a means of limiting future MIC risk and associated hazards, such as the release of toxic metal ion species and growth of opportunistic pathogens.

REFERENCES

[1] L. Hall-Stoodley, J.W. Costerton and P. Stoodley, Nat. Rev. Microbiol., 2 (2004) 95.
[2] W.G. Characklis and K.C Marshall (eds.), Biofilms, John Wiley & Sons, New York, 1990.
[3] A.K. Camper, Biofilms: recent advances in their study and control, L.V. Evans (ed.), Harwood Academic Publishers, London, 2000, pp. 311-332.

[4] I.B. Beech and C.L.M. Coutinho, Biofilms in medicine, industry and environmental biotechnology – characteristics, analysis and control, P. Lens, A.P. Moran, T. Mahony, P. Stoodly and V. O'Flaherty (eds.), IWA Publishing of Alliance House, London, 2003, pp. 115-131.
[5] P. J.B. Scott, Mater. Perf., 43 (2004) 3.
[6] L.S. McNeill and M Edwards, J. - Am. Water Works Assoc., 93 (2001) 88.
[7] I.W. Sutherland, Trends Microbiol., 9 (2001) 222.
[8] I.B. Beech and J.A. Sunner, Curr. Opin. Biotechnol., 15 (2004) 181.
[9] J.C. Scully, Fundamentals of Corrosion, 3rd edn, Pergamon Press, Oxford, UK, 1990.
[10] A.K. Lee and D.K. Newman, Appl. Microbiol. Biotechnol., 62 (2003) 134.
[11] W.A. Hamilton, Biofouling, 19 (2003) 65.
[12] I.B. Beech, Encyclopaedia of Environmental Microbiology, G. Bitton (ed.), John Wiley, London, 2002, pp. 465-475.
[13] L. Hanjangsit, I.B. Beech, R.G.J. Edyvean and C. Hammond, Microbial corrosion, A.K Tiller and C.A.C. Sequeira (eds.), Proc. 3rd International EFC Workshop, Portugal, 1994, European Federation of Corrosion Publication Nr. 15, Institute of Materials, London, 1995, pp. 322-327.
[14] I.B. Beech, R.G.J. Edyvean, C.W.S. Cheung and A. Turner, Microbial Corrosion, A.K Tiller and C.A.C. Sequeira (eds.), Proc. 3rd International EFC Workshop, Portugal, 1994, European Federation of Corrosion Publication Nr. 15, Institute of Materials, London, 1995, pp. 328-337.
[15] G.G. Geesey, I.B. Beech, P.J. Bremmer, B.J. Webster and D. Wells, Biofilms II: Process Analysis and Applications, J. Bryers (ed.), Wiley-Liss Inc., London, 2000, pp. 281-326.
[16] B.V. Kjellerup, B.H. Olesen, J.L. Nielsen, B. Frolund, S. Odum and P.H. Nielsen, Water Sci. Technol., 47 (2003) 117.
[17] B.J. Little, P.Wagner and F. Mansfeld, Microbiologically influenced corrosion. Corrosion Testing Made Easy, vol.5, NACE International, Houston, TX, (1997), pp. 29-52.
[18] B.J. Little and P. Wagner, Mater. Perform., 36 (1997) 40.
[19] H.T. Dinh, J. Kuever, M. Mussmann, A.W. Hassel, M. Stratmann and F. Widdel, Nature, 427 (2004) 829.
[20] A.L. Neal, K.M. Rosso, G.G. Geesey, Y.A. Gorby and B.J. Little, Geochim. Cosmochim. Acta, 23 (2003) 4489.
[21] J.S. Lee, R.I. Ray, E.J. Lemieux, A.U. Falster and B.J. Little, Biofouling, 20 (2004) 237.
[22] C.S. Chan, G. de Stasio, S.A. Welch, M. Girasole, B.H.Frazer, M.V. Nesterova, S. Fakra and J.F. Banfield, Science, 303 (2003) 1656.
[23] I.B. Beech and J.A. Sunner, Rev. Microbiol., 2005, in press.
[24] J.W. Moreau, R.I. Webb and J.F. Banfield, Am. Mineral., 89 (2004) 950.
[25] D.K. Newman, Science, 292 (2001) 1312.
[26] K.H. Nealson and D.A. Saffarini, Annu. Rev. Microbiol., 48 (1995) 311.
[27] B.J. Little, P. Wagner, K. Hart, R. Ray, D. Lavoie, K. Nealson and C. Aguilar, Proc. NACE Corrosion '97, NACE International, Houston, TX, (1997), Paper N° 215.
[28] M.W. LeChevallier, Heterotrophic Plate Counts and Drinking-water Safety. J. Bartram, J. Cotruvo, M. Exner, C. Fricker and A. Glasmacher (eds.), World Health Organization (WHO), IWA Publishing, London, UK, 2003, p. 177-197.
[29] C.N. Haas, M.A. Meyer and M.S. Paller, J. - Am. Water Works Assoc., 75 (1983) 139.
[30] A.K. Camper, M.W. LeChevallier, S.C. Broadway and G. McFeters, Appl. Environ. Microbiol., 52 (1986) 434.
[31] C.D. Norton and M.W. LeChevallier, Appl. Environ. Microbiol., 66 (2000) 268.
[32] D.R. White and M.W. LeChevallier, J. - Am. Water Works Assoc., 85 (1993) 112.
[33] D. Schoenen, Biotechnology vol. 8, H.J. Rehm and G. Reed (eds.), VCH Verlagsgesellschaft, Weinheim, Germany, 1986.
[34] M. Clement, R. Seux and S. Rabarot, Water Res., 34 (2000) 219.
[35] M. Edwards and A. Dudi, J. - Am. Water Works Assoc., 96 (2004) 69.
[36] R. Vilagines and P.Leroy, Bulletin de L'Academie Nationale de Medecine, 176 (1995) 1393.

[37] C.M. Davidson, N.J. Peters, A. Britton, L, Brady, P.H.E. Gardiner and B.D. Lewis, Water Sci. Technol., 49 (2004) 49.
[38] A.S. Templeton, A.M. Spormann and G.E. Brown, Environ. Sci. Technol., 37 (2003) 2166.
[39] Y.M. Nelson, L.W. Lion, M.L. Schuler and W.C. Ghiorse, Limnology and Oceanography, 44 (1999) 1715.
[40] B.J. Webster, D.B. Wells and P.J. Bremer, Proc. NACE Corrosion '96, NACE, Houston, TX, Paper No. 29, 1996.
[41] J.R. Myers and A. Cohen, Mater. Perform., 34 (1995) 60.
[42] D.B. Wells, B.J. Webster, P.T. Wilson and P.J. Bremer, Materials Performance Technologies Report 78076.31, Wellington, New Zealand, 1998.
[43] J.P. Busalmen, M. Vázquez and S.R. de Sánchez, Electrochim. Acta, 47 (2002) 1857.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 15: Adhesion analysis of scaling systems

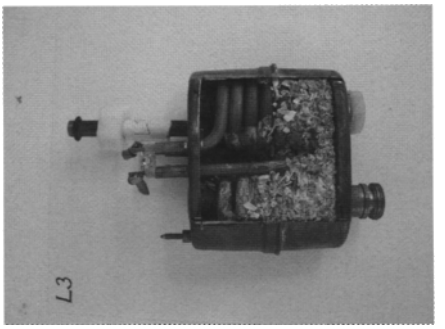
Simon A Parsons and Bruce Jefferson

Cranfield University
Building 39, School of Water Sciences
MK43 0AL
UK

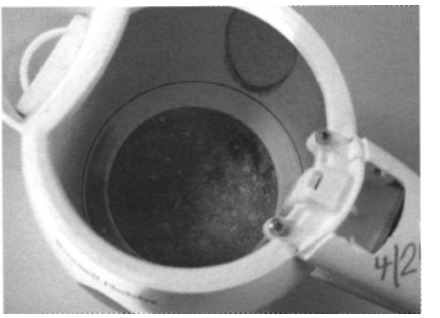
1. INTRODUCTION

The formation of scale is a widespread problem found in domestic, commercial and industrial areas where water, or wastewater, comes into contact with a surface. We are all probably most familiar with the problems associated with the scaling of heat transfer equipment by inverse solubility salts, known as scaling or crystallisation fouling. If your domestic water supply is hard (for instance, over 60% of households in England and Wales are supplied with hard water) then when this is heated in domestic appliances such as kettles, boilers and showers, a hard white scale which is typically calcium carbonate ($CaCO_3$) will form. This scale can lead to the loss in heat efficiency of any appliances heating water, such as kettles, washing machines, showers etc. and subsequently increasing gas and electricity consumption and so costs. Two examples of this type of scale are shown below for an electric shower and a kettle (Fig. 1a and b). The most significant issue though is for the energy production industry where a recent estimate of the overall cost of fouling is US$ 26,850 million per year [1]. Calcium carbonate scales do not form just because of heating though, and there are a number of other examples where evaporation or concentration can lead to both aesthetic and operational problems such as in toilet urinals and membranes (see Fig. 1 c and d).

The movement of water through pipelines can also cause significant problems in both water and wastewater treatment works. In the UK, the dosing of lime to correct pH and control corrosion is common place. Lime slurry is prepared from dry hydrated lime and mixed with a carrier water, typically treated water from the plant. If this water contains carbonate then a hard calcium carbonate scale will form (see Fig. 1e) which can cause a range of operational problems such as: pipelines and dosing equipment may require further maintenance, de-scaling and in some circumstances replacement. Replacing pipe work is expensive and can be disruptive, particularly if most of the pipework is located below ground. Similar problems are observed in wastewater treatment plants but in this case it is a different scale, struvite, that is the cause of the problem (Fig. 1f). Struvite or magnesium ammonium phosphate ($MgNH_4PO_4.6H_2O$) is a common scale in wastewater treatment plants that biologically remove phosphorus. Here, it is not only pipes that are affected, pumps, centrifuges and aerators are also susceptible to accumulation of struvite [2].



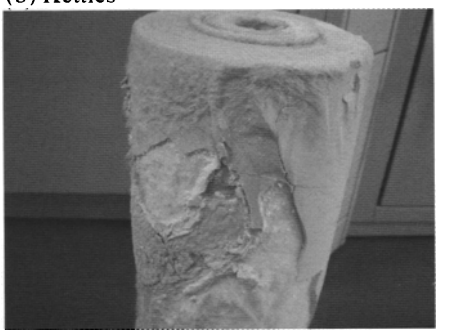
(a) Domestic shower



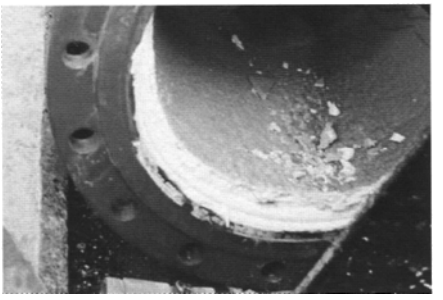
(b) Kettles



(c) Toilet urinals



(d) Pre filter for reverse osmosis membrane



(e) Lime slurry pipe line



(f) Pipe fouled with struvite

Fig.1. Examples of scaling

Traditional methods for scale control involve chemical alteration of the solution to achieve either pre-precipitation of the scale formers, inhibition of scale formation or replacement of scale formers [1]. These methods are effective but substantially change the solution chemistry, can be prohibitively expensive and are perceived as a poor environmental solution [3]. In addition to water chemistry, scale formation is affected by the surface properties of the materials on to which the scale will form through a combination of their surface roughness and their surface energy. For instance, Keysar et al. [4] found that the energy required to remove a scale deposit was as much as 30 times greater on a rough surface

(average roughness = 30 μm) compared to a smooth surface (average roughness = 0.1 μm). Similarly, Müller-Steinhagen et al. [5] found that the heat transfer coefficient of a standard pipe fell from 0.75 to 0.1 $kW.m^{-2}.K^{-1}$ over 120 minutes compared to no measurable change when using an electro polished pipe. MacAdam and Parsons [1] reported a 30 % increase in scaling rate as the average surface roughness of a heating surface increased from 0.2 to 0.8 μm. Doyle et al. [2] extended this idea to investigate how material selection may influence scaling rates during the precipitation of magnesium ammonium phosphate from wastewater. Both roughness and surface finish were observed to influence the scaling rate with smooth low energy surfaces scaling the least.

Forster and Bohnet [6] investigated the influence of surface energy on the induction period on calcium sulphate fouling in a heat exchange system. The induction period for copper (surface energy = 50 $nN.m^{-1}$) was significantly shorter than for stainless steel (surface energy = 40 $nN.m^{-1}$). Müller-Steinhagen and Zhao [7] investigated how coatings can adapt fouling properties of stainless steel and showed that low energy coatings could reduce scaling by 70% for pool boiling and almost eliminate the issue for convective heat transfer. Similar results have been reported with copper docosanoic acid and PTFE coatings which could significantly reduce induction time [8-9] and polycarboxyic acid coatings which reduced overall scale formation [10].

Whilst the subject of scaling and its control are established new environmental drivers are reducing the perceptive benefits of chemical solutions to the problem. Consequently, increased attention is now being placed on controlling the induction phase through better understanding and management of the influence of surface properties on the overall process. The problem then becomes an issue of understanding the fundamental colloidal and interface science that then can provide both methods to assess potential issues and insights for material selection and development. The ultimate aim is to produce low or even non fouling surfaces. The aim of this chapter is to cover the important basic principles involved in such a process together with describing both the traditional and recently developed methods that assist in the process. The approach will be demonstrated through a case study looking at how material properties impact on eth scaling rate in domestic heating systems.

2. THEORY

The scaling of surfaces can be described as a two stage process [1]. In the first phase, induction, stable crystalline nuclei are formed, this stage is followed by a fouling phase where material accumulates at the surface alternating the response of the system either as a change in radius of a pipe, accumulation of mass, or a reduction in the heat transfer properties of the surface (Fig. 2). During the induction period a minimum quantity of material accumulates on the surface and so the performance of the system does not alter significantly. However, during this phase the surface is effectively being conditioned as the surfaces become covered with a thin layer of the foulant material which then enables rapid fouling to take place.

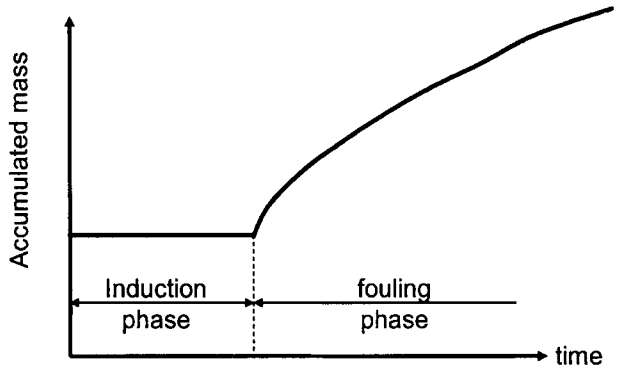


Fig.2. Schematic of two stage scaling process

The fouling phase is normally described in terms of a mass transfer expression coupled with a chemical reaction and a number of well established theories exist which accurately predict growth rates during the fouling phase [6] and so will not be expanded on here. The development of scale during the induction period is a balance between deposition and removal of material at the solid liquid interface. Deposition is typically viewed as a heterogeneous nucleation process where the roughness on the surface of the material about to be scaled acts as nucleation sites. Consequently theoretical expressions are generated by adaptation of classical homogeneous nucleation theory such that the process is viewed in terms of the free energies associated with the precipitation process. In homogeneous nucleation this is expressed by:

$$\Delta G = \Delta G_s + \Delta G_v \tag{1}$$

where the overall excess free energy, ΔG, is equal to the sum of the excess free energy between the surface or the nuclei and the bulk of the particle (ΔG_S), and the excess free energy between a large particle and the solute in solution (ΔG_V). Generally, ΔG_S is positive and proportional to the radius of the nucleus squared, r^2, and ΔG_V is negative and proportional to r^3. Thus:

$$\Delta G = 4\pi r^2 \gamma + \tfrac{4}{3}\pi r^3 \Delta G_v \tag{2}$$

where ΔG_V is the free energy change of the transformation per unit volume and γ is the interfacial tension which represents the free energy change when the surface area of a medium is increased by unit area. The difference in dependency on radius causes the right hand side to pass through a maximum indicating a minimum stable nucleus size (r_c) below which it will redissolve:

$$\Delta G_{crit} = \frac{-4\pi\gamma_c^2}{3} \tag{3}$$

Heterogeneous nucleation requires a suitable surface to provide alternative, pre existing, nucleation sites which permit nucleation to occur with a lower overall free energy change than for homogeneous nucleation such that:

$$\Delta G_{Crit}^{'} = \phi \Delta G_{crit} \tag{4}$$

where ϕ tends from zero to unity depending on the affinity of the surface. The interfacial tension term includes three components denoted by γ_{cl} (between the solid crystal and the liquid), γ_{sl} (between the solid surface and the liquid) and γ_{cs} (between the crystal and the solid surface) (Fig. 3). Resolving these forces in a horizontal directions yields the classical term:

$$\cos\theta = \frac{\gamma_{sl} - \gamma_{cs}}{\gamma_{cl}} \tag{5}$$

where θ is the contact angle between the crystalline deposit and the solid surface. The heterogeneous factor (ϕ) can be expressed in terms of the contact angle by:

$$\phi = \frac{(2+\cos\theta)(1-\cos\theta)^2}{4} \tag{6}$$

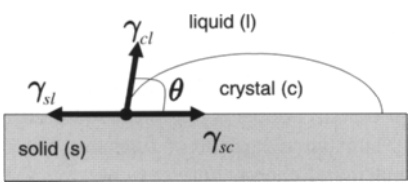


Fig.3. Interfacial energies at the boundary of the three phase contact point

Consequently, the contact angle is generally considered a measure of the affinity of the deposit phase with the substrate surface. Large contact angles indicate a low affinity and conversely small contact angles indicate a high affinity. The rate of nucleation can be linked to the critical number of nuclei formed per unit time and volume (J^{het}) by expression of the form [6]:

$$J^{het} = C_2 \exp\left(-\frac{\Delta G_{crit}^{het}}{k_B T}\right) \tag{7}$$

Under the conditions that normally apply the rate of nucleation is directly proportional to the mass deposition rate during the induction period:

$$\dot{m}_d = C_3 \exp\left[-\phi \frac{16\pi}{3} \frac{\gamma_{cl}^3 v^2}{k_B^3 T^3 \ln^2\left(c/c_s\right)}\right] \tag{8}$$

Where v is molecular volume, k_B is the Boltzmann constant, C_2 and C_3 are proportionality constants and c is the concentration of the precipitating material in the bulk and at the surface. The above equation provides a relationship between the mass deposition rate and the heterogeneous factor which demonstrates the importance of the interfacial energies on the initial deposition rate. Since γ_{cl} is fixed for any given liquid and γ_{sc} and θ result from thermodynamic equilibrium, one way to influence the deposition rate is to adjust the interfacial free energy between the substrate material and the liquid phase (γ_{sl}).

The work of adhesion in a liquid between two solid surfaces, also known as the free energy of adhesion (ΔG_{adh}) is related to the surface free energies of the system according to:

$$W_{slc} = \gamma_{sl} + \gamma_{cl} - \gamma_{sc} = \gamma_{cl}(1 + \cos\theta) \tag{9}$$

This is the generalised form of the well known Young-Dupré equation. Traditionally the surface energy has been expressed as the sum of the components due to the combination of dispersive (γ^{LW}) and polar (γ^{P}) forces [11]. The approach has recently been adapted to more clearly relate to the chemical nature of the phase by using a Lewis acid-base (AB) term instead of the polar component [12] such that the total surface energy can be described by:

$$\gamma_{sl} = \gamma_{sl}^{LW} + \gamma_{sl}^{AB} \tag{10}$$

The role of acid-base interactions in the interfacial behaviour of condensed phases is now largely recognised and is the preferred approach adopted by most researchers [13]. The definition of acids and bases in this regard is very general in that acids are considered as any substance available to accept electron density (e.g. organometallic compounds or electrophilic carbon) and bases being any substance capable of furnishing electron density (e.g. oxygen or aromatic rings) [14]. The interfacial energy between the two phases is expressed in terms of the two components of each phase known as the combining rule [15]:

$$\gamma_{sl}^{LW} = \left(\sqrt{(\gamma_s^{LW} \gamma_l^{LW})}\right)^2 \tag{11}$$

and

$$\gamma_{sl}^{AB} = 2\left(\sqrt{\gamma_s^+} - \sqrt{\gamma_l^+}\right)\left(\sqrt{\gamma_s^-} - \sqrt{\gamma_l^-}\right) \tag{12}$$

Combining the component part enables the work of adhesion between the three phases to be described in terms of the surface energies of each component in relation to its dispersion and acid base properties [16]:

$$W_{slc} = 2\begin{bmatrix} \sqrt{\gamma_s^{LW}\gamma_l^{LW}} + \sqrt{\gamma_c^{LW}\gamma_l^{LW}} - \sqrt{\gamma_s^{LW}\gamma_c^{LW}} - \gamma_l^{LW} + \\ \sqrt{\gamma_l^{+}}\left(\sqrt{\gamma_s^{-}} + \sqrt{\gamma_c^{-}} - \sqrt{\gamma_l^{-}}\right) + \\ \sqrt{\gamma_l^{-}}\left(\sqrt{\gamma_s^{+}} + \sqrt{\gamma_c^{+}} - \sqrt{\gamma_l^{+}}\right) - \sqrt{\gamma_s^{+}\gamma_c^{-}} - \sqrt{\gamma_s^{-}\gamma_c^{+}} \end{bmatrix} \quad (13)$$

The liquid components of Eq. 13 are well known and easily accessible in the literature (Table 1). The components for the solids (both the crystal and the surface under investigation) have to be determined through a series of contact angle experiments with the liquids described in Table 1. Contact angles are measured as the tangent to the drop at the three phase boundary point between the solid, liquid and gas phases at the moment when the drop stops advancing. As described by van Oss [16] only the advancing contact angle should be used in such analysis as it guarantees that the line at the liquid-solid-gas border has not previously been wetted by the drop. Today the process is conducted in devices such as drop shape analysers which automatically determine the contact angle through image processing and convert it to a surface energy. The components of the surface energies are resolved through a simplified version of the expanded Young-Dupré equation:

$$\frac{1}{2}(1+\cos\theta)\gamma_L = \sqrt{\left(\gamma_s^{LW}\gamma_l^{LW}\right)} + \sqrt{\left(\gamma_s^{+}\gamma_l^{-}\right)}\sqrt{\left(\gamma_s^{-}\gamma_l^{+}\right)} \quad (14)$$

Table 1
Surface energy parameters ($mJ.m^{-2}$) for standard wetting liquids (adapted from [16, 17])

Liquid	γ_l	γ^{LW}_l	γ^{-}_l	γ^{+}_l
Methylene iodide (MI)	50.8	50.8	0	0
1-bromonapthalene (BRMN)	44.4	44.4	0	0
Water	72.8	21.8	25.5	25.5
Formamide (F)	58	35.6	39.6	2.28
Ethylene glycol (EG)	48	29	47	1.92
Glycerol (G)	64	34	57.4	3.92
Chloroform (C)	27.1	27.1	3.8	0

The choice of appropriate wetting agents is still under discussion [15, 18, 19] but initially an apolar liquid is used such as methylene iodide or 1-bromonapthalene as they have no acid base components. This enables the dispersive component to be measured directly from the contact angle and the surface tension of the apolar liquid:

$$\gamma_s^{LW} = \gamma_l \frac{(1+\cos\theta)^2}{4} \quad (15)$$

This is then followed by using two polar liquids which enables the equations to be solved simultaneously. The polar liquids are generally taken from water and then one of: formamide, glycol, ethylene glycol, dimethylsulfoxide, chloroform. The choice of liquids is usually discussed in terms of the condition number of the matrix solution to provide as robust a solution as possible [14, 19]. Based on this approach suitable combinations of liquids appear to be water with either ethylene glycol, bromonapthalene, diiodomethane or formamide.

McCafferty [17] offers another method by dividing Eq. 14 by either $\sqrt{\gamma_l^-}$ or $\sqrt{\gamma_l^+}$ such that the LHS of the equation contains only known parameters, by either measurement or tabulated sources. By using data from a number of different wetting liquids a linear graph can be plotted with a slope of $\sqrt{\gamma_s^+}$ and an intercept of $\sqrt{\gamma_s^-}$ (Fig. 4.). Reported use of this approach indicates that dividing by $\sqrt{\gamma_l^-}$ provides the most reliable results as it spreads the data points more evenly on the graph. Comparison with the matrix approach indicates broad agreement and provides a more interactive approach.

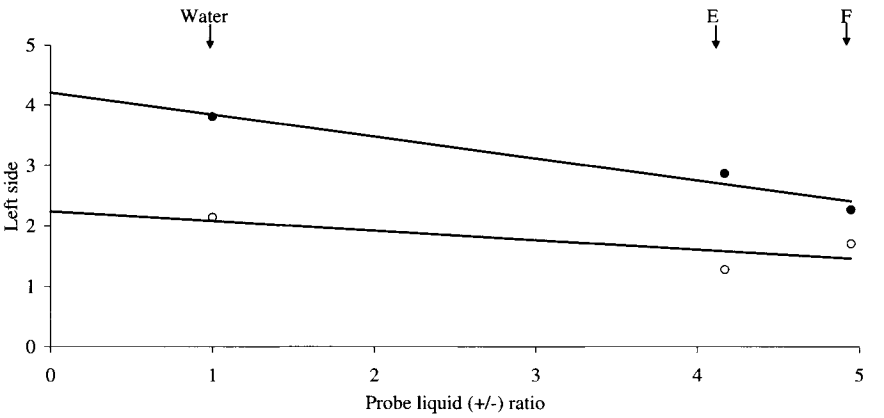


Fig.4. Example plot to determine the acid base component of the surface energy of the test material

$$\text{x-axis} = \sqrt{\left(\frac{\gamma_L^+}{\gamma_L^-}\right)}, \text{ y-axis: } \frac{\frac{(1+\cos\theta)\gamma_L}{2} - \sqrt{\left(\gamma_s^{LW}\gamma_L^{LW}\right)}}{\sqrt{\left(\gamma_L^+\right)}}$$

3. ATOMIC FORCE MICROSCOPY

The development of the scanning probe microscope (SPM) has provided an alternative approach to characterise surface properties of materials and as such is becoming an important tool for studies into adhesion. SPMs are a family of techniques that utilise a range of physico-chemical interactions between a tip and a surface to image surface topography of both insulating and conductive material in both gaseous and liquid environments without the need for any additional surface preparation. Perhaps the most important member of the family with regards to the current discussion is the atomic force microscope (AFM) which can generate topographical and direct force measurements of a surface. AFM has been successfully used to image a wide variety of surfaces such as semiconductors, biological systems and polymers as well as investigate inter and intra molecular forces including adhesion phenomena. A number of good reviews and texts exist which provide detail on the equipment and its application such as Leite and Herrmann [20], Ohnesorge and Binning [21]and Braga and Ricci [22]. AFM is also described in Chapter 12, so only a relatively brief description will be provided here.

The basic components of an AFM include a very sharp tip, flexible cantilevers, a sensitive deflection sensor and a piezoelectric scanner to move the tip/sample (Fig. 5). The tip

and cantilever are commonly manufactured from silicon or silicon nitride due to the uniformity of production and the reproducibility that it generates. AFMs usually resolve the deflection of the cantilever down to picometer resolution by either using optical lever or beam bounce methods. In the latter, a laser beam is reflected from the top of the cantilever onto a photo detector of at least two photodiodes. Movement of the cantilever tilts the reflected beam and the difference between the two photodiodes determines the position of the cantilever. Normally a feedback circuit is applied between the detector and the scanner so that the cantilever deflection remains constant.

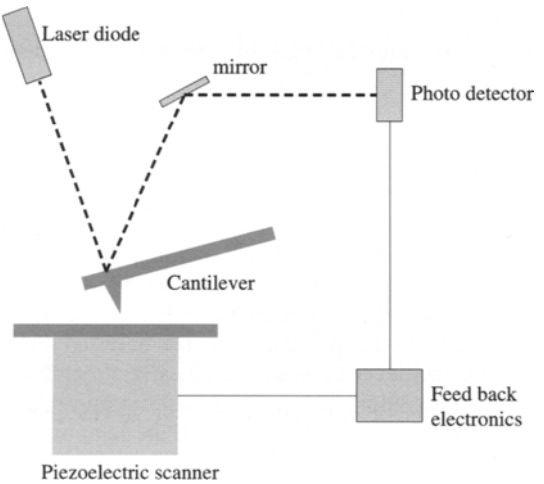


Fig.5. Schematic of an AFM

There are three fundamental modes of operation to generate an image of the topography of the surface: static or contact mode, intermittent contact (tapping), and non contact mode. In all cases, the tip interacts with the surface as a result of the different force components of the interaction between the two surfaces (Fig. 6). In contact mode, the tip is drawn across the surface much like a record player stylus and provides a lateral resolution typically of 0.2-0.3 nm down to true atomic resolutions [21]. As the tip is in contact, short range forces control the response of the system such as van der Waals in gases and Coulomb and structural forces in liquids [20]. In tapping mode the cantilever oscillates as it is scans over the surface. Perturbations in the amplitude are monitored and related back to the position of the surface. Non contact mode extends this by never actually touching the surface. In this case the resonating tip interacts with the relatively long range forces which cause a change in the amplitude of the oscillation and hence the height of the surface. The tip is scanned across the surface such that a 3 dimensional image of the surface can be generated and the data analysed to provide a detailed topographical account of the scanned area.

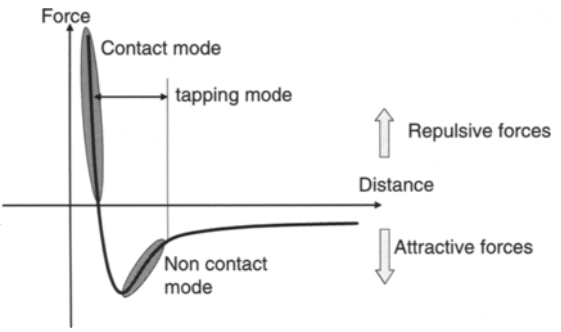


Fig.6. Force vs. distance plot showing the types of interactions used during the imaging mode operation

When using the microscope in force-distance mode the tip is brought towards the surface under investigation and the deflection of the cantilever is monitored and directly related to the strength of the interaction between the tip and the surface (Fig. 7). At some point as the tip approaches the surface the total force acting on the tip exceeds the stiffness of the cantilever and so the tip will jump into contact with the surface (point *b-c*). At point *d* the tip is in contact with the surface with the slope of the line providing an indication of the hardness of the sample. The tip is then retracted from the surface with the tip remaining in contact beyond its original position due to adhesive bonds between the two surfaces. At point *e* the bending moment of the cantilever exceeds the adhesive strength and so the tip snaps off sharply. Differences between the approaching and retracting portions of the curve whilst in contact provide information of the plastic deformation of the sample [23].

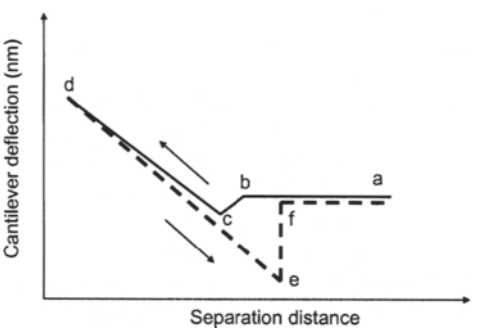


Fig.7. A typical cantilever displacement vs separation distance plot as obtained by AFM

The most interesting part of the curve in relation to the current topic is the jump off portion which is a direct measure of the adhesion force between the tip and the surface. The force is calculated according to Hooke's law by multiplying the maximum vertical deflection

by the spring constant of the cantilever. The minimum level of interaction that can be detected is thus limited to the spring constant of the cantilever in combination with the sensitivity of the photdector. Commercial cantilevers can be purchased with spring constants in the range of 0.03Nm to 30Nm, giving a range of detectable forces in the order of 10^{-14}N to 10^{-9}N. These values for forces are of the same order as van der Waals and similar inter species/inter atomic forces [24, 25]. The adhesive force measurements can be converted to a work of adhesion based on the theory of sphere-planar interactions derived by either Johnson, Kendall and Roberts [JKR] or Derjaguin, Muller and Toporov [DMT] [20]:

$$W_{adh} = \frac{F_{adh}}{-\alpha R} \tag{16}$$

Where R is the tip radius and α is a constant with values between $(3/2)\pi$ for soft materials [DMT] and 2π for hard materials [JKR]. Thus the AFM can be used to directly determine the work of adhesion for a given sample albeit in relation to the material of the tip. Researchers are attempting to overcome this by functionalizing tips, known as colloid probes, with materials of interest. Examples to date include: proteins, minerals, biological species, natural organic matter, DNA strains [20, 26] (Fig. 8). An important sub set of this is the use of modified tips with specific chemical species, known as chemical force microscopy, which has been used extensively to investigate molecular level forces [27]. The most common method of generating the functionalized tip is through the formation of thio self assembled monolayers on gold coated tips [28]. This has been very important in the development of the understanding of the role and management of conditioning films during biofilm formation [29, 30].

4. CASE STUDIES

4.1. Introduction

The approach outlined above has been used to assess the impact of surface properties on the rate of calcium carbonate scaling. Ten different surface finishes were assessed and compared to scaling rate measurements performed in a simple scaling test. The materials investigated included polymer and metal surfaces with low and high surface energies. Three metal surfaces were included, copper, aluminium and stainless steel. The first two are the most commonly used materials in the manufacture of showers and boilers with stainless steel used as both a benchmarking material and a surface to test how changes in roughness impacted on the scaling rate. The stainless steel was treated with three surface finishes, polished, sand blasted and mechanically roughened with sand paper. The other materials tested were a commercial kettle coating (KC), polytetrafluoroethylene (PTFE), titanium nitride (TiN), two types of diamond like carbon (DLC) coatings (*Graphit-iC* and *Dymon-iC*) and two sets of gold coating (0.1 and 0.3 μm thickness).

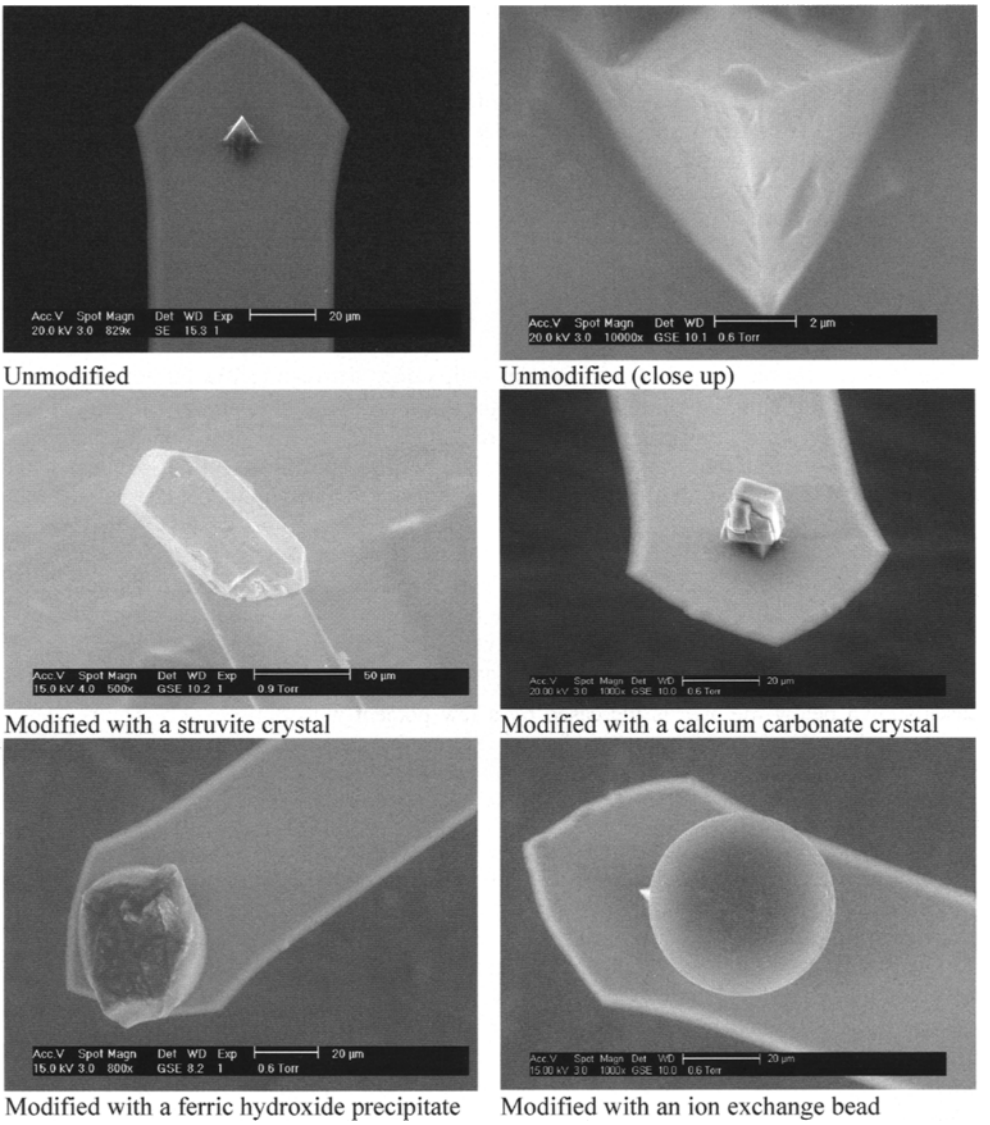


Fig. 8. Examples of unmodified and modified AFM tips

Details of the specifics of the experiments are available in the literature [31] and so will be covered only briefly here. Scaling rate was determined in a rapid scaling test rig which consisted of a 1 L tank, a submerged heating element, covered by a removable sleeve and a temperature controller. Scaling rate was measured by conducting 10 heating and cooling cycles which consisted of heating the test solution to 70°C for 45 minutes and then allowing it to cool for 15 minutes. After 10 complete cycles the sleeve was removed and acid soaked to remove all the deposited scale and analysed through ICP-AES measurement. Topographical and adhesive force measurements were performed on a Dimension D3000 AFM (Vecco

Instruments, CA). Force measurements were performed on ten separate sites on each surface with at least ten repeats at each point.

Modified tips were produced by introducing a cleaned cantilever into a droplet of synthetic hard water which had been pH adjusted as to precipitate out the excess calcium. After one hour the tip was removed and excess calcite removed by carefully contacting the unwanted material with a drop of acid.

4.2. Results

4.2.1. Topographical

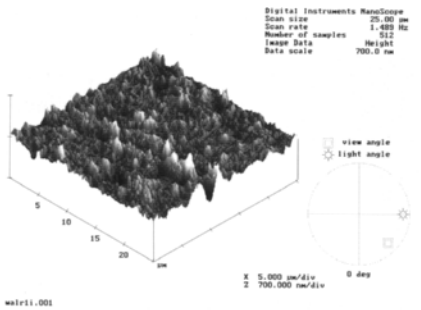
The measured roughness level of the materials ranged from relatively smooth surfaces such as DLC (RMS = 24 nm) and stainless steel (RMS = 13.6 nm) to relatively rough surfaces such as the kettle coating (RMS = 97 nm) and the roughened stainless steel (RMS = 292 nm) (Table 2). The roughness of the surface is important as it impacts on the available surface area for scale formation and potentially provides crevice environments suitable for nucleation. The latter has been successfully shown as the key to bubble formation [32] such that the kettle coating offers a 'mountain range' like surface, relative to the base roughness, such that bubble formation can be enhanced to reduce the noise generated during boiling. The majority of polymer or coated surfaces offer a regular roughness such that the peak height is relatively small such as in Dymon and PTFE (Fig. 9). Whereas, metals often demonstrate greater deviation across the surfaces especially when treated to some form of finish as shown in the cases of the mechanically roughened and sand blasted stainless steel samples (Fig.9).

The impact of the roughness on the available area for scale formation can be determined by calculating the percentage increase due to the topographical features (Table 2). In most cases shown below the area is increased by a few percent with the exception of the kettle coating and the mechanically roughened steel where a 20% increase in area is observed. This is due to the combination of a micro mountain range-like structure superimposed with a nano scale roughness. An important observation is that RMS values do not always match the increase in area generated by the topographical features due to differences in features and so the combination of parameters is often useful in describing topographical features.

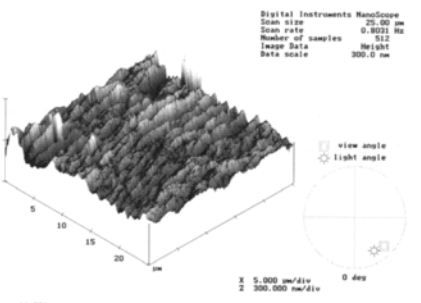
Table 2.
Summary of topographical data for the different surfaces.

Material	RMS (nm)	Surface area enhancement[1] (%)
Aluminium	24.26	103.0
Copper	19.95	102.1
DLC (Dymon)	24.07	104.0
DLC (Graphite)	35.77	101.6
Gold (0.1)	-	101.4
Gold (0.3)	11.35	100.8
Kettle coating	97.53	120.1
PTFE	28.11	101.1
SS (smooth)	13.61	101.2
SS (rough)	217.43	123.3
SS (coarse)	292.74	106.3
TiN	39.82	102.0

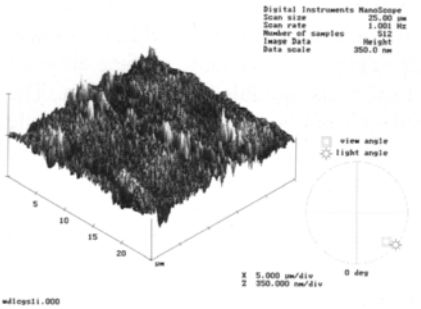
[1]when compared to a completely flat surface of the same dimensions



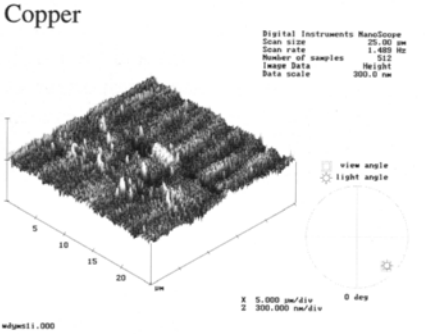
Aluminium



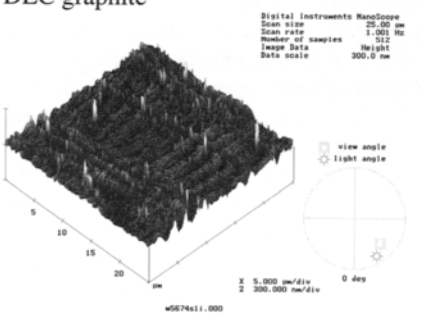
Copper



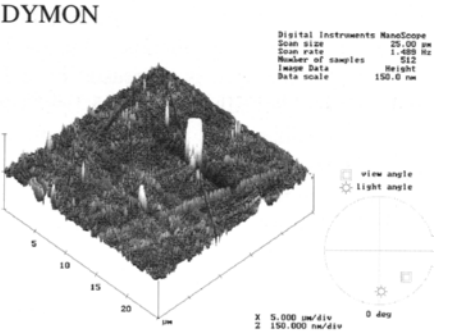
DLC graphite



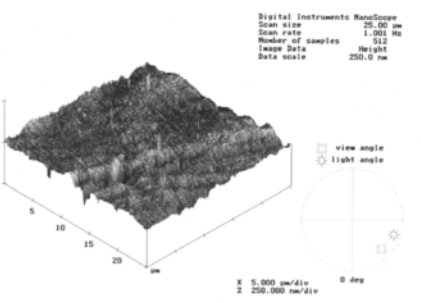
DYMON



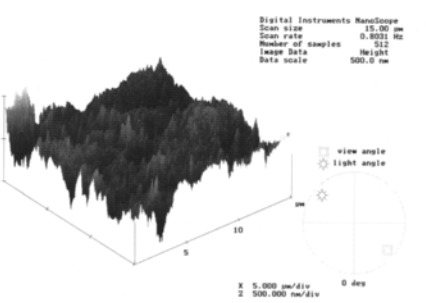
Gold coating (0.1 µm)



Gold coating (0.3 µm)



PFTE



Kettle coating

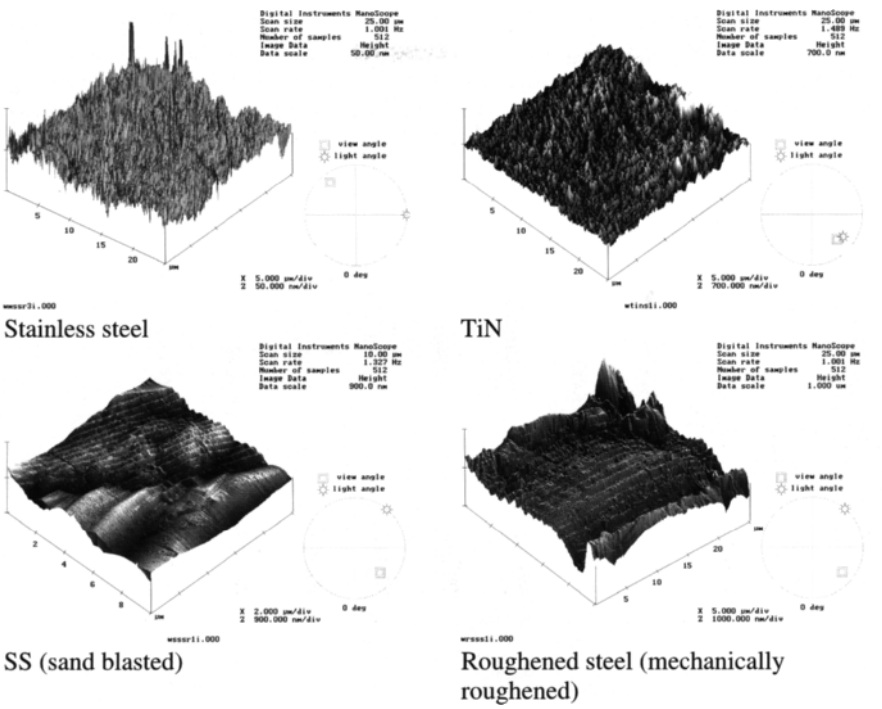


Fig. 9. Example surface topographical features of the different materials studied

4.2.2. Surface energy

The surface energy of the different materials was assessed in two ways. Initially surfaces were analysed through the traditional contact angle approach where data is obtained in different liquids to resolve dispersive and acid-base components of the surface energy of each surface which is then converted to a work of adhesion according to Eq. 13. In addition, the energy was measured with the AFM by taking 50 readings of adhesion force in ten locations on each surface, and these were converted to a work of adhesion according to Eq. 16.

Contact angle measurements in 1-bromonaphtalene were similar for most of the materials tested with average values between 18° and 23° (Fig. 10) with the exception of alum and k-Coating where angles over 35° were observed. Consequently, the dispersive component of the surface energy is lower for these last two surfaces (Table 3). The acid-base components of the surface energy were determined according to the graphical approach described by McCafferty [17]. The contribution of the acid-base component to the surface energy of the different materials varied from a low of 5-6 % for aluminium, gold and kettle coating samples to a high of 25-30% for the stainless steels. In line with similar reported observations, the basic component dominates especially for polymer surfaces (Table 3). This is due to the reference value for water used in the original Good-van OSS-Chaudhury (GvOC) theory and is discussed in detail in papers such as Della Volpe and Siboni [14]. The use of the acid-base component permits better comparison between materials but it must be remembered that direct comparison of the acid and base components of an individual material has no real meaning. Similar contact angle data and surface energy components for the materials are reported elsewhere although general variation between studies is common [18, 33, 34]. It is

important to note that the surface energy of the metal surfaces are significantly smaller than the theoretical values of 1000s mJ.m^{-2} that are predicted for fresh metal surfaces [35, 36]. This is due to a combination of effects such as oxide layer formation [37], CH_n- contaminating hydrocarbon layers formed when exposed to air [38] and contamination with carbon [39] which significantly reduces the surface energy.

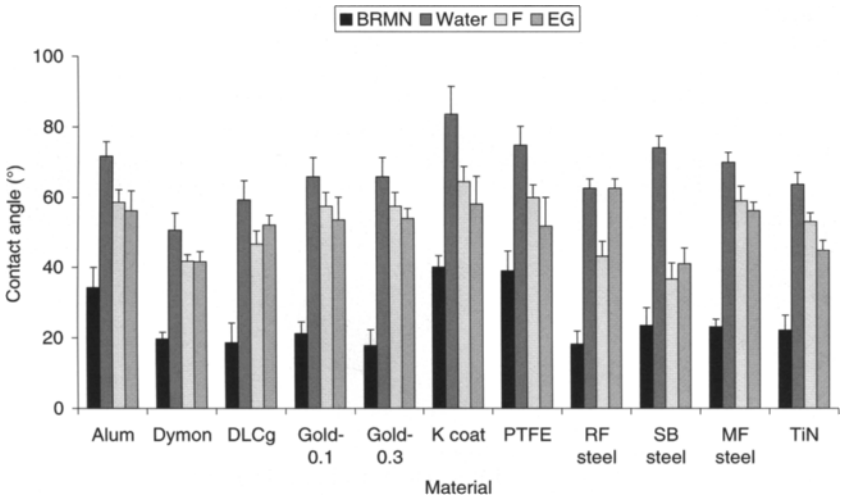


Fig.10. Contact angle data for the different materials

Table 3.
Summary of the surface energy components of the different materials by contact angle measurement

Material	γ^{LW} (mJ.m^{-2})	γ^{+} (mJ.m^{-2})	γ^{-} (mJ.m^{-2})	γ^{AB} (mJ.m^{-2})^{1}	γ^{total} (mJ.m^{-2})^{2}	W_{132} (mJ.m^{-2}) [x10^{-2}]
Aluminium	36.91	0.08	14.15	2.12	39.03	1.23
DLC (Dymon)	41.82	0.4	36.15	7.60	49.42	0.65
DLC (Graphite)	41.94	0.56	21.59	6.95	48.89	0.85
Gold (0.1)	41.39	0.22	18.2	4.00	45.39	1.23
Gold (0.3)	42.19	0.12	16.28	2.79	44.98	1.18
Kettle coating	34.56	0.74	5.82	2.07	36.63	2.85
PTFE	34.96	0.98	9.72	6.17	41.13	2.24
SS (smooth)	40.88	0.4	17.71	5.32	46.20	1.18
SS (rough)	42.13	5.47	15.31	18.7	60.83	2.60
SS (coarse)	40.67	6.19	7.79	13.88	54.55	3.76
TiN	41.09	0.02	16.74	1.15	42.24	1.00

$^{1}\ \gamma^{AB} = 2\sqrt{\gamma^{+}\gamma^{-}}$; $^{2}\ \gamma^{total} = \gamma^{AB} + \gamma^{LW}$

The alternative approach of using an AFM to determine adhesion force is shown in Fig. 11. Adhesion force values range from 270 to 4553 pN for unmodified tips and from 524 to 5433 pN for a modified tip with calcium carbonate crystals grown directly onto the cantilever.

The modified tips were calibrated according to the method of Cleveland et al. [40] which resulted in a spring constant of 0.0502 $N.m^{-1}$. Comparison of the data reveals broad agreement between the two tips with the modified tip generating larger adhesion forces with respect to most materials tested. The variation observed within each material reflects the surface heterogeneity in terms of topographical features and surface composition observed at the scale of investigation used here. More detailed analysis of the adhesion measurements made on each material suggests the level of variation to be due to the specific location chosen in terms of peak and trough portions of the structure. Peaks are known to produce local charge centres and troughs provide the possibility of interactions between the sides of the tip and the trough. Consequently, surfaces that generate roughness evenly across the surfaces through small peak heights such as Dymon and PTFE show much less variation in adhesive force compared to surfaces with more irregular roughness and corresponding higher peak heights such as TiN and alum (Fig. 12). The high variability in the metal samples is generally attributed to the above in combination with variation in the surface composition and localised contamination which is difficult to avoid in applied studies [35].

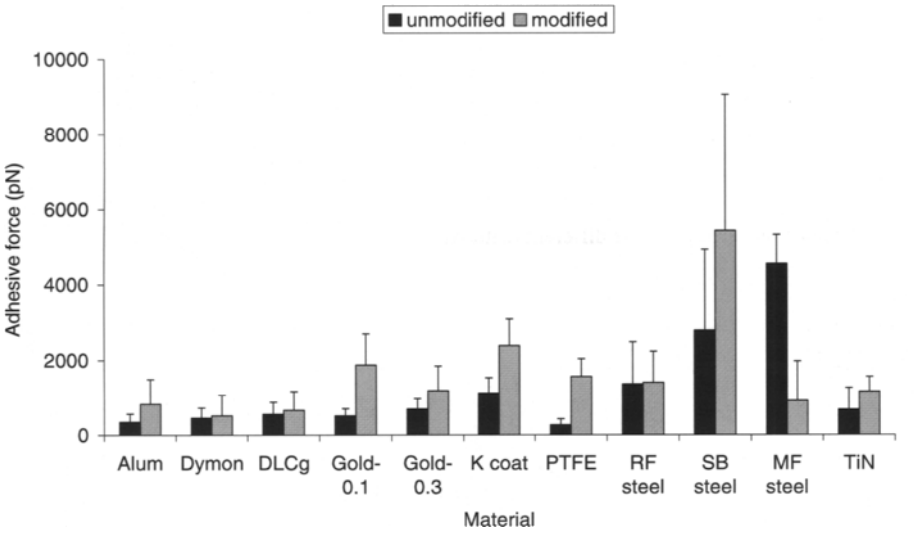


Fig.11. Adhesive force measurements for the different materials based on liquid testing with modified and unmodified cantilevers

Comparison of the work of adhesion calculated from the two methods indicates broad agreement (Fig. 13) with the straight line representing complete agreement between the two approaches. In general, the contact angle approach tends to estimate a lower work of adhesion compared to the AFM. However, when considering the observed variation in the approaches the AFM appears to provide a suitable method for rapid determination of the work of adhesion of a surface and more importantly in the ranking of potential materials. The figure presented is based on an unmodified tip as modified tips present irregular geometries which increase the complexity of the conversion between force and work. The importance of this is in the potential for the AFM to be used as a screening tool for surface selection and modification. Similar approaches have been demonstrated with bacterial adhesion studies in medical applications [22] and wastewater treatment.

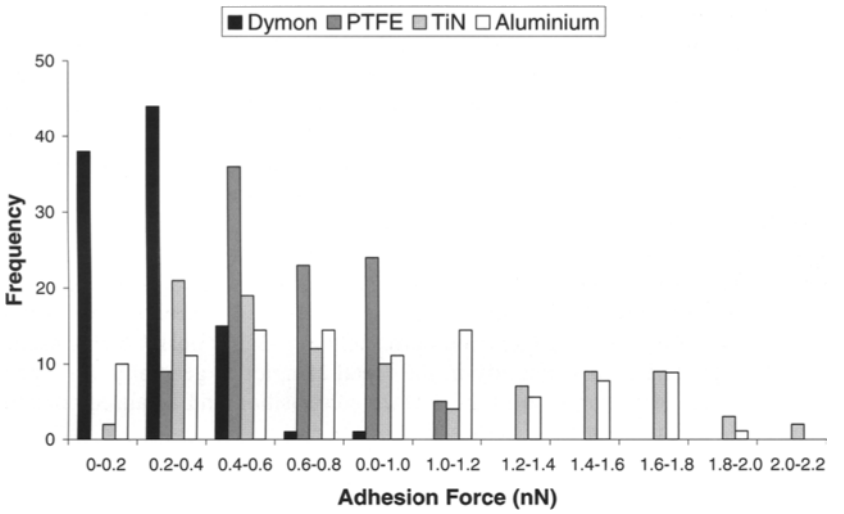


Fig.12. Frequency plots of adhesion force for examples of narrow and broad distributed materials

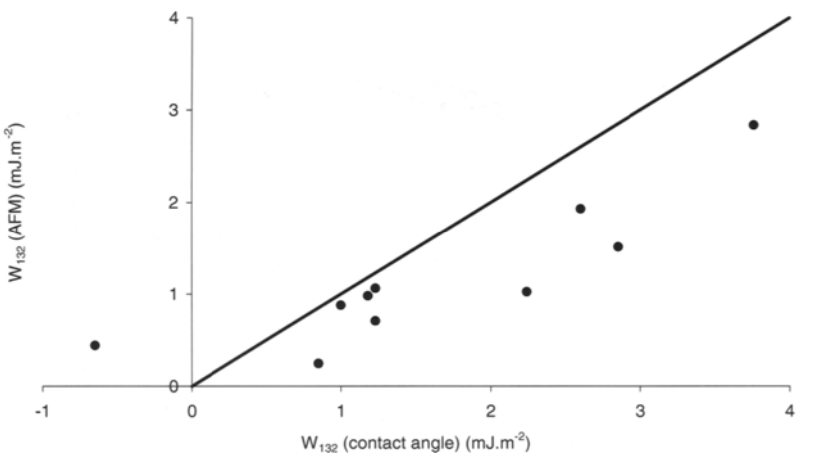


Fig.13. Comparison of the calculated work of adhesion from contact angle and AFM data (unmodified tip)

4.2.3. Correlations

The work of adhesion between calcite and the different materials shows a clear trend to the rate of scaling (Fig. 14). with higher scaling rates observed at larger values of the work of adhesion. Comparison between the two surface characterisation approaches reveals broad agreement with the AFM approach generating a steeper relationship between the two parameters and less scatter around the regression line. The relationship between scaling rate and adhesion force follows a similar line and for practical purposes represents the information match that is required to select materials for specific applications. Modification of the tip

enhanced the range of forces observed and interestingly suggests that there may be a limiting level beyond which surface energy is no longer as important (Fig. 15). Variation in the data at similar levels of either adhesion force or works of adhesion appear to be related to the variation observed within the surface samples with the data point furthest from the regression line representing those point with the highest level of variation seen in Fig. 11.

Attempts to correlate surface roughness, as either RMS or surface area enhancement, generated no overall correlation indicating that the surface energy of the system dominates over issues of roughness in determining the likely rate of scaling that may be observed. However, when considering a single material, the roughness correlates well with scaling rate. To illustrate, in the case of the stainless steel scaling, rates of 8.3, 14.1 and 18.81 $g.m^{-2}.h^{-1}$ were observed for RMS values of 13.61, 217 and 292 nm respectively. Similar conclusions have been reported for a similar system investigating the role of surface properties on struvite formation rate [41].

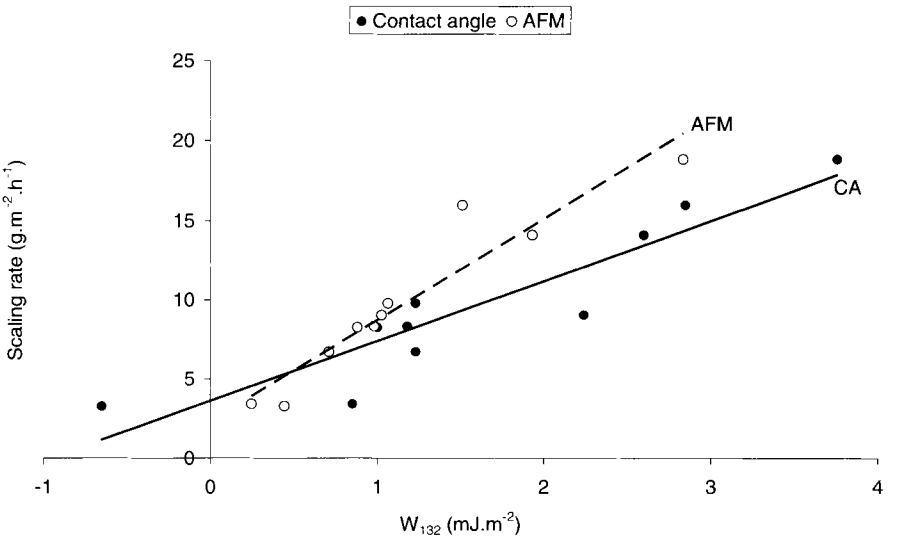


Fig. 14. Comparison of the relationship between work of adhesion and scaling based on both contact angle and AFM data

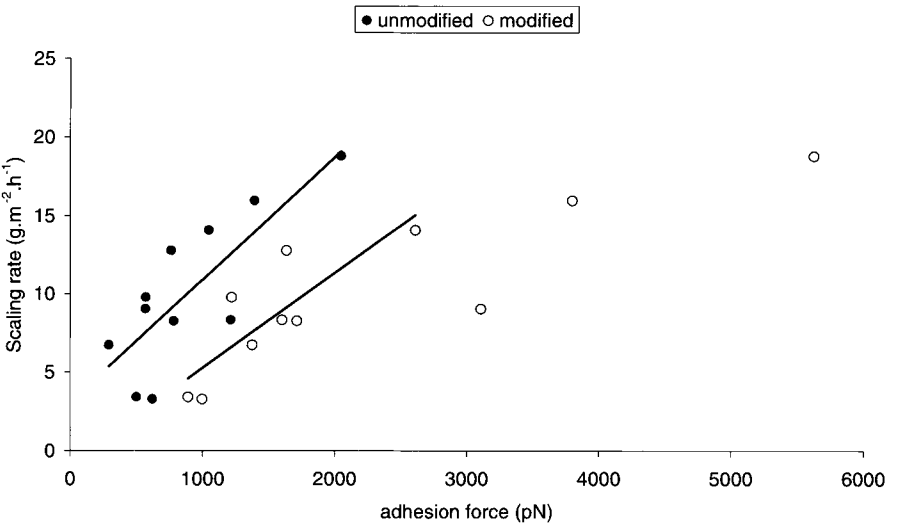


Fig.15. Correlation between adhesion force and scaling rate

The above case study demonstrates the potential for colloid and interface science to assist in the understanding of fouling and ultimately help develop the next generation of surfaces. Critically, the scaling rate of the system appears to be dominated by the work of adhesion between the crystal and the solid in water rather than topographical features. Consequently, the development need is towards reducing the work of adhesion through generating low energy surfaces, delivered uniformly across the material with as small a variation in properties as possible. The application of the AFM is crucial in this regard as it provides a relatively easy and rapid method of surface characterisation which can be used to assess overall properties as well as local variations. The application of tip modifications continues to provide potential to enhance the selectivity of the assessment made by the AFM but many issues remain which must be overcome before such an approach can be consider routine for industrial applications.

REFERENCES

[1] J. MacAdam and S.A. Parsons, Rev. Environ. Sci. Biotechnol.., 3 (2005) 159.
[2] J.D. Doyle, K. Oldring, J. Churchley and S.A. Parsons, Water Res., 36 (2002) 3971.
[3] D. McNaughtan, I.G. Winning and M. Sim, The evaluation of a new environmentally friendly combined product for application in the North Sea for effective scale and corrosion control, in: Proc. SPE Annual Technical Conference and Exhibition, Houston, USA, 2004.
[4] S. Keysar, R Semiat, D. Hasson and J. Yahalom, J. Colloid Interface Sci., 162 (1994) 311.
[5] H. Müller-Steinhagen, Q.Zhao, A. Helali-Zadeh and X.-G. Ren, Can. J. Chem. Eng., 78 (2000) 12.
[6] M. Förster and M. Bohnet, Int. J. Therm. Sci., 38 (1999) 944.
[7] H. Müller-Steinhagen and Q. Zhao, Chem. Eng. Sci., 52 (1997) 3321.
[8] Q. Yang, J. Ding and Z. Shen, Chem. Eng. Sci., 55 (2000) 797.
[9] Q. Yang, J. Ding and Z. Shen, Chem. Eng. Sci., 57 (2002) 921.
[10] A. Neville and A.P. Morizot, Chem. Eng. Sci., 55 (2000) 4737.
[11] F.M. Fowkes, Ind. Eng. Chem., 54 (1964) 40.

[12] C.J. van Oss, R.S. Good and M.K. Chaudhury, Chem. Rev., 88 (1988) 927.
[13] F.M. Fowkes, J. Adhes. Sci. Technol., 1 (1997) 7.
[14] C. Della Volpe and S. Siboni, J. Adhes. Sci. Technol., 14 (2000) 235.
[15] R.J. Good, J. Adhes. Sci. Technol., 6 (1992) 1269.
[16] C.J. van Oss, J. Adhes. Sci. Technol., 16 (2002).
[17] E. McCafferty, J. Adhes. Sci. Technol., 16 (2002) 239.
[18] C. Della Volpe and S. Siboni, J. Colloid Interface Sci., 195 (1997) 121.
[19] E. McCafferty, and J.P. Wightman, J. Adhes. Sci. Technol., 13 (1999) 1415.
[20] F.L. Leite, and P.S.P. Herrmann, J. Adhes. Sci. Technol., 19 (2005) 365.
[21] F. Ohnesorge and G. Binnig, Science, 260 (1993) 1451.
[22] P.C. Braga and D. Ricci (eds.) Atomic Force Microscopy: Biomedical Methods and Applications. Humana Press, New Jersey, USA, 2004.
[23] D.H. Gracias and G.A. Somorja, Macromolecules, 31 (1998) 1269.
[24] J.H. Hoh, J.P. Cleveland, C.B. Prater, J.P. Revel and P.K. Hansma, J. Am. Chem. Soc., 114 (1992) 4917.
[25] G.U. Lee, D.A. Kidwell and R.J. Colton, Langmuir, 10 (1994) 354.
[26] D. Christendat, T. Abraham, Z. Xu and J. Masliyah, J. Adhes. Sci. Technol., 19 (2005) 149.
[27] D.V. Vezenov, A Noy and P. Ashby, J. Adhes. Sci. Technol., 19 (2005) 313.
[28] C.D. Frisbie, L.F. Rozsnyai, A. Noy, M.S. Wrighton and C.M. Lieber, Science, 265 (1994) 2071.
[29] R. Avci, , M. Schweitzer, R.D. Boyd, J. Wittmeyer, A. Steele, J.Toporski, I. Beech, F. T Acre, B. Spangler, K.M. Cole and D.S. McKay, Langmuir, 20 (2004) 11053.
[30] P.K. Sharma and R. Hanumantha, Adv. Colloid Interface Sci., 98 (2002) 341.
[31] J. MacAdam and S.A. Parsons, Water Sci. Technol., 49(2) (2004) 153.
[32] G. Barker, B. Jefferson and S. Judd, Chem. Eng. Sci., 57 (2002) 565.
[33] C. Jacquot and J. Takadoum, J. Adhes. Sci. Technol., 15 (2001) 681.
[34] S.D Knorr, E.C. Combe, L.F. Wolff and J. S. Hodges, Dental Mater., 21 (2005) 272.
[35] L. Vitos, A.V.Ruban, H.L. Skriver and J. Collar, Surf. Sci., 411 (1998) 186.
[36] N.J. Gedes, E.M. Paschinger, D.N. Furlong, F. Caruso, C.L. Hoffmann and J.F. Rabolt, Thin Solid Films, 260 (1995) 192.
[37] R. Wang, L. Cong and M. Kido, Appl. Surf. Sci., 191 (2002) 74.
[38] M. Vinnichenko, T. Chevolleau, M.T. Pham, L. Poperenko and M.F. Maitz, Appl. Surf. Sci., 201 (2002) 41.
[39] Y. Watanabe, Y. Hara, T. Tozuda, N Kitazawa and Y. Nakamura. Surf. Eng., 16 (2000) 211.
[40] J.P. Cleveland, B. Anczykowski, A.E Schmid and V.B. Elings, Appl. Phys. Lett., 72(1998) 2613.
[41] S. Dunn, S.A. Impey, C. Kimpton, S.A. Parsons, J. Doyle and B. Jefferson, Water Sci. Technol. 49 (2) (2004) 183.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 16: Fate of particles in the distribution system

Jan Vreeburg

Technical University of Delft, Stevinweg, 12628CN Delft.
The Netherlands

1. INTRODUCTION

This chapter deals with the interactions of particles with a variety of surfaces encountered in the distribution system and with their ultimate fate. Particles in the distribution system play an important role in the aesthetics of water quality with the emphasis on discolouration. Of all the water quality issues in the network, discolouration is the most visible to the consumer and leads to the greatest number of complaints (Fig. 1).

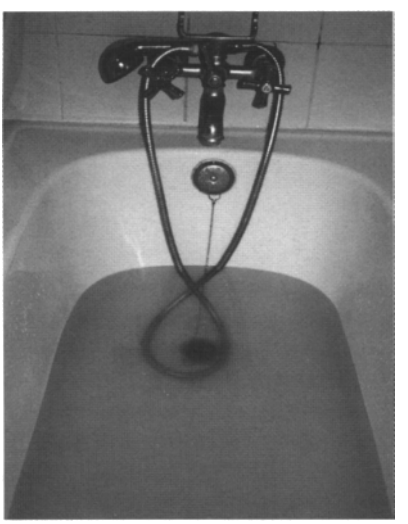


Fig. 1. Example of discoloured water leading to customer complaints

In general the quality of the finished water will change during the journey from the pumping station to the customer's tap. Particles present in the finished water that enter the network will settle, resuspend and then be transported or just roll along the bottom of the pipe. In addition to this the water is in contact with a range of different materials used in the distribution system depending upon its age and location. The materials can be ferrous or cement based or plastic. All these materials can interact with the drinking water affecting the quality.

Not only will the interactions with the materials change the water quality, but biological processes will also have an effect. Regrowth of bacteria results in higher colony counts, and also the development of a biofilm layer on the wall of pipes.

Finally the fluid dynamics in the distribution system influence the water quality in the network. The fluid or hydraulic dynamics in a system are 'caused' by the demand patterns of consumers. A high demand will result in higher velocities and higher sheer stresses, resuspending any settled particles, while periods of lower velocity give opportunities for settable material to actually accumulate on the bottom of the pipes.

The phenomenon of discoloured water is a difficult process to study, because it is often a temporary situation. Fig. 2 gives a typical discolouration event, monitored through the turbidity of the water. Due to a hydraulic incident, in this case cracking a fire hydrant, the velocity in the pipes increases dramatically, resuspending loose particles from the bottom of the pipes and causing the discolouration. After a short while the turbidity level has decreased again following closure of the fire hydrant. The particles resettle and it is impossible to get a sample of the water during the incident.

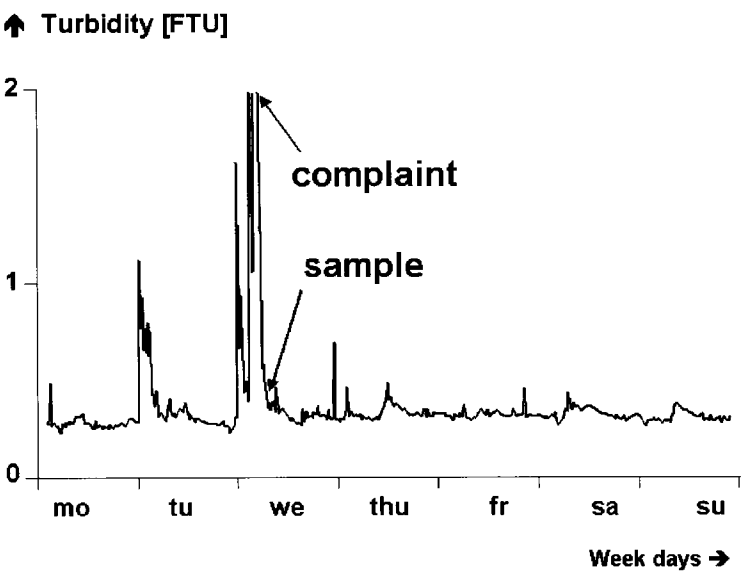


Fig. 2. Typical discolouration event

This chapter deals with an analysis of the origin of the particles, using a mass balance, the assessment of the discolouration risk using the Resuspension Potential Method and the possible countermeasures to deal with the problem such as cleaning protocols or the redesign of the network.

2. MASS BALANCE

2.1. Introduction

The framework for the analysis of the origin and dynamics of particles is a mass balance model for a network. Within the model, the different processes dealing with particles can be

quantified and analysed. The mass balance can be applied to the complete network, but can also be applied to selected parts of the networks.

The model of the mass balance is used to sum up the relevant processes involved in particle input and particle generation in the network. All these processes are interconnected and interdependent. Nevertheless for the sake of this study the processes will be separated as much as possible to simplify the analysis.

Next to the input and the generation of the particles, the kinetics of the particles' interaction are also important because they determine the fate of the particles in the network. Are they just transported through the network and supplied to the customers in low concentration or do they accumulate in the network adding to the mass content of the pipes? The balance between accumulated and mobile particles primarily determines the risk of discolouration.

Fig. 3 represents the mass balance schematically. The common discolouration event is a result of a high concentration of (coloured) particles in the water, disturbing the clarity to an extent that people notice it. Thus discolouration is the result of several processes in the water involving particles.

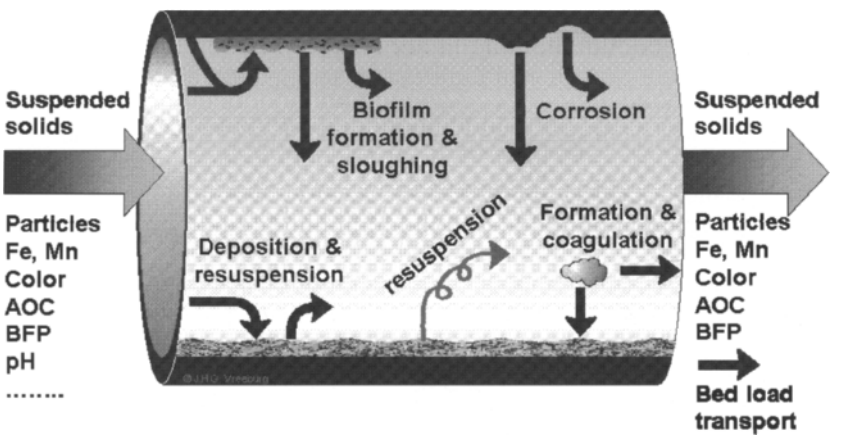


Fig. 3. Mass balance model

The basic assumption is that accumulated sediment in combination with a hydraulic disturbance will lead to discoloured water. Consequently managing the transport and accumulation of the particles is the key to minimising of water quality deterioration.

The origin of the accumulated particles is multiple, as demonstrated in the above figure. To quantify the influence of the various processes this chapter is focussed on the concept of a mass balance model over the network or part of the network.

The basic elements of the mass balance model are described with respect to their contribution to the storage of particles. The storage of particles plays an important role in the generation of discolouration problems.

Storage = Mass in – Mass out + Mass produced

The term 'Mass in – Mass out' represents the transport phenomena in the pipes, while the term 'Mass produced' represents the particles generated as a result of biological or chemical

processes such as biofilm growth, corrosion or coagulation, all of which produce mass and particles dependent on the quality of the bulk water.

In formula:

$$\begin{aligned} \Delta\text{Storage} = \; & Q*T*Cp(in) - && \text{(Mass in)} \\ & Q*T*Cp(out) - \text{Bed load out} + && \text{(Mass out)} \\ & CP*\alpha*A*T + BP*\beta*A*T + Q*T(BPP + CPP) && \text{(Mass produced)} \end{aligned} \quad (1)$$

with

Q	Volume flow
T	Time scale used
Cp(in)	Concentration particles in the incoming water in the form of suspended solids, particles etc
Cp(out)	Concentration particles in the outgoing water in the form of suspended solids, particles etc
Bed load out	Transport of particles through bed load transport
A	Inner surface of the pipe
α,β	Factors applied to the active part of the surface giving corrosion products or biological parts respectively
CP	Corrosion production of particles from the pipe wall and released to the water
BP	Biological production of particles from the biofilm and released to the water
BPP	Biological particle production potential in the bulk fluid
CPP	Coagulation particle production potential in the bulk fluid

'Mass in' and 'Mass out' are part of the water quality as represented by the suspended solids, particles etc. entering the control volume. These water quality parameters represent a direct input of mass into the control volume. Other parameters of the water quality also have an impact on the capability of production of mass. If aggressive water enters the control volume then interaction with cast iron will produce corrosion products. A high assimable organic carbon (AOC) content will result in biofilm formation, but probably also in the formation of particles [1]. 'Mass out' is measured in the same parameters as 'mass in' when looking at the water leaving the control volume. Mass will not only leave the control volume by the 'supernatant' water, but also through the bed load transport: sediment rolling on the bottom of the pipe. The 'mass storage' or ΔStorage is the net result of the mass balance. The behaviour of the storage is decisive for discolouration.

2.2. Mass in

In the total analysis of discoloured water the role of the incoming mass is often neglected. Basically the incoming mass is the amount of suspended solids formed by the oxides of the elements Fe, Mn and Al and in the Netherlands this can reach levels of up to 1.0 mg $L^{-1}$of these oxides. Thresholds for Fe, Mn and Al are respectively 200, 50 and 200 $\mu g\ L^{-1}$. As the elements are in the form of the oxides ($Fe(OH)_3$, MnO_2 and $Al(OH)_2$) the actual mass load is 382, 79 and 578 $\mu g\ L^{-1}$. respectively. In total this is 1039 $\mu g\ L^{-1}$. In normal practice this will be much lower as water rarely meets all the thresholds on these elements, still the amount of material can amount to a few hundred $\mu g\ L^{-1}$. Other elements that contribute to the amount of incoming mass can be small sand or carbon particles from the filterbeds. Some other

treatment processes can also contribute to the total suspended solids, for instance conditioning with marble filtration or pallet softening.

Suspended solids can also have organic characteristics. Gaulthier et al. [2], found in several samples an organic fraction of the suspended matter of 77%. Relating this to the inorganic load of up to 1.0 mg L^{-1} this would result in and organic load of up to 4,5 mg L^{-1}. But again in practice this will rarely be reached. The contribution of suspended solids originating from the treatment plant is often neglected in literature. The water quality parameter dealing with particle load is turbidity. In Dutch legislation the maximum value is set on 1 NTU. The background for this boundary for turbidity is however more based on the disinfection efficiency than on particle load.

The suspended solids level is the amount of mass that actually enters the network and it has the potential to add to the sedimentary deposits in the pipe. Determining the suspended solids is not yet a routine measurement. As the quantities of the suspended solids will be very low, in the order of 50 to a few hundreds $\mu g\ L^{-1}$, filtering of water over small mesh filters must be done with great care or with large quantities. The dilemma is to filter a relative small sample of water over a small mesh filter that would quickly clog and 'catch' as the suspended solids or to filter a larger quantity of water (200 to 300 litre) over a coarser filter that would not clog. The first method requires very sensitive equipment while the second one would be more robust.

2.2.1. Measuring mass in

The particle load directly from the treatment plant can best be evaluated using continuous monitoring techniques. At first, this will be the turbidity of the water and the particle counts and particle shapes. The turbidity or particle counts of the water leaving the treatment plant can show irregularities. Experience has shown that a large part of the sediment load to a network has its origin in these irregularities.

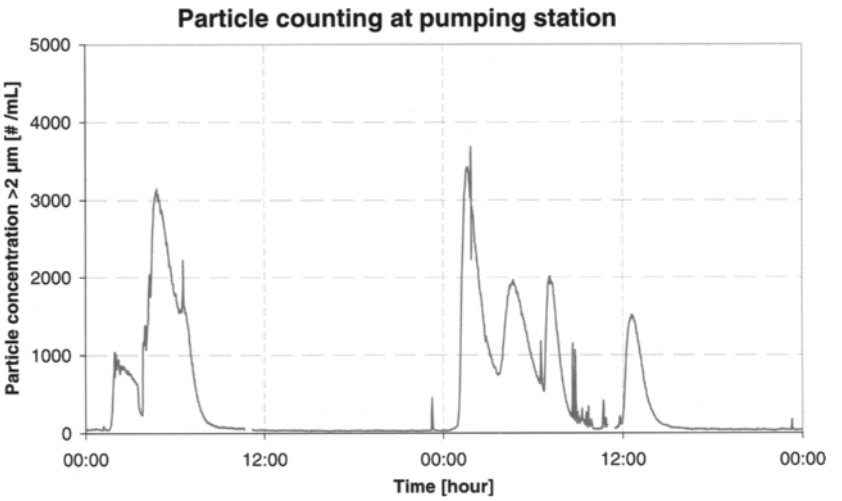


Fig. 4. Particle counts at pumping station

The particle counts of the treated water at a pumping station, shown in Fig. 4, shows a regular pattern of high counts, related to the backwash program. These irregularities are responsible for almost 75% of the total particle load to the system.

In addition to the continuous monitoring, samples can be taken from the incoming water and analysed for different components. Given that the irregularities in the treatment are largely responsible for the particle load, this makes the interpretation of the data obtained for grab samples, often taken at office hours, difficult, particularly in respect to any generally applicable conclusions.

Continuous monitoring of particles and turbidity can give an indication of the continuity of the treatment process and the associated particle or mass load to the system. The combination of the sample analyses of the dry solids and flow data enables the calculation of the total mass directly entering the system.

Samples of the water analysed for components such as Fe, Mn, dry solids, ATP etc. can be used to estimate the amount of mass that is introduced to the network. However these components are difficult to monitor on a continuous base. The relationship of these parameters with parameters that can be analysed on a continuous base such as turbidity and particle counts can help characterise the water quality and its ability to accumulate sediment.

2.3. Mass produced

Three different processes are recognised as being involved in the production of mass: corrosion of cast iron, biofilm formation and dying and coagulation of dissolved matter.

2.3.1. Corrosion of cast iron

There is a large amount of literature focussing on the relation between water quality and corrosion, and on the mechanism of corrosion as well as on the build up of corrosion layers [3, 4]. However, little is known on the effect of corrosion on the particle content of the water and on the contribution to the amount of sediment. In the Netherlands corrosion velocities of approximately between 0.00 and 0.13 mm/yr were reported in the main sections in service areas of several treatment plants. A clear relation between the corrosion velocity and the corrosion index of the water however was not found. The objective of the study was to determine the decrease in the strength of a main due to cross sectional material loss. The weakest point in the surface is decisive for the residual strength of the pipe. The study demonstrated that corrosion is a local process resulting in pits and not in a global attack of the material. For an estimation of the contribution of corrosion to the particle load the corrosion surface is estimated to be 5 to 10% of the total surface. The iron (Fe) wall material will corrode with the formation of Fe_2O_3. The concentration of corrosion products in drinking water can be calculated by using the equation:

$$c_c = \frac{\upsilon_{cor} \cdot A_c \cdot \rho_{Fe} \cdot \alpha}{Q} \qquad (2)$$

C_c	Concentration of corrosion product (kg m-3)
v_{cor}	Corrosion velocity (m h-1)
A_c	Corroding wall surface (m2) (5 to 10% of total wall)
ρ_{Fe}	Specific weight iron (7900 kg m-3)
α	Element to oxide factor (Fe_2O_3 => 1,43)
Q	Discharge ($m^3\ h^{-1}$)

Assuming a 1 km 100 mm diameter main with an average velocity of 0.1 m s^{-1} (2,8 $m^3\ h^{-1}$) results in a mass load of 0.24 to 0.47 mg L^{-1}. This is only a rough approximation since the actual mass load contribution depends on local conditions such as actual corrosion velocity, diameter of the cross section and the rate of discharge. In practice the results vary

significantly and indicate that the above range of values is the maximum found in practice. However it does show that the contribution of corrosion to the particle load can be of the same order as that from the treatment plant, being up to a few hundred micrograms per liter. Considering the local nature of the corrosion and the stabilization of the process over years, the contribution of the treatment plant can very well be dominant.

2.3.2. Biofilm formation

Biological processes generate material from the AOC that is present in the water. Primarily this will materialise in the biofilm that is formed within the system. The biofilm formation potential can quantify the amount of biofilm that is formed in pipes. For further reading on this topic Kooij [1], is recommended. The equilibrium biofilm will be dynamic with a constant dying and regenerating of organisms. Within this dynamic equilibrium the biofilm will produce material that can add to the mobile part of the sedimentary deposits. Quantifying this part of the mass balance is still a challenge.

2.3.3. Coagulation of suspended and dissolved solids

In the finished water colloidal particles are present that can coagulate to larger particles that have finite settling rates. In some preliminary tests in the Netherlands, using a special developed flow-trough jar test, this phenomenon has been postulated, but it has not yet been proven [5]. Research on the presence of aluminium in scales in water distribution systems showed a higher Al-content in scales in networks supplied with water treated with Al-based coagulants[6]. Future research is still required to substantiate the hypothesis of coagulation in the network.

2.4. Mass out

'Mass out' is separated into two different processes. Firstly, some material stays in suspension and leaves the control volume with the water. Secondly, there is the settled material that rolls along the bottom of the pipe in the form of a bed load transport. The several 'mass production' processes such as corrosion and biofilm formation and stability will contribute both to the suspended particles and to the settled particles.

By measurement of particle counts or turbidity as a general parameter for particles, the net result for storage can be evaluated. Fig. 5 shows the result of particle counts both at the pumping station and in the distribution network.

Bed load transport is very difficult to measure and although this is probably of overall significance in the transport of a large number of loose particles from the transportation lines to the reticulation system, it has yet to be accurately determined.

2.5. Storage

Storage is the net result of the mass balance. The mobility of the accumulated particles largely determines the discolouration risk. Disturbing the status quo will resuspend the particles and discolour the water to such a level that it becomes noticeable to the customer. Prevention of the resusupension is the key to better water quality in the network. In normal circumstances there is bound to be equilibrium between the 'mass in' and 'mass produced' with the 'mass out'. Disturbing this equilibrium, for instance through an increased velocity, will resuspend the sediment and cause discolouration. This phenomenon is also deliberately used to remove loose particles by flushing pipes with a high velocity surge. The net discolouration risk is consequently linked to both the amount of mobile sediment in the equilibrium and the likelihood of disturbances of the flow velocity.

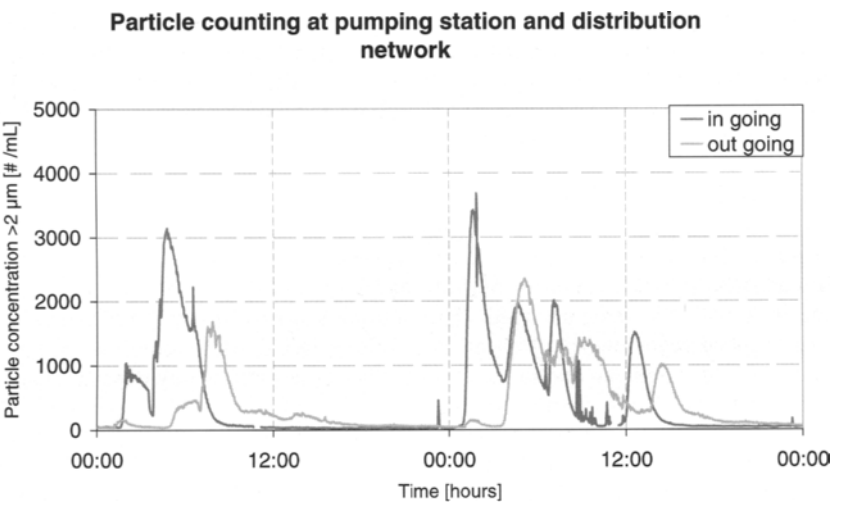


Fig. 5. Incoming and outgoing particle counts

The first action which is already available to water companies to minimise hydraulic disturbances is a ban on the use of fire hydrants other than by the fire-fighting department. The frequent but irregular use of fire hydrants for all kind of purposes (watering public gardens, helping laying streets, cleaning of outer walls of houses etc) often caused unexpected discolouration complaints.

Although a lot of the measures to improve the quality of the water in the distribution network have to do with controlling the storage, it is difficult to estimate thresholds for the absolute amount of particles needed to prevent discolouration problems. However, measurement methods such as the Resusupension Potential Method are capable of quantifying the discolouration risk in such a way that action levels can be determined.

2.6. Controlling the mass balance: a three way approach

Many processes contribute to the mass balance in a piped system. However at this moment not all the processes can be quantified properly. Current research is intended to quantify the contributions of the different processes with the goal of identifying the most important contributions and then dealing with these first. For instance in a network with a lot of unlined cast iron, the corrosion contribution can be dominant and one solution is the lining or replacement of the cast iron pipes. Where the particle load from the treatment plant is dominant, lowering this load will improve the water quality.

In contrast, the storage part of the mass balance is the most important part when it comes to discolouration risk. Discoloured water is caused when hydraulic forces create a disturbance and dealing with the storage is the major challenge in operating a network. A three-way approach is recommended:

Prevent the sediment from entering or being formed in the network. Basically this is controlling the different processes in the mass balance as described earlier. Attention must be paid to the production process, as this can be one of the major contributors to the mass load. Also prevention of biological regrowth and corrosion are other possibilities to decrease the particle load in a system. Prevent the sediment from accumulating to an unacceptable level. This comes down to regular cleaning and maintenance of the network.

Prevent the sediment from accumulating in the first place by adopting a new approach to designing and constructing networks (especially towards the supply of fire fighting water).

Before analysing these options what is needed is a measurement method for assessing the discolouration risk formed by the accumulated particles. This is the resusupension potential method (RPM).

3. THE RESUSUPENSION POTENTIAL METHOD (RPM)

Irrespective of their origin, it is the presence and mobility of loose deposits within the distribution system that determines the discoloration risk. The RPM, developed in the Netherlands, is based on measuring the capability of the sediment in a network to resuspend and create visually noticeable turbidity levels.

The RPM is based on a continuous monitoring of turbidity during a controlled and normalised increase of the velocity in a pipe. The hydraulic shear stress as result of the increased velocity causes sediment to resuspend. One practical restriction in this procedure is of course that the change in hydraulic regime can only be allowed to cause a moderate increase in shear stress because too much of an increase in turbidity levels would lead to actual complaints from the consumers. Fig. 6 illustrates the protocol of the method:

- Isolate the pipe for which the discoloration risk is to be assessed by closing the valves that connect to other feeding pipes;
- Induce acceleration of the flow by opening a fire hydrant in such a way that the velocity in the pipe is increased by an additional 0.35 m s^{-1} on top of the normal velocity and maintain the higher velocity for fifteen minutes. The velocity of 0.35 m s^{-1} is empirically determined. For a 100 mm internal diameter pipe this requires an extra volume flow of 2.78 L s^{-1} (10 m^3 h^{-1}), which leads to a moderate disturbance. The increased levels of turbidity are measurable, and in most cases do not exceed acceptable levels.
- Monitor the turbidity in the pipe
- After fifteen minutes, reduce the velocity to normal and monitor the turbidity until it is back to the starting or initial level.

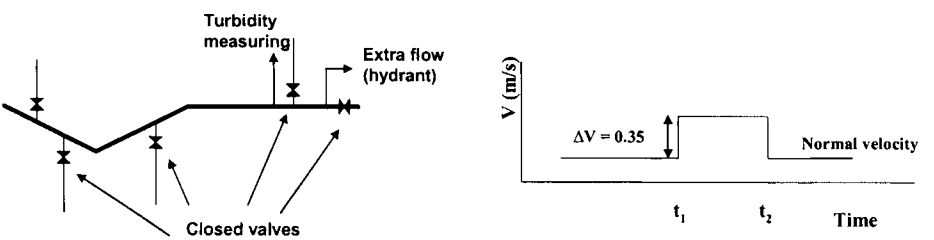


Fig. 6. Principle of the resusupension potential method: increased velocity in the isolated pipe

3.1. Practical application

The RPM has been developed for smaller water mains ranging from 80 to 200 mm internal diameter. The velocity needed to induce the required flow is dependent on the internal diameter of the pipe and ranges from 1.76 L s^{-1} (6.3 m^3 h^{-1}) for an 80 mm pipe to 11.0 L s^{-1} (40 m^3 h^{-1}) for a 200 mm pipe.

A typical value for a 100 mm pipe is 2.78 L s^{-1} (10 m^3 h^{-1}). Opening a hydrant creates the extra flow needed for the required increase in velocity. Fig. 7 shows the standpipe developed for this purpose. The flow is restricted using calibrated plates in the standpipe, which then limit the flow to the desired value without the risk of an extra volume flow when the hydrant is opened. In normal practice a hydrant with a standpipe is 'cracked' with opened orifices. This causes an instantaneous increase in the flow, which is in most cases higher than the desired flow for the RPM causing an extra velocity and an extra resusupension of the sediment.

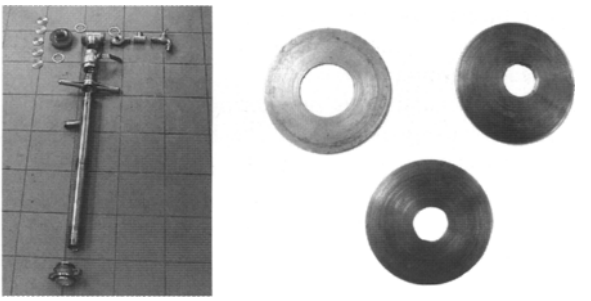


Fig. 7. Standpipe and calibrated plates for standardized disturbances

For measuring the turbidity several flow-through turbidity meters that use different measurement principles are commercially available. A comparison of these led to the conclusion that they measure comparable turbidity patterns in turbulent solutions, but that the absolute values of turbidity are at different levels [7]. This should be carefully taken into account when interpreting the results of the RPM measurements.

The flow through the standpipe responsible for the disturbance in the pipe can be monitored using a flow meter or a water meter. These flows should be recorded because any deviation from the standard disturbance will cause a different sheer stress on the sediment and consequently a different effect on the turbidity. Using well-calibrated plates in the standpipe can eliminate the need for flow measurement.

3.2. The RPM-curve

A typical turbidity curve of the RPM is shown in Fig. 8. For rating the discoloration risk four elements of the curve are important. These elements are:

- Initial increase in turbidity at the start of the hydraulic disturbance
- Development of turbidity during the hydraulic disturbance
- Resettling time and pattern in the return to the base (initial) turbidity level

3.2.1. Base turbidity level

The base turbidity level is the level preceding the hydraulic disturbance. Base line turbidity can be linked to the turbidity of the source water and give some insight in the source of the sediment. The base line turbidity is also needed in the measurement of the time needed for the turbidity to resettle after the increased velocity stopped.

3.2.2. Initial increase in turbidity

Following the actual disturbance the turbidity will rise immediately to a certain level. This initial increase gives the instantaneous mobility of the sediment, providing a peak turbidity,

and is due to the presence of loosely attached material. The initial increase is an indication of the maximum turbidity caused by a hydraulic incident.

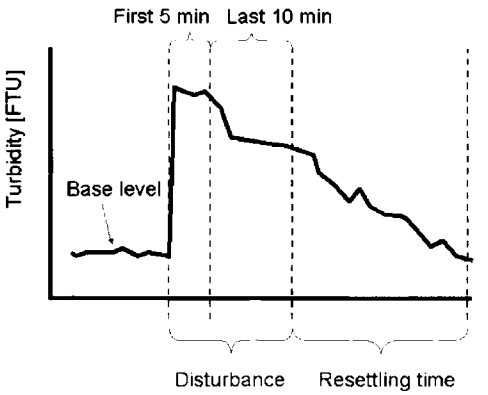


Fig. 8. Typical RPM-curve

3.2.3. Development of turbidity during the hydraulic disturbance

The hydraulic disturbance is maintained for 15 minutes, allowing the turbidity to develop to a stable level. If the turbidity remains at almost this level for the first five minutes, then the amount of sediment is considerable and the composition of the sediment is homogeneous. This indicates a high discoloration risk.

In many cases the turbidity drops during the 15 minutes disturbance time. Three phenomena can explain this:

- A relatively small amount of heavy sediment is present in the pipe. The extra forces that accelerate the flow also promote the initial resuspension of this heavy sediment [8]. The significance of this sediment is limited because it is settling even during the disturbed hydraulic conditions, lowering the discoloration risk. The chance that the initial discoloured water will actually reach a tapping point, at which it can be visually identified, is small.
- An overly short length of isolated pipe can be the second reason for a declining turbidity during the 15 minutes disturbance. If the length is less than 315 meters, then water may be drawn from pipes upstream of the isolated pipe. This water originates from the looped network or from pipes with larger diameters and is much less disturbed than water from within the intended study area.
- A non-homogeneous deposit over the length of the tested pipe, for instance in a hilly area with concentrations of sediment in the depression of a pipe.

In all cases however the level of turbidity following the first peak determines the continued discoloration risk.

3.2.4. Resettling to base level

After closing the hydrant, it takes a certain time for the turbidity to settle back to the base level. The time needed is important for the discoloration risk or actually the complaint risk. If the turbidity stays high during a longer period, the risk of noticing the turbidity in an application as filling a white basin (bathtub, washing bin, bathroom sink, etc) is larger.

3.3. RPM and discoloration risk

The RPM is interpreted with respect to five aspects that are rated equally:

- Absolute maximum value of turbidity during first five minutes of disturbance (20%)
- Average value of turbidity during first five minutes of disturbance (20%)
- Absolute maximum value of turbidity during last ten minutes of disturbance (20%)
- Average value of turbidity during last ten minutes of disturbance (20%)
- Time needed to resettle again to initial turbidity level. (20%)

For each aspect a validation on a scale of 3 is made: 0 is the lowest or best rating and 3 the highest or worst rating. The lowest value equivalent with 'no discoloration risk' is 0 (zero) and the highest value or 'maximum discoloration risk' is 15. For the rating per aspect a scale must be made, that is calibrated to the turbidity equipment used. Also site-specific elements can be taken into account. If for instance the intuitive feeling is that the discoloration risk of a network is moderate, then the rating scale can be adjusted to this level.

Changes in the discoloration risk that may occur when for instance the treatment is improved or a cleaning program is started can be related to the adjusted level. The discoloration risk established in this way is a relative figure that can be company-specific or even area-specific, allowing the assessment of changed circumstances for the network. For every situation and measuring equipment, a ranking table can be established, depending on the type of turbidity measuring equipment and local circumstances. Tables 1 and 2 give the values for discoloration risk for two different turbidity meters. The turbidity meters differ in the wavelength used to measure the turbidity.

Table 1
Example of ranking RPM for discoloration risk using turbidity meter A

	0	1	2	3
Absolute max first 5 min	<0.3 FTU	0.3-1.0 FTU	1.0-2.4 FTU	>2.4 FTU
Average first 5 min	<0.3 FTU	0.3-1.0 FTU	1.0-2.4 FTU	>2.4 FTU
Absolute max last 10 min	<0.3 FTU	0.3-1.0 FTU	1.0-2.4 FTU	>2.4 FTU
Average max last 10 min	<0.3 FTU	0.3-1.0 FTU	1.0-2.4 FTU	>2.4 FTU
Time to clear	< 5 min.	5-15 min	15-60 min	>60 min

Table 2
Example of ranking RPM for discoloration using turbidity meter B

	0	1	2	3
Absolute max first 5 min	<3 FTU	3 – 10 FTU	10-40 FTU	>40 FTU
Average first 5 min	<3 FTU	3 – 10 FTU	10-40 FTU	>40 FTU
Absolute max last 10 min	<3 FTU	3 – 10 FTU	10-40 FTU	>40 FTU
Average max last 10 min	<3 FTU	3 – 10 FTU	10-40 FTU	>40 FTU
Time to clear	< 5 min.	5-15 min	15-60 min	>60 min

The ranking tables are meter type specific and for several types the tables have been developed [9, 10]. The ranking tables give a certain flexibility to tailor the results for a specific application. If the RPM is used to prioritise areas for an initial cleaning program, the maximum values can be adjusted to the outcomes of the measurements. For instance most of the measurements end up with a total value larger than 12, because the average turbidity over the whole period is larger than 40 FTU. By adjusting the ranges it is possible to distinguish

very high and high degrees of discoloration risk enabling the desired prioritisation. In most case a ranking table can be made using known areas with high and low complaints levels. Experience built up over time will help building the tables.

3.4. Typical RPM-curve

Fig. 9 gives an example of a RPM with a high discoloration risk. The measurement has been done with turbidity meter A, so Table 1 is applicable. The base turbidity is quite constant and low (0.22 FTU). The initial turbidity during the first five minutes as well as during the following 10 minutes is high and above 2.4 FTU, which is the maximum turbidity that can be measured in this setting. The time to clear is several hours, so on all the items the maximum score is reached, resulting in an overall score of 15. Despite the fact that the base line turbidity is very constant and low, this location will experience a high turbidity incident with any hydraulic event

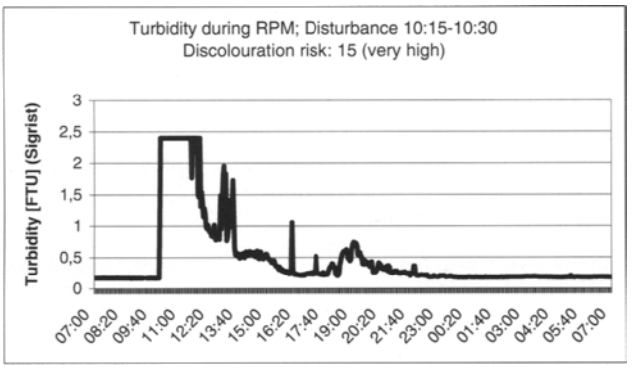


Fig. 9. Example of RPM, measured with turbidity meter A, indicating a high discoloration risk.

3.5. Alternative analysis of the RPM

A full evaluation of the RPM requires a measuring time of at least a few hours, mostly consumed by determining the initial base level and pattern and the resettling time. It also requires turbidity measuring equipment that can be costly and requires skill to operate. The implementation of the principle of the RPM during the last few years has led to some alternatives in the analysis methods. The first adjustment is to limit the measuring time of the base level turbidity to some minutes prior to the disturbance or the time needed to isolate the pipe. Limiting the time used to measure the resettling time to a specific time, for instance 30 minutes is another adjustment. The examples, shown in Fig. 6, use this abbreviated version of the RPM. As seen the resettling time is not clear in the first graph. However the overall impression is that the discoloration risk is so high (>12) action is required.

Limitation of the measuring time allows for more measurements in a working day. It is possible to do 4 measurements in an 8-hour working day. Replacing continuous monitoring with 5 to 8 grab samples taken during and after the disturbance limits the total measuring time even further. All the adjustments however cause loss of information, specifically on the base level and the resettling time. For the assessment of the trigger level towards the discoloration risk, this is less important as the examples of Fig. 6 show. The increase in number of measurements and the decrease of labour costs however can compensate for this.

3.6. Cleaning of networks

Along with the prevention of particles coming into the network, the regular removal of accumulated particles is also critical to the reduction of discolouration risk. The basic principle for the cleaning of pipes is the application of high sheer stress to the settled particles. The most efficient and cost-effective way is to use high velocity water, this is known as flushing. Other cleaning methods include combined flushing with water and air to produce even greater turbulence or to use foam pigs pushed through the pipes with water pressure.

The method of flushing with water has proven to be efficient and cost-effective if done with great care. This means that the following conditions have to be met:

- Flushing velocities of at least 1.5 m s-1
- Total volume flushed is to be two to three times the content of the pipe
- Water is drawn from clean pipes (using a so-called clear water front)

These standards seem to be simple, but in practice it takes quite some effort to actually prove all the conditions are met.

The RPM as described in the former paragraph is a powerful tool to establish the efficiency of any proposed cleaning action: The RPM is determined pre- and post cleaning and the difference between the two gives a good indication of the efficiency of the cleaning.

Fig. 10 gives the result of a RPM before and after cleaning by flushing a selected section of pipe.

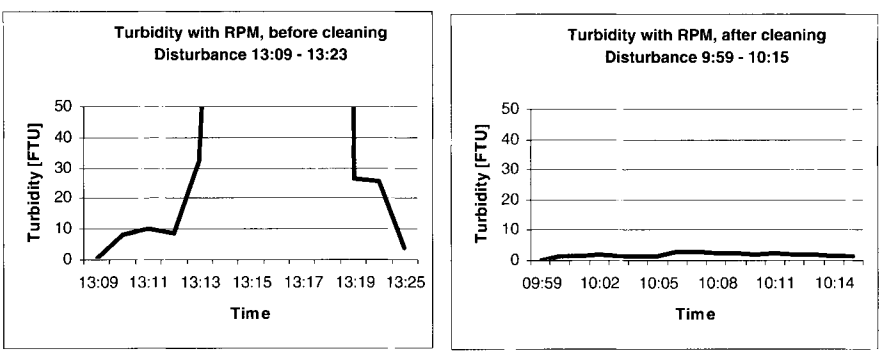


Fig. 10. Left: pre-cleaning, discoloration risk is high (>12); post cleaning discoloration risk is very low (0), turbidity meter (type B) and ranking as in Table 2

4. DESIGN OF DRINKING WATER DISTIBUTION NETWORK

4.1. Introduction

Particles accumulate in areas with low velocities. Typically in distribution networks these conditions occur because the major design criterion for a small distribution network is compliance with fire fighting demands. These supersede the water demand that normally occur in the network, leading to an over-designed distribution network. This situation is generic across the developed world. In the Netherlands a fresh approach to the design of distribution networks is being developed and applied.

Historically a lot of water companies originate from municipalities. Fire fighting departments also originate from municipalities and in the Netherlands, as in almost all of the European countries, fire fighting still is a municipal responsibility. The interaction of the fire flow requirements and the drinking water network is very obvious and probably the only viable way of supplying these large amounts of water. However this has an impact on the design of distribution systems and water quality. Although not uniformly defined, the usual capacity of a fire hydrant is 30 to 60 $m^3 h^{-1}$ and there should be a hydrant within 40 to 50 meter from every building, which dictates that fire hydrants are placed within every 80 - 100 meters or every 20 to 25 houses in a network. The maximum flow for 25 houses is 5 $m^3 h^{-1}$ while the needed fire flow is 30 to 60 $m^3 h^{-1}$. Thus fire flow requirements dominate normal drinking water demand.

4.2. Conventional distribution network

The structure and capacity of conventional distribution networks are determined in the Netherlands by fire fighting requirements. This results in distribution networks largely consisting of pipes with diameters of 110 mm or more, arranged in a ring structure. An example of a conventional distribution network is shown in Fig. 11.

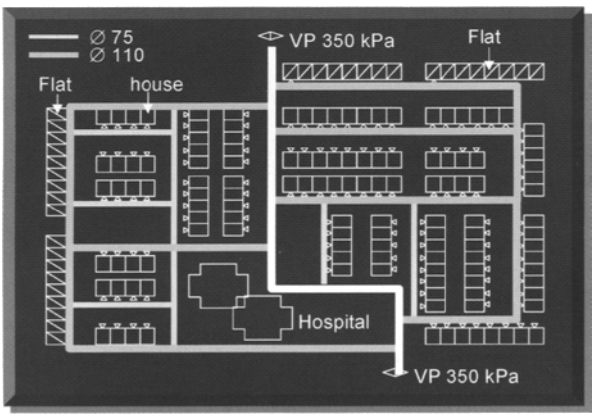


Fig. 11. Conventional distribution network

4.3. Modern distribution network

The awareness of the adverse effects of fire fighting demands on the water quality has initiated a new approach towards the design of new distribution networks. The basic requirement is a rethinking of the fire fighting demand. In close cooperation with national fire fighting agencies in the Netherlands this has led to lower fire fighting demands 30 $m^3 h^{-1}$ in residential areas in which buildings meet modern fire fighting codes. Practically all post 1970-buildings meet those requirements. Consideration is also being given to even further decreases in the fire fighting demand by installing residential sprinklers.

Having reached this agreement the design of new systems can start. The starting point for a modern distribution network is that in normal drinking water demand situations a velocity of at least 0.4 m s^{-1} is required. This velocity is sufficient to prevent large accumulation of sediment. The main characteristics of a modern distribution network are a branched system with pipes with a relatively small diameter. The process of (re)designing a distribution network is briefly demonstrated in the next section.

4.3.1. Actual drinking water demand; maximum flow

The demand pattern of one house is much more erratic than the demand pattern of 1000 houses. The pattern for one house will include a number of spikes throughout the day, while the demand pattern of a cluster of houses will be much smoother. The method used to determine the maximum demand is the so-called q $\sqrt{n}$ -method:

$$q_{m.n} = 0{,}083\sqrt{n * TU_{house}} \quad (3)$$

with

q_{mn} : Maximum instantaneous demand of n houses in l s^{-1}
n : Number of houses
TU_{house} : Number of taps per house
0.083 : Capacity of 1 TU in L s^{-1} (300 L h^{-1})

The method was originally developed to determine the maximum instantaneous demand of a single household for the design of the in-house plumbing system. Every house will have a number of tapping points with a varying number of taps. A bath or shower mixer tap will have a different maximum flow to a toilet cistern tap. A toilet cistern tap for instance is equal to 0.25 TU, while a kitchen sink tap or a bath mixer tap will have 4 TU. In Table 3 a typical Dutch single-family home installation is given

Table 3
Typical house installation

Tap point per house	Number of TU
Toilet cistern tap1	0.25
Toilet washbasin1	0.25
Toilet cistern tap 2	0.25
Toilet washbasin 2	0.25
Kitchen sink	4
Dish washer	4
Bath/shower mixer tap	4
Washbasin mixer tap (bath room)	1
Wash basin tap (bed room)	4
Washing machine tap	4
Total per house	22

The maximum flow for this house will be

$q_{max} = 0.83\sqrt{1 * 22} = 0.389$ L $s^{-1} = 1.4$ m^3 h^{-1}

The maximum flow for a set of 9 of these houses will be

q_{max} $0.083\sqrt{9 * 22} = 1.1681$ L $s^{-1} = 4.2$ m^3 h^{-1}

and so on.

In Fig. 12 three maximum demand curves are given for several types of dwelling.

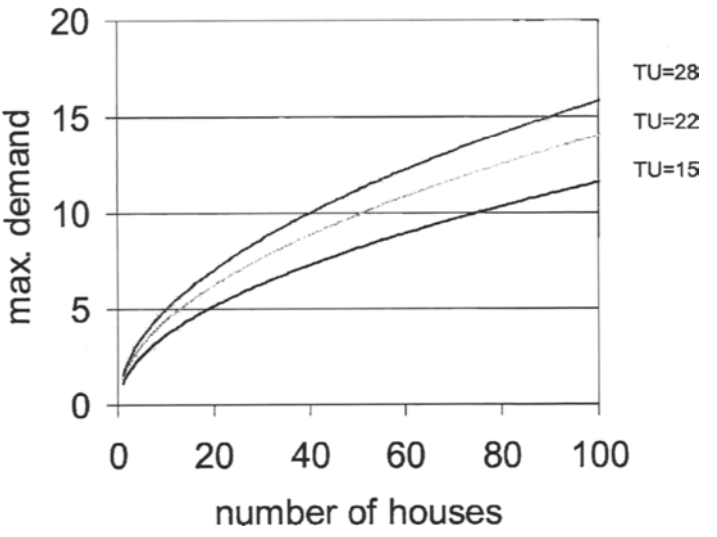


Fig. 12. Maximum instantaneous demand for three types of dwellings

In the above example the demand per unit of houses can be determined by counting the number of houses and the maximum instantaneous demand calculated.

4.3.2. Design branched system

The pipes connecting the houses to the main structure are branched pipes with a diameter that will decrease downstream. The maximum flow in the pipes will be determined with the $q\sqrt{h}$ - method. A table can be composed to determine the maximum number of houses that can be supplied with a range of diameters. The table is based on the assumption that the minimum velocity in the pipe is 0.4 m s^{-1} and the maximum velocity is 1.5 m s^{-1}.

$$n = \left(\frac{v \cdot \frac{1}{4} \cdot \pi \cdot D^2}{0{,}083 \cdot 10^{-3} \sqrt{TU_{house}}} \right)^2$$

with

v	velocity, $0.4 < v < 1.5$ m s^{-1}
n	number of houses
D	Diameter pipe [m]
TU	Number of tap units

Based on 22 tap units per house the following Table 4 can be composed.

Table 4
Maximum number of houses that can be supplied with a range of diameters

Outside Diameter pipe	Inside diameter pipe	Min# houses (v=0.4)	Pressure drop [mwc km^{-1}]	Max# houses (v=1.5)	Pressure drop [mwc km^{-1}]
40	35	1	4.66	14	65.53
50	45	3	3.62	38	50.97
63	57	7	2.86	97	40.24
75	67	13	2.43	185	34.23
90	81	28	2.01	394	28.32
110	101	68	1.61	953	22.71

With this table the branched system or section can easily be composed and will give a system such as that shown in Fig. 13. Normally a range of maximum 3 diameters will be used; more diameters simply complicate maintenance and repair. If for instance the range 40-63-110 is chosen, the network will be similar to that shown in Fig. 13. Other ranges could be 40-75-90, etc.

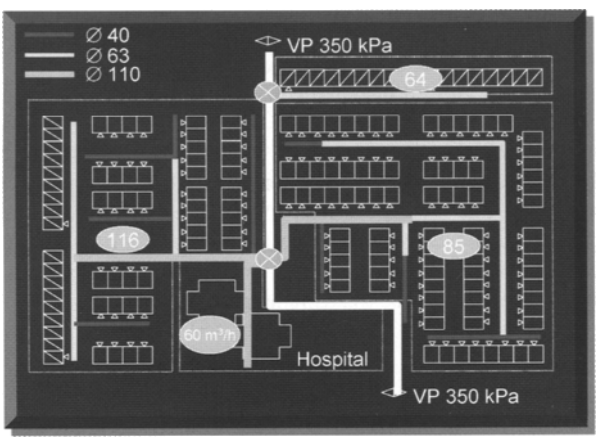


Fig. 13. Modern branched network

The result of this exercise is a network that complies with the demands for a drinking water network. The last step is to see what fire flows can be delivered and how the network can be adjusted to accommodate that demand without compromising the initial requirements. Normally a 63 mm pipe can accommodate a fire hydrant with 30 m^3 h^{-1} assuming that is enough. Sometimes these pipes have to be extended to allow for the special fire requirements of schools, hospitals and other larger buildings.

When all the fire hydrants are located the pressure drop should be checked. During fire extinguishing the pressure in the network may drop below the minimum pressure of 200 kPa. The probability that the fire hydrant will be used is very low, less than once per twenty years. The risk introduced by lowering the pressure is back flow from the house installations with possible contaminations. The barrier for prevention of this happening is the check-valve in the house connection.

4.4. Possibilities for controlling the mass balance in a network

Particles in the distribution network are the key to the aesthetic water quality in the network. There are several ways to control the amount of particles in the network that are graphically summarized in Fig. 14.

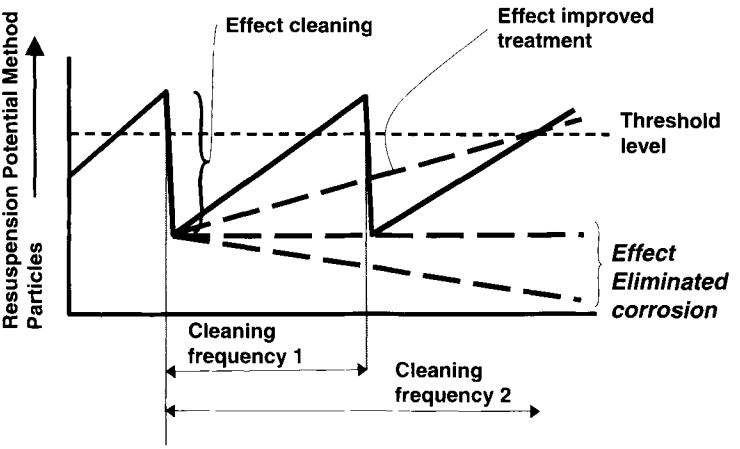


Fig. 14. Possibilities to control the discolouration risk in a network

The solid line represents the (re)charging of the network with particles, expressed in the RPM. If a certain threshold is reached remedial action is needed to clean the network, reducing the RPM to acceptable levels. If nothing further is done, the network will recharge with sediment resulting in the need for further cleaning. This procedure gives the cleaning frequency for this network. A further possibility for improving the quality or prolonging the cleaning frequency is reducing the particle load through the treatment. Eliminating all the particle sources eventually will eliminate the actual particle load and leave the network without any need for cleaning and without customer complaints about discolouration.

REFERENCES

[1] D. van der Kooij, "Managing regrowth in drinking water distribution systems" Kiwa BTO report, 139, Kiwa Water Research, Nieuwegein, The Netherlands June, 2002 (www.kiwa.nl).

[2] V. Gaulthier, J.M. Portal, H. Yvon, C. Rosin, J.C. Block, V. Lahousisine, S. Benabdallah, J. Cavard and S. Fass, Water Sci. Technol.: Water Supply, 1 (2001) 89.

[3] P. Sarin, V.L. Snoeyink, D.A. Lytle and W.M. Kriven, J. Eviron. Eng., ASCE, 4 (2004) 364.

[4] P. Sarin, V.L. Snoeyink, J. Bebee. K.K. Jim, M.A. Beckett, W.M. Kriven and J.A. Clement, Water Res., 38 (2004) 1259.

[5] M. den Boer, Post coagulation and Sedimentation in Distribution Networks, Kiwa Report BTO.024 in Dutch, BSc-thesis 2004.

[6] V.L. Snoeyink, M.R. Shock, P Sarin, L. Wang, A.S.C. Chen and S.M Harmon, J. Water Supply, Res. Technol. - Aqua, 52 (2003) 455.

[7] P.G.G. Slaats, L.P.M. Rosenthal, W.G. Siegers, M. van den Boomen, R.H.S. Beuken and J.H.G. Vreeburg, Processes Involved in the Generation of Discoloured Water" Awwa Research Foundation, Denver, USA (2000).

[8] E.J.M. Blokker and J.H.G.Vreeburg,, Proc. SCE Conf. 173, 34, May 15-19 2005.

[9] J.H.G. Vreeburg, P. Schaap and J.C. van Dijk, Proc. Leading Edge Conference on Water and Wastewater Treatment, Prague, 2004.
[10] J.H.G. Vreeburg, P.G. Schaap, M. Den Boer, J.C. van Dijk, Proc. 4th IWA World Water Congress, Marakech, 2004.

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 17: Natural organic matter

Ronald Beckett[a] and James Ranville[b]

[a]Water Studies Centre, Monash University (Clayton Campus), Monash University, Melbourne, AUSTRALIA 3800

[b]School of Chemistry and Geochemistry, Colorado School of Mines, Golden, CO, USA

1. INTRODUCTION

This chapter provides a brief review of the physical and chemical characteristics of aquatic natural organic matter (NOM). The presence of NOM influences a diverse range of processes occurring in natural waters, including contaminant speciation and bioavailability, suspended particle stability, foaming, and aquatic redox reactions. Of particular importance in this book is the influence of NOM on the performance of water treatment plants.

It would appear that NOM is ubiquitous in all environments, being present to some extent in soils, sediments, waters and even the atmosphere. This chapter will concentrate on aquatic NOM, which has classically been divided into dissolved and particulate forms (DOM and POM respectively), with recent studies also considering an intermediate colloidal state (COM). However, the exact definition of these terms does vary depending on the method of fractionation used (e.g. the filter size or centrifugation conditions used).

NOM consists of very complex mixtures of substances that all arise from an initial biological source. Although the components of NOM can be classified, no single "molecule" of NOM exists. Rather we think of different NOM samples as having a set of average properties, which depend on the source material, the degree of subsequent degradation reactions, and transport-associated fractionation processes. Differences in these average properties can have a marked effect on processes involved in water treatment and this fact provides the impetus for the following introduction to NOM characteristics and reactivity.

It should be noted at the outset that although we will focus on the term NOM, much of our information comes from the analysis of total or dissolved organic carbon (TOC or DOC). TOC analysis is generally based on the oxidation of NOM to, and subsequent detection of, carbon dioxide. The efficiency of this conversion, and thus the accuracy of TOC determination, is influenced by NOM characteristics and potential side reactions of the oxidant. It is common to see the terms NOM and TOC or DOC used interchangeably. In most cases the organic carbon content of NOM is considered to be roughly 50%, although exceptions can occur [1].

1.1. Sources of NOM and seasonal effects

The initial source of molecules making up NOM includes higher plants, animals, and microorganisms. NOM produced by the decomposition of aquatic organisms is termed *aquogenic* whereas that derived from terrestrial sources is termed *pedogenic*. This latter term reflects the fact that most decomposition of terrestrial organisms occurs within soils. Given

that most biomass is contained in plant material, this source is presumed to dominate aquatic NOM. In most natural aquatic systems NOM is primarily of a *pedogenic* origin, the usual exceptions being eutrophic freshwater environments and marine water bodies where in-situ biological production can result in the formation of *aquogenic* NOM. Limnologists define NOM transported into a water body from the surrounding watershed as *allochtonous* and that produced within the water body as *autochtonous*.

Any life processes such as photosynthesis, organism metabolism, microbial degradation, and others lead to the formation of the biomolecules that are the precursors to NOM. We generally restrict our view to the process of organic matter degradation when considering NOM. Microbial degradation can lead to two different NOM "pools" one being fairly refractory organic matter that is resistant to further degradation. This material is generally classified as humic substances, or in the case of soils, humus. This material generally cannot be further decomposed and as a result accumulates in aquatic systems. Due to their complexity, humic substances are generally not identifiable at a molecular level. Humic substances often comprise more that 50 % of the NOM present. The other products of microbial degradation are simpler, identifiable molecules such as low molecular weight acids (acetate, citrate), sugars, and proteins that are classified as non-humic substances. These compounds can be further degraded, are thus more transient than humic substances, but can reach steady-state concentrations that allow them to be present at significant concentration is some systems. Examples of this situation are streams whose flow is dominated by sewage effluents. The term effluent organic matter (EfOM) has been used to describe the NOM in this situation where humic substances are secondary in importance to more hydrophilic non-humic NOM. Also present in these human-influenced waters are a myriad of xenobiotic compounds, which do not contribute significantly to the TOC but are potentially biologically and chemically reactive. Some examples include pesticides, detergents, solvents, and endocrine disrupting chemicals (EDCs) but discussion of these compounds is beyond the scope of this chapter.

1.2. Levels of OC in aquatic environments

The concentrations of NOM found in natural waters vary enormously depending on the environmental conditions [1]. DOC in precipitation, ground waters and seawater is usually less than 1 mg L^{-1}, lakes and rivers most commonly vary from 2-10 mg L^{-1} whereas waters originating from wetlands and bogs may contain as mush as 50 mg L^{-1}. DOC levels are typically up to 10 times that for particulate organic carbon (POC), while little quantitation of COM has been performed.

The DOC in interstitial water of topsoil may be 10-30 mg L^{-1} but this usually decreases rapidly with depth in the soil profile due to adsorption on soil particles and microbial degradation. In lakes and the ocean the DOC is elevated in the surface euphotic zone and may decrease by a factor of 2-5 in the bottom water.

There is usually a surface microlayer at the air/water interface that contains high concentrations of NOM, trace elements and other contaminants [2]. The DOC in this microlayer may be an order of magnitude greater than the underlying water.

In uncoloured waters the DOC is usually fairly equally divided between humic substances (hydrophobic acids) and hydrophilic (non-humic) substances, although in coloured waters and soil pore waters the humic substances may be higher (e.g. 60-80%). In most aquatic systems the humic substances fraction is comprised of about 90% fulvic acid and 10% humic acid (defined below), which reflects the higher water solubility of fulvic acid apparent from the operational definition (soluble at both high and low pH conditions).

2. CHARACTERISATION

2.1. Isolation and fractionation of humic and non-humic NOM

The organic compounds that make up NOM consist of molecules having a wide range of chemical characteristics. One means of NOM classification is by describing the degree to which the NOM components display a hydrophobic or hydrophilic character. This property is an aggregate of the chemical functionality and physical characteristics of the NOM components. Properties to consider include aromatic and aliphatic structures, polar or ionizable functional groups (alcohol, carboxylic acids, esters), hetero-ions (N, S, P), molecular weight, and compound structures. Although characterization of NOM focuses on the organic components, inorganic constituents such as complexed ions (Fe^{3+}, Al^{3+}, Ca^{2+},etc.) and nanoparticles such as metal oxides (FeO_x, AlO_x, MnO_x) will influence the "hydrophobicity" of a given NOM sample.

Table 1
General property differences between the two operationally-defined classes of humic substances, humic and fulvic acid. Adapted from Stevenson [3]

	Fulvic acid		Humic acid
Color	Yellow		Brown-black
Polymerization	Lower		Higher
Molecular weight	< 2000		Up to 300,000
Carbon content	45%	← Range →	62%
Oxygen content	48%		30%
Aromaticity	Lower		Higher
Exchange acidity	1,400 (meq/100g)		500 (meq/100g)

A classic approach to an operational characterization of NOM based on its degree of hydrophobicity is to utilize pH-dependent solubility. Classically, dissolved NOM has been separated into humic (HA) and fulvic acid (FA) by precipitating the NOM at pH < 2. At this pH most ionizable functional groups are protonated, lessening the water solubility of the NOM components and HA is insoluble whereas FA remains soluble. Centrifugation or filtration can separate these two fractions. The ratio of HA to FA can be determined gravimetrically or by total organic carbon analysis (TOC). This fractionation produces two operationally defined materials that differ in properties such as aromatic content, molecular weight, acidity, and a number of other characteristics. Although these characteristics depend on several factors including NOM source, degree of biodegradation, and natural fractionation, some generalities can be made about HA and FA acid properties. These are summarized in Table 1 [3].

The simplicity of this fractionation scheme has proved useful for describing the effect of source on NOM characteristics and for describing many environmental and ecological impacts of NOM. HA and FA fractions tend to dominate the total amount of organic matter in natural aquatic samples, however other components can be present in significant amount [1].

The operational approach based on pH-dependant solubility has been further refined to provide more fractions having a narrower range of properties. Soil scientists further differentiate HA into hymatomelanic acid (alcohol-soluble), grey HA (precipitated from an alkaline HA solution by adding electrolyte) and brown HA (soluble in alkaline solution after electrolyte addition), but these categories are not widely used in aquatic NOM investigations. More commonly, aquatic NOM is further fractionated by retention characteristics on nonionic acrylic resins [4]. Most of these approaches are based on passing the sample through a resin filled column. Various studies have developed procedures that differ in the resin used, the eluent composition and pH, and resin to solution ratio. The latter variable leads to the commonly defined term k′, which describes the hydrophobicity of the isolated NOM. Because of the larger degree of possible isolation conditions, this approach, while providing a greater degree of NOM characterization, does not give a consistent set of NOM fractionation across the multitude of studies. These more complex fractionation methods are however finding use in the investigation of NOM issues in drinking water including membrane fouling and trihalomethane formation.

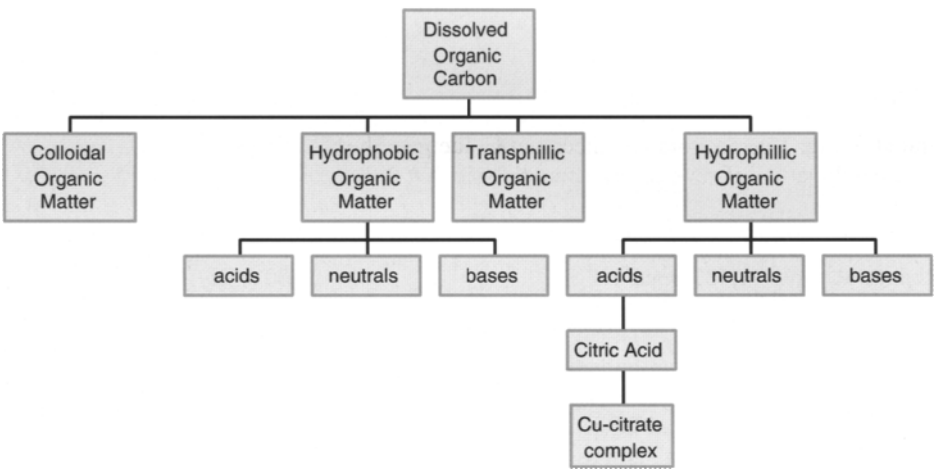


Fig. 1. A comprehensive classification scheme for natural organic matter based principally on resin fractionation and dialysis. The level of compound specific information increases as one proceeds further along the scheme, for example hydrophilic acid to citric acid to Cu citrate complex [6]

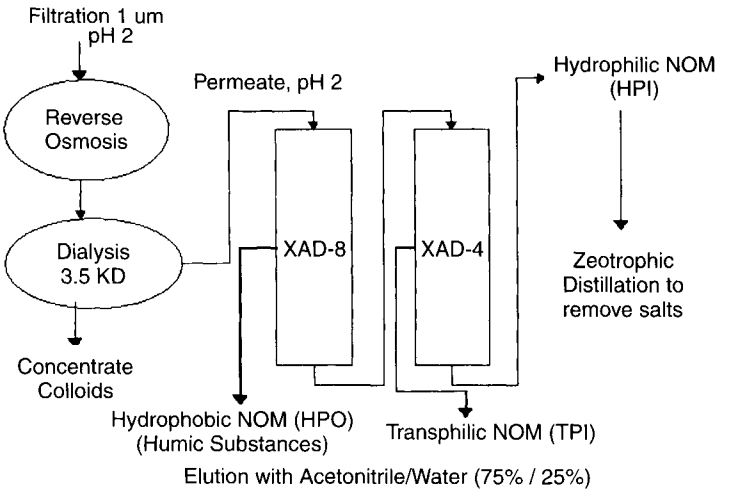


Fig. 2. Procedures and operationally-defined fractions associated with a comprehensive NOM isolation procedure proposed by Croue [7]

A comprehensive isolation and fractionation scheme was recently proposed by Leenheer et al. [5]. This scheme takes a tiered approach that, depending on the number of stages employed, produces up to seven operationally-defined fractions. These fractions can be subsequently used in compound-class specific characterization procedures. This scheme is illustrated in Fig. 1 taken from Leenheer and Croue [6]. The isolation scheme is based on the use of a combination of reverse osmosis, dialysis, XAD-8 and XAD-4 resin adsorption, and desalting by zeotrophic distillation [6, 7]. The use of the "first" tier results in four NOM fractions classified as: colloidal (COM), hydrophobic (HPO), transphilic (TPI), and hydrophilic (HPI). The experimental approach is illustrated in Fig.2 taken from Croue [7].

Although variation in the properties of the fractions will occur depending on NOM source and the isolation procedures employed, some general observations on their chemical nature have been made [7]. As the dissolved fraction becomes more hydrophobic (HPI<TPI<HPO) the C/O, C/H, and C/N atomic ratios increase. This is due to greater amounts of nitrogenous and oxygenated functional groups in the more hydrophilic dissolved fractions and the COM.

A number of recent investigations have used complete or simplified versions of this scheme to investigate NOM characteristics and their influence on various water treatment processes. Several studies have focused on wastewater effluent organic matter (EfOM) whose characteristics can differ significantly from source water NOM due to the greater influence on microbial derived organic matter. Jarusutthirak et al., [8] examined the fouling of nano (NF) and ultra (UF) filtration membranes. The colloidal fraction caused significant flux decline arising from pore blockage. The HPO and TPI fractions showed less fouling due to the smaller size and higher charge density that promoted electrostatic interactions with the membranes. Further characterization of the fouled membranes implicated polysaccharides and amino-sugars from the colloidal fraction as the cause of membrane flux decline. The behaviour of these fractions and soluble microbial products (SMPs) in oil-aquifer treatment of EfOM was examined by Rauch and Drewes [9]. Both COM and HPI appeared to be subject to microbial removal while both HPO and SMPs appeared recalcitrant to microbial degradation.

COM also exhibited additional physical removal through either adsorption or filtration. The authors concluded that physical removal of recalcitrant HPO, HPI, and SMPs would be minimal and that significant subsurface transport would be possible. The formation of disinfectant by-products (DBPs) from various NOM fractions and the role of water treatment was examined by Hwang et al. [10]. Their approach resulted in fractions that contained both acid and neutral components in the HPI fraction. The highest yields of trihalomethanes (THMs) during chlorination were from the HPI acid + neutrals (HPIA+N) fraction. The lowest yields were from the COM fraction. The HPI fractions had the highest yields of haloacetic acid (HAA) and haloaetonitriles (HANs).

Factors such as NOM source and reactivity during treatment (filtration, aquifer treatment, disinfection) contribute to the composition of NOM as investigated through comprehensive fractionation. As a whole the studies illustrated here and other studies show that comprehensive NOM fractionation can greatly improve our understanding of NOM behaviour in water treatment.

2.2. Size of molecules and aggregates

The previous discussion focused on chemical fractionation of NOM that was fundamentally based on NOM solubility behaviour. Another important fractionation approach involves size separation, generally utilizing filtration or less frequently, dialysis. NOM can be considered to exist of a continuum of molecular weights or hydrodynamic diameters. This continuum can be simplified to three operationally defined fractions: dissolved NOM, colloidal NOM, and particulate NOM. Although the extremes of this continuum can be easily visualized, for example POM occurring as a suspended bacterial cell and DOM occurring as dissolved low molecular weight compounds, the intermediate COM is more poorly defined.

The choice of filtration method and its degree of success depends on whether size fractionation alone or concentration and isolation are also desired. Simple filtration-based fractionation approaches can be successful if small sample volumes and large pore sizes (> 1-5 micrometer) are used. The most common example being the calculation of POC by measurement of TOC (unfiltered sample) and DOC (filtered sample). However simple "dead-end" filtration is highly subject to clogging. If large volumes need to be processed then some form of tangential-flow filtration (TFF) is employed [11, 12]. In TFF, flow is principally across the membrane surface with only a small percentage of the flow at any given time actually passing through the membrane. This leads to a high degree of shear over the membrane surface, effectively limiting the formation of a fouling layer. Filtration configurations usually consist of hollow-fiber or flat plate membranes. Size cutoffs are defined by the nominal molecular weight (MWCO) of the membrane. Classifications of membranes having increasingly finer pores are: microfiltration (MF), ultrafiltration (UF), and nanofiltration (NF), the latter category also commonly referred to as dialysis. As previously described, incorporation of dialysis with resin-based adsorption methods has been recently applied to more comprehensively characterized NOM [6].

There has been considerable controversy about the molecular weight of dissolved natural organic matter. Some results give the average MW as 100-200 kDa, however, most recent studies report values considerably less than 20 kDa. Reasons for the discrepancy may stem from the differences between the isolation/concentration methods used or from the fact that the molecules may aggregate under certain conditions (low pH or high ionic strength).

The humic substances (hydrophobic acids) fraction appears to have the most detailed MW information available. It is readily isolated and has a strong chromophore that can be used for detection in elution methods such as flow field-flow fractionation (FlFFF) [13] and

size exclusion chromatography (SEC) [14]. At high pH (>8) these molecules do not appear to aggregate and provided the correct carrier matrix is chosen, there are no strong stationary phase interactions [15]. For these reasons consistent results have been obtained using FFF, SEC, fluorescence [16 - 18] and colligative properties [1]. The conclusions from these studies is that the average MW of fulvic acid, which is the dominant fraction in freshwater, is quite low (<2000 Da). Recent results for humic acid tend to suggest an only slightly higher average MW for this fraction (2-10 KDa). Both FFF and SEC analyses that utilized UV absorbance (λ generally = 254 nm) for detection tend to show a rather featureless distribution that has been modelled as a log normal distribution [19]. The actual size of the species in the complex NOM mixture present in natural waters is likely to be quite different from the isolated fractions due to interactions between the different types of molecules

The MW of other NOM fractions has not been studied extensively. They do not generally have any strong chromophores and hence are difficult to detect by UV-visible absorption measurements. Recent work has extended the scope of SEC by including online TOC and fluorescence detection [20 - 22]. TOC detects all organic carbon and thus is not biased by differences in the optical properties of various NOM components. Both SEC and FFF are sufficiently sensitive to allow MW characterization of NOM in water samples without isolation/concentration. This reduces the possibility of alteration of the MW distribution by either aggregation or by preferential isolation of any specific MW fractions. On-line fluorescence detection, through the choice of different excitation/emission wavelengths, can be used to examine MW distributions of humic substances and protein-derived material. This latter material is expected to be a major component in waters impacted by effluent from sewage treatment plants. The use of these additional detectors, particularly in conjunction with SEC, has shown a much more complex MW distribution for unfractionated NOM in water. In a number of studies protein-rich NOM was shown to result in an additional peak at higher MW (20-40 kDa) which can significantly increase the weight average MW of the NOM. Furthermore low molecular weight peaks (<1 kDa) having low UV absorbance have been observed using SEC with TOC detection. An example for a groundwater sample where UV detection is compared to TOC detection is shown in Fig.3 taken from Her et al. [22].

In addition to the dissolved and aggregated NOM species, considerable amounts of organic carbon may be associated with organic-mineral colloids. An example is the association of iron hydroxide colloids with NOM. By including ICP-MS detection in FFF and SEC separations the association of NOM with inorganic colloids has been investigated by a number of researchers. The MW of iron oxide-NOM colloids has been quantified in rivers by Benedetti et al. [23] using FFF-ICP-MS and TEM. The MW distributions of other elements have also been investigated by FFF-ICP-MS and SEC-ICP-MS [24]. Uranium associated with inorganic colloids and dissolved NOM was differentiated in a study of soil leachates by Jackson et al. [25]. This study demonstrated that use of both SEC-ICP-MS and FFF-ICP-MS allowed examination of elemental MW distributions over a larger range than possible using only a single technique.

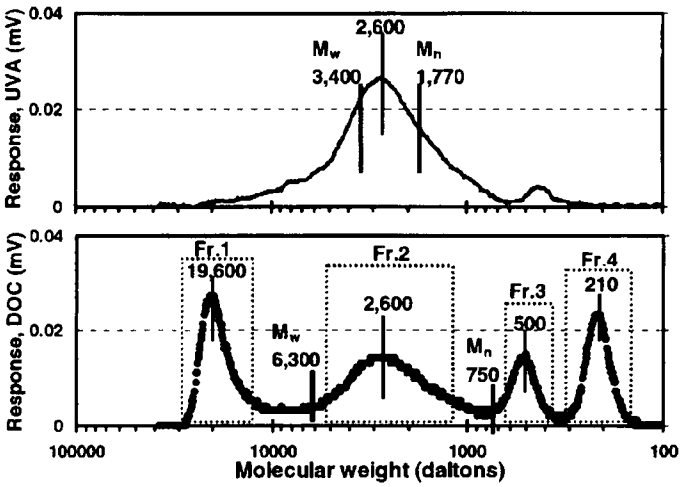


Fig. 3. Molecular weight characterization of NOM by size exclusion chromatography, illustrating the use of both UV and online TOC detection. Detection of humic substances at a molecular weight of approximately 2-3 kD is achieved by both detectors whereas non-humic substance of both high and low molecular weight are detected by on-line TOC [22]

2.3. Proposed molecular structures for NOM

As a consequence of the importance of NOM, the search for chemical structures to describe this material began more than a century ago. It should be emphasised that NOM is a complex mixture of molecules that can never be adequately be described by a single structure. However, suggestions of typical structures present are sometimes useful.

Early structures of humic substances resembled coal with large aromatic regions and many carboxylic acids and other oxygen containing polar functional groups rendering the molecules soluble (Fig. 4) [26]. The application of modern spectroscopic techniques has seen much progress being made on structure elucidation. The molecular structures now being advanced are much more aliphatic with perhaps only 20-30% of the carbons being aromatic [27]. The functional groups are still present with carboxylic acids dominating.

There is little difference between the structure of humic and fulvic acid, which is not surprising considering the similar manner in which they are defined. Humic acids have slightly greater MW and fulvic acids have a greater proportion of polar functional groups.

pH titrations show that there is a wide spread in pKa of the acid functional groups present [28]. Most acid groups appear to have a pKa greater than 2 and the substances are not fully hydrolysed until the pH is above 10. This reflects the different acid functional groups, which include carboxylic acids and phenols, as well as the different structural environments of these groups within the molecule. The polyelectrolyte effect may also play a role, whereby even the same functional group on a given molecule becomes increasingly more difficult to ionise as the number of anionic groups per molecule increases.

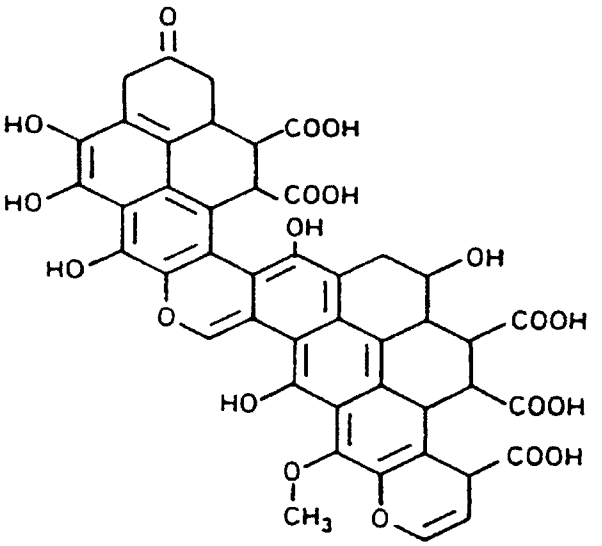


Fig. 4. Proposed structure for humic substances

Optical properties of NOM, especially UV absorbance, provide information on NOM characteristics. Absorbance at a wavelength of 254 nm when normalized to the DOC is defined as the specific UV absorbance (SUVA) and has been correlated to NOM aromaticity [29]. SUVA has also been shown to be positively correlated to average MW [30]. SUVA has in turn been related back to NOM source. SUVA appears to be a very useful property in predicting membrane fouling potential, perhaps suggesting a hydrophobic mechanism in this process [31]. SUVA was found not to be a good predictor of chemical reactivity however [29].

3. PROPERTIES AND ROLE OF NOM IN AQUATIC PROCESSES

3.1 . Surface activity

NOM displays some properties similar to those of a weak surfactant. This is usually attributed to the humic fraction. The most obvious manifestation of this is the appearance of foams in turbulent waters such as in surf and at the bottom of waterfalls or rapids. In contrast pure water will not foam at all. This may have some significance transferring contaminants from water to the atmosphere, where long range transport can occur [32].

Humic substances dissolved in water will lower the surface tension of water slightly. For example 100 mg L^{-1} of humic acid may result in a decrease in surface tension of water from 72 mJ m^{-2} to about 60 mJ m^{-2} [33]. Larger decreases have been reported occasionally for humics extracted from marine sediments [34]. The ability to lower the surface tension of water is usually attributed to amphipathic structures that have a hydrophobic region in the molecules separated from some hydrophilic groups. This is never apparent in the molecular structures suggested for typical NOM and perhaps this indicates an oversight in the structure predictions.

Strong synthetic surfactants display a surface tension versus concentration curve that first decreases and then reaches a plateau after the critical micelle concentration. There has been some controversy about whether this should also apply to humic substances. Some humic samples have been reported to display the sharp plateau but the concentrations required are of the order of 1000 mg L^{-1} which is well above the levels normally encountered in natural systems [34]. Wershaw [35] has argued for the micelle concept to be applied to humic substances but there is not widespread support for this. However, there is no doubt that DOM does tend to aggregate at lower pH and higher ionic strength, especially if divalent cations are involved, but whether the mechanism should invoke the hydrophobic effect is questionable. Perhaps hydrogen bonding and ion bridging is a more appropriate explanation. The formation of humic aggregates is probably necessary to understand the ability of humic and fulvic acid to "solubilise" nonpolar compounds in water [36] but a surfactant-like molecular structure may not be essential.

3.2 . Contaminant binding to dissolved NOM

3.2.1. Metal complexation and toxicity

It has long been realised that NOM can complex trace metals. Reference to the structure of humic substances (Fig. 4) reveals the presence of many oxygen-containing functional groups (carboxylic, phenolic) that should act as good ligands. Some studies also have suggested that heteroatoms (S, N) can contribute to metal complexation, especially for non-humic fractions that may be largely proteinaceous such as the transphilic neutral fraction [37]. From the 1970s there have been many studies on the binding capacity and strength of metals to NOM. Electrochemical methods such a ion selective electrodes and anodic stripping voltammetry and other experiments can be designed to evaluate the lability of metal-DOM complexes and often enable the stability constant (K_{stab}) to be determined [38]. Values for K_{stab} between 10^4 and 10^9 are typical [39]. Generally the binding strength of transition metals follows the classical Irving-Williams series supporting the ligand complexation model. A number of computational approaches to describing the metal complexing properties of NOM have been developed [40 - 42]

From the 1970's it was also recognized that water composition plays an important role in metal aquatic toxicity and bioavailability. The binding of metals to NOM has particularly important consequences. It is widely held that the uncomplexed or "free" forms of metals are most bioavailable, and hence most toxic at high concentration. Some of the earliest evidence for the role of organic complexation in mitigating metal toxicity was provided by Zitko et al. [44], Pagenkopf et al. [45] and Anderson and Morel [46]. These and other studies lead to the development of conceptual and computational models including the free ion activity model (FIAM), the gill surface interaction model (GSIM) and more recently the biotic ligand model (BLM). The BLM is currently viewed by EPA as the best general approach to incorporate the role of water composition into water quality criteria (WQC) for metals [43] and its essential components are illustrated in Fig. 5.

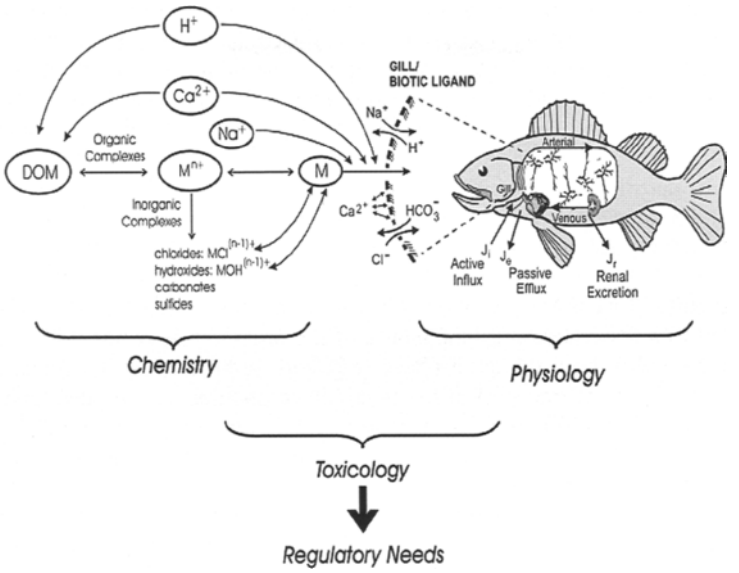


Fig. 5. Essential elements of the Biotic Ligand Model for metal toxicity to aquatic organisms, including the role of dissolved organic matter (DOM) in toxicity reduction through the formation of non-bioavailable metal-organic matter complexes [43]

An underlying assumption for metal toxicity to aquatic organisms is that toxicity is related to a critical level of metal accumulation at the site of toxic action, which in the case of many aquatic organisms is the gill surface. It is viewed that water composition will influence the amount of dissolved metal required to reach this critical gill metal level. In addition to quantitatively describing the role of metal-NOM complexes, the BLM describes the influence of hardness (Ca^{2+}, Mg^{2+}) and pH on metal toxicity. The role of hardness and pH in aquatic toxicity is explained primarily by a competition between these ions and the metals for the gill binding sites. The effect of hardness is incorporated into current WQC through the use of hardness-dependent equations. The role of NOM is empirically accounted for in the regulatory process by setting site specific WQC using the water-effects ratio (WER) approach. WERs are determined by measuring metal toxicity to a test species in water collected from the site and comparing the results to laboratory water prepared to contain an equivalent hardness. WERs greater than one generally implicate metal complexation, most often due to the presence of NOM. This is commonly observed where high DOC effluents, such as wastewater effluents, are discharged to surface waters.

3.2.2. Partitioning of anthropogenic organic compounds

Many toxic organic contaminants find their way into natural waters. Of most concern are the sparingly degradable nonpolar compounds which persist for long periods, tend to concentrate in the bottom sediments at deposition zones in rivers (e.g. reservoirs and estuaries) and can strongly bioaccumulate up a food chain. Examples include certain pesticides (herbicides and insecticides), phenols, polychlorinated biphenyls, polycyclic aromatic hydrocarbons and

trihalomethanes. Considerable success in explaining the environmental behaviour of these compounds has come from the recognition that this behaviour can be thought of as an equilibrium partitioning process [47]. Any organic compound can be thought to partition between the aqueous phase and an organic matter phase, which is described by;

$$K_p = C_{om}/C \quad (1)$$

where: K_p is the partition coefficient

C_{om} is the concentration in the organic matter phase

C_w is the concentration in the aqueous phase

The magnitude of the K_p is most influenced by the hydrophobicity of the organic compound and to a lesser extent by the nature of the organic matter phase [48]. Since NOM can be associated with both the solid phase (e.g. soil, sediment, aquifer materials) and the water phase, its role in the transport of hydrophobic compounds is complex. One observation that strongly supports the role of NOM in mobilizing hydrophobic compounds is the fact that addition of humic acid to water increases the solubility of nonpolar compounds. Chiou et al. [36] reported increases in the solubility of DDT and PCBs by factors of about 5 when 100 mg L^{-1} of humic acid was added to water. The suggested mechanism involves an argument similar to micellar solubilisation by surfactants, however, the structure of the humic aggregates responsible may not necessarily resemble traditional micelles. The situation with polar and ionisable compounds is much more complex than for hydrophobic compounds and their behaviour depends on water composition as well as the amount and type of NOM present. It has been observed that nonpolar organics and other contaminants are concentrated in surface microlayer films on natural waters [2]. However, it is not certain whether this is due to DOM or POM and other particulates accumulating at the water surface.

3.3. Organic coatings on particles

NOM also accumulates on solid surfaces. This can be observed as a biofilm on rocks, and boats, but is also apparent as soon as particles are exposed to natural waters containing even very small amounts of DOM (>0.2 mg L^{-1}). This has been demonstrated using electrophoretic mobility measurements of the particles [49-52]. The major observation is that with very few exceptions (e.g. freshly precipitated iron hydroxide in waters with low DOC concentrations) the particles in a natural water sample all display a negative surface charge with a relatively narrow distribution in the magnitude of the mobility (typically the relative standard deviation is about 20%). This is rather unexpected in view of the heterogeneous mineralogical composition of suspended particles in aquatic environments. The surface charge for samples collected from different freshwater sites may vary considerably, with the electrophoretic mobility usually being in the range - 0.5-2.5x10^{-8} m^2 V^{-1} s^{-1}. The major variable controlling this is the concentration of divalent cations present in the water.

The implication is that all natural aquatic particles, irrespective of the core mineral content, are covered with a coating of NOM rendering their surfaces uniform. The negative charge results from the acidic functional groups, which are at least partially hydrolysed at the pH of natural waters. This NOM surface coating model of aquatic particles is represented schematically in Fig. 6.

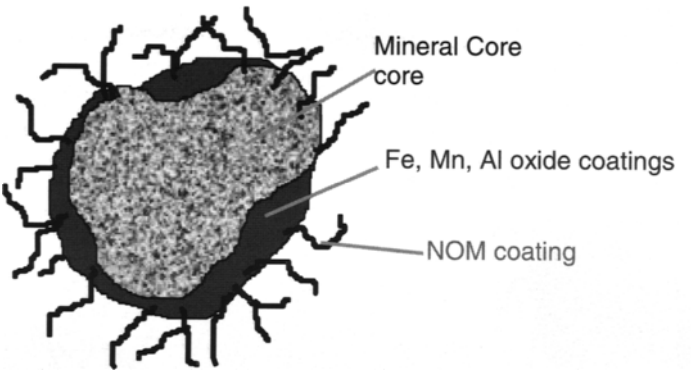


Fig. 6. Model of an NOM coated aquatic particle

Note that the surface coating is also considered to contain hydrous metal oxides (of Fe, Al, Mn, etc), which is important in the uptake and release of contaminants particularly during changes in redox and pH conditions. However, the surface charge characteristics of the particles are controlled by the organic coatings otherwise the isoelectric point of the metal oxyhydroxides would also play a role and this does not appear to be the case. The only exception to the observation of ubiquitous negative surface charge has been in streams heavily impacted by iron-rich acidic mine waters [53]. The formation of abundant iron oxides in these waters that are generally low in NOM and are acidic, allows for the presence of positively charged particles and thus presumably at least partially bare oxide surfaces.

In general the colloidal particles in fresh waters are stable implying they do not aggregate or settle. This is due to the charge repulsion between the negative particles and the small size (<1 μm). Coagulation may be induced if the ionic strength increases which often occurs due to saline ground water intrusion or during estuarine mixing with sea water. DLVO theory predicts that the critical coagulation concentration for divalent cations should be a factor of about 64 less than for monovalent cations ($1/Z^6$). For this reason, even in sea water, where the major ion is Na^+, it is the hard water ions Mg^{2+} and Ca^{2+} that are responsible for coagulation and clarification of a natural water. Less than about 1 mg L^{-1} of either divalent ion is required to lower the surface charge sufficiently for coagulation to occur [54].

3.4 Contaminant association with particle surfaces

The major forms of a pollutant in an aquatic system are the freely dissolved species, which could include inorganic complexes of a metal (e.g. MOH^+, MCl^+) as well as the free ion (M^{z+}), dissolved species associated with DOM, and pollutants associated with suspended particles. This is illustrated in Fig. 7. It is also useful to distinguish between colloids and micron size particles because larger particles have the potential to settle out if the flow is low enough, whereas fine colloids will not settle unless aggregation occurs. This simple speciation scheme is useful in predicting the bioavailability, transport and fate of contaminants in aquatic environments.

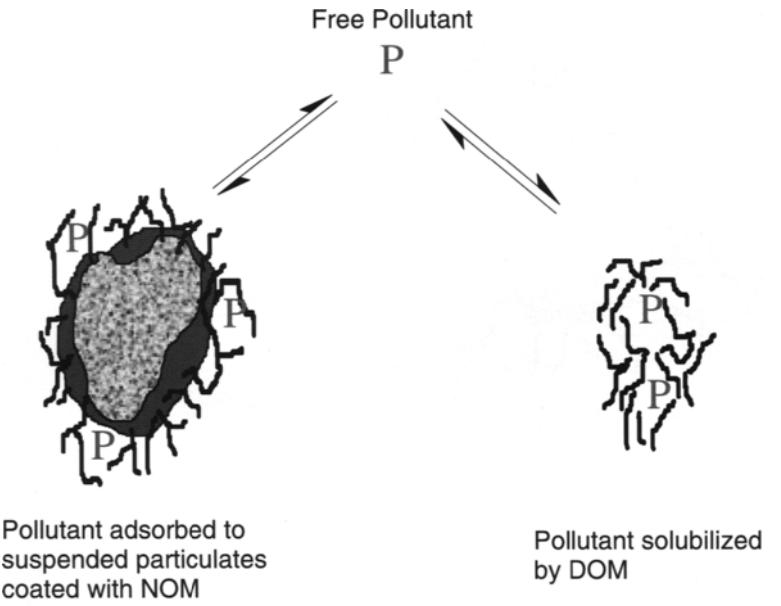


Fig. 7. Influence of NOM on pollutant speciation

The NOM concentration can influence the proportion of a pollutant found in the dissolved or particulate forms. This is illustrated by the experiments of Day et al. [55] who found that adding enough NOM to coat goethite particles greatly increased the adsorption of γ-BHC (atrizine) but further additions of NOM resulted in a decrease in the amount of herbicide adsorbed. This can be understood by referring to Fig. 7. Excess NOM added to the DOC pool thus solubilizing the γ-BHC and reduced the amount adsorbed to the suspended particles.

Lyven et al. [24] used flow field-flow fractionation with ICP-MS detection to show that metals in an aquatic system could be distributed between both DOM, detected as carbon by the ICP-MS and colloidal nanoparticle forms, detected as Fe and Al. Fig. 8 shows the element size distributions which show that Cu and U is mainly complexed to DOC whereas Pb is bound to Fe/Al rich suspended particles. Zn was distributed between both the DOM and colloidal phases. In this experiment the free dissolved species were not measured as they are lost through the channel membrane during the separation step. They further demonstrated that when the water sample was spiked with either Cu^{2+} or Pb^{2+}, these metals were also bound almost exclusively to either the DOM or colloidal matter respectively.

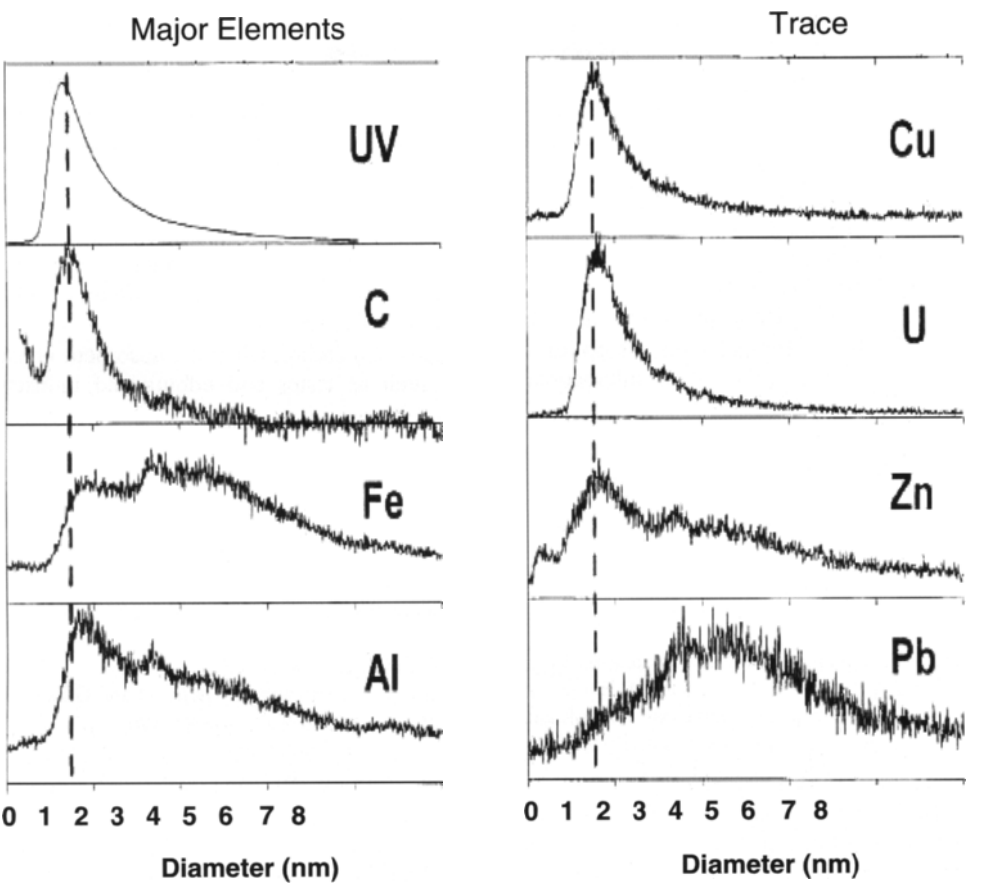


Fig. 8. Trace metal contaminants associated with DOM and colloidal nanoparticles as determined by FFF-UV-ICP-MS [56]

4. INFLUENCE OF NOM ON WATER TREATMENT

The specifics of NOM behaviour in water treatment are the subject of several chapters of this book. Therefore this chapter is intended to be a reference work and has focused on a discussion of the fundamental properties and processes involving NOM. However a few general points concerning the importance of NOM in drinking water treatment are worth mentioning:

- Although NOM can be effectively removed in treatment plants, high levels can cause undesirable consequences. High levels of NOM leads to the excessive use of treatment chemicals and foam formation can be a physical nuisance
- The feed water in treatment plants should ideally be continuously monitored in order to optimise the dosage of coagulants and flocculants. Variations in the concentration

and composition of the NOM lead to different dosage requirements making it more difficult to control the plant.

- In drinking water treatment where chlorination is usually employed for final disinfection trihalomethane (THM) and other organochlorine compounds may be produced. Although the levels are low it is asserted that long term exposure may result in health problems. This has led to calls for chlorination to be replaced by alternative technologies such as UV irradiation, ozonation or membrane filtration.
- NOM results in clogging of membrane filters requiring more frequent cleaning and shorter lifetimes for reverse osmosis and ultrafiltration membranes. The relative amounts of dissolved versus colloidal NOM are important in that different fouling mechanisms may be related to these different types on NOM.
- NOM competes for adsorption sites on activated carbon when the adsorbent is used for the removal of microcontaminants such as tastes and odours and industrial pollutants.

REFERENCES

[1] E.M. Thurman, Organic Geochemistry of Natural Waters. Martinus Nijhoff/Dr. W. Junk Publishers, 1986.
[2] K.A. Hunter, Mar. Chem., 9 (1980) 49.
[3] F.J. Stevenson, Humus Chemistry. Wiley Interscience Publications. John Wiley and Sons, New York (1982).
[4] E.M. Thurman and R.L Malcolm, Environ. Sci. Technol. 15 (1981) 463.
[5] J.A. Leenheer, J.P. Croue, M. Benjamin, G.V. Korshin, C.J. Hwang, A. Bruchet and G.R. Aiken, Comprehensive isolation of natural organic matter from water for spectral characterizations and reactivity testing. In: Natural Organic Matter and Disinfection By-Products, S.E. Barrett, S.W. Krasner and G.L. Amy (eds), ACS Symposium Series 761, 2000.
[6] J.A. Leenheer, J.P. Croue, Environ. Sci. Technol. 37 (2003) 18.
[7] J.P. Croue, Environ. Monit. Assessment, 92 (2004) 193.
[8] C. Jarusutthirak, G. Amy and J.P. Croue, Desalination, 145 (2002)247.
[9] T. Rauch and J.E. Drewes, Water Sci. Technol., 50 (2004) 245.
[10] C.J. Hwang, M.J. Sclimenti and S.W. Krasner, Disinfection by-product formation reactivities of natural organic matter fractions of a low-humic water, S.E. Barrett, S.W. Krasner and G.L. Amy (eds.), Natural Organic Matter and Disinfection By-Products, ACS, Washington, DC, 2000,173.
[11] B.G. Whitehouse, R.W. McDonald, K. Iseki, M. Yunker and F. McLaughlin, Mar. Chem., 26 (1989) 371.
[12] G.B. Douglas, R. Beckett and B.T. Hart, Hydrological Proc., 7 (1993)177.
[13] R. Beckett, J. Zhang and J.C. Giddings, Environ.Sci. Technol., 21 (1987) 289.
[14] Y.P. Chin and P.M. Gschwend, Geochim. Cosmochim. Acta, 55 (1991) 1309.
[15] M. De Nobili and Y. Chen, Soil Sci., 164 (1999) 825.
[16] C. Pelekani, G. Newcombe, V.L. Snoeyink, S. Assemi and R. Beckett, Environ. Sci. Technol., 33 (1999) 2807.
[17] J.R. Lead, K.J. Wilkinson, E. Balnois, B.J. Cutak, C.K. Larive, S. Assemi and R. Beckett, Environ. Sci. Technol. 34 (2000) 3508.
[18] S. Assemi, G. Newcombe, C. Hepplewhite and R. Beckett, Water Res., 38 (2004) 1467.
[19] S.E. Cabaniss, Q. Zhou, P.A. Maurice, Y. Chin and G.R. Aiken, Environ. Sci. Technol., 34 (2000) 1103.
[20] S. Huber and F. Frimmel, Int. J. Environ. Anal. Chem., 49 (1992) 49.
[21] N. Her, G. Amy, D. Foss and J. Cho, Environ. Sci. Technol., 36 (2002) 3393.
[22] N. Her, G. Amy, D. Foss, J. Cho, Y. Yoon and P. Kosenka, Environ. Sci. Technol., 36

(2002)1069.
[23] M. Benedetti, J. F. Ranville, T. Allard and N. Menguy, Colloids Surf., A, 217 (2003) 1.
[24] B. Lyven, M. Hassellov, D.R.. Turner, C. Haraldsson and K. Andersson, Geochim. Cosmochim. Acta., 67 (2003) 3791.
[25] B.P. Jackson, J.F. Ranville, P.M. Bertsch and A. Sowder, Environ. Sci. Technol., 39 (2005) 2478.
[26] W. Fuchs and H. Huminsaurin, Kolloid Z., 52 (1930) 124.
[27] C. Steelink, Implications of element characteristics of humic substances, G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy (eds.), Humic Substances in Soil, Sediment and Water, Wiley Interscience, New York, 1985.
[28] G. Newcombe, M. Drikas, S. Assemi, and R. Beckett, Water Res., 31 (1997) 965.
[29] J.L. Weishaar, G.R. Aiken, B.A. Bergamaschi, M.S. Fram and R. Fuji, Environ. Sci. Technol., 37 (2003) 4702.
[30] Y. Chin, G. Aiken and K. Danielson, Environ. Sci. Technol., 31 (1997) 1630.
[31] J. Cho, J. Sohn, H. Choi, I.S. Kim and G. Amy, J. Water Supply: Res. Technol. – Aqua, 51 (2002) 109 .
[32] K.A. Hunter and P.S. Liss, Organic sea surface films, E.K. Duursma and R. Dawson (eds.), Marine Organic Chemistry, Elsevier, Amsterdam, 1981.
[33] R. Le and N.P. Beckett, Colloids Surf., 44 (1990) 35.
[34] K. Hayase and H. Tsubota, Geochim. Cosmochim. Acta, 47 (1983) 947.
[35] R.L. Wershaw, J. Contam. Hydrology, 1 (1986) 29.
[36] C.T. Chiou, R.L. Malcolm, T.I. Brinton and D.E. Kile, Environ. Sci. Technol., 20 (1986) 502.
[37] J.P. Croue, M.F. Benedetti, D. Violleau and J.A. Leenheer, Environ. Sci. Technol., 37 (2003) 328.
[38] T.M. Florence, Electrochemical approaches to trace element speciation in waters. A review, Analyst, (1986).
[39] B.T. Hart, Environ. Technol. Lett., 2 (1981) 95.
[40] E. Tipping, Aquat. Geochem., 4 (1998) 3.
[41] J.C. Westall, , J.L. Zachary and F.M.M. Morel. 1976. MINEQL: A computer program for the calculation of chemical equilibrium composition of aqueous systems. Tech. Note 18. Ralph M. Parsons Lab. Dep. Civil Eng., Massachusetts Inst. Technol., Cambridge, MA.
[42] M. Benedetti, C.J. Milne, D.G. Kinniburgh, W.H. Van Riemsdijk and L.K. Koopal, Environ. Sci. Technol., 29 (1995) 446.
[43] P. Paquin, V. Zoltay, R.P. Winfield, K.B. Wu, R. Mathew, R.C. Santore and D.M. Di Toro, Comp. Biochem. Physiol. C., 133 (2002) 305.
[44] P. Zitko, W.V. Carson and W.G. Carson, Bull. Environ. Contam. Toxicol., 10 (1973) 265.
[45] G.K. Pagenkopf, R.C. Russo and R.V. Thurston, J. Fish. Res. Board Can., 31 (1974) 462.
[46] D.M. Anderson and F.M.M. Morel, Limnology Oceanography, 23 (1978) 283.
[47] S.W. Karickhoff, D.S. Brown and T.A. Scott, Water Res., 13 (1979) 241.
[48] R.P. Schwarzenbach, P.M. Gschwend and D.M. Imboden, Environmental Organic Chemistry, John Wiley and Sons, New York, 1993.
[49] R.A. Niehof and G.I. Loeb, Limnology Oceanography, 17 (1972) 7.
[50] K.A. Hunter and P.S. Liss, Nature, 282 (1979) 823.
[51] E. Tipping, Mar. Chem., 18 (1986) 161.
[52] R. Beckett, The surface chemistry of humic substances, R. Beckett (ed.), Surface and Colloid Chemistry in Natural Waters and Water Treatment, Plenum, new York, 1990.
[53] P.P. Newton and P.S. Liss, Limnology Oceanography, 32 (1987)1267.
[54] M.R. Grace, T.M. Hislop, B.T. Hart and R. Beckett, Colloids Surf., 120 (1997) 123.
[55] G.M. Day, B.T. Hart, I.D. McKelvie and R. Beckett, Environ. Technol., 18 (1997) 769.
[56] M. Hassellov, B. Lyven, C. Haraldsson and W. Sirinawin, Anal. Chem., 71 (1999) 3497.

This Page Intentionally Left Blank

Interface Science in Drinking Water Treatment
G. Newcombe and D. Dixon (editors)
© 2006 Elsevier Ltd. All rights reserved.

Chapter 18: Fundamentals of particle stability

C. R. O'Melia

Department of Geography and Environmental Engineering
The Johns Hopkins University
Baltimore, MD 21218, U.S.A.

A chemist does not think, indeed does not live, without models
The Periodic Table, Primo Levi

1. INTRODUCTION

In most natural aquatic environments, particle stability and the kinetics of the aggregation and deposition of aquasols (fine particles in aquatic systems) are controlled by surface and solution chemistry. In many technological environments, a task of the engineer is to reduce or eliminate the retarding effects of surface and solution chemistry on these kinetics in order to accomplish aggregation or deposition in economically feasible treatment systems. In this chapter we consider why these are so.

We begin with a presentation of the relationship between interaction force and interaction energy. This is followed by descriptions of van der Waals interactions which occur between all colloidal particles, of the origins of surface charge, and of electrostatic interactions between particles. Because of its possible importance in fresh waters, the secondary minimum in the interaction energy between two solids is discussed. The adsorption and conformations of polymers and polyelectrolytes at solid-water interfaces are reviewed, as are the properties and effects on colloidal stability of macromolecular natural organic matter. Finally, use of these concepts in studying deposition phenomena is introduced.

In this chapter we take the blurred boundary between physics and chemistry and include as aquasol chemistry much of what can also be described as physics. This includes, for example, coulombic interactions in DLVO theory and the dipole-dipole, dipole-induced dipole, and fluctuating dipole interactions in van der Waals forces. The term specific chemical interactions will be used primarily to refer to noncoulombic interactions.

2. FORCE, ENERGY, AND THE DERJAGUIN APPROXIMATION

Chemists often characterize interactions between particles with calculated interaction *energies*. This is because experimental data for molecular systems often have a thermodynamic base and are more readily expressed as chemical interaction energies such as the Gibbs and Helmholtz free energies. Physicists and engineers frequently describe interactions between macroscopic bodies by *forces*, an example of which is the gravity force. These forces are often experimentally accessible to direct measurement. Colloidal particles fit within the size range from molecules to macroscopic dimensions and both energies and forces have been used to describe the interactions between them, with chemical measurements used

to obtain interaction energies and physical measurements used to determine interaction forces. Since force is the negative of the derivative of energy with respect to distance, the conversion between energy and force would at first seem to be a simple matter. Often, however, the mathematics of the interactions and the geometries of the systems being considered make this a difficult task. A very useful approach for some systems was introduced by Derjaguin in 1934 [1]. Because of its widespread use in colloid chemistry, its simplicity, and its application subsequently in this chapter, it is summarized here. The development is taken from Israelachvili [2] and Derjaguin [3].

Let us begin by considering two spheres with radii a_1 and a_2 and make two assumptions. First, the spheres are separated by a small distance, h, which is much smaller than the radii of the spheres; $h << a_1$ *and* a_2. Second, in addressing interaction forces between the spheres, we limit our consideration to those forces that decay sufficiently rapidly with distance so that the contributions to the total force, $F(h)$, are insignificant very far from the point of closest approach denoted by h. When these conditions are met, the total interaction force can be calculated approximately by summing the forces between different small sections of the spheres. This is called the Derjaguin approximation. For these two spheres the following is obtained:

$$F(h) = 2\pi\left(\frac{a_1 a_2}{a_1 + a_2}\right) V_a(h) \tag{1}$$

Here $V_a(h)$ is the interaction energy of the two surfaces separated by the distance h

Together with Eqs. (2) and (3) given below, Eq. (1) is valid for any type of interaction force between two spheres. It gives the force between two spheres in terms of the interaction energy per unit area between two flat plates. It is valid when both the separating distance, h, and the range of the force are small relative to the radii of the two spheres. It is useful because it is often easier to determine the interaction energy between two flat plates than the force between two curved surfaces. Eq. (1) has been applied to aggregation processes of Brownian particles.

Similar relationships can be derived for other geometries. For a spherical particle interacting with a flat plate, the following is obtained:

$$F(h) = 2\pi a V_a(h) \tag{2}$$

where a is the radius of the particle. This expression has been used extensively to describe interactions in particle deposition on surfaces including porous media. Finally, for two cylinders with radii a_1 and a_2 crossed at right angles, Derjaguin's approximation yields

$$F(h) = 2\pi\sqrt{a_1 a_2}\, V_a(h) \tag{3}$$

This expression has been used to compare force measurements made with crossed cylinders having mica surfaces with predictions of interaction energies for flat plates in systems with similar chemistry [2].

The Derjaguin approximation can also be used to relate the interaction energy between two bodies to the interaction energy per unit area between two flat plates ($V_a(h)$). The following are obtained:

(a) for two spheres of radii a_1 and a_2,

$$V(h) = 2\pi\left(\frac{a_1 a_2}{a_1 + a_2}\right)\int_h^\infty V_a(h)\,dh \tag{4}$$

(b) for a sphere of radius a interacting with a flat plate,

$$V(h) = 2\pi a \int_h^\infty V_a(h)dh \tag{5}$$

and (c) for two cylinders with radii a_1 and a_2 crossed at right angles,

$$V(h) = 2\pi\sqrt{a_1 a_2}\int_h^\infty V_a(h)\,dh \tag{6}$$

3. VAN DER WAALS INTERACTIONS

Van der Waals interactions are ubiquitous, existing between atoms, molecules, colloidal particles, and macroscopic objects. They have an important role in hydrodynamic retardation or the lubrication effect between solids; without them, hydrodynamic drag would prevent interparticle contacts from occurring. They have a role in many other phenomena in gases, liquids, and solids including the aggregation and deposition of aquasols in aquatic environments. For this reason an overview of their origins and mathematical expressions is provided here. In this summary we will focus on van der Waals interactions among colloidal particles in water. For additional information the reader is referred to the following, listed in approximate order of increasing coverage: Gregory [4], Russel et al. [5], Parsegian [6], Lyklema [7], Israelachvili [2] and Hunter [8].

In 1873 the Dutch scientist J. D. van der Waals modified the ideal gas law and developed the following equation of state:

$$\left(P + \frac{a}{V^2}\right)(V - b) = RT \tag{7}$$

where P is the pressure, V is the molar volume, R is the gas constant, and T is the absolute temperature. The parameters a and b introduced into the ideal gas law have physical significance. The term b is subtracted from the molar volume to account for the finite volume of the molecules in the system. This is called the excluded volume effect. Finally, the term a/V^2 is added to the pressure to account for attractive interactive forces among neutral molecules. These forces have come to be known as van der Waals forces.

Van der Waals forces or interaction energies between molecules and colloidal particles are considered to arise from three types of interactions, resulting in what are known as Keesom-van der Waals, Debye-van der Waals, and London-van der Waals forces. These are also termed orientation, induction, and dispersion forces, respectively. Interactions between polar molecules are termed Keesom forces. Interactions between nonpolar molecules that can develop a dipole moment in an electric field are termed Debye forces. The Keesom and Debye interactions have been characterized using classical electrostatics. The third interaction, the London-van der Waals interaction, is quantum mechanical in origin. It acts between all atoms and molecules including totally neutral ones such as carbon dioxide. The resulting forces are often called dispersion forces because of their relationship to the dispersion of electromagnetic radiation including ultraviolet and visible light. The London-van der Waals forces between two atoms are the result of fluctuations in the charge densities of the electron clouds surrounding the nuclei of the atoms. Stated another way, although for a neutral atom the time average of the dipole moment is zero, there is a finite dipole moment from the instantaneous positions of the electrons about the nucleus. These instantaneous dipole moments produce electric fields that polarize any nearby neutral atom. The resulting interaction between the two dipoles produces an instantaneous attractive force between the two atoms, the time average of which is not zero.

For most molecules, dispersion forces dominate the van der Waals interactions in a gas. For example, they contribute essentially all of the van der Waals interactions between methane molecules in a gas phase [2]. For a polar molecule such as water, however, interactions between the permanent dipoles (orientation effects) contribute some 69 % of the theoretical interaction energy, with dispersion forces providing 24 % [2].

There are two theoretical approaches to determining the van der Waals interactions between colloidal particles or between colloidal particles and larger objects. In the older, classical, or microscopic approach, the interaction between two such bodies in a medium is obtained by a pairwise summation of all of the relevant intermolecular interactions. This approach results in relationships that contain a geometrical component and a constant, A, which depends on characteristics of the two bodies and the medium. The development of this approach was primarily by Hamaker [9] and the constant (A) is termed the Hamaker constant. The second approach was suggested by Lifshitz [10] and is called the macroscopic approach. In it the assumption of pairwise additivity is avoided. The total interaction is obtained from considering the macroscopic electromagnetic properties of all media. Each of these approaches is summarized briefly here.

3.1. Hamaker

The Hamaker or microscopic approach to determining van der Waals interactions is based on the assumption that the attractive force between two particles, each of which consists of many atoms, is the sum of the interactions between all pairs of atoms on the two particles.

The total interaction energy per unit area for two flat plates comprised of the same material and separated by the distance h across a vacuum is as follows:

$$V_a(h) = -\frac{A_{11}}{12\pi h^2} \tag{8a}$$

For two flat plates comprised of different materials (plate 1 and plate 2), the result is

$$V_a(h) = -\frac{A_{12}}{12\pi h^2} \tag{8b}$$

For the Hamaker constant of two different phases interacting across a vacuum, the following geometric assumption is frequently made:

$$A_{12} = (A_{11}A_{22})^{1/2} \tag{9}$$

The corresponding relationship for the van der Waals interaction force per unit area is

$$F_a(h) = -\frac{dV_a(h)}{dh} = \frac{A_{12}}{6\pi h^3} \tag{10}$$

The Hamaker constants for these systems, A_{11} and A_{12}, depend on the types of atoms comprising the materials and the atom density in each material. In principle, Hamaker constants can be calculated from molecular properties of the materials; however, these properties of the materials are often not known in sufficient detail. Hamaker constants for most condensed phases interacting with identical phases across a vacuum are in the order of 0.4 to 4 x 10^{-19} J. Hamaker constants of some aquatic systems of interest will be considered subsequently.

Van der Waals interaction energies and forces have been developed for other geometries. The interaction between a sphere and a plate is important in such systems as the deposition of colloidal particles in filtration technologies and the transport of aquasols and associated pollutants in aquifers. Using Derjaguin's approximation, we combine Eqs. (5) and (8b):

$$V(h) = 2\pi a \int_h^\infty V_a(h)dh$$

$$= -\frac{A_{12}a}{6h} \tag{11}$$

Note that the total van der Waals interaction energy (J) between a sphere and a flat plate is directly proportional to the radius of the sphere. Note also the energy decays in inverse proportion to the separating distance.

For two spheres of different size, the total interaction energy can be obtained using Derjaguin's approximation and combining Eqs. (4) and (8b). The result is as follows:

$$V(h) = -\left(\frac{a_1 a_2}{a_1 + a_2}\right)\frac{A_{12}}{6h} \tag{12a}$$

For two spheres of identical radii the result is

$$V(h) = -\frac{A_{12}a}{12h} \tag{12b}$$

These relationships are useful in considering aggregation processes in water and wastewater treatment systems and in surface waters including rivers, lakes, estuaries and oceans. They indicate that the van de Waals interaction energy depends on the size of the particles and decreases inversely with the separating distance between them. We see also that, for the aggregation of identical spheres (Eq. (12b)), the interaction energy is one-half of that between a sphere and a plate in particle deposition in porous media (Eq. (11)). Eqs. 12a and b are subject to the restrictions of the Derjaguin assumption, so that $h < a_1$, a_2, and a.

For interactions of two phases, 1 and 2, across a liquid such as water, the Hamaker constant is determined as follows:

$$A_{132} = A_{12} + A_{33} - A_{13} - A_{23} \tag{13}$$

Here, A_{132} denotes the Hamaker constant for phases 1 and 2 interacting across medium 3. For two identical phases 1 interacting across medium 3, Eq. (13) becomes

$$A_{131} \approx \left(A_{11}^{1/2} - A_{33}^{1/2}\right)^2 \tag{14}$$

From Eq. (14) we see that van der Waals interactions are always attractive for two particles of the same material interacting across a liquid medium. For dissimilar particles interacting across a liquid, A_{132} can be negative (e.g., when $A_{11} < A_{33} < A_{22}$). For most aquasols, however, the net van der Waals interaction is attractive. Finally we note that these expressions for the Hamaker constant and the van der Waals interaction do not consider the retardation of the London-van der Waals force as the separating distance increases between the particles.

3.2. Lifshitz

In the macroscopic or Lifshitz approach, the atomic structure of the colloids is not considered; solids and fluids are treated as continuous media. Interaction forces and energies between particles are derived from bulk properties of the media such as their dielectric constants and refractive indices. General expressions for interactions, such as Eqs. (8) to (12), developed previously, remain valid but the method of determining the Hamaker constant is changed.

The theory is physically and mathematically complex; only one analytical expression obtained with the use of simplifying assumptions is given here. When the separating distance is sufficiently small that the interactions are not retarded and the main electronic absorption frequency (ν_e) of all media is the same (typically in the UV at around 3 x 10^{15} s^{-1}), the following approximate expression for the Hamaker constant of two solids of identical composition 1 interacting across medium 3 is obtained:

$$\begin{aligned} A_{131} &= A_{\nu=0} + A_{\nu>0} \\ &= \frac{3}{4}kT\left(\frac{\varepsilon_1 - \varepsilon_3}{\varepsilon_1 + \varepsilon_3}\right)^2 + \frac{3\bar{h}\nu_e}{16\sqrt{2}}\frac{\left(n_1^2 - n_3^2\right)^2}{\left(n_1^2 + n_3^2\right)^{3/2}} \end{aligned} \tag{15}$$

Here ε_1 and ε_3 denote the dielectric constants of the solid phase and the separating medium, n_1 and n_3 are their respective refractive indices, and $\bar{h}$ is Planck's constant (6.626 x 10^{-34} J s). The first term on the right-hand-side of Eq. 14, termed the zero frequency term, represents the Keesom and Debye interactions; the second term arises from the London dispersion interactions.

Consideration of Eq. (15) yields the following:

[1] The van der Waals force between two particles of identical composition is always attractive. Similarly, the van der Waals interaction between any two phases in a vacuum or in air is positive ($\varepsilon_3 = 1$ and $n_3 = 1$; A is positive).

[2] The purely entropic zero-frequency component, $A_{\nu=0}$, can not be larger than $\frac{3}{4}kT$ or 3 x 10^{-21} J at 20 ^{0}C. This contribution is negligible in air ($\varepsilon_3 = 1$) but can be important in a polar medium such as water ($\varepsilon_3 = 80$ when the ionic strength is less than 0.1 M.

[3] Metals and metal oxides have high polarizabilities and, as a result, have high dielectric constants and refractive indices. Because of their high refractive index, these materials have high Hamaker constants in both air and water.

Example 1

Estimate the Hamaker constants for two alumina phases interacting across air and also water at 20 °C. Compare these results with two hydrocarbon phases (n-octane) interaction across these same media. Following are data for these materials [2].

Substance	Dielectric constant (ε)	Refractive index (n)	Absorption frequency (ν_e, 10^{15} s^{-1})
Water	80	1.333	3.0
Al_2O_3	11.6	1.75	3.0 (est.)
n-octane	1.95	1.387	3.0

We will use Eq. (15), the approximate expression for the unretarded Hamaker constant of two identical materials acting across a third medium. It is interesting first to determine the coefficients for the zero frequency and higher frequency terms, $\frac{3}{4}kT$ and $(3h\nu_e)/(16\sqrt{2})$.

$$\frac{3}{4}kT = \frac{3}{4}\left(1.381 \times 10^{-23}\ J\ K^{-1}\right)(293\ K) = 3.03 \times 10^{-21}\ J$$

$$\frac{3h\nu_e}{16\sqrt{2}} = \frac{3}{16\sqrt{2}}\left(6.626 \times 10^{-34}\ J\ s\right)\left(3 \times 10^{15}\ s\right) = 2.64 \times 10^{-19}\ J$$

A comparison of these coefficients illustrates again the dominance of the London dispersion force (non-zero frequency terms) over the Keesom and Debye interactions.

For Al_2O_3 in air or a vacuum we obtain

$$A_{131}(air) = 14 \times 10^{-20}\ J$$

Similarly, for Al_2O_3 in water,

$$A_{131}(water) = 4.2 \times 10^{-20}\ J$$

Comparison of these two results indicates (1) dispersion interactions dominate the Hamaker constant and the van der Waals interaction of alumina in both air and water and (2) the Hamaker constant in water is significantly less than in a vacuum or in air.

For n-octane in air or a vacuum we find that:

$$A_{131}(air) = 4.6 \times 10^{-20}\ J$$

Finally, for n-octane in water,

$$A_{131}(water) = 0.35 \times 10^{-20}\ J$$

From these calculations we see again that the Hamaker constant is larger in air than in water. In addition, the zero frequency term (Keesom and Debye interactions) is much larger than the dispersion interaction for this organic substance in water.

3.3. Retardation and electrostatic screening

We have considered three types of van der Waals interactions: Keesom, Debye, and London, or orientation, induction, and dispersion interactions. Each of these interaction energies between phases decreases with increasing separating distance. The Keesom and Debye interactions are derived with classical electrostatics. The inverse dependence of these interaction energies with distance (Eqs. (8a,b), (11), and (12)) holds to large separations. The London forces, in contrast, are electromagnetic and involve the time taken for the instantaneous electric field of one atom to reach a second atom and return. For an atom with a characteristic orbiting frequency of $3 \times 10^{15}\ s^{-1}$, the distance traveled by light during one rotation is $(3 \times 10^{8}\ m\ s^{-1})/(3 \times 10^{15}\ s^{-1}) \approx 100$ nm. As a result the returning field can find the direction of the instantaneous dipole to be different than before; the resulting interaction is less attractive and the dispersion energy between the two atoms decays faster as the distance increases. This is called the *retardation effect*. Note again that it influences only the dispersion interaction, the orientation and induction interaction energies remain non-retarded. For interactions between two molecules in a gas phase, retardation becomes important only at distances where the van der Waals interactions are already very weak so the effect is generally not significant. However, between colloidal particles in water, retardation can be significant and is often taken into account. There are several approaches in doing this.

When the Hamaker constant is evaluated using the microscopic (Hamaker's) approach, the result includes only the London-van der Waals dispersion force and is subject to retardation that causes A to vary with separating distance, h. There is no simple description of this retardation that describes the interaction energy over the entire range of separating distances. A number of approximate analytical expressions have been developed to describe retarded interactions for different geometries, particle sizes, and separating distances [11]. Some of these are presented in Table 1 together with companion expressions for non-retarded interactions.

The expressions in Table 1 for retarded interactions (Eqs. (2), (4), and (6)) all contain λ, the characteristic wavelength for the dispersion interaction, often assumed to be 100 nm. Eqs. (3) to (6) include the Derjaguin approximation in their development and so incorporate the requirement that the range of the interaction be smaller than the size of the particle; $h << a$. This can limit their accuracy for particles at the lower end of the colloidal size range.

The macroscopic or Lifshitz approach to calculating the van der Waals interaction includes the orientation, induction, and dispersion forces and also incorporates the retardation effect. As in the microscopic method, with the Lifshitz approach there again is no analytical expression relating the retarded van der Waals interaction force and the separating distance between two objects that is accurate at all separations. The Hamaker constant is not actually a constant; it can be viewed more rigorously as the Hamaker function, $A(h)$, that varies with the separating distance. Numerical solutions for $A(h)$ are available [12] and several approximate analytical expressions have been also developed. For non-retarded interactions, the Hamaker approach and some applications of the Lifshitz theory give the same geometric distance dependency so Eqs. (1), (3) and (5) in Table 1 can be used to describe geometric effects with both approaches.

The ionic strength of an aqueous solution has essentially no effect on the dispersion or London-van der Waals interaction. The zero frequency contribution (the Keesom and Debye forces), in contrast, are electrostatic and can be screened by the free ions in the solution. At high ionic strength ($I > 0$), their effective range is less than 1 nm [2] so that attraction between most materials in such solutions is determined by the dispersion force. For biological materials this may also be weak, so that van der Waals attractive interactions between biological materials in estuarine and marine systems may be quite small.

4. THE ELECTRICAL DOUBLE LAYER

An aquasol does not have a net electrical charge. The particle charge arising from surface complex formation, isomorphous substitution, and other reactions is termed the *primary* charge of the particle. This charge must be counterbalanced in the aquasol system and this occurs in the interfacial region adjacent to the particle. Fig. 1 is a schematic representation of a negatively charged particle in water with a cloud of ions (*diffuse layer*) surrounding it. Because the particle in the sketch has a net negative primary charge, an excess of ions of opposite (positive) charge accumulates in the interfacial region adjacent to it. The diffuse layer is formed by the electrostatic attraction of ions of opposite charge to the particle (counterions), by electrostatic repulsion of ions with the same charge as the particle (similions or coions), and by thermal or molecular diffusion that acts against the concentration gradients produced by the electrostatic effects. The primary charge and the diffuse layer, taken together, form an *electrical double layer.*

Table 1
Van der Waals interaction energies for selected geometries

Geometry	Expression	Number	Conditions*	Ref.
Plate - plate	$V_{A(a)} = -\frac{A}{12\ \pi\ h^2}$	(1)	unretarded, infinite thickness	[13]
Plate - plate	$V_A = -\left(\frac{A}{12\ \pi h^2}\right)\left(1+\frac{5.32h}{\lambda}\right)^{-1}$	(2)	retarded, LVDW only	[11]
Sphere - plate	$V_A = -\frac{Aa}{6h}$	(3)	unretarded, DA, $h \ll a$	[11]
Sphere - plate	$V_A = -\left(\frac{Aa}{6h}\right)\left(1+\frac{14h}{\lambda}\right)^{-1}$	(4)	retarded, LVDW only, DA, h << a	[11]
Sphere - sphere	$V_A = -\frac{A}{6}\left(\frac{a_1\ a_2}{a_1+a_2}\right)$	(5)	unretarded, DA, $h \ll a_i$	[9]
Sphere - sphere	$V_A = -\left(\frac{A}{6h}\right)\left(\frac{a_1 a_2}{a_1+a_2}\right)\left[1-\frac{5.32h}{\lambda}ln\left(1+\frac{\lambda}{5.32h}\right)\right]$	(6)	retarded, LVDW only, DA, $h \ll a$	[11]

* LVDW only = London van der Waals dispersion force only (Keesom and Debye van der Waals forces not included); DA = Derjaguin approximation.

4.1. The Gouy-Chapman diffuse layer

Electrical double layers at aquasol interfaces usually form spontaneously; they are ubiquitous in aquatic environments. They affect surface speciation and so, in turn, can contribute in varying degrees to chemical weathering, pollutant adsorption, surface catalyzed redox reactions, and a variety of other environmentally relevant processes. Their impact on adsorption and surface speciation will be illustrated in this section. We will also, in Section 6, consider the effects of the electrical double layer on the adsorption of polyelectrolytes and natural organic matter and on the conformations of these substances at solid-water interfaces. The major application of this section, however, is in considerations of the kinetics of aquasol aggregation and deposition processes.

The description of the electrical double layer presented here is taken from the classical work of Verwey and Overbeek [13]. This includes the earlier developments of the diffuse layer by Gouy [14] and Chapman [15]. A comprehensive and current presentation of this material is provided by Lyklema [16].

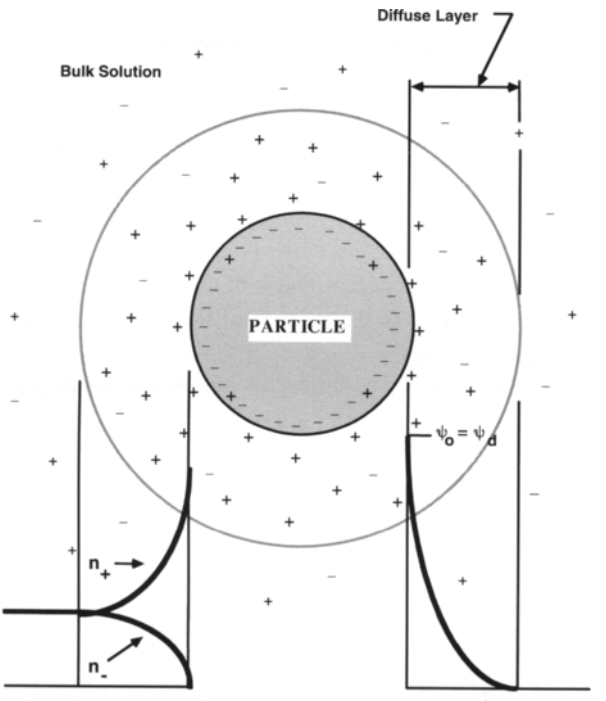


Fig. 1. Schematic representation of the electrical double layer (Gouy-Chapman model) surrounding a solid particle in solution

The primary surface charge and the diffuse layer charge expressed per unit of surface area are denoted here by σ_o and σ_d, respectively, both in units of C m^{-2}. Since the double layer is electrically neutral overall we can write

$$\sigma_o + \sigma_d = 0 \tag{16}$$

We will focus on a flat surface for which the diffuse layer charge varies only with x, the distance from the surface into the solution. Three relationships are used: (i) electroneutrality for the particle and its surrounding interfacial region as expressed in Eq. (16), (ii) Poisson's equation relating potential and space charge, and (iii) Boltzmann's expression relating ion concentration and potential. Results have been used, for example, in electrostatic models for coulombic interactions affecting proton transfer, metal adsorption, and anionic complex formation reactions.

Poisson's equation, based on electrostatics, is written for one dimension (x) as

$$\frac{d^2\psi(x)}{dx^2} = -\frac{\rho(x)}{\varepsilon_o \varepsilon} \tag{17}$$

Here $\psi(x)$ is the potential (V) at a distance x (m) from the flat particle surface, ε_o is the permittivity of free space (8.854 x 10^{-12} C V^{-1}) and ε is the relative dielectric constant of water (78.5 at 25 °C).

Boltzmann's equation relating ion concentration and electric potential is as follows:

$$n_i(x) = n_i(\infty) exp\left(-\frac{z_i e\psi(x)}{kT}\right) \tag{18}$$

Here $n_i(x)$ *and* $n_i(\infty)$ are the number concentrations of ion i at distance x from the surface and in the bulk solution, respectively, z_i is the valence of the ion (including sign), and e is the elementary charge (1.602 x 10^{-19} C).

Combining Eqs (17) and (18) and assuming a symmetrical electrolyte $(z_{i+} = z_{i-} = z)$ yields the *Poisson-Boltzmann (PB) equation* for this case:

$$\frac{d^2\psi(x)}{dx^2} = -\frac{zen}{\varepsilon_o \varepsilon}\left\{exp\left(-\frac{z_+ e\psi(x)}{kT}\right) - exp\left(-\frac{z_- e\psi(x)}{kT}\right)\right\} \tag{19}$$

Some analytical solutions to the PB equation are available for simple geometries and other constraints; numerical solutions are common. Here we will summarize a few of the more common analytical solutions.

It is useful to summarize the principal assumptions in the PB equation (19). The following compilation is adapted from Sposito [17] and Lyklema [16]:

- The surface is an infinite plane with a uniformly distributed charge density, σ_o.
- The ions in the solution do not occupy any volume; they are point charges.
- The electrolyte is symmetrical; $z_+ = z_- = z$.
- The water is a uniform continuum whose effects are through its dielectric constant, e, for which the value in the bulk fluid is used.
- The attractive or repulsive interaction of the ions with the surface is purely electrostatic and described by $z_i e\psi(x)$ for an ion located at x.
- We note that the potential in Poisson's equation is a mean electrical potential in a homogeneous field; an incoming unit charge is not seen by the field; it acts as a probe. Conversely, the electrical potential in the Boltzmann distribution considers ion-ion correlations; it is termed a potential of mean force. In this case an incoming unit charge causes rearrangement of the collection of charges in the field. In developing the Poisson-Boltzmann equation, these two electrical potentials are assumed to be equal. The significance of this assumption is not clear.

An interesting and useful result is obtained by considering a surface with a low surface potential, say $\psi_o < 25$ mV so that $ze\psi_o/kT < 1$. The following are obtained:

$$\frac{d\psi}{dx} = -\kappa\psi \tag{20}$$

and, considering the boundary condition that $\psi = \psi_o$ at $x = 0$,

$$\psi = \psi_o \exp(-\kappa x) \tag{21}$$

The term κ is defined as follows (units of m^{-1}):

$$\kappa = \left(\frac{e^2 \sum_i n_i z_i^2}{\varepsilon_o \varepsilon kT} \right)^{1/2} \tag{22}$$

For a symmetrical electrolyte,

$$\kappa = \left(\frac{2e^2 z^2 n}{\varepsilon_o \varepsilon kT} \right)^{1/2} \tag{23}$$

With this result we see that the electrical potential of a flat double layer decreases exponentially with distance from the surface when the surface potential is small. We see also that the reciprocal of κ, termed the Debye length (κ^{-1}), is a characteristic dimension or thickness of the diffuse double layer. The surface potential decays to a value of ψ_o / e at $\kappa x = 1$ when the surface potential is small. For large surface potentials, the decrease in potential with distance is somewhat greater than the exponential decrease calculated for large potentials using Eq. (21).

Considering the relationship between surface charge and surface potential, the following is obtained without restricting the surface potential too low values:

$$\sigma_o = -\sigma_d = \sqrt{8 \varepsilon_o \varepsilon kTn} \sinh\left(\frac{ze\psi_o}{2kT} \right) \tag{24}$$

When the surface potential is low, the relationship between surface charge and surface potential is as follows:

$$\sigma_o = -\sigma_d = \varepsilon_o \varepsilon \kappa \psi_o \tag{25}$$

It is often useful to express the Debye length, κ^{-1}, in terms of the ionic strength of the bulk solution, I, defined as:

$$I = \frac{1}{2} \sum_i z_i^2 C_i \tag{26}$$

This yields the following:

$$\kappa^{-1} = \left(\frac{\varepsilon_o \varepsilon kT}{2 N_A e^2 I} \right)^{1/2} \tag{27}$$

where κ^{-1} has dimensions of m. Considering water at 25 °C, using units of mol L^{-1} for I and expressing κ^{-1} in nm, Eq. (26) can be written as

$$\kappa^{-1} = \frac{0.30}{I^{1/2}} \tag{28}$$

Example 2

From experimental results obtained with alkalimetric titrations of hematite suspensions [18], the following information is obtained: (i) at pH 6.5 and in 0.001 M $NaNO_3$, the surface charge is +0.053 C m^{-2} and (ii) at pH 4.0 and in 0.1 M $NaNO_3$, the surface charge is +0.28 C m^{-2}. Assume the temperature is 25 °C. Using the Gouy-Chapman double layer model, for each of these two cases determine the double layer thickness (κ^{-1}), the diffuse layer charge (σ_o), the surface potential (ψ_o), and the concentrations of Na^+ and NO_3^- at the interface.

Case (i): pH = 6.5, I = 0.001 M, s_o = +0.053 C m^{-2}.

- Debye length (Eq. (28)) $\kappa^{-1} = 9.5\ nm$
- Diffuse layer charge (Eq. (16)) $\sigma_d = -\sigma_o = -0.053\ C\ m^{-2}$
- Surface potential (Eq. (24)) $\psi_o = +172\ mV$
- Interfacial ion concentrations at $x = 0$ (Eq. (18)) $[Na^+]_{x=0} = 1.2 \times 10^{-6}\ M$; $[NO_3^-]_{x=o} = 0.81\ M$

Case (ii): pH = 4.0, I = 0.1 M, σ_o = +0.28 C m^{-2}

- Debye length (Eq. (27)) $\kappa^{-1} = 0.95\ nm$
- Diffuse layer charge (Eq. (15)) $\sigma_d = -0.28\ C\ m^{-2}$
- Surface potential (Eq. (23)) $\psi_o = +140\ mV$
- Interfacial ion concentrations at $x = 0$ (Eq. (17)) $[Na^+]_{x=0} = 4.3 \times 10^{-6}\ M$; $[NO_3^-]_{x=o} = 23.2\ M$!

These two cases were selected for examination in part because of their differences. The first case (pH = 6.5, I = 0.001 M, σ_o = +0.053 C m^2) has a low surface charge and also a low ionic strength; conversely, the second case (pH = 4.0, I = 0.1 M, σ_o = +0.28 C m^2) has a high surface charge and a high ionic strength. These are *experimentally measured* or *determined* conditions. Despite these differences, the *calculated* surface potentials using the Gouy-

Chapman model are not dramatically different. In fact, the surface potential calculated at low surface charge and ionic strength (+ 172 mV) is somewhat larger than that calculated at high surface charge and ionic strength (+140 mV). As expected, similions [Na^+] are repelled from the positively charged surface and counterions [NO_3^-] are electrostatically attracted to it. The electrostatic attraction is sufficiently strong in case (ii) that concentrations of nitrate in excess of 20 M are predicted at the surface by this model. Such high concentrations are commonly predicted with the Gouy-Chapman approach. They arise because of the assumption in developing the Poisson-Boltzmann equation that all ions are point charges. While the assumption is physically unrealistic, this is not usually a problem. However, some of the calculated results are physically unrealistic and this *can* be a problem, usually when surface potentials are high and ionic strengths are also substantial. The difficulty was recognized long ago and was addressed by Stern [19]. We turn next to this modification of the double layer model.

4.2. The Stern Layer

Following after Stern [19], it is convenient to introduce an artificial subdivision on the solution side of the solid-water interface. Several such subdivisions are in common use. Here we begin by partitioning the solution side into two parts: an inner part or empty *Stern layer* which is devoid of ions and an outer part, a *Gouy-Chapman* or *diffuse layer* as described previously. The empty Stern layer is introduced to account for the finite size of the ions in the solution; it is assumed that counterions cannot approach the surface closer than their hydrated diameter. Hence the thickness of the empty Stern layer (δ) is in the order of the size of a hydrated counterion. Except for the restriction at close approach stemming from their finite size, counterions are assumed to be entirely mobile and have only coulombic interactions with the surface. Such a double layer system (actually, a triple layer system) is sketched in Fig. 2. Since it is considered empty, the charge in the Stern layer (σ_s) is zero and Eq. (16) still applies:

$$\sigma_o + \sigma_d = 0 \tag{16}$$

Equations developed previously for the double layer continue to pertain with the exception that the distance from the surface (x) is replaced by the distance from the Stern layer $(x - \delta)$. The boundary between the empty Stern layer and the diffuse layer is sometimes termed the *outer Helmholtz plane* (OHP). The potential at the inner boundary of the diffuse layer (the OHP) is no longer the surface potential; here we denote it as the diffuse layer potential, ψ_d.

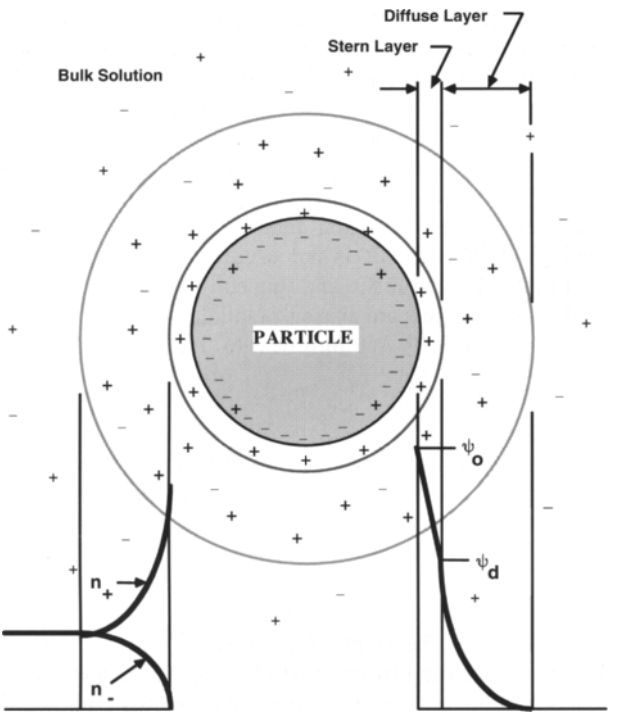


Fig. 2. Schematic representation of the electrical double layer (Stern model) surrounding a solid particle in solution

For the empty Stern layer illustrated in Fig. 2, we consider first the case in which the case the potential decay is linear with distance so we can write

$$\sigma_o = C_s^i(\psi_o - \psi_d) \text{ or } \frac{\sigma_o}{(\psi_o - \psi_d)} = C_s^i \tag{29}$$

Here C_s^i is the integral capacitance in the empty Stern layer (C V^{-1} m^{-2}; F m^{-2}).

Stern also proposed a model in which the inner or Stern layer contains ions that can be bound to the surface by electrostatic, van der Waals, or other forces strong enough to counter thermal diffusion from the surface. In this case, the Stern layer is charged and the overall electroneutrality of the interface can be written as

$$\sigma_o + \sigma_s + \sigma_d = 0 \tag{30}$$

in which σ_s is the charge in the Stern layer (C m^{-2}). The "adsorbed" ions may be partially dehydrated, so that the thickness of a charged Stern layer is expected to be smaller than that of an empty one. The outer boundary of the charged Stern layer is sometimes termed the *inner Helmholtz plane* (IHP).

Stern assumed a Langmuir type isotherm to describe the equilibrium between ions adsorbed in the charged Stern layer and those in the adjacent layer of solution. For details, see Hunter [20]. For dilute solutions the following is obtained:

$$\sigma_s \cong z_i e X_i N_s \exp\left[-\frac{z_i e\psi_s + \phi}{kT}\right] \tag{31}$$

Here X_i is the mol fraction of the ion of interest, I, in the bulk solution, N_s is the maximum site density in the Stern layer, and ϕ is the noncoulombic energy of interaction of the ions with the surface.

Example 3.

Consider again the experimental results for hematite examined in Example 2. At pH = 6.5 and $I = 10^{-3}$ M, σ_o = 0.053 C m^{-2}; at pH = 4.0 and I = 0.1 M, σ_o = 0.28 C m^{-2}. The temperature is 25 °C. Determine the surface and diffuse layer potentials $(\psi_o \text{ and } \psi_d)$, the Stern and diffuse layer charges $(\sigma_s \text{ and } \sigma_d)$, and the counterion and similion concentrations at the inside boundary of the diffuse layer using (A) the empty Stern layer-Gouy-Chapman model and (B) the charged Stern layer-Gouy-Chapman model. We will assume a capacitance for the Stern layer at the hematite-water interface of 1.7 F m^{-2} after Hiemstra and Van Riemsdijk [21]. For the charged Stern layer we will assume that the noncoulombic interaction energy, ϕ, is zero and that the maximum site density in the Stern layer, N_S, is 2.5/nm^2 or 2.5 x 10^{18} m^{-2}.

(A) Empty Stern Layer Model

Case (i): pH = 6.5, I = 0.001 M, σ_o = +0.053 C m^{-2}.

- Stern layer charge — $\sigma_s = 0$ (as defined in model)
- Diffuse layer charge (Eq. (16)) Similar to Example 2. — $\sigma_d = -\sigma_o = -0.053\ C\ m^{-2}$
- Diffuse layer potential Eq. (24)) Similar to Example 2. — $\psi_d = +172\ mV$
- Surface potential (Eq. (29)) — $\psi_o = +203\ mV$
- Interfacial ion concentrations at $x = \delta$ (Eq. (18)) Similar to Example 2) — $[Na^+]_{x=0} = 1.2 \times 10^{-6}\ M$, $[NO_3^-]_{x=o} = 0.81\ M$

Case (ii): pH = 4.0, I = 0.1 M, σ_o = +0.28 C m^{-2}

- Stern layer charge — $\sigma_s = 0$ (as defined in model)
- Diffuse layer charge (Eq. (16)) (Similar to Example 2) — $\sigma_d = -\sigma_o = -0.28\ C\ m^{-2}$
- Diffuse layer potential (Eq. (24)) (Similar to Example 2) — $\psi_d = +140\ mV$
- Surface potential (Eq. (29)) — $\psi_o = +305\ mV$
- Interfacial ion concentrations at $x = \delta$ (Eq. (18)) (Similar to Example 2) — $[Na^+]_{x=0} = 4.3 \times 10^{-6}\ M$; $[NO_3^-]_{x=o} = 23.2\ M$!

As in Example 2, *experimental* values of surface charge density, pH, and ionic strength are available and we *calculate* some other characteristics of the interfacial region based on an electrostatic model, in this case using a model that includes an empty Stern layer and a Gouy-Chapman diffuse layer. Comparison of the results of the two examples indicates that the use of the empty Stern layer leads to calculations of surface potentials that are even higher than with the Gouy-Chapman approach and, at the same time, does not resolve the problem of the high counterion concentrations calculated at the inner boundary of the diffuse layer arising from the point charge assumption in the Poisson-Boltzmann equation. To examine this issue further, we proceed to the case of the charged Stern layer.

(B) Charged Stern Layer Model

We need the following equations to address this problem:

$$\sigma_o + \sigma_s + \sigma_d = 0 \tag{30}$$

$$\sigma_d = -\sqrt{8\varepsilon_o \varepsilon kTn}\sinh\left(\frac{ze\psi_d}{2kT}\right) \tag{24}$$

$$\sigma_s \cong z_i e X_i N_s \exp\left[-\frac{z_i e\psi_s + \phi}{kT}\right] \tag{31}$$

We have three unknowns $(\sigma_s, \sigma_d, and\ \psi_d)$ in these three independent equations. Note that z in Eq. (24) is the charge (without sign and = 1 in these two cases) of the ions in the symmetrical indifferent electrolyte while z_i in Eq. (31) is the charge (with sign and = -1 in these two cases) of the counterion in the charged Stern layer.

Case (i): pH = 6.5, I = 0.001 M, σ_o = +0.053 C m^{-2}. Results are as follows:

- Diffuse layer potential $\psi_d = +168\ mV$
- Stern layer charge $\sigma_s = 0.005\ C\ m^{-2}$
- Diffuse layer charge $\sigma_d = 0.049\ C\ m^{-2}$
- Surface potential $\psi_o = 199\ mV$
- Interfacial ion concentrations at x = δ $[Na^+]_{x=\delta} = 1.4 \times 10^{-6}\ M$
  $[NO_3^-]_{x=\delta} = 0.69\ M$

Compared to the Gouy-Chapman model (Example 2) and to the empty Stern layer-Gouy-Chapman model (part A of this example), the charged Stern layer in this case contains about 10 % of the total charge, reduces the diffuse layer potential by 4 mV, and slightly lowers the accumulation of counterions at the inside of the diffuse layer.

Case (ii): pH = 4.0, I = 0.1 M, σ_o = +0.28 C m^{-2}

Calculations using Eqs. (23), (29), and (30) yield the following results;

- Diffuse layer potential $\psi_d = +122\ mV$
- Stern layer charge $\sigma_s = -0.083 \frac{C}{m^2}$
- Diffuse layer charge $\sigma_d = -0.197 \frac{C}{m^2}$
- Surface potential $\psi_o = +287\ mV$
- Interfacial ion concentrations at x = δ $[Na^+]_{x=\delta} = 8.7 \times 10^{-4}\ M$
  $[NO_3^-]_{x=\delta} = 11.5\ M$

For these conditions (high ionic stength, low pH, high surface charge), the charged Stern layer-Gouy-Chapman model predicts that about 30 % of the surface charge is balanced in the charged Stern layer. The diffuse layer potential is reduced by about 18 mV compared to the diffuse layer model and the concentration of counterions calculated at the inside of the diffuse layer is reduced by over 50 % but is still very high (11.5 M).

The results of Examples 2 and 3 are summarized and compared in Table 2. Four conclusions are noted here. First, for the conditions examined, coupling a Stern layer with the Gouy-Chapman model increases the surface potential with the increase being particularly large at high ionic strength. Second, the differences between the cases with the empty and the charged Stern layers are not large. Third, the use of an uncharged Stern layer has no effect on

the high counterion concentrations calculated to exist at the inside boundary of the diffuse layer. Finally, the use of a charged Stern layer lowers but does not eliminate these high counterion concentrations in the diffuse layer

Table 2

Comparison of three double layer models for hematite in $NaNO_3$. Measured system characteristics,

Case 1: pH = 6.5; $I = 10^{-3}$ M; $\sigma_o = +0.053$ C m^{-2}

Case 2: pH = 4.0; $I = 10^{-1}$ M; $\sigma_o = +0.28$ C m^{-2}

Calculated system characteristics	Gouy- Chapman	Empty Stern-Gouy-Chapman	Charged Stern-Gouy-Chapman
Case 1			
ψ_o (mV)	+172	+203	+199
ψ_d (mV)	+172	+172	+168
σ_s (C m^{-2})	0	0	-0.005
σ_d (C m^{-2})	-0.053	-0.053	-0.049
$[Na^+]_{x=0,\delta,\gamma}$ (M)	1.2 x 10^{-6}	1.2 x 10^{-6}	1.4 x 10^{-6}
$[NO^-]_{x=0,\delta,\gamma}$ (M)	0.81	0.81	0.69
Case 2			
ψ_o (mV)	+140	+305	+287
ψ_d (mV)	+140	+140	+122
σ_s (C m^{-2})	0	0	-0.083
σ_d (C m^{-2})	-0.28	0.053	-0.197
$[Na^+]_{x=0,\delta,\gamma}$ (M)	4.3 x 10^{-4}	4.3 x 10^{-4}	8.7 x 10^{-4}
$[NO^-]_{x=0,\delta,\gamma}$ (M)	23.2	23.2	11.5

We conclude by noting a fourth model of the electrical characteristics of solid-water interfaces. If we consider that some ions in a Stern layer are partially dehydrated while others are not, we can subdivide the Stern layer into two regions, an inner layer of physically adsorbed and partially dehydrated ions and an outer layer where fully hydrated counterions reside. The result has been termed the *triple layer model.* It is, of course, a quadruple layer model but is usually not named as such. The plane where the physically adsorbed ions are located is termed the IHP and the plane of closest approach of the diffusing ions (the inside of the diffuse layer) is termed the OHP as before. This perspective was proposed by Grahame [22] and has been called the Stern-Grahame model. It has been applied to chemical speciation at interfaces by Yates et al. [23] and Davis and coworkers [24, 25], among others. Several new parameters are introduced in order to formulate the model and to use it including, for example, capacitances for both of Stern layers. Useful more recent reviews are provided by Davis and Kent [26] and Hayes and Katz [27].

5. DLVO THEORY

When two similar particles and their associated electrical double layers approach each other, their diffuse ion atmospheres overlap and a coulombic repulsive force is produced. The classic DLVO theory [13, 28] considers this electrostatic repulsion together with the van der

Waals attractive interaction in assessing the interaction forces [28] or the interaction energies [13] between the particles. Here an approach to determine the coulombic interaction force for a simple system will be outlined and used to determine the coulombic interaction energy. This will then be combined with the van der Waals interaction energy (Section 3) to calculate net interaction energies. Finally, these results are applied to the problem of aquasol stability. For additional information the reader is referred to the classical presentation by Overbeek [29] the recent summary by Gregory [30] and the extensive description by Hunter [8].

5.1. Charge and potential between two flat plates

Suppose two similarly charged flat surfaces approach each other. Let us consider the distribution of charge and electric potential between the two overlapping and interacting flat double layers as the distance between the particles changes. We assume Gouy-Chapman diffuse layers at each surface with $\psi_o = \psi_d$. We also assume a *constant surface potential interaction* between the surfaces, i.e., the surface potentials of the two surfaces are assumed to remain constant as the two diffuse layers overlap. Because the two surfaces are similar, the potential at any separation reaches a minimum value, ψ_m , midway between the surfaces. The separation distance is given by h and the midway plane between the surfaces is located at d. This is illustrated in Fig. 3.

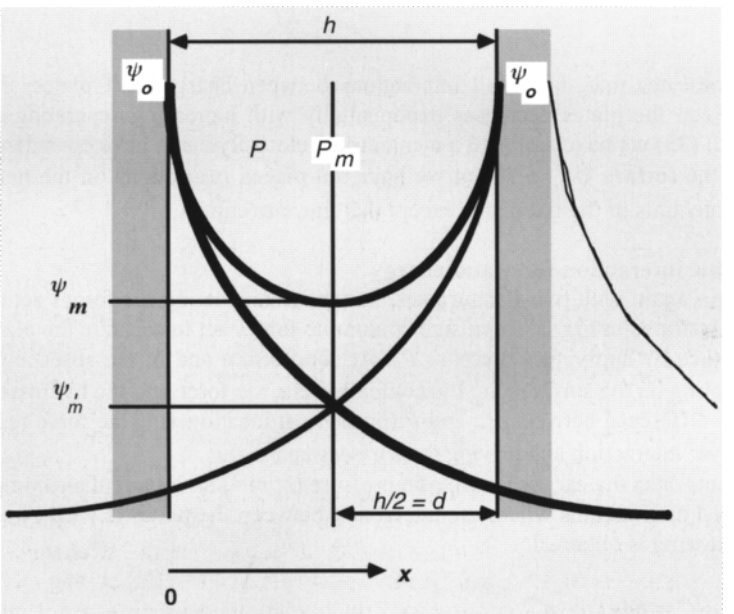


Fig. 3. Interacting Gouy-Chapman diffuse layers of two similar flat plates

The electrical potential profile resulting from the interaction will satisfy the Poisson-Boltzmann equation. An analytical solution describing potential as a function of separating distance for any value of the surface potentials has not been obtained. Numerical results for this constant surface potential case are given by Verwey and Overbeek [13] and Overbeek [29]. The following has been obtained for charge as a function of separating distance,:

$$\sigma_o = \sqrt{2\varepsilon\varepsilon_o kTn}\left[2\cosh(y_o) - 2\cosh(u)\right]^{1/2} \tag{32}$$

5.1.1 Small interaction

Here we consider the case where the coulombic interaction between the two flat plates in Fig. 3 is small because the distance separating the plates (h) is large. Under these conditions, the electrical potential due to the double layer of one plate will be negligible near the surface of the second plate. In these circumstances it can be assumed that the electric potential at the midway point (ψ_m) is simply the sum of the two separate and equal electric potentials of the unperturbed diffuse layers at this point.

$$\psi_m = \delta\left(\frac{kT}{ze}\right)\gamma e^{-\kappa d} \tag{33}$$

in which

$$\gamma = \frac{e^{y_o/2} - 1}{e^{y_o/2} + 1} \qquad \text{and} \qquad y_o = \frac{ze\psi_o}{kT}$$

This result indicates that, for small interactions between charged flat plates, the potential midway between the plates decreases exponentially with increasing separating distance. In arriving at Eq. (33) we have assumed a symmetrical electrolyte and have considered locations distant from the surface $(\kappa x >> 1)$ but we have not placed restrictions on the magnitudes of the surface potentials of the flat plates except that they are equal.

5.2. Coulombic interaction force and energy

We begin again with two flat surfaces, this time considering the forces acting on them. The case is illustrated in Fig. 3. Repulsive coulombic forces act to separate the plates and they are kept in place by applying a pressure P_o. At equilibrium and in the absence of flow, the only forces acting on the surfaces are the coulombic electric force and the hydrostatic pressure gradient. The difference between P_m and P (pressure at location x) is the force resulting from the double layer interaction and driving the surfaces apart.

We denote π as the excess or osmotic pressure (N m^{-2}) at the midplane compared to the bulk solution. For situations where the interaction between the plates is weak, i.e., when d is large, the following is obtained:

$$\pi = 64kTn\gamma^2 e^{-2\kappa d} = 64kTn\gamma^2 e^{-\kappa h} \tag{34}$$

This repulsive pressure is also often termed a *disjoining pressure.*

With the coulombic interaction *force* in hand, we turn now to the coulombic interaction *energy*. The repulsive potential energy per unit area between the two plates is the integral of the excess pressure between them from an infinite separation $(h = \infty)$ to the separating distance h (= $2d$).

$$V_{R(a)} = -\int_{\infty}^{h} \pi dh = -2\int_{\infty}^{d} \pi d(d)$$

$$= -64kTn\gamma^2 \int_{\infty}^{d} e^{-2\kappa d} d(2d)$$

Here $V_{R(a)}$ denotes the electrostatic repulsive energy per unit area (J m^{-2}). The well known result is

$$V_{R(a)} = \frac{64kTn}{\kappa}\gamma^2 e^{-2\kappa d} = \frac{64kTn}{\kappa}\gamma^2 e^{-\kappa h} \tag{35}$$

As stated previously, when the separating distance is large so that the interaction between the flat plates is weak, the overlap of the two diffuse layers has negligible effects on the region near the surfaces of the plates. Eqs. (34) and (35), while developed for a constant potential interaction, are also valid for the constant charge interactions discussed subsequently, as long as these interactions are weak.

A selection of expressions for the electrostatic interaction between interacting double layers is given for three different geometries in Table 3. Eq. (35) is listed as Eq. (1) in this table.

5.3. Types of electrostatic double layer interactions

Some particles have a *constant charge surface*, or almost so, so that their primary charge is independent of solution composition. Smectite clays are an example of this particle type. The surface potential of these particles can be expected to decrease as the ionic strength of the solution is increased (Eqs. (24), (25)). Other particles have a surface potential that is determined by the activity of potential determining ions in solution; these are said to have *constant potential surfaces*. The silver halide sols (e.g., AgI) are examples of this particle type. The surface charge of these particles will increase as the ionic strength of the solution is increased (Eqs. (24), (25)). These two descriptions are traditional in the colloid chemical literature. Metal oxides such as hematite are neither. They tend to behave as constant potential surfaces at low ionic strength and at pHs near their pzc while acting more like constant charge surfaces at high ionic strength and at pHs far removed from their pzc.

It is also common to speak of *constant charge* and *constant potential interactions*. These expressions are not equivalent to the terms constant potential and constant charge surfaces. At least two particles are considered in a particle-particle interaction; only one particle is needed to discuss surface properties. As we will see subsequently, for example, two constant potential surfaces can have a constant charge interaction with each other.

Some analytical expressions for repulsive double layer electrostatic interaction energies are given for constant potential and constant charge interactions for three different geometries in Table 3.

In fact, several types of double layer interactions are possible during an encounter between two charged particles. Lyklema [34] has illustrated important possibilities during a Brownian encounter. For a particle with a diameter of 1 μm in a 0.1 M solution of 1:1 electrolyte, the time for an non-retarded Brownian encounter is about 10^{-5} s. Considering hydrodynamic retardation, Lyklema raises this estimate to 10^{-4} s. This time is considerably longer than the equilibration or relaxation time of ions in the interacting diffuse layers,

estimated to be about 10^{-8} s. Rates of adjustment of surface layers (e.g., Stern layers) to the encounter can be considerably longer. For an AgI sol, Lyklema estimates a relaxation time for a Stern layer of 10^{-3} s; this may be the time to transfer a charge determining ion across the interface. It is considerably longer than the Brownian encounter. As a result, two AgI particles with constant potential surfaces may undergo a constant charge interaction during a Brownian encounter because the adjustment time for their surface and/or Stern charges is longer than the Brownian encounter time.

If the surface potential of a solid is determined by the adsorption equilibrium of potential determining ions, then a fully relaxed interaction between two particles (during which equilibrium is maintained at the two surfaces) will result in the surface potential remaining constant while a reduction in the surface charge occurs as potential determining ions are desorbed. If the interacting double layers do not relax or equilibrate at all, then all potential determining ions remain at the surface and the surface potential increases during the interaction. The results reported by Lyklema [34] suggest that this latter interaction mode can occur with AgI sols and probably some others. In general, the type of electric double layer interaction will depend on the kinetics of ion and particle transport. Constant charge interactions can produce the largest repulsive interaction energies; constant potential interactions yield the lowest calculated repulsive interaction energies [35]. Intermediate interaction modes produce intermediate repulsive interaction energies.

5.4. Net DLVO interaction energies

The Derjaguin-Landau-Verwey-Overbeek (DLVO) theory considers van der Waals interactions which are attractive for aquasols together with electrostatic double layer interactions which may be attractive or repulsive. The total interaction energy (V_T) is written as the sum of these two interactions. No other forces are considered.

$$V_T = V_A + V_R \tag{36}$$

Here V_A and V_R refer to the van der Waals and electrical double layer interaction energies, respectively. Some analytical equations for V_A and V_R are given in Tables 2 and 3 together with indications of their limitations.

Let us consider an application of this approach to the deposition of colloidal particles in a bed of porous media. Geometrically, the system is approximated as a sphere-plate interaction since the size of the aquasol particle in suspension is much smaller than the size of the "collector" grain in the porous media. The following example is taken from Hahn et al. [36].

Representative interaction energy curves for the sphere-plate case are presented in Fig. 4. The total interaction energy depends on the individual interaction energies that reflect both the system of interest and the assumptions made in describing it. Nevertheless, the shapes of the curves in Fig. 4 are typical of a broad variety of applications.

Table 3
Double layer electrostatic interaction energies

Geometry/ Expression		Conditions*	Reference
Plate-plate/ $V_{R(a)} = \frac{64\ k\ T\ n}{\kappa} \gamma_1\ \gamma_2\ e^{-\kappa h}$	(1)	LSA, large κh	[13]
Plate-plate/ $V_{R(a)} = \frac{\varepsilon\ \varepsilon_o\ \kappa}{2}\left\{(\psi_1^2 + \psi_2^2)(1 - coth\kappa h) + 2\ \psi_1\psi_2\ cos\ \ ech\kappa h\right\}$	(2)	CPI, LPB, small ψ_i	[31]
Plate-plate/ $V_{R(a)} = \frac{\varepsilon\ \varepsilon_o\ \kappa}{2}\left\{(\psi_1^2 + \psi_2^2)(coth\kappa h - 1) + 2\ \psi_1\psi_2\ cos\ \ ech\kappa h\right\}$	(3)	CCI, LPB, small ψ_i	[32]
Sphere-plate/ $V_R = 64\ \pi\ \varepsilon\ \varepsilon_o\ a\left(\frac{k\ T}{z\ e}\right)^2 \gamma_1\ \gamma_2\ e^{-\kappa h}$	(4)	LSA, DA, large κh	[13]]
Sphere-plate/ $V_R = \pi\ \varepsilon\ \varepsilon_o\ a\left\{2\psi_1\ \psi_2\ ln\left(\frac{1+e^{-\kappa h}}{1-e^{-\kappa h}}\right) + (\psi_1^2 + \psi_2^2) ln(1 - e^{-2\kappa h})\right\}$	(5)	CPI, LPB, $\kappa a > 5$, $\psi_i < 60$ mV	[31]
Sphere-plate/ $V_R = \pi\ \varepsilon\ \varepsilon_o\ a\left\{2\psi_1\ \psi_2\ ln\left(\frac{1+e^{-\kappa h}}{1-e^{-\kappa h}}\right) - (\psi_1^2 + \psi_2^2) ln(1 - e^{-2\kappa h})\right\}$	(6)	CCI, LPB, $\kappa a > 5$, $\psi_i < 60$ mV	[33]
Sphere-sphere/ $V_R = 64\ \pi\ \varepsilon\ \varepsilon_o\ \left(\frac{a_1\ a_2}{a_1 + a_2}\right)\left(\frac{k\ T}{z\ e}\right)^2 \gamma_1\ \gamma_2\ e^{-\kappa h}$	(7)	LSA,DA, large κh	[13]
Sphere-sphere/ $V_R = \pi\ \varepsilon\ \varepsilon_o\left(\frac{a_1\ a_2}{a_1 + a_2}\right)(\psi_1^2 + \psi_2^2)\left\{\frac{2\psi_1\ \psi_2}{\psi_1^2 + \psi_2^2} ln\left(\frac{1+e^{-\kappa h}}{1-e^{-\kappa h}}\right) + ln(1 - e^{-2\kappa h})\right\}$	(8)	CPI, DA, LPB, small ψ_i, $\kappa a_i < 1$, $h < a_i$	[31]
Sphere-sphere/ $V_R = \pi\ \varepsilon\ \varepsilon_o\left(\frac{a_1\ a_2}{a_1 + a_2}\right)(\psi_1^2 + \psi_2^2)\left\{\frac{2\psi_1\ \psi_2}{\psi_1^2 + \psi_2^2} ln\left(\frac{1+e^{-\kappa h}}{1-e^{-\kappa h}}\right) - ln(1 - e^{-2\kappa h})\right\}$	(9)	CCI, DA, LPB, small ψ_i, $\kappa a_i < 1$, $h < a_i$	[33]

*All systems assume a symmetrical electrolyte and point charges; LSA = linear superposition approximation; LPB = linearized Poisson-Boltzmann equation; CPI = constant potential interaction; CCI = constant charge interaction; DA = Derjaguin approximation.

Eq. 4 in Table 2 was used for the retarded van der Waals interaction; Eq. 5 in Table 3 was used for the constant potential electrical double layer interaction. With these assumptions, the net DLVO interaction energy is as follows:

$$V_T = -\left(\frac{Aa}{6h}\right)\left(1+\frac{14h}{\lambda}\right)^{-1} + \pi\varepsilon\varepsilon_o a\left\{2\psi_1\psi_2 ln\left(\frac{1+e^{-\kappa h}}{1-e^{-\kappa h}}\right)+\left(\psi_1^2+\psi_2^2\right)ln\left(1-e^{-2\kappa h}\right)\right\} \quad (37)$$

Parameters used in the calculations are given in the legend to Fig. 4. These have been selected to represent situations of interest in aquasols. The particle's radius, a, is set at 50 nm, in the lower part of the colloidal size range. The ionic strength (I) is set at 0.1, typical of estuarine waters and commonly considered to promote the coagulation of many colloids. A value of 10^{-20} J is chosen for the Hamaker constant (A); this is in the range of many aquasols and is also appropriate for the polystyrene latex-water-glass system often used in laboratory studies of particle deposition. The characteristic wavelength of the retarded van der Waals interaction (λ) is assumed to be 100 nm. The surfaces of the suspended particles and the stationary flat plate are both given negative surface potentials of -30 mV, representative of surface properties of some aquasols at this ionic strength. The shapes of the curves presented in Fig. 4 are sensitive to these parameters.

Considering the interaction energy curves presented in Fig. 4, the attractive force dominates at very small separation distances and an infinitely deep well in potential energy is observed. This is termed the *primary well.* At still closer separations, a Born repulsion occurs. It is due to the interpenetration of ionic atmospheres, provides a strong repulsion at the surface and, with the DLVO forces, forms a primary energy minimum rather than a primary well. At larger separating distances, electrostatic repulsion between overlapping diffuse layers dominates van der Waals attraction and a repulsive energy barrier develops. The maximum energy of this barrier is denoted as $V_{T,max}$ and, for the conditions in Fig. 4, has a value of about 7 kT. Since the average one-dimensional kinetic energy of a Brownian particle is only 0.5 kT, this barrier is a significant impediment to attachment of the aquasol to the plate.

At still larger separating distances, about 4 nm in Fig. 4, a second attractive region occurs. This is termed the *secondary minimum* and it is much shallower than the primary well or minimum. It is denoted as V_{2min}. For the conditions represented in Fig. 4, the attractive secondary minimum has a depth of 2.8 kT, somewhat larger than the one-dimensional thermal energy of 0.5 kT. Association between the particle and the plate is possible for separating distances in this attractive region. For two surfaces to be attached, they must overcome the energy barrier and be held in the deep primary minimum or they can remain associated at larger separating distances in the secondary attractive region. Due in part to the differences in magnitude of these two attractive minima, contact in the primary minimum is thought to be irreversible while association in the secondary minimum can be reversible [37].

Aggregation and deposition rates, assuming attachment in the primary well, will depend on the height of the energy barrier, $V_{T,max}$. When such a net repulsive interaction energy exists, the interaction is termed unfavorable and the kinetics of aggregation and deposition are termed slow. Conversely, when there is no energy barrier to attachment in the primary minimum, the interaction is termed favorable and the kinetics are termed fast.

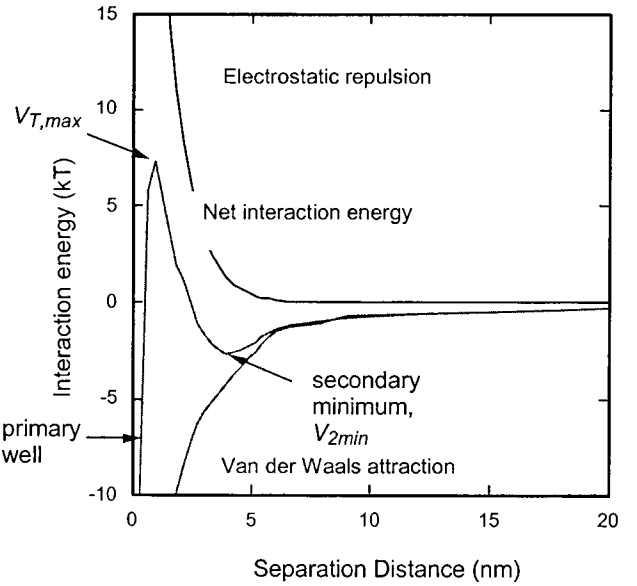


Fig. 4. Electrostatic, van der Waals, and net interaction energy curves for the sphere-plate case, constant potential double layer interaction, retarded van der Waals interaction: a = 50 nm, I = 0.1, A = 10^{-20} J, λ = 100 nm, $\psi_1 = \psi_2$ = -30 mV, T = 298 K. From [36]

Electrostatic repulsive forces and energies are very sensitive to the ionic strength of the solution. Ions in solution screen the coulombic interactions between the charged surfaces. This reduces the diffuse layer thickness (κ^{-1}, Eq. (38)), thereby shortening the range of the repulsive interaction so that attractive van der Waals forces can be more effective. At low ionic strength (for example, the ionic strengths of fresh waters), a net repulsion can predominate over intermediate separating distances between aquasols. At high ionic strength (for example, estuarine and marine waters), this electrostatic energy barrier can disappear.

Effects of ionic strength on the total or net interaction energy between a sphere and a plate are illustrated in Fig. 5 from [36]. In these results, ionic strength is assumed to affect only the diffuse layer thickness; the surface potentials of the particle and the plate remain unchanged. Effects of ionic strength on the electrostatic terms in the Hamaker constant and the van der Waals interaction have been neglected. This is a good assumption for inorganic aquasols such as clays and metal oxides. Three ionic strengths are considered; other conditions are similar to those used in Fig. 4. Increasing the ionic strength from 10^{-3} to 10^{-1} lowers the height of the maximum in the net interaction energy curve ($V_{T,max}$) from 46 to 9 kT. At an even higher ionic strength this energy barrier would disappear, leading to fast deposition in the primary well. Negligible secondary minima (V_{2min} < 0.2 kT) are formed at ionic strengths of 10^{-3} and 10^{-2} M, while a significant secondary minimum is formed at I = 10^{-1}. This secondary minimum would disappear as the ionic strength was increased further.

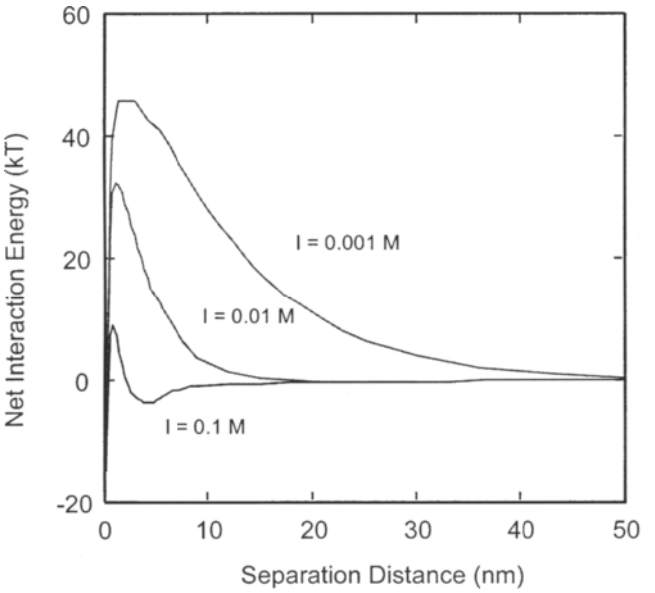


Fig. 5. Net interaction energy curves as a function of ionic strength for the sphere-plate case, constant potential double layer interaction, retarded van der Waals interaction: a = 50 nm, $A = 10^{-20}$ J, λ = 100 nm, $\psi_1 = \psi_2$ = -30 mV, T = 298 K. From [36]

In addition to ionic strength, the electrostatic double layer interactions depend on the surface potentials of the particles (see Eq. (37) and also Table 3).

An important result in Eq. (36) is that the attractive and repulsive interaction energies and, consequently, the net interaction energy, all scale directly with the size of the particle, a. This result is true for any other force that may be considered between the two interacting solids. In addition to the sphere-plate interaction typical of particle deposition and described by Eq. (37), it also applies to sphere-sphere interactions in aggregation phenomena. Large colloidal particles have deeper primary minima, higher energy barriers, and deeper secondary attractive minima than smaller ones, and the differences are directly proportional to the sizes of the particles.

Results presented in Fig. 6 illustrate the effect of particle size on net interaction energy curves for the sphere-plate case. The effects of particle size on the height of the energy barrier $V_{T,max}$ are observed best in Fig. 6b while particle size effects on the depth of the secondary minimum V_{2min} are seen best in Figure 6a. The height of the energy barrier scales in direct proportion to the particle size, increasing from 43 to 114 to 290 kT as the particle radius increases from 23 to 60.5 to 154 nm. DLVO theory is seen to predict a very strong dependence of particle stability on particle size when stability is viewed as the difficulty in achieving contact in the primary well. Large colloids are predicted to be much more stable than small ones. The depth of the secondary well also scales in direct proportion to particle size. The depth of this minimum increases from 0.11 to 0.28 to 0.72 kT as particle radius is increased from 23 to 60.5 to 154 nm. DLVO theory predicts a dependence of reversible

attachment in the secondary well on particle size with this attachment increasing significantly with increasing particle size. It is useful to note that the separating distances at which the energy barriers and the secondary minima are located (0.9 and 19 nm, respectively) are not affected by particle size.

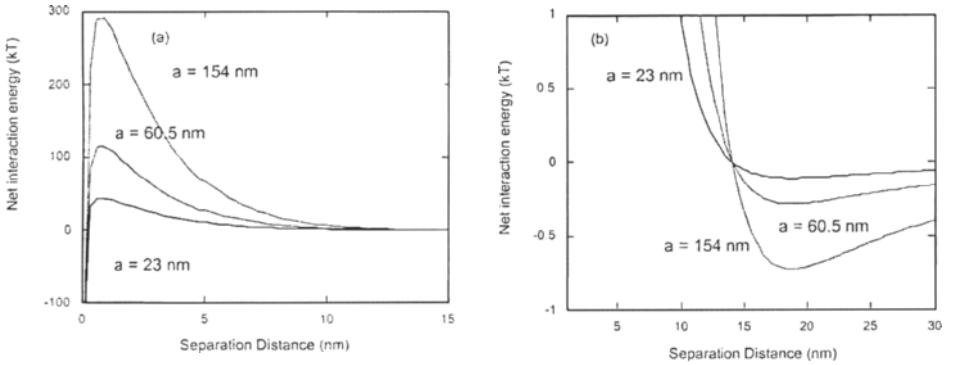


Fig. 6(a) Net interaction energy curves as a function of particle size for the sphere-plate case, constant potential double layer interaction, retarded van der Waals interaction: $I = 0.018$, $A = 10^{-20}$ J, $\lambda = 100$ nm, $\psi_1 = -40$ mV, $\psi_2 = -55$ mV, T = 298 K. (b) Enlargement of (a) to show the secondary minima. From [36]

Let us conclude this discussion of the effects of different parameters on DLVO interaction energies with a consideration of the effects of the Hamaker constant of a sphere-water-plate system on the net interaction energy of this system. Some calculated results are presented in Fig. 7 [36]. Three values of the Hamaker constant are considered: (1) $A = 3.5 \times 10^{-21}$ J, perhaps representative of some biological aquasols, (2) $A = 10^{-20}$ J, representative of the polystyrene-water-silica system and many natural systems, and (3) $A = 3 \times 10^{-20}$ J, illustrative of oxides in water. The electrostatic double layer interaction is independent of A while the van der Waals interaction is directly proportional to it (Eq. (36)). We see that the net interaction energy curve becomes more attractive as A is increased, with the energy barrier becoming smaller and the secondary minimum deeper as the Hamaker constant becomes larger. As the Hamaker constant is increased from 3.5×10^{-21} to 3×10^{-20}, the height of the energy barrier is decreased from 140 to 67 kT and the depth of the secondary minimum is increased from 0.1 to 1.2 kT. The location of the energy barrier is shifted outward slightly and the secondary minimum moves inward as the Hamaker constant is increased. DLVO theory predicts that inorganic aquasols such as oxides, aluminosilicates, silver halides and metals can be less stable than organic colloids with smaller Hamaker constants. They are more likely to associate in the secondary minimum and also in the primary well.

5.5. Stability factors and sticking probabilities

Interactions between solid particles have been conveniently segregated into two sequential steps: transport and attachment. The results of these interactions are given such terms as coagulation, aggregation, deposition, and filtration. Particle transport depends upon hydrodynamics and external forces such as gravity; it is primarily a physical process. The attachment of two particles can be dominated by the surface properties of the solid particles and by solution chemistry; it is primarily a chemical process and is discussed herein. This

distinction between transport and attachment, or physics and chemistry, is not perfectly sharp. Some colloidal forces (van der Waals forces) must be present to overcome hydrodynamic retardation and allow transport of one particle to the surface of another. Similarly, surface and solution chemistry does not always influence attachment; physical forces are sufficient to describe "favorable' interactions. In this chapter we introduce stability factors or sticking probabilities that are used to describe chemical aspects of these processes.

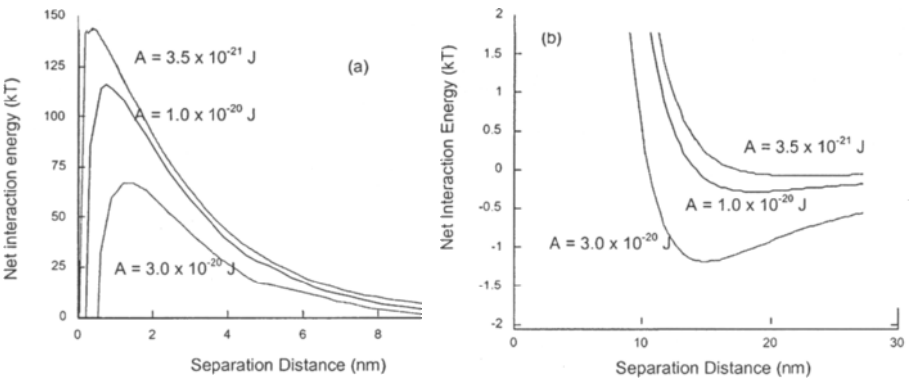


Fig. 7. (a) Net interaction energy curves as a function of Hamaker constant for the sphere-plate case, constant potential double layer interaction, retarded van der Waals interaction: a = 60.5 nm, I = 0.018, λ = 100 nm, ψ_1 = -40 mV, ψ_2 = -55 mV, T = 298 K. (b) Enlargement of (a) to show the secondary minima. From [36]

5.5.1. Aggregation

Following after Smoluchowski [38], the collision rate of a small colloidal particle with other particles of the same size is

$$N_{c,fast} = 16\pi Dan_{\infty} \tag{38}$$

in which n_{∞} is the bulk concentration of Brownian particles. Based on the work of Fuchs [39], the Brownian collision rate of a colloidal particles with other particles of the same size ($N_{c,slow}$) in the presence of a repulsive energy barrier (V_T) is given as follows:

$$N_{c,slow} = \frac{16\pi Dan_{\infty}}{2a\int_{2a}^{\infty} exp\left(\frac{V_T(r)}{kT}\right)\frac{dr}{r^2}} \tag{39}$$

The subscript *slow* indicats that the collision rate is reduced by the repulsive interaction energy. The numerator in Eq. (39) is the collision rate in the absence of a repulsive interaction between the particles and the denominator describes the effect of the repulsive interaction in reducing this collision rate.

Following after Fuchs [39], we define a stability factor for aggregation, W_a, as follows:

$$W_a = \frac{contact\ rate\ without\ repulsive\ interaction}{contact\ rate\ with\ repulsive\ interaction} = \frac{N_{c,fast}}{N_{c,slow}} \tag{40}$$

It follows that

$$W_a = 2a \int_{2a}^{\infty} exp\left(\frac{V_T(r)}{kT}\right)\frac{dr}{r^2} \tag{41}$$

To a first approximation, the Fuchs stability factor for aggregation can be written as

$$W_a \cong \frac{1}{2\kappa a} exp\left(\frac{V_{T,max}}{kT}\right) \tag{42a}$$

Finally, for a heterodisperse suspension, the result is

$$W_a \cong \frac{1}{\kappa\left(a_i + a_j\right)} \exp\left(\frac{V_{T,\max}}{kT}\right) \tag{42b}$$

An alternative way to describe the effects of repulsive interaction energies on the kinetics of aggregation is with the use of an attachment probability, α_a.

$$\alpha_a = \frac{rate\ at\ which\ particles\ attach}{rate\ at\ which\ particles\ collide} \tag{43}$$

In this approach the rate at which particles collide is taken as the collision rate in the absence of an energy barrier (fast aggregation) and the rate at which they attach is the rate at which they collide in the presence of an energy barrier (slow aggregation). Consequently,

$$\alpha_a = 1 / W_a$$

The Fuchs' stability factor as described by Eqs. (42a) and b) is based on interparticle transport by Brownian diffusion in the presence of a repulsive interaction energy. It does not, however, include the effects of hydrodynamic retardation on the thermal diffusion process. For perikinetic encounters (Brownian diffusion) between two particles with different sizes experiencing a repulsive electrostatic interaction, Spielman [40] has incorporated hydrodynamic retardation. Analytical solutions are not available; numerical approaches are required. Attractive van der Waals forces are needed overcome to the drag forces that approach infinity as separating distances approach zero, hence the term hydrodynamic retardation. Both chemical and hydrodynamic information are needed; a clean separation of mass transport (physical) from attachment (chemical) can not made.

5.5.2. Deposition

In this section we describe the mass transport of colloidal particles in porous media in the presence of a repulsive interaction energy. Our approach to particle deposition in porous media is analogous to the approach described previously for particle aggregation. The removal of particles from suspension in a porous medium is considered to involve two steps: the transport of a suspended particle to the immediate vicinity of the solid-water interface presented by the filter media, and then the attachment of the particles to that surface [41]. There are useful analogies in both the transport and the attachment aspects (also called, respectively, the hydrodynamic and colloid chemical aspects) of aggregation and deposition.

In aggregation processes, the transport models of Smoluchowski [38] and Findheisen [42] are often used. They predict that interparticle transport of particles larger than about 1 μm in size is accomplished by fluid motion (velocity gradients) and by differential settling (gravity) while, for smaller or colloidal particles, Brownian diffusion is more effective. As we will see in the presentation that follows, particle transport in deposition by bulk fluid flow, analogous to transport in aggregation by velocity gradients, and particle transport in deposition by gravity, analogous to transport in aggregation by differential settling, dominate colloid transport to the surfaces of filter media for aquasol particles larger than about 1 μm in size. Smaller or colloidal particles are transported in porous media by convective diffusion, analogous to the effects of Brownian diffusion in perikinetic flocculation.

The deposition of particles from a suspension to the surface of a porous medium can be written as follows:

$$\frac{dn}{dL} = -k_d n \tag{44}$$

Here n is the number concentration of particles (number/m^3) in suspension at location L that is the distance along the length of the porous medium (m) and k_d is a pseudo-first order rate constant (m^{-1}) that depends upon hydrodynamic and colloid chemical properties of the system. Deposition or filtration is actually a second order process involving collisions between suspended particles (n) and stationary collectors in the porous medium; the number of these stationary collectors is included in the distance L. Here we focus on k_d and, specifically, on the hydrodynamic contributions to it.

For the deposition of an aquasol on a stationary surface such as a porous medium we write

$$k_d \propto \alpha_d \eta \tag{45}$$

in which η is a dimensionless mass transport coefficient that is primarily hydrodynamic and frequently determined theoretically. As with α_a in aggregation, α_d in deposition is a dimensionless sticking coefficient or attachment probability that is primarily colloid chemical and often measured experimentally.

It is not possible to obtain a general solution for the transport of aquasols towards stationary surfaces because of the large number of variables that are involved. For Brownian particles in the absence of an energy barrier and neglecting hydrodynamic retardation, the following has proved useful [43]:

$$\eta_{bd} = 4.0 A_s^{1/3} Pe^{-2/3} \tag{46}$$

in which η_{bd} is the single collector efficiency for transport by convective diffusion. The term A_s incorporates the effects of neighboring grains or collectors on transport to a single grain. It is given by

$$A_s = \frac{2(1-p^5)}{2-3p-3p^5-2p^6} \tag{47}$$

where $p = (1-f)^{1/3}$ and f is the porosity of the porous medium.

5.6. Critical coagulation concentrations

The charge of the counterion has long been observed to have important effects on the concentration of electrolyte required to destabilize an aquasol completely and to bring about rapid aggregation and deposition in which the kinetics of these processes are determined only by mass transport. This concentration is termed the *critical coagulation concentration* or the c.c.c. Based on the *experimental* observations of Schulze [44, 45] and Hardy [46], the *Schulze-Hardy rule* states that the c.c.c. is determined primarily by the valence of the counterion and is in the order of 25 to 150 x 10^{-3} M for monovalent counterions, 0.5 to 2 x 10^{-3} M for divalent counterions, and 0.01 to 0.1 x 10^{-3} M for trivalent counterions. For strongly hydrolyzing metal ions such as Al(III) and Fe(III), it is important to note that the Schulze-Hardy rule applies directly only when the free aquometal ion dominates the metal ion species in solution. For Al(III), for example, this requires that the pH be less than about 4 to insure that the Al^{+3} species predominates in solution. Hydrolyzed Al(III) and Fe(III) species may have charges that are larger or smaller than Al^{+3} and Fe^{+3} but their effectiveness as coagulants is more closely related to their ability to form surface complexes and to adsorb on particle surfaces than it is to their charge.

Perhaps the most widely known test of the DLVO theory is in its use to predict the effect of the counterion charge on the critical coagulation concentration and the successful comparison of these results with the empirical Schulze-Hardy rule. The DLVO approach was applied to prediction of the c.c.c. by Verwey and Overbeek [13]. They considered the case of two infinite flat plates interacting by van der Waals attraction and coulombic repulsion. The net interaction energy has a slope of zero at its maximum value, i.e., $dV_T/dh = 0$ at $V_T = V_{T,max}$ (see Fig. 4). At the critical coagulation concentration, the energy barrier will just vanish, so that, at some separating distance, $V_T = V_{T,max} = 0$. The authors used the van der Waals interaction energy written in its unretarded form (Eq. 1 in Table 1) and, for the double layer interaction, Eq. 1 in Table 3. They obtained the following result:

$$c.c.c. \propto \frac{1}{z^6} \tag{48}$$

DLVO theory predicts that the c.c.c. is proportional to the reciprocal of the sixth power of the valence of the counterion, so that it will vary by 1:1/64:1/729 as the valence of the counterion increases in the order 1:2:3. This theoretical result is in excellent agreement with the empirical Schulze-Hardy rule stated above.

The Schulze-Hardy rule and the use of DLVO theory to support it have proved to be useful tools in a diverse array of laboratory studies. However, in aquatic environments, both natural and technological, they have been found to have limited application. This is because the Gouy-Chapman model for the diffuse double layer considers only electrostatic interactions of point charges with a surface. Many solutes in aquatic environments react "specifically" with aquasol surfaces. "Specific" is used here to mean interactions with surfaces in addition to coulombic ones and includes a range on interactions from hydrogen bonding to the formation of chemical bonds at the aquasol surface. An important consequence of these specific interactions is that they lead to adsorption of the coagulating species at the solid-water interface. Consequently, the c.c.c. of a specifically adsorbing solute will depend on the surface area concentration of the solid phase, a phenomenon termed the "stoichiometry of coagulation" [47].

6. POLYMERS AND POLYELECTROLYTES AT INTERFACES

Macromolecular adsorption has been the subject of extensive theoretical investigation. Simha [48] presented one of the first statistical models of the adsorption of polymers that emphasized their macromolecular nature. Currently several advanced models of polymer and polyelectrolyte adsorption are based on the work of Scheutjens and Fleer [49, 50]. This approach has been termed SF theory. A theoretical extension to weak polyelectrolytes has been provided by Böhmer et al. [51]; some experimental validation is provided by Blaakmeer et al. [52]. Among other useful improvements are the papers by van der Schee and Lyklema [53], Lyklema [54], Evers et al. [55], Cohen Stuart et al. [56], Scheutjens et al. [57], Fleer et al. [58], Hoogeveen et al. [59], and Vermeer et al. [60]. For applications to aquatic systems see Tiller and O'Melia [61] and Au et al. [18, 62]. In the following, brief descriptions are provided of Scheutjens-Fleer (SF) theory for polyelectrolyte adsorption, the Stumm-Schindler (SS) surface complexation model used primarily for monomer adsorption and surface speciation, and the hybrid SS/SF model for NOM adsorption developed by Au et al. [18, 62].

6.1. SF theory

SF theory is a statistical thermodynamic lattice model. It takes into account conformational entropy effects, enthalpic interactions, and electrostatic interactions among surface sites, polyelectrolyte functional groups, and electrolytes in solution including protons. In this theory, the solid/liquid interface is considered as a three-dimensional lattice having M layers (Fig. 8). The solid surface is an infinite flat plate, homogeneous in physical and chemical properties, and described by either a constant surface charge or surface potential. The macromolecules are linear, flexible, monodisperse, and, in most applications, homopolymers. Each macromolecule is an assembly of segments or monomer units. Each segment occupies one lattice site. Other macromolecular properties include its bulk concentration, valence, dielectric permittivity and, for weak polyelectrolytes, an acidity or basicity constant. Other species include solvent (water), H^+, OH^-, cations (e.g., Na^+) and anions (e.g., Cl^-); each of these species occupies one lattice site. Bulk concentrations of these species are specified as inputs to the model (e.g., pH and ionic strength). Divalent cations such as Ca^{2+} are not included in the present models. Lattice properties include size and type, with lattice type corresponding to cubic or hexagonal geometry.

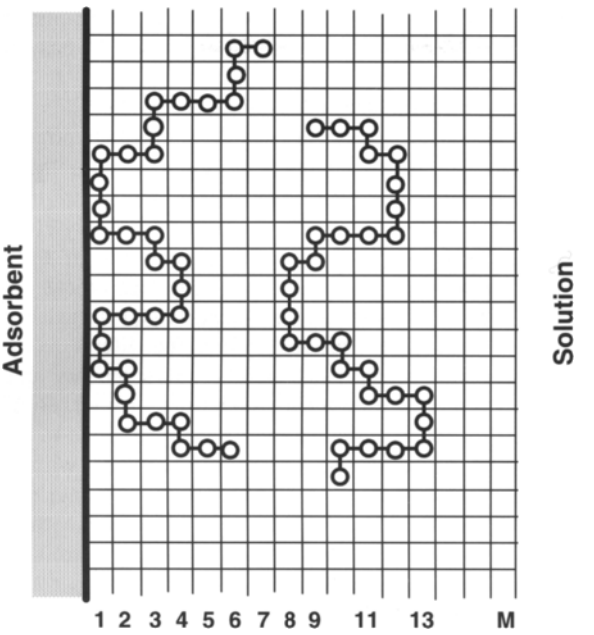


Fig. 8. Schematic diagram of polymer adsorption at an interface as depicted with a lattice model with M layers. Two polymers are shown, one adsorbed at a solid-solution interface and the other in solution. After Lyklema [54]

Interactions between polymer segments and solvent molecules are characterized by the Flory-Huggins parameter [63], χ. This is the net enthalpy change (in units of kT) of an exchange process in which a segment in a pure monomer phase is exchanged for a water molecule in a pure water phase. High values of χ indicate unfavorable interaction between the polymer and the solvent, i.e., a poor solvent and a hydrophobic effect.

Nonelectrostatic interactions between polymer segments and surface sites are described by an interaction energy, after Silberberg [64]. This is defined as the net enthalpy change (units of kT) of an exchange process in which a segment bound to the surface is exchanged for a solvent molecule in the solution. In this case, high values indicate favorable interaction between the polymer and the surface. Ligand exchange between a surface site and a polyelectrolyte segment is an example of such an interaction.

Electrostatic interactions in the interfacial region are described by a multi-Stern layer model that is a lattice version of the Poisson-Boltzmann equation. This approach was taken by Böhmer et al. [51] in extending the original SF theory for uncharged polymers to apply it to the adsorption of weak polyelectrolytes on charged surfaces.

SF theory has been developed to determine the equilibrium distribution of all polymer or polyelectrolyte segments at an interface. Model output includes not only the extent of polymer or polyelectrolyte adsorption at an interface but also the conformations of the adsorbed macromolecules, the consequent electric potential profile at the interface, and the interaction energy resulting from two approaching surfaces with adsorbed macromolecules. It has been used successfully for a wide variety of problems and applications in colloid science.

With the exception of the works of Tiller [65], Yang [66], and Au [67], publications derived from these theses, and research by Vermeer and coworkers [68-70], it has not been applied extensively to aquatic systems.

A number of parameters are typically needed as inputs to the SF model for its use. These are lattice type, lattice size, surface charge density or surface potential, the solution pH and salt concentrations, the number of segments per polymer chain, the valence of the segments, the bulk concentration of the polymer, the dielectric permittivity, the acidity or basicity constant of the polymer segments where appropriate, the Flory-Huggins parameter for solvent quality describing the hydrophilicity or hydrophobicity of the segment-water interactions, and the adsorption energy parameter. Most of these parameters are experimentally accessible for a given polyelectrolyte system (e.g., polygalacturonic acid and hematite in water) or can be estimated reliably from the literature. For NOM the determination of some of these parameters (e.g., number of segments and the dielectric permittivity can involve uncertainty. However prior simulations [18, 62, 66-67] indicate that the model results for NOM adsorption on hematite are not very sensitive to these parameters. An exception is the interaction energy of the polyelectrolyte segments with the surface sites on the solid. For the systems of interest in most prior research, this interaction energy has been used as a fitting parameter. This difficulty has been addressed with a hybrid SS/SF model [18, 62, 67] using surface complexation modeling, summarized hereinafter.

6.2. Surface complexation (Stumm-Schindler, SS) modeling

Excellent presentations of surface complexation modeling are available elsewhere (e.g., [21], [71-79]). Following is a summary of aspects of SS modeling that have been used in developing the SS/SF hybrid model outlined in Section 6.3.

We begin by considering the complex formation of a simple ligand (L^-) with a surface site on an oxide (>MeOH):

$$>MeOH + L^- \Rightarrow >MeL + OH^- \tag{49}$$

In terms of the mass law,

$$\frac{\{>MeL\}[OH^-]}{\{>MeOH\}[L^-]} = K_L^s(app) \tag{50}$$

Here { } denotes activities and [ ] denotes concentrations. The apparent equilibrium constant for ligand exchange at the surface, $K_L^s(app)$, is a conditional stability constant that depends on characteristics of the system.

It is not possible to separate the chemical and electrostatic interactions without making non-thermodynamic assumptions. Nevertheless, it is useful and common to separate operationally the overall interaction energy into a chemical part and a coulombic part (in this respect, the SS model is similar to SF theory for polyelectrolyte adsorption):

$$\Delta G_{tot}^o = \Delta G_{chem}^o + \Delta G_{coul}^o \tag{51}$$

where ΔG_{chem}^o is the "intrinsic" free energy term and ΔG_{coul}^o is the variable electrostatic or coulombic term.

The coulombic term is often written as follows:

$$\Delta G^{o}_{coul} = \Delta Z e \Delta\psi \tag{52}$$

where ΔZ is the change in charge of the surface species due to the adsorption reaction. The term ΔG^{o}_{coul} reflects the electrostatic work in transporting ions through the interfacial potential gradient, $\Delta\psi$.

For a surface complexation reaction, the intrinsic (primarily chemical) equilibrium constant can then be expressed in terms of the conditional or apparent equilibrium as follows:

$$K^{s}(int) = K^{s}(app)exp\left(\frac{\Delta Z e \Delta\psi}{kT}\right) \tag{53}$$

There are several models for $\Delta\psi$ and the electrostatic term. Three models in common use are termed the diffuse layer, constant capacitance, and triple layer models. Suffice it to say that the speciation of surface sites on an oxide can be modeled as a function of pH and, to a lesser extent, ionic strength using intrinsic equilibrium constants and an appropriate electrostatic correction.

In addition to Eqs. [49] and [50], these models use proton transfer equilibria at the surface and in solution; the following are examples:

$$>MeOH_2^{+} => MeOH + H^{+} \qquad K^{s}_{a1}(app) \tag{54}$$

$$>MeOH => MeO^{-} + H^{+} \qquad K^{s}_{a2}(app) \tag{55}$$

$$HL = L^{-} + H^{+} \qquad K_{a}(app) \tag{56}$$

The approach to surface chemistry and surface charge outlined above for surface complexation modeling contains some basic assumptions that facilitate understanding while at the same time containing simplifications that should be recognized in its applications. One such limitation is the use of three surface functional groups, $>MeOH_2^{+}$,$>MeOH$,*and*$>MeO^{-}$, to describe all of the possible surface sites at a solid-water interface. This has been described as a single site, two-pK approach. Based on its simplicity, its utility, and its demonstrated applicability, the single site, two pK model of Schindler, Stumm and coworkers was selected for surface complexation modeling in the hybrid SS/SF model.

6.3. Stumm-Schindler/Scheutjens-Fleer (SS/SF) hybrid model

The SF model incorporates non-coulombic surface site-polyelectrolyte segment interactions into a single parameter, χ_s; it also includes hydrophobic polyelectrolyte segment-water interactions through the Flory-Huggins parameter as a quantitative measure of solvent quality. The surface complexation (SS) model typically stresses chemical interactions and focuses on ligand exchange as the principal reaction between surface sites and anionic ligands. It does not address hydrophobic effects explicitly. The combined SS/SF model improves on the SF model by incorporating surface and ligand speciation in χ_s and improves

on the SS model by including solvent quality and by addressing the macromolecular properties of NOM. The hybrid model improves both approaches by combining important features of each.

In the hybrid (SS/SF) model for polyelectrolyte adsorption, natural organic matter such as Suwannee River humic acid (SRHA) is treated as a weak polyelectrolyte and its adsorption on an oxide such as hematite is considered to involve a specific chemical reaction between one type of ligand group on the polyelectrolyte with one type of proton-reactive surface species on hematite. The following reaction was found to describe well the adsorption of SRHA on hematite:

$$>MeOH + L^- => MeOHL^- \tag{57}$$

Here $>MeOH$ denotes a neutral site on an oxide surface and L^- denotes a deprotonated ligand group on NOM. Speciation of the oxide surface and the NOM ligand groups is calculated from experimental measurements of oxide and polyelectrolyte or NOM charge obtained from alkalimetric titrations at several ionic strengths (e.g. [66]).

In developing the hybrid SS/SF model, it is assumed that the reaction of a surface site with a ligand group in a reaction such as Eq. (56) involves a species-specific nonelectrostatic interaction parameter, $\chi_s(ads)$. Furthermore, the joint probability that the surface site is $>MeOH$ and the polyelectrolyte ligand site is L^- is given by the product $f_{MeOH} \cdot f_{L^-}$, where f_{MeOH} is the fraction of the proton-reactive surface species present as $>MeOH$ and f_{L^-} is the fraction of ligand sites on the polyelectrolyte that is deprotonated. If, as in the SF model, an average χ_s is assigned to all surface sites regardless of their chemical speciation, this average χ_s is termed $\chi_s(hyb)$ in the SS/SF hybrid model. It is determined as follows [18, 62, 67]:

$$\chi_s(hyb) = f_{MeOH} \cdot f_{L^-} \cdot \chi_s(ads) \tag{58}$$

Here, $\chi_s(ads)$ depends upon the specific reactants involved (e.g., Eq. (57)) and is constant, while f_{MeOH} and f_{L^-} vary with solution conditions such as pH and ionic strength. Changes in surface and polyelectrolyte speciation thus affect the average nonelectrostatic interaction parameter $\chi_s(hyb)$.

The adsorption of macromolecules at interfaces/aquafaces involves configurational entropic changes, several coulombic interactions, diverse noncoulombic or specific chemical interactions, and possible hydrophobic effects. Consideration of the adsorption of a humic acid on hematite over a broad pH range provides a useful example [18]. Typical results are presented in Fig. 9, in which the adsorption density of Suwannee River humic acid is plotted as a function of pH at an ionic strength of 0.01 M. Filled circles are experimental measurements. The dashed and dotted lines are model simulations using two different values of χ, the Flory-Huggins parameter incorporating hydrophobic effects. Configurational entropic changes for the adsorption of humic acid are always unfavorable under all of these conditions. These simulations indicate that SRHA adsorption is not sensitive to χ, indicating that hydrophobic effects are not important in this system. At pHs where bare hematite is positively charged (i.e., at pHs below 8.1, the oppositely charged case), coulombic interactions between anionic humic acid and positively charged hematite can dominate adsorption. Adsorption is accompanied by charge reversal [18], indicating that additional

interactions, non-coulombic in nature, are also involved. Humic acid is also adsorbed on hematite at pHs above 8.1 where the base oxide has a negative charge (the similarly charged case). With configurational entropic effects and all coulombic interactions being energetically unfavorable and hydrophobic effects being negligible, specific chemical effects between ligand sites on the humic acid and individual surface sites on the hematite drive adsorption. The process is very dependent of the speciation of the hematite surface sites and of the humic acid ligands and hence, in turn, on solution chemistry. *Since most particles in water have a negative primary charge and since most macromolecules in aquatic systems also have a net negative charge, it follows that specific or noncoulombic interactions can dominate macromolecular adsorption in natural and technological aquatic environments and affect all of the processes listed above as well as many others.*

7. CONTAMINANTS AND COAGULANTS

7.1. Natural organic matter, NOM

Natural organic matter is ubiquitous in natural aquatic and terrestrial environments. Prior to about 1970, interest in NOM in water treatment derived primarily from aesthetic considerations, viz., color removal. Since that time NOM has emerged as the constituent in raw water supplies that most often dominates the design, operation, and performance of water treatment systems. In the early 1970s, investigators in the Netherlands and the United States demonstrated that NOM is a precursor in the formation of hazardous disinfection byproducts (DBPs) such as trihalomethanes. This discovery continues to lead to dramatic changes in the practice of disinfecting water with accompanying complex interactions involving protection from hazardous chemicals and microbially induced disease that can be described as trade-offs among cancer, cryptosporidiosis, and costs. The important and often dominant role of NOM in establishing coagulant requirements has been recognized and became the driver of USEPA regulation such as the enhanced coagulation rule. NOM can consume the capacity of activated carbon beds used in potable water treatment and reduce their ability to remove organic micropollutants.

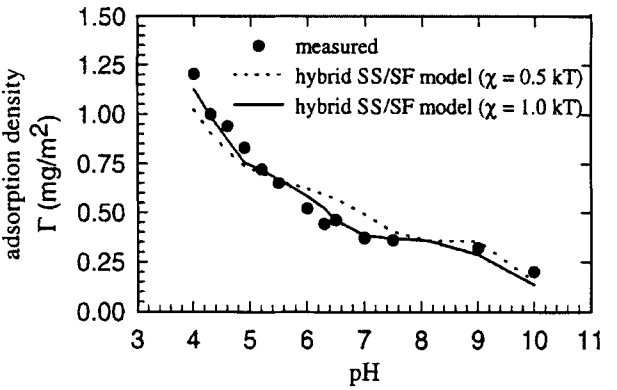


Fig. 9. Comparisons of measured and modeled adsorption density of Suwannee River humic acid (SRHA) on hematite as a function of pH. I = 0.01. Simulations are for two different assumed values of the Flory Huggins parameter, χ. From Au et al. [18]

NOM contributes significantly to the fouling of membranes in all membrane technologies. Finally, NOM can serve as a carbon and energy source for microbial fouling and regrowth in water distribution systems, problems that can be addressed by incorporating biological processes into drinking water treatment systems.

7.2. Aluminum and iron(III) salts.

These materials have been used for many decades as coagulants and their chemistry has been studied for at least as long a period. A consideration of our present understanding can begin with the work by Stumm and Morgan in 1962 [80]. These authors stressed that "coagulant metal ions are hydrolyzed in aqueous solution." These hydrolyzed species are responsible for the effectiveness of these metal salts as coagulants. Addition of alum to water results in the rapid, essentially instantaneous formation of monomeric hydroxoaluminum species. Depending on solution pH and alum dosage, cationic polyelectrolytes are formed, followed by the precipitation of an amorphous aluminum hydroxide. All of these reactions liberate protons, consume alkalinity and lower pH. Only the polyelectrolytes and the hydroxide precipitates are important as coagulants.

In the treatment of high turbidity waters that are low in NOM, the cationic polymers formed from alum adsorb on negatively charged particles and destabilize them, enabling aggregates to form when the particles are brought into contact. This is sometimes termed "charge neutralization". Coagulant requirements increase with raw water turbidity; a "stoichiometry" is observed. For waters that are low in turbidity and NOM, sufficient alum must be used to form a considerable amount of amorphous hydroxide precipitate that can enmesh the particles and increase interparticle collision rates. This has been termed "sweep floc" coagulation. Coagulant requirements are high in order to precipitate targets for collisions. Coagulant dose does not increase with raw water turbidity and stoichiometry is not observed. NOM reacts with metal coagulant species in two ways: (1) NOM-cationic polyelectrolyte precipitates can be formed, and (2) metal hydroxide precipitates can act as adsorbents for NOM. The relative importance of these two reaction pathways depends significantly on solution pH, with the importance of metal hydroxide precipitates acting as adsorbents for NOM increasing as the pH is increased above about 6 when Al(III) is used. These precipitation and adsorption reactions are different from the charge neutralization and sweep floc processes involved in particle or turbidity removal. Waters containing high concentrations of NOM require high coagulant dosages regardless of the removal process; stoichiometry is involved in both.

8. PACKED BED FILTRATION

A description of packed bed filtration in water treatment is provided here, summarized from some previous publications by the author and coworkers [81, 82].

8.1. Some filtration theory

A mathematical description of the removal process in packed bed filtration was presented by Iwasaki [83] and has been developed and used extensively since that time:

$$\frac{dC}{dL} = -\lambda_d C \tag{59}$$

Here C is the concentration of particles in suspension in the filter at depth L and λ_d is a filter coefficient that is a function of several physical properties of the system including particle size, media size, filtration rate, bed porosity, and temperature. The filter coefficient also reflects the influence of solution and surface chemistry on the attachment probability or sticking coefficient, α.

Several different approaches have been taken to relate the filter coefficient to filtration parameters quantitatively. The following approach introduced by Yao et al. [43] is adapted from the work of Friedlander [84].

$$\frac{dC}{dL} = -\frac{3}{2}(1-f)\alpha(p,c)\eta(p,c)\frac{C}{d_c} \tag{60}$$

Here $\eta(p,c)$ denotes the transport efficiency of a single collector (c) or media grain in the filter bed for receiving particles (p) from the suspension flowing past it. Analyses of mass transport summarized subsequently permit theoretical calculation of $\eta(p,c)$. This "single collector efficiency" can be considered as the ratio of the rate at which suspended particles are transported to the surface of a single grain of filter medium to the rate at which they approach it. The attachment probability or sticking coefficient of particles to the collector is given by $\alpha(p,c)$. It is the ratio (dimensionless) of the rate at which particles actually attach to a single collector to the rate at which they are transported to its surface. For a completely destabilized system, $\alpha = 1$; for a completely unstable one, $\alpha = 0$. In general, $0 \leq \alpha \leq 1$. With the exception of C, the remainder of the terms on the right-hand-side of Eq. (60) describe the number of single collectors in the differential length dL; d_c is the radius of a media grain (collector) and f is the porosity of the filter bed.

The first comprehensive treatment of particle transport in water filtration and the single collector efficiency was made by Yao et al. [43]. Significant improvements were made by Rajagopalan and Tien [85]. Recently Tufenkji and Elimelech [86] published the following relationship for the single collector efficiency in a clean porous media:

$$\eta_T = 2.4A_s^{1/3}N_R^{-0.081}N_{Pe}^{-0.715}N_{vDW}^{0.052} + 0.55A_sN_R^{1.675}N_A^{0.125} + 0.22N_R^{-0.24}N_G^{1.11}N_{vDW}^{0.053} \tag{61}$$

Here η_T is the total single collector efficiency resulting from transport by convective diffusion, interception, and gravity; A_s is described by Eq. (46). The other dimensionless groups in Eq. (60) are defined as follows:

$$N_R = \frac{d_p}{d_c} \qquad N_{Pe} = \frac{d_cU}{D} \qquad N_{vDW} = \frac{A}{kT}$$

$$N_A = \frac{A}{4\pi\mu d_p^2 U} \qquad \text{and} \qquad N_G = \frac{d_p^2(\rho_p - \rho_w)g}{18\mu U}$$

Here d_p denotes the size of the suspended particles; U is the superficial advective velocity, also termed the superficial filtration rate; D is the Brownian diffusion coefficient of the particles; A is the Hamaker coefficient for the particle-water-collector system; μ is the fluid viscosity, g is the gravity acceleration and ρ_p and ρ_w are the densities of the particle and the water respectively.

The first term on the right-hand side of Eq. (61) describes particle transport by convective diffusion; the second term represents transport by fluid flow/interception; the third term characterizes transport by gravity. These results by Tufenkji and Elimelech include the effects on particle transport to a single collector by neighboring collectors and also hydrodynamic retardation or the lubrication effect.

8.2. Conventional water treatment plants

A diagram showing the components of a conventional potable water treatment plant treating a surface water supply is presented in Fig 10. Solid-liquid separation is emphasized; disinfection processes, essential in these systems, are omitted for simplicity.

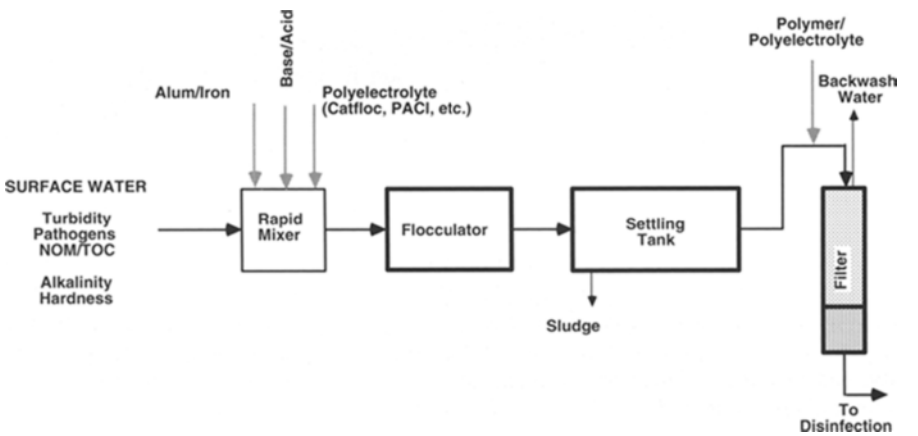


Fig. 10. Flow diagram of a conventional water treatment facility treating a surface water supply

Surface waters are depicted in Fig. 10 as containing turbidity, pathogens, NOM (natural organic matter) and TOC (total organic carbon), together with alkalinity (largely bicarbonate ions) and hardness (primarily calcium and magnesium ions). The principal objectives of a potable water treatment system are to remove or inactivate pathogens, to remove other particles and associated particle-reactive pollutants, often measured by light scattering as turbidity, and to remove natural organic matter (NOM) or total organic carbon (TOC). These natural organic substances are precursors of many of disinfection by-products (DBPs) formed during disinfection processes for pathogen inactivation and which have adverse health effects. Toxic contaminants such as herbicides, pesticides, and heavy metals are not addressed here.

Treatment units in Fig. 10 that are depicted in thick solid lines are units that are designed to produce and control mass transfer or interparticle collisions; these are depicted as flocculation tanks, settling tanks, and packed bed filters. The chemistry of the system is controlled in large part by the chemicals (coagulants) added in the rapid mixer, and by the reactions of these coagulants with water, particles, and NOM.

Particles and NOM in raw water supplies are removed primarily in sludge from sedimentation basins in conventional water treatment plants. Flocculation tanks provide many of the collisions needed to produce aggregates of settleable size. Additional collisions occur in the settling facilities. When NOM removal or high raw water turbidity requires coagulant

for NOM precipitation and adsorption or for particle destabilization, stoichiometry is involved. For waters low in NOM and turbidity, precipitation of metal hydroxides is needed to provide collisions in the flocculation process. In all cases, coagulant requirements and sludge production are high regardless of raw water quality if these plants are operated as designed.

Packed bed filters remove particles from the suspensions that flow through them. These particles may have been present in the raw water supply or have been formed in pretreatment. Particles in raw water supplies vary widely in their origin, concentration, composition, and size. Some are constituents of land-based or atmospheric inputs (clays, microorganisms including pathogens, asbestos fibers, etc.) and others are produced by chemical or biological processes in the water source (algae, organic detritus, precipitates of $CaCO_3$, FeOOH, etc.). Soluble natural organic substances (e.g., humic substances) are present in all surface water supplies; these can originate on land or in the water. Synthetic organic substances and many trace and toxic metals are particle reactive and are associated with particulate materials in the water source. The addition of chemical coagulants aggregates particles present in the water supply, adsorbs and precipitates humic substances, and forms new solid particles ($Al(OH)_3(s)$, etc. The particles reaching a filter in a water treatment system reflect all of these inputs and processes.

Effective particle removal in filtration depends on physical and chemical factors. Particles must be transported to the surfaces of filter media and they must attach there for removal to occur. Both hydrodynamics and chemistry are important; filtration is a physicochemical process. When biodegradable organic matter is present in the filter influent, biofilms can develop on the surfaces of the filter media and the process can have a substantial biological component. Design criteria most often specify filtration rate, media size and uniformity, and bed depth. These are pertinent to filter performance but are not sufficient to assure filter effectiveness and usually have less impact on the process than several other properties of the system.

Pretreatment chemistry is the most important factor affecting particle removal in packed bed filters. If the solution chemistry of the water to be treated, the surface chemistry of the particles to be removed, and the surface chemistry of the media used to accomplish removal are not properly controlled, efficient particle removal just will not occur, regardless of the filtration rate, bed depth, and media size installed in a filter system. In practice, this is accomplished by proper selection of the type and dosage of pretreatment chemicals. These may include aluminum or iron(III) salts, polymeric inorganic coagulants such as polyaluminum chloride (PACl) and activated silica, synthetic organic polymers and polyelectrolytes, simple acids and bases, and even oxidants such as ozone.

The principal factors affecting head loss development in packed bed filters are the concentration and size of the particles applied to the filters. Media size, bed depth and porosity, and filtration rate have important effects, but the physical properties of the particles applied to the filters dominate head loss development in those filters. Small particles produce large head losses; submicron particles induce rapid head loss development when treated by conventional packed bed filters.

The size of the particles in the water being filtered also affects the removal efficiency of a clean filter bed. A critical suspended particle size exists, in the region of 1 μm, for which the clean bed removal efficiency is a minimum. Particles in this size range have the fewest opportunity for contact with the filter media and subsequent removal from suspension. Smaller particles are transported predominantly and effectively by convective diffusion and larger ones by fluid flow (interception) and gravity settling.

When filtration is successful in the early stages of a filter run, i.e., when the hydrodynamics of particle transport and the chemistry of particle attachment combine to produce effective particle removal in a clean filter bed, these retained particles can act as filter media during later stages of the run. This produces a ripening of the filter; the efficiency of removal of particles or turbidity improves with time and can be substantially greater than the "clean" bed in place when the filter run is begun. Particles removed from suspension by the filter become sites, collectors, or media for the deposition of additional material. The important result is that in many and perhaps most of the packed bed filters used in water and wastewater treatment, the actual filter media operative during most of a run are not the sand, coal, or other media specified by the designer. Rather, they are the particles present in the water applied to the filter bed and removed in the filter. Stated another way, the actual filter media active during most of a packed bed filter run are formed from particles present in the raw water supply and altered during pretreatment processes and from new particles formed by precipitation during pretreatment. Hence the "design" of filter media really involves the selection of the raw water source and the design of the pretreatment facilities that are installed ahead of the filters. The design of filter media is accomplished "upstream" of the filters by the concentration, size, and surface properties of the particles applied to the filters and the solution chemistry of the aquatic phase.

REFERENCES

[1] B.V. Derjaguin, Kolloid Zeitschrift , 69 (1934) 155.
[2] J.N. Israelachvili, Intermolecular and Surface Forces, 2nd Ed., Academic Press, London, 1992.
[3] B.V. Derjaguin, Theory of the Stability of Colloids and Thin Films, translated by R.K. Johnston, consultants Bureau, New York, 1989.
[4] J. Gregory, Critical Reviews in Environmental Control, C.P. Straub (ed.), CRC Press, Boca Raton, Florida, 1989.
[5] W.B. Russel, D.A. Saville and W.R. Schowalter, Colloidal Dispersions, Cambridge University Press, Cambridge, 1989.
[6] V.A. Parsegian, Physical Chemistry: Enriching Topics from Colloid and Surface Science, H. van Olphen and K.J. Mysels (eds.), Theorex, La Jolla, 1975.
[7] J. Lyklema, Fundamentals of Interface and Colloid Science, Vol. I: Fundamentals, Academic Press, London, 1991.
[8] R.J. Hunter, Foundations of Colloid Science, Vol. I, Clarendon Press, Oxford, 1987.
[9] H.C. Hamaker, Physica, 4 (1937) 1058.
[10] E.M. Lifshitz, Theory of molecular attractive forces, Soviet Physics JETP, 2 (1956) 73.
[11] J. Gregory, J.Colloid Interface Sci., 83 (1981) 138.
[12] J. Mahanty and B.W. Ninham, Dispersion Forces, Academic Press, London, 1976.
[13] E.J.W. Verwey and J.Th.G. Overbeek, Theory of the Stability of Lyophobic Colloids, Elsevier, New York, 1948.
[14] G. Gouy, J. Phys. Radium, 9 (1910) 457.
[15] D.I. Chapman, Phil. Mag., 25 (1913), 475.
[16] J. Lyklema, Fundamentals of Interface and Colloid Science, Vol. II, Solid-Liquid Interfaces, Academic Press, London, 1995.
[17] G. Sposito, The Surface Chemistry of Soils, Oxford University Press, New York, 1984.
[18] K.-K. Au, A.C. Penisson, S. Yang and C.R. O'Melia, Geochim. Cosmochim. Acta, 63 (1999) 2903.
[19] O. Stern, Z. Electrochem., 30 (1924) 508.
[20] R. Hunter, Zeta Potential in Colloid Science, Academic Press, London, 1981.
[21] T. Hiemstra, W.H. van Riemsdijk and G.H. Bolt, J. Colloid Interface Sci.,133 (1989) 91.
[22] D.C. Grahame, Chem. Rev., 41 (1947) 441.

[23] D.E. Yates, S. Levine and T.W. Healy, J. Chem. Society, Faraday Trans. I, 70 (1974) 1807.
[24] J.A. Davis, R.O. James and J.O. Leckie, J. Colloid Interface Sci., 63 (1978) 480.
[25] J.A. Davis and J.O. Leckie, J. Colloid Interface Sci., 67 (1978) 90.
[26] J.A. Davis and D.B. Kent, Mineral-Water Interface Geochemistry, M.F. Hochella, Jr and A.F. White (eds.), Mineralogical Society of America, Washington, D.C. 1990.
[27] K.F. Hayes and L.E. Katz, Physical Chemistry of Mineral Surfaces, P.V. Brady (ed.), CRC Press, Boca Raton, 1996.
[28] B.V. Derjaguin and L.D. Landau, Acta Physicochemica URSS, 14 (1941) 633.
[29] J.Th.G. Overbeek, Colloid Science Volume 1, Irreversible Systems, Ch. VI, H.R. Kruyt (ed.), Elsevier, Amsterdam, 1952.
[30] J. Gregory, Critical Reviews in Environmental Control, C.P. Straub (ed.), CRC Press, Boca Raton, 1989.
[31] R. Hogg, T.W. Healy and D.W. Fuerstenau, Trans. Faraday Soc., 62 (1966) 1638.
[32] S. Usui, J. Colloid Interface Sci., 44 (1973) 107.
[33] G.R. Weise and T.R. Healy, Trans. Faraday Soc., 66 (1970), 490.
[34] J. Lyklema, Pure Appl. Chem., 52 (1980) 1221.
[35] D. Chan and D.J. Mitchell, J. Colloid Interface Sci., 95 (1983) 193.
[36] M.W. Hahn, D. Abadzic and C.R. O'Melia, Environ. Sci. Technol., 38 (2004) 5915.
[37] J.A. Long, D.W.J Osmond and B. Vincent, J.Colloid Interface Sci., 42 (1973) 545.
[38] M. von Smoluchowski, Zeitschrift für physicalische Chemie, 92 (1917) 129.
[39] N. Fuchs, Zeitschrift für Physik, 89 (1934) 736.
[40] L.A. Spielman, J.Colloid Interface Sci., 33 (1970) 562.
[41] C.R. O'Melia, Proc. ASCE, J. SED, 91, SA2 (1965), 92.
[42] W. Findheisen, Meteorligische Zeitschrift, 56 (1939) 365.
[43] K.-M. Yao, M.T. Habibian and C.R. O'Melia, Environ. Science Technol., 5 (1971) 1105.
[44] H. Schulze, J. Prakt. Chem., 25 (1882) 431.
[45] H. Schulze, J. Prakt. Chem., 27 (1883) 320.
[46] W.B. Hardy, Proc. Royal Soc. London, 66 (1900) 110.
[47] W. Stumm and C.R. O'Melia, J. - Am. Water Works Assoc., 60 (1968) 514.
[48] R. Simha, H.L. Frisch and F.R.Eirich, J. Phys. Chem., 57 (1953) 584.
[49] J.M.H.M. Scheutjens and G.J. Fleer, J. Phys. Chem., 83 (1979) 1619.
[50] J.M.H.M. Scheutjens and G.J. Fleer, J. Phys. Chem., 84 (1980) 178.
[51] M.R. Bohmer, O.A. Evers and J.M.H.M. Scheutjens, Macromolecules, 23 (1990) 2288.
[52] J. Blaakmeer, M.R. Bohmer, M.A. Cohen Stuart, and G.J. Fleer, Macromolecules, 23 (1990), 2301.
[53] H.A. van der Schee and J. Lyklema, J. Phys. Chem., 88 (1984) 6661.
[54] J. Lylema, Flocculation, Sedimentation, and Consolidation: Proceedings of the Engineering Foundation Conference, B. M. Moudgil and P. Somasundaran (eds.), Engineering Foundation, New York, 1985.
[55] O.A. Evers, G.J. Fleer, J.M.H.M. Scheutjens and J. Lyklema, J. Colloid Interface Sci., 111 (1986) 446.
[56] M.A. Cohen Stuart, T. Cosgrove and B. Vincent, Adv. Colloid Interface Sci., 24 (1986) 143.
[57] J.M.H.M. Scheutjens, G.J. Fleer and M.A. Cohen Stuart, Colloids Surf., 21 (1986) 285.
[58] G.J. Fleer, M.A. Cohen Stuart, J.M.H.M. Scheutjens, T. Cosgrove and B. Vincent, Polymers at Interfaces, Chapman Hall, London, 1993.
[59] N.G. Hoogeveen, M.A. Cohen Stuart and G.J. Fleer, J. Colloid Interface Sci., 182 (1996) 133.
[60] A.W.P. Vermeer, F.A.M. Leermakers and L.K. Koopal, Langmuir, 13 (1997) 4413.
[61] C.L. Tiller and C.R. O'Melia, Colloids Surf. A, 73 (1993) 89.
[62] K.-K. Au, S. Yang and C.R. O'Melia, Environ. Sci. Technol., 32 (1998) 2900.
[63] P.J. Flory, Principles of Polymer Chemistry, Cornell University Press, Ithaca, 1953.
[64] A. Silberberg, J. Chem. Phys., 48 (1968) 2835.
[65] C.L. Tiller, Colloidal Stability in Natural Waters: Experimental and Model Studies of the Role of Natural Organic Matter, PhD dissertation, Johns Hopkins University, Baltimore, 1993.

[66] S. Yang, Effects of Adsorbed Natural Organic Matter on Colloidal Stability in Aquatic Systems, PhD dissertation, Johns Hopkins University, Baltimore, 1996.
[67] K.-K. Au, Natural Organic Matter at Solid/Liquid Interfaces: Complexation, Conformation, and Colloidal Stabilization, PhD dissertation, Johns Hopkins University, Baltimore, 1998.
[68] A.W.P. Vermeer, W.H. van Riemsdijk and L.K. Koopal, Langmuir, 14 (1998) 2810.
[69] A.W.P. Vermeer and L.K. Koopal, Langmuir, 14 (1998) 4210
[70] A.W.P. Vermeer and L.K. Koopal, J. Colloid Interface Sci., 212 (1999) 176.
[71] W. Stumm, C.P. Huang and S.R. Jenkins, Croatica Chemica Acta, 42 (1970) 223.
[72] J.A. Davis and J.O. Leckie, Environ. Sci. Technol., 12 (1978) 1309.
[73] W. Stumm, R. Kummert and L. Sigg, Croatica Chemica Acta, 53 (1980) 291.
[74] J. Westhall and H. Hohl, Adv. Colloid Interface Sci., 12 (1980) 265.
[75] P.W. Schindler, Adsorption of Inorganics at Solid-Water Interfaces, M.A. Anderson and A.J. Rubin (eds.), Ann Arbor Science, Ann Arbor, 1981.
[76] T. Hiemstra, W.H. van Riemsdijk, and G.H. Bolt, Journal of Colloid and Interface Science, 133 (1989) 105.
[77] D.A. Dzombak and F.M.M. Morel, Surface Complexation Modeling: Hydrous Ferric Oxide, John Wiley and Sons, New York, 1990.
[78] W. Stumm, Chemistry of the Solid-Water Interface, John Wiley and Sons, New York, 1992.
[79] J.C. Westall, Materials Research Society Symposium Proceedings, Materials Research Society, 353 (1995) 937.
[80] W. Stumm and J.J. Morgan, J. - Am. Water Works Assoc., 54 (1962) 971.
[81] C.R. O'Melia, ASCE J.Environ. Eng. Div., 111 (1985) 874.
[82] C.R. O'Melia, J. Water SRT-Aqua, 40 (1991) 371.
[83] T. Iwaski, J. - Am. Water Works Assoc., 29 (1937) 1591.
[84] S.K. Friedlander, Ind. Eng. Chem., 50 (1958) 1161.
[85] R. Rajagopalan and C. Tien, AIChE J., 22 (1976) 523.
[86] N. Tufenkji and M. Elimelech, Environ. Sci. Technol., 38 (2004) 529.

Index

A

Activated carbon, 72, 82, 126, 133–145, 148–151, 155–162, 167–172, 174, 189, 201, 314, 355
Adsorption isotherm, 135, 140, 142, 148, 157, 170
Agglomeration, 58, 60, 75
Aggregation, 5, 13, 17, 25–29, 33–37, 53, 55, 58, 76, 113, 118, 122, 123, 305, 311, 317–319, 322, 326, 342, 344, 345, 347, 348, 349
Algal toxin, 133, 137, 147, 148
Arsenic, 17, 193–196, 198, 199–205
Atomic force microscopy (AFM), 222

B

Biocorrosion, 246, 248, 249, 252, 253
Biofilm growth, 113, 252, 282
Biofouling, 113, 252
Bridging, 5, 40, 65–67, 72, 76, 308
Brownian diffusion, 26, 94, 96, 106, 347, 348, 357

C

Cake resistance, 59, 112, 125, 227
Cationic polyelectrolytes, 72, 228, 356
Centrifugation, 74, 75, 209, 225, 227, 229, 230, 241, 299, 301
Charge neutralisation, 6, 58, 59, 65–67, 72, 76, 228
Chlorination, 66, 80–82, 84, 143, 184, 185, 188, 203, 207, 304, 314
Clarification, 1–3, 5, 6, 18, 64, 89, 103, 196, 225, 230, 236, 238, 311
Clarifier, 17, 18, 32, 75, 76, 225, 238–240
Coagulant, 6–10, 13, 14, 16–22, 33, 40, 59, 63, 65, 70–74, 76, 77, 83, 96, 98, 104, 122, 124, 127, 188, 196, 202, 225, 226, 228, 235–237, 241, 355, 356, 359
Colour, 64, 66, 70, 71, 76, 179, 228, 232, 235–237
Compressibility, 60, 227–230, 232, 235–241
Contact zone, 89–96, 98–100, 102, 104, 106
Corrosion, 2, 9, 64, 246–250, 252, 253, 257, 282, 284–286
Crossflow, 110–113, 116, 123, 124, 127, 128
Cryptosporidium, 73, 84, 104, 207–209, 217, 222

D

Deadend Filtration, 110–113, 116, 123, 124, 127, 128, 304
Destratification, 184, 185, 190
Dewatering, 5, 59, 63, 70, 74–76, 183, 204, 205, 225–229, 232, 234–238, 240–242
Diffusivity, 229, 234–237
Direct Filtration, 21, 70, 72, 83, 103, 185, 200
Discolouration, 184, 279–282, 285–287, 292, 297
Disinfection by-products, 17, 64, 76, 82
Dissolved Organic Carbon (DOC), 17, 64, 140, 208, 209
Distribution network, 189, 245, 285, 286, 292, 293, 297
DLVO theory, 4, 208, 216–219, 222, 311, 317, 336, 340, 344, 345, 349, 350

E

Electrical double layer, 3, 208, 209, 213, 216–218, 220, 325, 326, 340
Electrophoretic mobility, 76, 210, 212, 213, 219, 310
Electrostatic repulsion, 3, 4, 17, 66, 118, 120, 121, 202, 214, 325, 336, 342
Enhanced coagulation, 17, 73, 228, 355
Extracellular polymeric substances, 117, 246

F

Filtration, 1–3, 5, 6, 17, 18, 21, 55, 59, 63, 70–74, 76, 82, 83, 94, 96, 103, 105, 111, 112, 114, 116, 117, 119, 121, 122, 124, 125, 127, 137, 138, 180, 181, 185–190, 196, 199, 200, 202, 203, 207, 222, 225–230, 232–234, 236, 240, 241, 283, 301, 303, 304, 314, 321, 345, 348, 356, 357, 359, 360
Floc, 8, 18, 89, 94, 100, 102
Flocculation, 1, 2, 5, 6, 13, 25, 30–35, 37, 38, 40, 43, 55, 58, 59, 63–66, 70, 75, 76, 78, 89, 96, 103–105, 126, 139, 180, 183–185, 202, 227, 230, 348, 359
Flotation, 1, 2, 5, 63, 70, 72–74, 89, 92–96, 98, 100, 102
Flux decline, 114, 116, 119–122, 124, 129, 130, 303
Fractal, 16, 29, 34–37, 39, 45–50, 52, 53–55, 58–60
Fractal aggregates, 16, 35, 36
Fractal dimension, 35–37, 39, 46–50, 52, 54, 55, 58–60
Fractionation, 2, 71, 114, 140, 299, 301–304, 312
Fulvic acid, 8, 300, 301, 305, 306, 308

G

Granular activated carbon, 137, 171, 201
Gravity belt filtration, 241

H

Humic acid, 64, 73, 77, 118, 120–122, 124, 140, 188, 300, 305, 307, 310, 354, 355
Hydrolysis, 5, 6, 11, 12, 22, 31, 68–70, 79, 80, 193, 196
Hydrolysis constants, 6, 11, 12
Hydrophilicity, 119, 158, 167, 173, 352
Hydrophobicity, 59, 77, 114, 118–122, 128, 149, 301, 302, 310, 352

I

Image analysis, 46, 47, 53, 54, 58, 59
Ion exchange resin, 126

J

Jar tests, 9, 18, 19, 66, 73, 76

L

Light scattering, 17, 49, 53, 54

M

Manganese, 64, 179–190, 198, 199, 203, 248
Membrane, 5, 18, 59, 60, 64, 72, 73, 109–130, 181, 189, 190, 201–203, 233, 302–304, 307, 312, 314, 356
Membrane fouling, 59, 109, 112, 115, 117–122, 124–127, 129, 130, 302, 307
Microfiltration, 110, 117, 121, 124, 196, 202, 203, 304
Mixing, 5, 7, 10–12, 14, 19, 33–35, 41, 66, 71, 72, 75, 76, 104, 138, 202, 232, 311
Molecular size, 2, 149, 161, 174

N

Nanofiltration, 110, 116, 118, 121, 129, 130, 204, 205, 304
Natural Organic Matter (NOM), 1, 59, 63, 65, 71, 109, 114, 118, 129, 155, 171, 203, 213, 218, 219, 221, 222, 226, 228, 241, 267, 299, 304, 317, 326, 354, 355

O

Oocyst, 74, 209, 211, 212, 214–220, 222
Oxidation state, 180, 182, 183, 186

P

Particle collisions, 5, 27, 32, 33
Particle size, 25, 27–29, 37, 45, 72, 96, 137, 227, 344, 345, 357, 359
Particle stability, 1, 3, 25, 299, 317, 344
Pathogens, 89, 109, 253, 359

Permeability, 59, 60, 82, 112, 113, 121, 126–128, 130, 227, 228, 235–237, 239–241
Plate and frame filtration, 241
Point-of-use, 189, 190
Polarization, 110, 112–114, 121, 125
Polymer toxicity, 76, 84
Polymerization, 7, 128–130
Polymers, 5–9, 11, 13, 16, 40, 59, 63, 65–84, 97, 103, 126, 128, 129, 208, 216, 250, 264, 317, 350, 356, 359
Pore structure, 118, 120, 133–135, 138, 142, 155, 157, 172, 174
Potential determining ions, 17, 339, 340
Powdered activated carbon, 126, 137, 171, 189
Precipitation, 5, 9–13, 16, 20, 31, 37, 43, 64, 66, 196, 198, 203, 204, 228, 236, 259, 260, 300, 356, 359, 360

R

Regeneration, 137, 138, 199–201, 204
Regrowth, 41, 43, 53, 280, 286, 356
Residuals, 8, 76, 183, 202, 204, 205, 245, 250
Resuspension potential method, 280
Reticulation, 2, 179, 181, 184, 186, 188, 189, 285
Rise velocity, 92, 93, 95, 99–101

S

Scale formation, 113, 259, 269
Scaling, 46, 226, 257, 259, 267, 268, 274–276
Sedimentation, 2, 5, 26, 28, 36, 52, 63–65,70–75, 83, 89, 102–104, 185, 186, 226–228, 231
Sequestration, 188
Settling, 19, 28, 29, 46, 52–54, 59, 63, 72–74, 94, 96, 98, 103–106, 137, 227–232, 234, 235, 237–240, 285, 289, 348, 359
Settling rate, 28, 29, 227, 238, 239
Settling velocity, 19, 52
Shear rate, 28, 29, 31–34, 38, 40
Sludge, 9, 37, 50, 51, 54, 59, 63, 64, 67, 70, 72, 74–76, 78, 102, 103, 137, 183, 188, 189, 198, 199, 202, 204, 205, 226, 228, 229, 232–237, 239–241, 359
Streaming current, 4, 5, 19, 21, 76
Surface charge, 3, 13, 14, 65, 66, 72, 75, 76, 118, 120, 122, 128, 136, 140, 156, 158, 160, 161, 170, 171, 174, 201, 202, 208, 209, 212, 214, 222, 310, 311, 317, 327, 329–331, 334, 335, 339, 340, 350, 352, 353
Surface forces, 97, 208, 215, 222
Surface modification, 128, 129
SUVA, 17, 307
Sweep floc, 8, 18

T

Taste and odour, 64, 72, 133, 137, 142, 143, 150, 189, 245
Thickener, 238–240
Thickening, 225, 227–230, 238, 241
Topography, 208, 209, 211, 264, 265
Turbidity, 4, 17–19, 31, 58, 63, 65, 66, 70, 71, 73, 74, 76, 78, 82, 84, 89, 103, 109, 110, 128, 228, 235–237, 280, 283–285, 287–291, 356, 359, 360

U

Ultrafiltration, 110, 116–118, 120, 121, 125, 127, 140, 304, 314

V

Van der Waals forces, 3, 4, 97, 98, 317, 319, 320, 343, 346, 347

Z

Zeta potential, 4, 5, 66, 76, 120, 122, 208

This Page Intentionally Left Blank