Synthesis of zwitterionic compounds for aquatic toxicity testing for QSAR correlation studies.

Davies, Joanna

SYNTHESIS OF ZWITTERIONIC COMPOUNDS FOR AQUATIC TOXICITY TESTING FOR QSAR CORRELATION STUDIES

by

Joanna Davies, B.Sc. (Hons)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy by the University of Wales, Swansea

Declaration

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed.....(candidate)
Date.....17/06/03.....

Statement 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed.....(candidate)
Date.....17/06/03.....

Statement 2

I hereby give my consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed.....(candidate)
Date.....17/06/03.....

Acknowledgement

I would like to thank my academic supervisor Dr. Robert S. Ward for giving me the opportunity to pursue a research degree at the University of Wales, Swansea. I would especially like to express gratitude for his enthusiasm, continuous support and invaluable guidance throughout the course of the project. I would also like to express thanks to my industrial supervisor, Dr. Geoff Hodges, for his visits to Swansea for interesting discussion, for the free lunches and for making me feel welcome at SEAC. I would especially like to thank Dr. David W. Roberts for his teaching of the 'back-of-envelope' approach to log P calculations, Suzanne Murphy for her help with toxicity testing at SEAC (Port Sunlight) and both Helen David and Jayne Costello for their kindness and assistance at SEAC (Sharnbrook).

A big thank you goes to my research group (past and present) – Steve, Irene, Rose, Gareth. T, Big Dave, Shalina, Christian, Rainer, Markus and Michaela, for their valuable help, support, humour and gossip. I must not forget Stan – the indispensable technician for his invaluable assistance and Myron for providing such an efficient NMR service. I would also like to thank the National EPSRC Mass Spectrometry Centre for providing mass spectra and Jonathan Jones of the Mass Spectrometry Research Unit at Swansea for helping with the detection of the alkylammonium sulfobetaines and for providing the CCC data. I would also like to thank friends such as Owen, Louise, Swadipa, Rhian, Alun, Lara, Gareth. P, Matt, James, Pat, Mas, John. T, Cath, Big Chris, Kath, Michiko and Mami for their company at lunch-times, tea-breaks and evenings out. Thanks also to Friday afternoon' s basketball buddies!

I would like to take this opportunity to thank my parents for being supportive of my ambitions, and for their continued love and understanding over the years. I must also thank my brother, Carl, for not thinking that his big sister was too weird! Furthermore, thank you to my partner, Paul, for his love, encouragement and understanding.

Finally, I express gratitude to the EPSRC and Unilever Research for their financial support, without which the research would not have been possible. Thanks also to Unilever for conferences abroad (QSAR 2000 –Bourgas, Bulgaria; QSAR 2002 – Ottawa, Canada) and for 3 months experience of research in industry.

Abstract

22 zwitterionic compounds (10 short-chain surfactants; 12 non-surfactants) were synthesised obeying the general formula $R\text{-}N^{+}(CH_3)_2(CH_2)_nSO_3^{-}$, where n = 2 to 4, by reacting the corresponding N, N-dimethylamines with either sodium-2-chloroethane sulfonate (n = 2), 1, 3-propanesulfonate (n = 3) or 1, 4-butanesulfonate (n = 4). The R group varied from a C6 to C12 alkyl chain, to a phenylalkyl unit bearing a C1 to C4 chain and finally to a phenylpropyl unit with a C4 to C6 *para*-substituted alkyl group.

Octanol/water partition coefficients of the 22 sulfobetaines were determined by a conventional stir-flask procedure. The amount of solute in both the octanol and water layer was quantified using a reverse-phase HPLC technique. A UV detection mechanism was employed for those sulfobetaines that possessed a suitable chromophore for UV detection and an electrospray ionisation mode of detection was used for the analysis of those sulfobetaines that lacked a chromophore suitable for UV detection.

Acute aquatic toxicity to the aquatic invertebrate, *Daphnia magna* Straus, was reported as log ($1/EC_{50}$). The EC_{50} values were determined experimentally using a standard *Acute Immobilisation Test* recommended by the OECD Guideline 202 and the internal Unilever document, Ecotoxicology SOP 019 11.

A log P-based QSAR was then derived which was found to be analogous to the standard polar narcosis equation, suggesting that zwitterionic sulfobetaines act as polar narcotics.

Experimental log P determined by the stir-flask procedure and the rules of Rekker and Roberts for the calculation of log P for quaternary ammonium compounds of the cationic type, were then used in the derivation of key fragment values and interaction factors for use in log P calculations of sulfobetaines using the Leo and Hansch approach. Furthermore, log P predictions provided by KowWin, a computerised program developed by the Syracause Research Cooperation, were suitably amended to take into consideration our experimental results.

Finally, an investigation into the suitability of the phospholipophilicity parameter, log $k'_{IAM,}$ for defining aquatic toxicity was performed and the efficacy of using this parameter and log P to predict aquatic toxicity was compared. In addition, other chromatographic methods for estimating log P were investigated. These include the indirect reverse-phase HPLC method, the direct reverse-phase HPLC method and Counter Current Chromatography.

Contents

Abbreviations

Ac	acetyl
AchE	acetylcholine esterase
A. niger	aspergillus niger
ASTM	American Society for Testing and Materials
ATPase	adenine triphosphatase
AU	absorption units
Av.	average
BCF	bioconcentration factor
BTMAC	benzyltrimethyl ammonium chloride
Calcd.	calculated
CCC	countercurrent chromatography
CE	capillary electrophoresis
CI	chemical ionisation
ClogP	computerised program for calculation of log P using rules of Leo and Hansch
CoMFA	comparative molecular field analysis
Conc.	concentration
CPC	centrifugal partition chromatography
CMC	critical micelle concentration
CNDO	complete neglect of differential overlap
CNS	central nervous system
C_w	surfactant concentration in water
C8	ocylsilane bonded phase
C18	octadecylsilane bonded phase
d	doublet
DCA	direct current amperometry
DCCC	droplet countercurrent chromatography
DCCI	1,3-dicyclohexylcarbodiimide
D. magna	*Daphnia magna*
DMF	dimethylformamide
DO	dissolved oxygen
d.p	decimal places
D. pulex	*Daphnia pulex*
EC	electrochemical
E. coli	*escherichia coli*
EC_{50}	median effective concentration (concentration estimated to produce a specified effect in 50 % of a test population)
EI	electron impact
ELSD	evaporative light scattering detection
EPA	environmental protection agency
E_s	steric substituent constant
ESI	electrospray ionisation
EtOAc	ethylacetate
EtOH	ethanol
Expt.	Experiment
F	structural factor
f	fragment constant
FAB	fast atom bombardment
FD	field desorption
FDA	Food and Drug Administration
GC	gas chromatography
GLP	good laboratory practice
h	hours
H-bond	hydrogen bond
HDES	hydrodynamic equilibrium system
HMG-CoA	3-hydroxy-3-methyl-glutaryl-coenzyme A
HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSCCC	high speed countercurrent chromatography

HSES	hydrostatic equilibrium system
IAM	immobilised artificial membrane
ICs	isolated carbon atoms
IC_{50}	median inhibitory concentration
ID	internal diameter
INDO	intermediate neglect of differential overlap
IR	infrared
IS	ionspray
J	coupling constant
k′	retentivity (or capacity factor)
K_α	hydrogen bond donor ability
K_β	hydrogen bond acceptor ability
k'_{C8}	retentivity measured on octylsilane column
k'_{C18}	retentivity measured on octadecylsilane column
$k'_{OC/C8}$	retentivity measured on octanol-coated octysilane column
k'_{IAM}	retentivity measured on immobilised artificial membrane column
K_m	solute intestinal absorption
k'_w	retentivity measured using mobile phase of 100 % water
KowWin	computerised program used to calculated log P developed by the syracause research cooperation
LAB	linear alkylbenzene
L and H	Leo and Hansch
LAS	linear alkylbenzene sulfonate
LC	liquid chromatography
LC_{50}	median lethal concentration (concentration estimated to produce lethality in 50 % of a test population)
LD	laser desorption
Lit.	literature
Log D	log of the octanol/water distribution coefficient
Log P	log of the octanol/water partition coefficient
LOEC	lowest observed effect concentration
LOEL	lowest observed effect level
LRMS	low resolution mass spectrometry
LSDR	lime soap dispersant requirements
LSIMS	liquid secondary-ion mass spectrometry
LUMO	lowest unoccupied molecular orbital
m	multiplet
MALDI	matrix-assisted laser desorption
MATC	maximum acceptable toxicant concentration
MEEKC	microemulsion electrokinetic chromatography
MeOH	methanol
MNDO	modified neglect of differential overlap
MO	molecular orbital
m.p	melting point
MR	molar refractivity
MS	mass spectrometry
m/z	mass/charge ratio
NICs	non-isolated carbon atoms
NMR	nuclear magnetic resonance spectroscopy
NOEC	no observed effect concentration
NOEL	no observed effect level
NP	normal phase
NTP	national toxicology programme
Octanol	*n*-octanol
ODP	octadecyl-polyvinylalcohol
OECD	Organisation for Economic Cooperation and Development
P	octanol/water partition coefficient
PC	phosphatidylcholine
PD	plasma desorption
PDBF	position dependant branch factor
Ph	phenyl

PS-DVB	polystyrene-divinylbenzene
PTFE	polytetrafluoroethylene
q	quartet
q^5	quintet
QAC	quaternary ammonium compound
QUATS	quaternary ammonium salts
QSAR	quantitative structure-activity relationship
r^2	correlation coefficient
RI	refractive index
RF	response factor
R_f	retention factor
R_m	free energy-based constant
RP	reverse phase
r.t	room temperature
s	singlet
s^6	sextet
S. aureus	*staphylococcus aureus*
SAR	structure-activity relationship
SDS	sodium dodecyl sulfate
SEAC	Safety and Environmental Assurance Centre
SMILES	simplified molecular input line entry system
SRC	Syracause Research Cooperation
SS	squalene synthase
t	triplet
t_o	dead time (elution time of unretained compound)
***t*-BuOH**	tertiary butylalcohol
TH	total hardness
THF	tetrahydrofuran
TLC	thin layer chromatography
t_R	retention time of analyte
TSP	thermospray
UV	ultraviolet
V_{aq}	volume of aqueous phase
V_e	volume of eluent
V_{oct}	volume of octanol
V_R	retention volume of analyte
V_s	volume of stationary phase
δ	nuclear magnetic resonance chemical shift
ΔC	change in constitution of a compound
$\Delta\phi$	change in physiological action
γ_{AW}	air-water interfacial surface tention
λ	wavelength
π	hydrophobicity substituent constant
π^*	dipole term
σ	electronic substituent constant
2-D	two dimensional
3-D	three dimensional
12-MO-silica	12-hydroxdecylaminopropyl silica
12-OH-silica	12-methoxydodecanoic acid on aminopropyl silica

Examples of structure abbreviations for sulfobetaines:

C6C3	$C_6H_{13}N^+(CH_3)_2(CH_2)_3SO_3^-$
PhC2C4	$C_6H_5(CH_2)_2N^+(CH_3)_2(CH_2)_4SO_3^-$
C4PhC3C3	p-$C_4H_9C_6H_4(CH_2)_3N^+(CH_3)_2(CH_2)_3SO_3^-$

Chapter 1

General Introduction

1.1 Surfactants

1.1.1 Introduction to Surfactants

As their name suggests, surfactants are surface-active agents, indicating that they adsorb strongly at various interfaces. e.g. air-water, oil-water, water-solid, oil-solid etc... to lower the interfacial surface energy (or tension). They most commonly take the form of a hydrocarbon portion and a polar/ionic portion as illustrated schematically by Figure 1.1

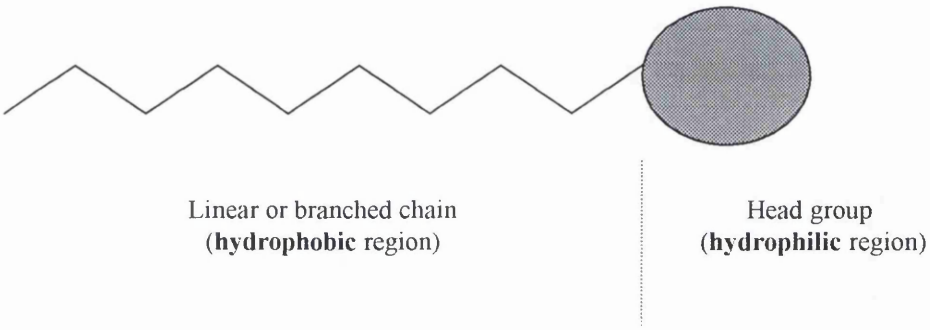


Figure 1.1 Schematic diagram of a surface-active agent

The hydrocarbon portion, which may be linear or branched, interacts only weakly with water molecules in an aqueous environment. Moreover, the strong interactions between water molecules, arising from dispersion forces and hydrogen bonding, act co-operatively to 'squeeze' the hydrocarbon out of water. In contrast, the polar/ionic portion of the molecule, often called the head group, interacts strongly with water *via* dipole-dipole or ion-dipole interactions and is thus solvated. Consequently the head group is referred to as the hydrophilic region. It is generally the balance between the hydrophobic and hydrophilic regions of the molecule that gives rise to the special properties that we associate with surfactants.[1]

1.1.2 Classification of Surfactants

Surfactants are usually classified according to the nature (or charge) on the head group. This head group may be non-ionic, anionic, cationic or zwitterionic. Table 1 shows a common example of each class of surfactant.[1]

Table 1.1 A table to illustrate each class of surfactant by example.

Class	Nature of head group	Example
Non-ionic	Uncharged	Alcohol Ethoxylates e.g. $C_{12}H_{25}[OCH_2CH_2]_6OH$
Anionic	Negatively charged	LinearAlkylbenzene Sulfonates e.g. $C_6H_5(CH_2)_{10}SO_3^-Na^+$
Cationic	Positively charged	Quaternary Ammonium Compounds e.g. $C_{12}H_{25}N^+(CH_3)_3Br^-$
Zwitterionic	Positively and negatively charged	Sulfobetaines e.g. $C_{12}H_{25}N^+(CH_3)_2(CH_2)_3SO_3^-$

1.1.3 Properties and Common Applications of Synthetic Surfactants

Surface-active chemicals are encountered in many spheres of our lives. They are used in ore flotation, as emulsifiers in oil drilling, as adjuvants in pesticide formulations and in textile processing. Furthermore, they are personal care products and, most importantly, detergents. The latter products account for the largest proportion of annual use of surfactants. Surfactants are far more versatile than the older detergents, soaps (such as sodium stearate) and now make up a billion dollar industry that encompasses the world. The global usage of surfactants (excluding soap) now stands at over 10 million tonnes.

The surfactant market can be subdivided into two main categories, namely cleaning products and process aids. Cleaning applications cover a variety of areas including household detergents, consumer products, personal care products and industrial/institutional cleaning. Process aids make use of one or more of a surfactant's surface-active properties. e.g. foaming, wetting and emulsification, to meet the performance requirements in a selected process. Examples of the second category include emulsifiers for producing the water-based polymer latex used in paints and inks and formulation aids for crop protection chemicals.[2,3]

In many cases potential uses are related to class. Firstly, it is important to note that all but the zwitterionic surfactants are used in detergents. Anionic surfactants are the most widely used class, particularly in detergent applications, where they display excellent cleaning and good foaming properties. Non-ionic surfactants are the other high volume class offering a broad spectrum of properties including wetting, dispersion, emulsification, antistatic properties, lubrication and detergency. The addition of a hydrophobic moiety to a conventional ethoxylate is often responsible for excellent low-foam and biodegradable non-ionics, useful in machine

dishwashing where foaming must be avoided. Cationic and zwitterionic surfactants are currently being produced on a very small scale in comparison with both anionics and non-ionics, mainly due to the fact that their environmental fate and effects have not been comprehensively investigated. Cationic surfactants are best known for their bactericidal properties and their application in disinfectants. However their antistatic and softening properties also make them ideal for use in fabric softeners. They are also finding use as deodorizers, as enzyme coatings for washing powders, as wood preservatives and as industrial lubricants. Further applications include corrosion protection, concrete production and leather production, and finally due to their ability to 'hydrophobe' a surface their use as emulsifiers (such as in bitumen for road repair) has also been appreciated. Zwitterionic surfactants are renowned for their mildness and have found application in toiletries, hair-care preparations and cosmetic formulations. They have also found industrial application in foam spray cleaners and other situations where a high stability foam profile is required.[3-6]

1.1.4 Critical Micelle Concentration

Due to the fact that surfactants displace water molecules at the interface of a biphasic system the number of hydrogen bonds per interface area is decreased in their presence. This can be measured macroscopically as a decrease in the interfacial tension, which can be defined as the work required to increase the unit area of the interface.[7]

By plotting a graph of the air-water interfacial tension (or surface tension, γ_{AW}) versus surfactant concentration, C_W, at reasonably low surfactant concentrations, it would be observed that γ_{AW} and C_W are inversely proportional. However, a point will be reached where for an increase in C_w we no longer see a change in γ_{AW} (Figure 1.2). The inflection in the curve indicates that the air-water interface has become saturated with surfactant molecules, and any additional surfactant molecules are forced into the water phase, since they are unable to adsorb to the interface. Interactions between the hydrophobic tails and water molecules are obviously energetically unfavorable, and so to counteract this effect surfactants will tend to aggregate with each other, so that hydrophilic head groups remain solvated and the hydrocarbon moieties can help shield each other from the aqueous environment through the formation of hydrophobic microenvironments. These spontaneously formed aggregates are called *micelles* and this is why we refer to the intersection of the two linear parts of the surface tension plot as the *critical micelle concentration (CMC)*. The CMC is an important consideration when considering surfactant

systems, since at concentrations above the CMC, bioavailability of the compound fails to increase and this effect is paralleled by the lack of increase in measurable soluble concentration.[2,7]

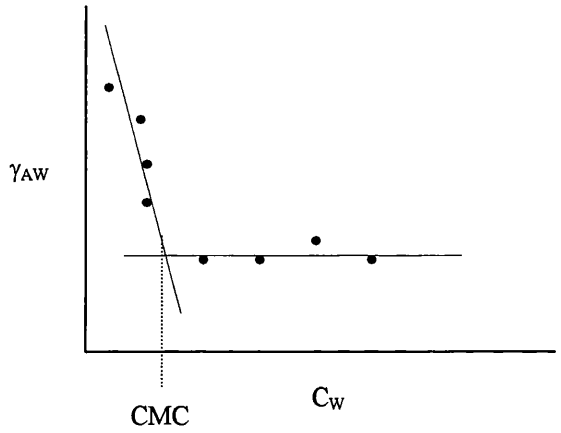


Figure 1.2 The value of concentration at the intersection of the two linear parts of a graph of surfactant concentration versus air-water interfacial tension is equivalent to the critical micelle concentration

1.1.4.1 Effect of Structure on CMC

Using a series of alkylsulfate surfactants, Rosen showed that within homologous groups of surfactants the CMC decreases with increasing hydrophobicity (or length) of the hydrocarbon chain. He generalised that regression of log CMC versus carbon chain length for all ionic surfactants should yield a plot with a slope of 0.3, but for non-ionic and zwitterionics the slope would become approximately 0.5.

Rosen also demonstrated the effect of a polar head group. He concluded that steric repulsion increased the CMC, from the observation that an increase in the number of oxyethylene groups in a non-ionic alcohol ethoxylate surfactant led to an increase in the CMC. He also compared CMCs of some non-ionic and ionic surfactants with approximately the same head group area, and found that the CMC of the non-ionic surfactant was the lowest. He concluded that electrostatic repulsion also increases the CMC.[8]

1.1.4.2 Influence of Environmental Factors on CMC

Rosen demonstrated that the presence of counterions had the potential to influence the CMC of ionic surfactants. He showed that an increase in ionic strength was accompanied by a decrease in the CMC, resulting from the shielding of charges (or decrease in electrostatic repulsion) of the equally charged head groups. The effect is significant for anionic surfactants with the ions Ca^{2+} and Mg^{2+}. Finally, Rosen highlighted the fact that CMC decreases with hardness in the order anionics > cationics > non-ionics.[2]

1.1.4.3 Prediction of CMC

Huibers *et al.*[9] described the use of a predictive model using topological predictors, which was found suitable for the estimation of log CMC for a wide variety of non-ionic surfactants. Attempts were also made by Zoeller and Blankschtein[10] to predict CMC based on molecular-thermodynamic theory. Muller[11] tried to predict the influence of temperature, Carale *et al.*[12] investigated the effects of ionic strength and Zoeller and Blankschtein studied the effects of mixing on the CMC.[9-12]

Within one class of surfactants, it is known that a relationship exists between CMC and LC_{50}. However, it is acknowledged that the relationship is often specific for the set of compounds studied and the water conditions employed, and so extrapolation to other surfactant classes and /or conditions is somewhat difficult.[2]

Roberts described a Quantitative Structure-Property Relationship (QSPR) for micellisation of anionic surfactants using hydrophobicity. It was found that the micellisation potential of anionic surfactants covering a diverse range of structures could be modeled by a combination of two parameters, π_H and L. These parameters are simple to calculate using a 'back of envelope' approach.

$$\mathbf{pCMC = 0.32\pi_H + 0.09L - 0.82} \qquad \textit{(Equation 1.1)}$$

pCMC = negative logarithm of the CMC; π_H = log P fragment value* for the hydrophobe (defined as the whole molecule minus the negatively charged fundamental fragment); L = length of the hydrophobe in C-C bond units

CMC data for a set of linear alkyl benzene sulfonate (LAS) isomers and homologues and a set of primary alkyl sulfonates (PAS) were found to fit the QSPR given by Equation 1.1, and the author intends to test this QSPR on further anionics, and also cationic surfactants.[13]

Parameters obtained from surface tension measurement, such as the surface area of the head group, have also been evaluated with regard to their relationship with CMC. Morall *et al.*[14] appreciated the advantage that they are not specific for just one class of surfactant.[2]

1.1.5 Surfactants as Water Pollutants

Surfactants are very common water pollutants, in part because they are primarily used as aqueous solutions, which are later introduced to the environment by consumers, through product usage and disposal (e.g. during the laundry wash cycle). Although surfactants are relatively non-toxic to mammals, they can be very toxic to aquatic organisms, such as fish and aquatic invertebrates,

* refer to Section 1.4.3.2

even in trace amounts. During the water treatment process, much of these surfactants are removed by biodegradation as well as by sorption to the biomass. Hence the amount discharged into surface waters, where it becomes a potential threat to our aquatic environment, is largely dependent on the efficiency of our water treatment process. Surfactant manufacturing companies, such as Unilever, are aware that aquatic life *will* be exposed to surfactants. Not even highly effective water treatment can prevent surfactants from reaching surface waters, therefore the question arises whether a given surfactant poses a risk to aquatic organisms. One aspect of that question is whether that surfactant is intrinsically toxic. Another aspect is whether it has the potential to bioaccumulate in organisms.[15,16]

1.2 Aquatic Toxicology

Aquatic toxicology is a relatively new and still evolving discipline, originating from concern for the safety, conservation and protection of aquatic environments. Scientists from various backgrounds, such as academic, industrial and government institutions, are making significant contributions to this multidisciplinary science and its many applications in managing toxic substances and complex waste.

The key principles of aquatic toxicology and some of the most recent advances in approach and understanding will be considered in this chapter.

1.2.1 The Aquatic Environment

The aquatic environment is highly complex and diverse. It includes several distinct ecosystem types – freshwater streams, lakes, ponds and rivers; estuaries; and marine coastal and deep ocean waters – that have many different biotic and abiotic components and unique characteristics. The *biotic* or living components consist of many combinations of plants, animals, and microorganisms that inhabit specific ecological niches in each ecosystem. The *abiotic* or non-living components include the physical environment (e.g. water, substrate-sediment, and suspended particulate material) within the boundaries of the ecosystem. Each aquatic ecosystem is thus a dynamic product of complex interactions of living and non-living components, with both constant and changing features in time and space.[17]

1.2.2 Aquatic Toxicity

Aquatic toxicity is a relative property reflecting a chemical's potential to have a harmful effect on a living aquatic organism. It is a function of the concentration and composition/properties of the

chemical to which the organism is exposed (i.e. externally and internally) and the duration of exposure.[17]

1.2.2.1 Toxicant

A toxicant is an agent that can produce an adverse response/effect in a biological system, seriously damaging its structure or function or causing death. The adverse response may be defined in terms of a measurement that is outside the 'normal' range for healthy organisms. A toxicant or foreign substance (i.e. xenobiotic) may be introduced deliberately or accidentally into the aquatic ecosystem, impairing the quality of the water and making it unfavourable for aquatic life.[17]

1.2.2.2 Introduction of Toxicant to Aquatic Ecosystem

Toxicants enter aquatic ecosystems from (1) *non-point* (diffuse) sources, such as agricultural run-off from land, contaminated ground water and bottom sediments, urban run-off, dredged sediment disposal, and atmospheric fallout, and (2) *point* (discrete, localised) sources such as discharges (effluents) from processing and manufacturing plants, hazardous waste disposal sites and municipal wastewater treatment plants. A chemical or oil spill on or near the surface of a body of water is also considered a point source. The effects of contamination on the aquatic environment from diffuse sources may be detrimental but are usually less obvious than those from point sources. This is because there is often no adjacent uncontaminated area with which comparisons can be made. Point sources can be characterised because the amount and location of the discharge can be measured more accurately than with non-point sources, which cover large areas or are a composite of many point sources.[17,18]

1.2.2.3 Direct and Indirect Toxicity

Toxicity can be divided into the broad categories direct and indirect. *Direct* toxicity results from a toxicant acting more or less directly at sites of action in and/or on organisms, causing internal biochemical changes. In contrast, *indirec*t toxicity occurs as a result of the influence of changes in the chemical, physical and/or biological environment (e.g. changes in quantity and/or quality of food organisms). Laboratory-oriented aquatic toxicity studies tend to focus on the examination of direct toxic effects, and field-oriented studies more often include the consideration of indirect effects and changes in communities and populations.[17]

1.2.2.4 Immediate and Delayed Toxic Effects

Toxic effects may be manifested *immediately* during exposure or directly following termination of exposure to a chemical. In many cases they may be *delayed* until some time after the exposure. Reaction time is determined by the properties of the chemical, its mode of toxic action and the ability of the organism to metabolise or biotransform the chemical. Chemicals that are susceptible to biotransformation have a short half-life in the organisms and are therefore excreted rapidly. As a result, these types of chemicals are not expected to produce delayed effects.[17]

1.2.2.5 Reversible and Irreversible Toxic Effects

Some toxic effects may be *reversible* by normal repair mechanisms, such as regeneration of damaged or lost tissue and recovery from narcosis. In many cases effects become reversible only if the organism can escape the toxic medium to a toxicant-free environment. Serious damage to an organism may be *irreversible*, eventually resulting in death. Reversibility or irreversibility of chemical effects can be highlighted in the laboratory. It involves transferring organisms from a toxic medium to a medium free of toxicant and this is known as a 'recovery study'.[17]

1.2.2.6 Categories of Toxic Effect

Two general categories of toxic effect can be distinguished. *Acute* toxic effects are those that occur rapidly as a result of short-term exposure to a chemical.[17] A large dose of toxicant of short duration is usually lethal (i.e. causing death by direct action). In fish and other aquatic organisms, effects that occur within a few hours, days or weeks are considered acute.

Chronic toxic effects may occur when a chemical produces detrimental effects as a result of a single exposure, but more often they are a consequence of repeated or long-term exposure to low levels of persistent chemicals, alone or in combination. There is often a relatively long latent period for the expression of these effects, especially if the exposure concentration is very low. Chronic toxicity may be *lethal* (causing death) or *sublethal* (below the level that directly causes death). An example of a lethal chronic effect is failure of the chronically exposed organisms to produce *via*ble offspring. Most common sublethal effects are behavioural (e.g. swimming), physiological (e.g. growth), biochemical and historical changes. Sublethal effects have the potential to lead indirectly to mortality. For example, a behavioural change, such as the inability to swim, may diminish the ability to find food or escape a predator, ultimately leading to death. Other sublethal effects may have little or no effect on the organism because they are rapidly reversible or simply cease with time.[17,19.]

1.2.2.7 Concentration-Response Relationships

In a supposedly homogeneous population, there may in fact be tangible differences among individual organisms e.g. not all organisms would respond quantitatively if challenged by the same concentration of toxicant or other potentially toxic stress. The effects of an exposure will be varied, as the result of biological variation, reflecting the genetic make-up, health and condition of the individual organisms. Therefore, in the determination of the toxicity of a xenobiotic, the object is to *estimate* as precisely as possible the range of concentrations that produce some selected, readily observable and quantifiable response in groups of the same test species (of the same age) under controlled laboratory conditions.[17]

The concentration-response relationship is a graded relationship between the concentration of the xenobiotic to which the organisms are exposed and the severity of the response educed. e.g. mortality. If the distribution of mortality versus concentration (dose) is drawn so that the cumulative mortality is plotted at each concentration a sigmoid dose-response curve similar to that given by Figure 1.3 is observed. Two parameters of this curve are used to describe it: (1) concentration (dose) that results in 50% of the measured effect and (2) the slope of the linear part of the curve that passes through the midpoint. Both parameters are necessary to accurately describe the relationship between chemical concentration and effect.[18]

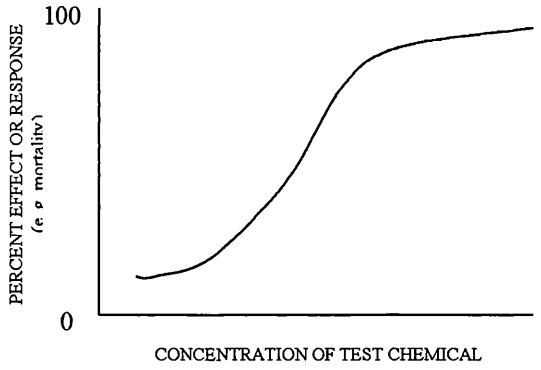


Figure 1.3 Typical form of the concentration-response curve.

1.2.2.8 Extrinsic Factors Affecting Toxicity

Extrinsic environmental factors have the potential to modify the acute toxic effect of pollutants. For example, variation in water qualities such as temperature, pH, hardness and dissolved oxygen concentration can make conditions unfavourable for aquatic life.

Temperature is an important consideration since it influences the metabolism and behaviour of organisms, which may then affect their response to the exposure of a pollutant. On

the other hand, it may alter the physical and/or chemical state of the pollutant and also the metabolic activation to toxic species. In general, it has been found that toxicity increases with temperature, as is the case for a variety of metal pollutants. However, there are many exceptions to the direct relationship. For example, the time to death of rainbow trout exposed to phenol was found to increase as temperature increased, but the LC_{50} decreased. Phenol acts as a toxicant by causing paralysis and cardiovascular congestion, leading to suffocation. The internal concentration of phenol is influenced by the relative rates of absorption and detoxification, both of which are directly proportional to temperature, but it is considered that temperature influences the rate of detoxification to a greater extent than the rate of absorption, at least at lower temperatures. Therefore it is likely that phenol accumulates to higher levels at low temperatures, accounting for the greater toxicity in a cold environment.[19]

The toxicity of many acids and bases is greatly affected by *pH*. For example, hydrogen cyanide is especially toxic in the molecular form, so that any variation in pH that reduces the degree of dissociation will increase the toxicity of the solution. It has also been reported that the toxicity of ammonia is affected by pH.

The chemical speciation of some metals is markedly affected by pH. Metal 'species' can be grouped into three phases, namely an aqueous phase (free ions and dissolved complexes), a solid phase (particles and colloids) and a biological phase (incorporated into cells or adsorbed on to biological surfaces). Generally, the ionic form of a metal is the most toxic. Campbell and Stokes[20] described two contrasting responses of an organism to metal toxicity with a decrease in pH. 'If there is little change in speciation and metal binding is weak at the biological surface, a decrease in pH will decrease toxicity due to competition for binding sites from hydrogen ions' and 'where there is a marked effect on speciation and strong binding of the metal at the biological surface a decrease in pH will increase metal availability'. Zinc and copper are examples of metals that highlight the first response and lead, the second.

Hardness is another important water quality that may have a marked effect on toxicity. As a general rule, the toxicity of pollutants is greater in soft waters. This is true for the metals lead, mercury, copper and zinc. Low calcium concentration in waters enhances the toxicity of metals to fish because the permeability of the gill membrane (to metals) is inversely proportional to the aqueous calcium concentration. Calcium competes with other metal cations for binding sites on the gill surface. The fish is thus protected in hard waters because the direct uptake of metal ions is reduced.[19] In contrast to the above statement that toxicity in generally greater in soft waters, it was found that LAS demonstrated increased toxicity to fish species and aquatic invertebrates in hard

water.[21,22] Trout and goldfish also appeared markedly more sensitive to sodium lauryl sulfate in hard water than in soft.[23]

The effects of many toxicants become more severe at reduced *dissolved oxygen* levels, simply because an increase in respiratory rate occurs, leading to an increase in the amount of toxicant which the organism becomes exposed to.

A number of chemicals have been found to be carcinogenic to fish, causing *neoplasms* or cancers. These include aflatoxins, azo- and nitroso-compounds, polynuclear aromatic hydrocarbons, polychlorinated biphenyls and a number of pesticides. Neoplasms, especially of the skin and liver, are found in wild fish, and in some areas incidence may be quite high.[19]

Harshbarger and Clark[24] divided neoplasms into two broad categories in relation to exposure to toxicants:

(a) not obviously associated with pollution – lesions in this category include haemic neutral pigment cell, connective tissue and gonadal neoplasms:

(b) associated with pollution – including epithelial neoplasms of the liver, pancreas and gastro-intestinal tract.

Bottom-living fish, for example, are most likely to be affected by group (b), since there is greater potential for contact with contaminated sediments. However, it is very difficult to secure a link between contaminants and neoplasms in the environment. Interspecific factors come in to play and confuse the relationship, as well as variables such as age, life history and feeding behaviour of the organism, and dose and duration of exposure to the toxicant. Furthermore, a polluted site is likely to have a number of compounds that could potentially induce neoplasms.

Damage to the immune system by pollutants is likely to be a major cause of the susceptibility of fish to the development of neoplasms.

Many compounds released into watercourses have the ability to undergo *transformation* under the environmental conditions that they experience. This can often lead to increased toxicity. For example, inorganic mercury (Hg^{2+}) can be transformed *via* bacteria and fungi to methyl and dimethyl mercury (CH_3Hg^+ and $(CH_3)_2Hg$ respectively). The methylation process may take place under aerobic or anaerobic conditions in the presence of different species of bacteria and/or fungi. It has been found that methyl mercury is 'exceptionally toxic' to many organisms.

Pesticides have also been shown to undergo transformations, but the effects tend to be minor. For example, aldrin can be converted to dieldrin in the environment, and both compounds are highly toxic.[19]

Effluents are often *mixtures* of chemicals that are potential poisons. If two or more poisons are present in an effluent they may exert a combined effect on an organism which is *additive*. An

additive effect occurs when the combined effect of two chemicals is equal to the sum of the effects of the individual chemicals applied alone. Alternatively, they may interfere with one another (*antagonism*), or the overall effect on the organism may be greater than if the two chemicals were acting alone (*synergism*).[17,19]

An example of an *additive* interaction is the combined toxicity of zinc and cadmium to fish. Calcium is *antagonistic* to lead, zinc and aluminium, while the toxicity of copper and the anionic surfactant, LAS, to rainbow trout is reported to be *synergistic.*[19]

1.2.2.9 Intrinsic Factors Affecting Toxicity

Species differ in their susceptibility to chemicals. This may be attributed to differences in accessibility, with certain species effectively excluding a toxic medium for short periods of time. In addition, rates and patterns of *metabolism* and excretion can greatly affect susceptibility. Differences in susceptibility to toxic agents among fish of different strains can also result from genetic factors.

Dietary factors also influence toxicity, by producing changes in body composition, physiological and biochemical functions, and nutritional status of the organism.

Developmental stage of the organism is also significant. Immature or young neonatal organisms often appear to be more susceptible to toxicants than the adult organism. This may be due to differences in the degree of development of detoxification mechanisms between the young and adult organism, but the difference in body size is also an important consideration. However, embryos may be less sensitive (more resistant) than adults because, at particular stages, they may have protective or impermeable membranes. Other considerations include previous experience of *stressors* and/or *disease parasites.*[17]

Some populations of organisms may develop the ability to *tolerate* the effects of a chemical, enabling them to survive in highly polluted environments. This is usually the result of previous exposure earlier in the life cycle. They can achieve this in one of two ways – either by functioning normally at high toxic loadings or by successfully removing pollutants by metabolism or detoxification. The specific mechanisms of tolerance to pollution are extremely complex, usually involving several metabolic systems, but species tend to develop tolerance to a particular pollutant in different ways.[19]

Chemicals with low solubility in water usually have an affinity for fatty tissues and thus can be stored and concentrated in tissues with high lipid content. Such hydrophobic chemicals may persist in water and demonstrate cumulative toxicity to organisms.

Bioconcentration is the process by which chemicals from water enter organisms through gills or epithelial tissue, and are accumulated. *Bioaccumulation* is a broader term and includes not only bioconcentration but also accumulation of chemicals through consumption of food. *Biomagnification* refers to the total process, including bioconcentration and bioaccumulation, by which tissue concentrations of accumulated chemicals increase as the chemical passes through several trophic levels. Organochlorine pesticides have been shown to biomagnify along the food chain, but biomagnification is rarely observed with metals, mercury being an exception.[17]

Bioconcentration occurs with many toxic pollutants, very high levels being accumulated in organisms from very low levels in water. The rate of accumulation of pollutants will depend on both external and internal factors to the organism. The concentration of the pollutant in water is clearly important, and many species carry higher loadings of pollutants when living in contaminated waters. e.g. metal concentrations in algae are significantly correlated with concentrations in water. Temperature influences absorption, detoxification and excretion rates of pollutants, but not necessarily to the same extent, so that the overall bioconcentration may vary with temperature.

Internal factors that influence bioconcentration include physiological condition. The concentration of lipophilic organochlorine compounds in different species of fish is directly related to the fat content of the fish. Periods of fast growth and/or higher metabolic rates are associated with the reduction in the level of contamination. Age, sex and the presence of competing pollutants in the water may also influence accumulation rates.[18]

The accumulation of a toxicant is a function of both uptake and elimination, and a generalised curve is shown in Figure 1.4. Assuming uptake is due solely to chemical diffusion, the process will continue until the internal level is equal to the level in the environment (point X, Fig. 1.4). Most pollutants can, however, be eliminated from the body and this is an active biochemical and physiological process, which cannot be described in simple diffusion terms.[19]

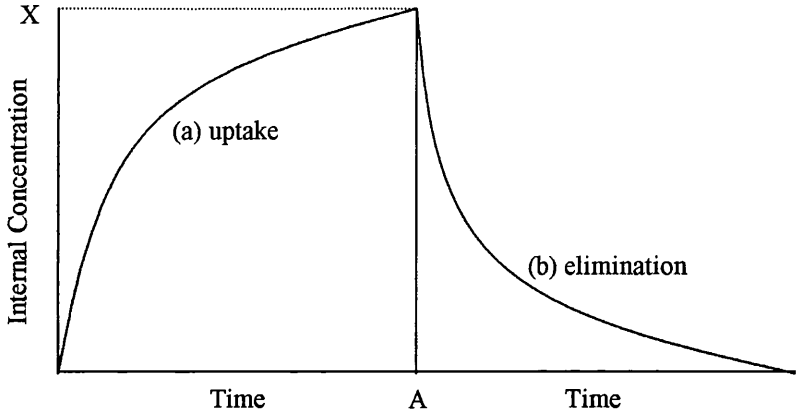


Figure 1.4 A generalised curve for the uptake and elimination of a pollutant. The external concentrations are reduced to zero at time A.

1.2.3 Aquatic Toxicity Testing

Aquatic toxicity tests are used to detect and evaluate the potential toxicological effects of chemicals on the organisms in our aquatic environment.

1.2.3.1 Criteria for Toxicity Testing

The criteria used to determine an appropriate standard test procedure have been established. Some of the following points have been taken into consideration:

- The test should be widely accepted by the scientific community.
- The tests should be standardised and carried out in compliance with defined protocols, so that reproducible results can be obtained from laboratory to laboratory.
- The data set should incorporate a range of concentrations to which a test species is subjected with realistic durations of exposure. The test result should also be quantifiable through graphical interpolation and/or statistical analysis.
- The test should be economical and easily conducted.
- The test should be as sensitive and as realistic as possible, and the test should have some field predictive capability for similar organisms.[17]

1.2.3.2 Standard Methods

Over the years a variety of test methods have been standardised. These protocols are available from the American Society for Testing and Materials (ASTM), the Organisation for Economic Cooperation and Development (OECD), and the National Toxicology Programme (NTP), and are available as the United States Environmental Protection Agency publications (the *Federal*

Register) and often from the researchers that developed the standard methodology. The protocols are frequently updated and amended versions made available.[19]

1.2.3.3 Good Laboratory Practices

Good Laboratory Practice (GLP) promotes the development of quality test data that is considered reliable and reproducible from laboratory to laboratory. GLP is concerned with the conditions under which laboratory studies, such as toxicity testing, are planned, conducted, monitored, recorded and reported. Adherence to GLP and quality assurance is critical with respect to regulatory and litigious matters. GLP regulations deal with all phases of aquatic toxicity testing including *personnel* – qualifications, responsibilities; *facilities* – maintenance/handling of test organism, chemical handling and storage; *equipment* – design, maintenance and calibration; *laboratory operations* – standard operating procedures; *protocols* – test methods, data collection and handling; *reports* and *disqualification of testing facilities.* Failure to meet the major provisions of the GLP requirements may result in a ruling by the Food and Drug Administration (FDA) or U.S. Environmental Protection Agency (EPA).[17]

1.2.3.4 General Approaches

Tests on single species in the laboratory (with a limited number of variables), and multispecies in model or natural ecosystems (*in situ*) are all important for evaluating the potential impact of a chemical on the aquatic environment.

Single species tests are designed to examine the response of a few individuals within a species. Most single species tests are conducted in the laboratory. This approach to testing can provide information on the concentration of chemicals and duration of exposure giving rise to mortality, changes in growth, reproduction etc. of organisms within species. These tests are simple to conduct, and can be standardised and replicated. Cause-and-effect relationships can be easily established due to the degree of control over laboratory conditions, however, results can rarely be used to assess the chemical impact above this level of organisation. Single species laboratory tests generally do not account for the adaptive ability of natural populations of organisms in the environment, and bioavailability will be different in laboratory tests due to the use of 'standard' laboratory waters. These waters do not typically contain realistic environmental concentrations of dissolved and particulate organics.[17]

Despite these disadvantages most aquatic toxicity tests to date have been laboratory studies on single species and many of these tests have yielded results that correlate well with observed effects of chemicals under natural conditions e.g. surfactants and pesticides.

Multispecies tests consist of at least two or more interacting species and range widely in both size and complexity. They can be conducted using model ecosystems in the laboratory. These studies usually involve 'laboratory microcosms'. These are small-scale enclosures (plastic or glass) containing samples from the natural ecosystem (water, sediment, invertebrates and plants). Their advantage is that effects beyond the level of a single species can be identified, providing information more predictive of ecological consequences of the chemical's release. Environmental influences are controlled, and so cause-and-effect relationships are more easily analysed than in natural systems. In contrast, microcosms also have several limitations since they are only simple simulations of natural ecosystems e.g. the impact of the physical environment (i.e. temperature and seasonal changes) may be very different from that represented by a microcosm.

Tests are more realistically carried out in the field (natural ecosystem) e.g. pond, stream, lake or estuary. Some influences and interactions of biotic and abiotic components that are not present in laboratory studies can be identified in such studies, and they may be more useful for predicting fate and effect of chemicals on the environment.[18]

However there are limitations to these approaches. Environmental variables are unstable and many undefined, and thus cause-and-effect relationships are particularly difficult to establish and experiments are difficult to monitor and impossible to replicate. Standardisation of these tests is therefore uncommon.[17]

1.2.3.5 Test Organisms

One of the most crucial aspects of a toxicity test is the suitability and health of the test organisms. Some of the criteria for choosing a test species are highlighted below: -

- The test species should be widely available (abundant) for toxicity testing
- The species should be amenable to routine maintenance in the laboratory and techniques should be available for culturing and rearing them
- The genetic composition and history of the culture should be known
- The sensitivity of the test species should be representative of the particular class or phyla that the species represents
- Wherever possible, species that are ecologically important or representative of the ecosystem that may receive the impact should be studied.[17,18]

Traditionally, tests have been performed with a variety of freshwater and saltwater test species representing algae, fish and invertebrates. These species are recommended because they tend to fit

the criteria highlighted previously. Species for a particular test must be from the same source and they should not be collected in a manner that may provoke stress in the organisms e.g. electroshock. Culturing of species in the laboratory tends to eliminate these problems.[25]

In the aquatic arena, an interesting publication by Doherty[26] compared four test species for sensitivity to a variety of compounds. These test species were rainbow trout, blue gill sunfish, fathead minnow and the water flea, *Daphnia magna*. The results discussed were very interesting. A high level of correlation ($r > 88\%$) was reported among the four species in all combinations. Of course three of the species are teleost fish, but *Daphnia* also fitted the pattern. The exceptions regarding correlation were compounds that contained chromium. In these cases, *D. magna* were found to be more sensitive than the fish species.[18]

1.2.3.6 General Test Design

Although details of a protocol may differ, the general test design is similar in each case. An aquatic toxicity test requires careful control of water qualities such as pH, temperature, dissolved oxygen concentration and hardness. In addition to this there must be photoperiod control. Test organisms are normally exposed to various concentrations of a test compound in water solutions. The effect (e.g. mortality) is then evaluated from a comparison between chemically exposed organisms and untreated organisms (as experimental controls). Controls must run concurrently to ensure that the effects observed are associated with or attributed to exposure to that concentration of test compound. There are three basic types of control: -

- *Untreated (negative) water control* – a group of organisms exposed to the same dilution water (in absence of test compound) and the same procedures and conditions related to that specific test. It is used to determine the inherent background effects such as health of test organisms and quality of test media, and provides a baseline for test results
- *Organic solvent or carrier control* – a group of organisms exposed to the same dilution solvent (in absence of test compound). This type of control is used in cases where the test compound is insoluble or poorly soluble in the normal dilution water and so the stock solutions are prepared *via* dilution in a solvent/carrier
- *Reference (positive) control* – a group of organisms exposed to a reference compound that is known to produce a defined effect. It is used to determine the health and sensitivity of organisms, to compare the relative toxicities of chemicals by using the control as an internal standard and to perform inter-laboratory calibrations

Wherever possible, chemical analysis/analytical support should be employed to measure the actual concentrations to which the test organisms are exposed.[17]

1.2.3.7 Exposure Systems

In a *static* test, the test solution is not replaced during the test. This methodology has the advantage of being simple and cost-effective. The quantity of test solution and the toxic waste generation is small; hence no special equipment is required for these tests. Since oxygen and toxicant concentrations generally decrease as a test progresses, this methodology is generally applied to short-term tests using smaller organisms.

Recirculating methodology is an attempt to maintain the water quality of the test solution without changing the concentration of toxicant. A filter is used to remove metabolic waste or a form of aeration may be used to maintain dissolved oxygen concentration at a specified level. A disadvantage linked with this type of methodology is the uncertainty that the water treatment methods employed do not alter the concentration of xenobiotic.

Another option is a *renewal* test. In this exposure scenario the test solutions of varying concentrations are replaced after a specified time period by a new test solution. This method has the advantage of replacing the solutions containing toxicant so that metabolic waste can be removed and dissolved oxygen concentrations can be returned to target levels. One disadvantage is that test organisms have to be transferred to the new media, potentially inflicting stress on the test organisms.

The best method for ensuring a precise exposure and water quality is the use of *flow-through* methodology. Continuous-flow methodology usually involves the application of peristaltic pumps, flow meters and mixing chambers, but is rarely used. More commonly an intermittent flow using a proportional diluter to mix the stock solution with diluent to obtain the required test concentrations is employed. The special equipment required is generally expensive.[18]

1.2.3.8 Classification of Toxicity Tests

There are five major classes of toxicity test; acute, chronic, short-term sublethal, early life stage and bioaccumulation tests.

Acute toxicity tests cover a relatively short period of an organism's life span. They may be *time-dependent* or *time-independent.* A time-dependent test has a predetermined length of time. For example, a test may be conducted to estimate the 96 h LC_{50} of fish (such as rainbow trout) or the 48 h EC_{50} of an invertebrate (such as *Daphnia magna).* In contrast, a time-independent test has a duration that is not predetermined. Exposure of the test organisms continues until the toxic

response manifested has ceased or economic or other practical considerations dictate that the test be terminated.

Chronic toxicity tests last for a significant portion of an organism's life expectancy. In a *full* chronic toxicity test, the test organism is exposed to (at least five concentrations of) the test compound for an entire reproductive cycle. Exposure is generally initiated with an egg or zygote and progresses through development and hatching of the embryo, growth and development of the young organism, attainment of sexual maturity, followed by reproduction to produce a second-generation organism. The duration of a chronic toxicity test will therefore vary depending on the species investigated. For example, in the case of *Daphnia magna* the test would take place over a period of approximately 21 days compared with the test for the fathead minnow, *Pimephales promelas*, which progresses over 275–300 days.[17]

The short-term sublethal test was designed to focus on the most sensitive stages of the life cycle of an organism. The test was introduced by the EPA to evaluate the toxicity of effluents to aquatic organisms. They commonly take place over a period of 7–9 days or less, and measure end-points such as change in growth, reproduction and survival.[1]

As the name suggests, the early life stage test involves continuous exposure of the early life stages (e.g. egg, embryo, larva and fry) of aquatic organisms to various concentrations of a chemical for 1-2 months, depending on the species. Although these tests do not provide total life cycle exposure and lack a full assessment of reproduction, they do include exposure during the sensitive life stages. They have been used to predict more accurately *maximum acceptable toxicant concentration* (MATC) values estimated in fish life cycle tests.[17,18]

Chemicals that are extremely hydrophobic can be stored and concentrated in tissues with high lipid content. These chemicals may persist in water and demonstrate cumulative toxicity to organisms as described in Section 1.2.2.9. These types of chemicals can be considered for bioaccumulation tests, which are designed to determine the *bioconcentration factor* (BCF). This parameter is defined as the ratio of the average concentration of a test chemical accumulated in the tissues of the test organisms under steady-state conditions to the average measured concentration in the water to which the organisms are exposed.[17]

1.2.3.9 End-points

To evaluate the safety of chemical substances, it is necessary to have a precise means of expressing the toxicity and a quantitative method of measuring it. Various end-points (otherwise known as criteria for effects) of toxicity may be used to compare chemically exposed organisms with unexposed organisms.

1.2.3.9.1 Commonly Used End-points

A measure of toxicity should be chosen that is unequivocal, clearly relevant, biologically significant, readily observable and hence is describable, measurable and reproducible.[17]

Measurement of lethality (mortality) can be useful for estimating the concentration and potency of a xenobiotic. Mortality and survival over a specific period of time are typical effect criteria in *acute* exposure tests. Data from lethality tests are *quantal*; which means there is an all or nothing response (organisms can only be dead or alive) during counts. An acute toxicity test is normally carried out to estimate the *median lethal concentration* (LC_{50}) of the chemical in the water to which test organisms are exposed. The LC_{50} is the concentration estimated to produce mortality in 50 % of a test population over a specific time period. The length of exposure is usually 24-96 h, depending on the species.[17,18]

When effects other than lethality are measured, the expression EC_{50} is used to denote the *median effective concentration*. Similarly, this is defined as the concentration of the chemical estimated to produce a *specific effect* in 50 % of a population of test species after a specified length of time (24-48 h). The *specified effect* can be behavioural or physiological and most common examples include immobility, a developmental abnormality/deformity, loss of equilibrium, failure to respond to an external stimulus or abnormal behaviour.

Whatever the effect or response chosen for measurement, the relationship between the degree of response and concentration almost always assumes a classic concentration-response form (Figure 1.3). Each point on the sigmoid curve represents an average cumulative response to the specific concentration, and each average has an associated error, which is the result of variable response due to individual organisms. The least variability is at the 50 % level of response and this therefore explains why the median lethal (or effective) concentration is usually recorded.[17,18]

End-points of chronic tests are described as *quantitative* (or graded); they are measured not in terms of incidence but in some unit of measured response (such as milligrams or centimetres). Growth (length and weight), number of normal embryos, morphological anomalies (e.g. double-headedness and deformed spines) and number of offspring are typical sublethal effect criteria applicable to chronic exposure. The MATC is often depicted from chronic toxicity data. The upper end of the MATC is represented by the *lowest observed effect concentration* (LOEC) and the lower end represented by the *highest no observed effect concentration* (NOEC). Two less frequently used terms to describe similar chronic effects are the *no observed effect level* (NOEL) and the *lowest observed effect level* (LOEL).[1]

1.2.3.9.2 Calculation of End-points

Obviously the LC_{50} or EC_{50} can be graphically determined, however there are a variety of statistical methods that may be employed to estimate these values for a given data set. For example, BMPDIN, a computer program modified by Stephan may be used to determine the median lethal/effective concentration of a test substance and the associated confidence limits.[27,28]

Ideally, the test should produce a monotonic response (i.e. an increasing effect with increasing concentration), but this does not always happen in practice due to the mechanism of toxicity of some test substances. This does not preclude statistically sound estimations of the LC_{50} or EC_{50} provided the appropriate method is chosen. The program generates three statistical analyses of the data using the moving average method, the probit method and the binomial method.[28]

Similarly, programs exist providing statistical means of calculating end-points of chronic toxicity tests. e.g. MATCs, such as LOEC and NOEC.[17]

1.2.4 Mechanisms of Action

The precise mechanism by which many environmental pollutants exert their toxicity remains to be elucidated, however four plausible mechanisms have been introduced.

A pollutant may exert its injurious effect on a specific organ by causing structural damage to its tissues. e.g. SO_2, a phytotoxin, even present in low concentrations, can injure epidermal and guard cells, leading to enhanced stomatal conductance and greater entry of the pollutant into the plant.

A toxicant may combine with a cell constituent to form a complex. This often leads to impaired function.

The basic function of enzymes is to increase the rate of reactions occurring within living cells. Optimum activity of many enzymes depends on the presence of co-factors. A toxic agent may inhibit an enzyme by inactivating its associated cofactor or by competing with it for the active site. Alternatively, the pollutant may deactivate the enzyme by combining with the active site. The activity of the enzyme may simply be inhibited by the presence of a toxic metabolite.

Finally, the presence of a pollutant in the living system may trigger the release of certain substances, injurious to cells.[1]

1.3 Quantitative Structure-Activity Relationships

1.3.1 Introduction

The ultimate aim of ecotoxicology is to monitor and predict the impact of potentially harmful chemicals on our natural ecosystems. Human activity is primarily responsible for a major increase in the levels of toxic chemicals in the environment and also an increase in the number of different commercial chemicals being used each year, with synthetic chemicals such as pesticides and plasticisers being of most concern. However, in addition, there has been increased contamination from combustion products, crude oils and algal toxins.

Experimentally, it would be impossible to assess the problem on a compound-by-compound, species-by-species, habitat-by-habitat basis, both from a labour and economical point of view, and thus greater pressure has been put on ecotoxicologists to use a 'tool' that would allow the prediction of toxicity without the need for profuse biological testing.[29-31]

All organic compounds are unique, but it is a known fact that many share common structural features and/or physicochemical properties that allow them to be grouped. By taking advantage of the fact that the toxicological behaviour of different compounds in a group vary in a systematic, hence predictable, manner in relation to these features/properties, mathematical models known as Structure-Activity Relationships (SARs) or Quantitative Structure-Activity Relationships (QSARs) can often be fitted. The possibility of SARs and QSARs have been appreciated for over a hundred years, but more recently their application has been extended to all aspects of the environmental behaviour of organic chemicals.[32]

1.3.2 History and Definition of SAR and QSAR

It has been recognised for thousands of years that different chemicals have different biological effects, but it was not until chemistry developed to the stage that chemical structures could be assigned to compounds that it became possible to speculate on the cause of such biological properties.

The ability to determine structure allowed early co-workers to establish SAR, which may be defined as simple observations used to describe the effect that a particular change in chemical structure has on biological activity. Such relationships are empirical and semi-quantitative, since the effect that arises from variation in structure is represented as an 'all or nothing' effect.

The earliest report of a quantitative relationship between activity and chemical structure was published by Crum, Brown and Fraser[33]. It was expressed as shown by Equation 1.2.

$$\phi = \mathbf{f}\ (\mathbf{C}) \qquad \textit{(Equation 1.2)}$$

ϕ = expression of biological response; C = measure of the 'constitution' of a compound.

It was suggested that a known change in the constitution of a compound, ΔC, would produce a change in its physiological action, $\Delta\phi$. By applying this principle to a sufficient number of compounds, it was possible to define what function ϕ is of C.

The obvious difference between SAR and QSAR, lies in the term quantitative. It is important to note that this term does not refer to the use of quantitative measures of biological response, but to the quantitative physicochemical descriptors employed in these models.[32]

The term QSAR includes all statistical methods by which biological activities (most often expressed as logarithms of molar activities) are related to physicochemical properties. There are situations where parabolic relationships exist between biological response and hydrophobicity. It is expected that this observation can be attributed to the existence of an optimum hydrophobicity. One interpretation postulates that many membranes may have to be transversed for compounds to get to the target site and compounds with greatest hydrophobicity will become localised in the membranes they initially encounter, therefore slowing their transit to the target site.[34] However, in its *simplest* form, a QSAR can be defined as a linear regression model that relates the variation in the biological activity (**Y**-matrix) within a homogenous series of compounds to their physicochemical and/or structural properties (**X**-matrix).

As already described, they are based on the principle (assumption) that each substructure of a molecule will contribute to its toxicity systematically, and so if considered reliable after significant validation, a given QSAR may be used in a screening process for new compounds.[35]

The major benefits of the use of QSARs lie in the fact that there is a dramatic reduction in the need for large scale synthesis and biological testing, in the early stages of any screening process. This means that the overall cost of labour and consumables is kept at minimum, time is not wasted on experimental evaluation of highly toxic compounds and there is a reduction in the total number of organisms destroyed by such tests.

Determination of the X-variables is usually faster, cheaper and easier (often by calculation methods), and so conventionally, QSARs are used to predict biological response from a known physicochemical property, but there has been recent interest in their use to establish an estimation of physicochemical properties from a known biological response.[35,36]

The most famous examples of early QSARs were proposed by Meyer and Overton in 1897-99.[37,38] They independently derived linear relationships between the narcotic action of organic compounds and their oil/water partition coefficients. However, probably the most striking instigation to modern QSAR development was Hansch and co-workers in the early 1960s, who derived of a single linear free energy equation, taking into account electronic (σ) and steric (E_s) substituent constants developed from studies of chemical reactivity, and a hydrophobicity substituent constant (π). The equation is generally presented in the form shown by Equation 1.3.

$$\mathbf{Log\ 1/C = -K_1\pi^2 + K_2\pi + K_3\sigma + K_4E_s + K_5} \qquad \textit{(Equation 1.3)}$$

C = concentration required to produce an effect; $K_{1\text{-}5}$ = constants

It was this equation that instigated the multiparameter regression approach (often termed 'the Hansch approach') which is in common use today, where a large and increasing range of descriptors are used covering the three main molecular features: electronic, steric and hydrophobic.[39]

1.3.3 Classification of Physicochemical Properties

An enormous range of physicochemical properties have been defined that fit within one of the three broad classes intimated previously. The key parameters and their associated symbols are highlighted in Table 1.2.

Table 1.2. A table to show the key physicochemical parameters used in QSAR development and the symbols commonly used to define them.[40]

Parameter	**Symbol**
Electronic Descriptors	
Hammett constants	σ, σ^-, σ^+
Taft's inductive (polar) constants	σ^*, σ_I
Swain and Lupton field parameter	F
Swain and Lupton resonance parameter	R
Ionisation constant	pK_a, ΔpK_a
Chemical shifts (^{13}C and ^{1}H)	δ
Theoretical Electronic Parameters	
Atomic net charge	q
Superdelocalisability	S^N, S^E, S^R
Energy of highest occupied molecular orbital	E_{HOMO}
Energy of lowest unoccupied molecular orbital	E_{LUMO}
Electrostatic potential	$V(r)$
Steric Descriptors	
Taft's steric parameter	E_s
Molar volume	MV
Molecular weight	MW
Van der Waals radius	r
Van der Waals volume	V_w
Molar refractivity	MR
Hydrophobic Parameters	
Octanol/Water Partition coefficient	log P
Substituent constant	π
Hydrophobic fragmental constant	f, f'
Distribution coefficient	log D
Apparent partition coefficient (fixed pH)	log P', log P_{app}
Capacity factor in HPLC	log k', log k'_w
Solubility Parameter	log S

1.3.3.1 Electronic Parameters

There are many types of intermolecular interaction forces including ion-ion, ion-dipole, ion-induced dipole, dipole-dipole, dipole-induced dipole, instantaneous dipole-induced dipole and hydrogen bonding, which depend on the electron distribution within a molecule or substituent and the ease with which this distribution can be modified.

Such interactions have been found to have a direct influence on the biological response that xenobiotics are able to induce at a receptor, and a wide range of parameters have been employed to model them. Electronic parameters may be broadly classified as *classical substituent parameters*, *whole molecule parameters* or *parameters derived from quantum chemistry*.

Classical substituent parameters include the *Hammett substituent constant, molar refractivity* and *hydrogen bonding*.[41]

Hammett[42] reasoned that the effect of a substituent, X, on the acid dissociation constant of benzoic acid could be given by Equation 1.4.

$$\rho\sigma_x = \log K_x - \log K_H \qquad \textit{(Equation 1.4)}$$

σ = Hammett substituent constant; ρ = series constant (unity for benzoic acid in water at 25°C); K_H = ionization constant of benzoic acid in water at 25°C; K_X = ionization constant of *meta*/*para* derivative in water at 25°C.

Ionisation of benzoic acid obviously results in the formation of a negatively charged carboxylate ion, and thus an electron-donating substituent will tend to disfavour this reaction giving a weaker acid (higher pKa) than the parent. In this case the value of σ will be negative. Similarly, an electron-withdrawing substituent tends to favour ionization giving a stronger acid (lower pKa) than the parent, with a positive value of σ. Hammett obtained σ values for *meta* and *para* substituents (which vary due to differing inductive and resonance contributions at the two positions), but he found that consistent values for *ortho* substituents could not be obtained, which he attributed to variable short-range effects such as steric hindrance and hydrogen bonding.[32]

This parameter has been widely applied in correlation studies, both physical-organic and biological. For example, from a biological point of view, Roberts described the correlation of σ_X for a series of nitrobenzenes with aquatic toxicity to the fathead minnow.[43]

It is perhaps also useful to note that ρ, the series constant, is a measure of the sensitivity of a reaction to the effects of substitution. The value of ρ will therefore vary with the type of solvent employed and conditions such as temperature. Traditionally these experiments were carried out in water at 25 °C. However σ constants are now determined in other solvents, with consequent

variation in values. This is one of the main reasons why there is now a confusing range of σ constants available for which some authors insist that average values should be taken.[44,45]

It is perhaps also useful to note that some parameters have since been derived from Hammett constants, including the Swain-Lupton parameters (F and R) which have become useful electronic field and resonance descriptors (Table 1.2).

Molar refractivity (MR) is the molar volume corrected by the refractive index. It represents the size and polarisability of a fragment or molecule. MR is conveniently defined by the Lorentz-Lorenz equation (Equation 1.5).

$$\mathbf{MR = (n^2 - 1 / n^2 + 1) . (M / d)} \qquad \textit{(Equation. 5)}$$

n= refractive index; M = relative molecular mass; d = density (normally of a liquid)

MR has the units of molar volume and has been shown to correlate well with electron polarisability, α_E, which can be described as the measure of the ability of electrons in an atom or molecule, to be polarised in an electric field. As a result MR has clearly proven to be an electronic as well as a steric property.[33,42] Grieco *et al.*[46] supported the use of MR as a parameter to model both electronic and steric effects, but reported that it is often a challenge to interpret the MR term in a QSAR correlation study. The authors make the general conclusion that when the parameter is used to model dispersive or semi-polar interactions a positive correlation will result. In contrast, one would associate a negative correlation with a steric effect.

Hansch and co-workers originally used MR extensively to model many aspects of biological activity. To a large extent, MR has now been displaced by the hydrophobicity parameter, log P, which is discussed in detail in Section 1.4. Despite the fact that both MR and log P are considered additive-constitutive properties and can readily be determined from fragment values and/or computer programs, log P appears to be the over-riding factor. This is thought to be partly due to the abundance of lipophilic sites in the living system, thus MR effects are outweighed by log P. This parameter, however, is now being looked upon with resurgent interest particularly for use in multi-parameter type approaches. However according to Livingstone other bulk parameters are also being considered such as corrected atomic radii.[32,41]

Hydrogen bonding is considered a significant interaction, but has had limited use so far in QSAR studies, which can be attributed to the fact that it is extremely difficult to establish quantitative measures of these effects. Hydrogen bonding effects are considered valuable to help

describe biological activity, since these types of interactions affect processes such as solubility, partitioning and receptor binding.

The most promising attempts at quantifying H-bonding interactions were reported by Abraham *et al.*[47] who devised hydrogen bond donor ability (K_α) and acceptor ability (K_β) scales for over 80 functional groups. The authors predicted that these types of H-bonding parameters would be better able to model the direct electronic influence of a substituent on a receptor than could the Hammett constant. However, these parameters have not fully been tested in QSAR studies so this remains only speculation at this stage.

Classical whole molecule parameters are widely used in QSAR when it is clear that substituent parameters should not be used to model an effect, for example, in a situation where a non-congeneric series is being tested. The most obvious of these parameters are *the acid dissociation constant, dipole moment, solvatochromic parameters and NMR chemical shifts.*[41]

The *acid dissociation constant*, pK_a, is used as a parameter to reflect electron-directing effects as well as controlling the ionisation of compounds, and examples exist for the use of pK_a in QSAR for both reasons. pK_a is closely related to the Hammett constant, since σ is actually derived from K_a. It is useful to note that in some cases it is not possible to identify the specific role of pK_a in the correlation. However Saarikoski *et al.*[45] were able to confirm the role of pK_a as controlling ionization in a QSAR to define toxicity of phenols to fish.[41]

It is useful to note that pK_a can be calculated for members of a congeneric series if the pKa for one member of this series is known along with appropriate ρ and σ values. However, Dearden stresses that experimental determinations are considered far more accurate and should always be used in preference to calculated values.

Dipole moments, μ, are often used in QSAR development in the pharmaceutical industry, since many drug-receptor and drug-solvent interactions require the presence of dipoles. Dipole moments are predominantly used as experimentally determined values in QSAR. However since μ is quite difficult to measure and has been shown to vary with the solvent employed, there has been great interest in the use of MO theory to calculate gas-phase dipole moments. e.g. using the MOPAC program.

Solvatochromic parameters are further possibilities. These are parameters that were originally derived from solvent effects on electronic spectra. Their use is based on the assumption that solubility in a given solvent is controlled by three factors, namely a volume term, a dipole term (π^*) and H-bond donor and acceptor terms (α and β respectively). These parameters were found to describe both aqueous solubility and octanol/water partition coefficients, and finally it

was discovered that solvatochromic parameters also correlate extremely well with many types of biological activity.[41,48] Solvatochromic parameters are available for several hundred compounds and Kamlet *et al.*[49] describe how it is possible to estimate them for other compounds, increasing the potential for use in QSAR.

The *NMR chemical shift* of a proton or other nucleus is a sensitive indicator of a local electronic effect within a molecule, and can be used to probe individual atomic or group interactions. This parameter is thought to deserve a wider application in QSAR correlation studies, however the use of chemical shifts has been somewhat limited, largely due to the fact that they must be experimentally determined.

Quantum chemical parameters are the last to consider. They are fundamental properties generated through quantum chemical calculations. They have been commonly used to provide a quantitative description of biological activity, which is hardly surprising since all properties of a molecule are related to its electron distribution and behaviour. The obvious reason that these types of parameter are dominant in QSAR analysis is that they can be obtained relatively easily *via* computational means in comparison with parameters that require experimental determination. The most extensively used quantum chemical parameters will be briefly discussed. These include *atomic charge, frontier electron density, HOMO and LUMO energies* and *superdelocalisability*. It is also important to note that another important quantum chemical parameter is the gas-phase dipole moment, briefly mentioned earlier in this section.

Atomic charge can be calculated by a number of MO methods including Complete Neglect of Differential Overlap (CNDO), Modified Neglect of Differential Overlap (MNDO) and Intermediate Neglect of Differential Overlap (INDO). For highest accuracy, it is recommended that *ab initio* calculations should be employed and charges should be obtained through calculation of electrostatic potential. However, if a charge is calculated for every atom in a molecule a data handling problem often emerges, and so the most usual way of using atomic charges in QSAR is to take the charge on a particular atom, that may be an atom common to the whole set of molecules being analysed or an atom from a substituent group. Other approaches include summing the modulus of atomic charges over whole, or part, of the molecule, to yield a measure of the polar interaction of which the molecule is capable or to use the difference of charge across a given bond (perhaps of a common functional group).

Frontier electron density should replace the atomic charge parameter in situations where very localised interactions are concerned, since this parameter is related to the outermost (frontier) orbital. Values relating to the Highest Occupied Molecular Orbital (HOMO) and Lowest Unoccupied Molecular Orbital (LUMO) can be calculated.

HOMO and *LUMO* are whole molecule properties. HOMO represents the ease with which an electron can be donated by the molecule and is therefore related to the ionisation potential. LUMO is then described as a measure of the ease with which a molecule will accept an electron. Both terms have proved useful for modeling intermolecular interactions and reactivities, and have been found particularly useful for modeling charge-transfer interactions.

Superdelocalisability may be defined as the sum of the frontier electron densities (HOMO/LUMO) on an atom divided by the HOMO and LUMO energies, and has also proven to be representative of reactivities or interactive abilities of molecules.[41]

1.3.3.2 Steric Parameters

The contributions of various electronic parameters have now been discussed in some detail. The second major factor that is often considered in QSAR development studies involves steric effects. Steric effects generally fit one of two broad classes, namely *bulk parameters* and *shape parameters*. The size and shape of molecules and substituents are important in several ways for the control of biological activity. A bulky substituent adjacent to a reactive site in a molecule can potentially hinder metabolism or even shield a polar group - hence reducing a molecule's affinity for water and/or increase its affinity for the lipid phase. To avoid steric constraints, twisting can often occur in aromatics with hindered substituents. For example, Dearden and O' Hara[50] found that 2-methyl substitution in 4-hydroxyacetanilide lowered the octanol/water partition coefficient. This observation was due to the acetamido group being forced out of the plane of the ring by the effects of 'steric twisting', causing loss of conjugation. Size and shape are particularly important in receptor binding. Xenobiotic and receptor should be complimentary in shape. Therefore, any change in conformation de*via*ting from the 'best fit' will result in weaker binding and concurrent loss of activity. In addition, size may be a barrier to the passage of molecules through aqueous channels in membranes. It is therefore hardly surprising that considerable effort has gone into the exploration of steric parameters for use in correlation studies.

Examples of commonly used *bulk* parameters used in QSAR include *molecular weight and volume*, *surface area* and terms related to *Van der Waals radius*.

Molecular weight is the simplest measure of size and this is why it has been used so abundantly in QSAR. *Molar volume*, defined as the molecular weight divided by density, is another frequently used parameter. Experimental determination is a tedious procedure and so approaches using calculation are often favored. One method involves summing the Van der Waals volume of the constituent atoms for either whole molecules or substituents. Another, and perhaps

more realistic method, involves the use of a computer program that rolls a water molecule over the molecular surface defined by the Van der Waals radii, to give a cavity surface volume.

As with molecular volume, the best approach to obtaining *molecular surface area* is to use a program that rolls a water molecule over the Van der Waals surface of an energy-minimized molecule, to yield a cavity surface area. This parameter has also had wide application in correlation analysis.[42]

Van der Waals radius is applicable only to substituents, but has little meaning to substituents that are obviously non-spherical. As a result, its use has been somewhat restricted. For studies involving reactivity of organic compounds, a radius-related steric parameter, E_s, was defined by Taft using the following equation: -

$$\mathbf{E_s = \log (K_x/K_H)_A} \qquad \textit{(Equation 1.6)}$$

The rate constants relate to the acid hydrolysis of esters, RCOOR'. This parameter was used extensively in early QSAR work. The range of substituents for which experimentally determined E_s values are available is very limited. However it is now possible to calculate values for substituents for which experimental values are not available.[34]

Another parameter that is related to molar volume and steric effects is MR, which has previously been mentioned in Section 1.3.3.1.

Many parameters encoding *shape* have also been used in QSARs. Of these, parameters that help to model the three-dimensional shape of a molecule, have recently found broadest scope for use in QSAR. Most recent 3-D treatment of steric effects, now being quite extensively applied to biological systems, is Comparative Molecular Field Analysis (CoMFA). This approach superimposes the conformations of molecules of interest and is an extension of ligand-based drug design. The technique was developed and first appreciated by Cramer *et al.*[51] to correlate biological activity with 3-D property fields.[41]

1.3.3.3 Hydrophobicity Parameters

Towards the end of the Nineteenth century, Berthelot and Jungfleisch[52] undertook the study of partitioning as a purely physicochemical phenomenon. They were first to prove that the ratio of concentrations of small solutes remained constant when distributed between water and an immiscible solvent, even when the solvent ratios varied widely.

Following this discovery, Meyer and Overton[37,38] showed that the narcotic action of many simple organic molecules was reflected by their oil/water partition coefficients, and it was this phenomenon that initiated the use of the hydrophobicity parameter in biological structure-activity relationships.

Since then, the octanol/water partition coefficient has dominated all other hydrophobicity parameters to become the key parameter used to define (or aid the definition of) many aspects of biological activity, such as aquatic toxicity. A comprehensive discussion of this parameter, including methods of experimental determination and calculation, comprises Section 1.4.[53]

1.3.4 QSARs for Aquatic Toxicity

QSARs are currently being applied in many disciplines, with many pertaining to drug design and environmental risk assessment. For the purpose of this thesis those that allow the prediction of aquatic toxicity for use in risk assessment are considered.

It is a widely accepted phenomenon that hydrophobic chemicals inhibit almost all biochemical processes due to their ability to penetrate the lipid membrane. Donkin,[54] in addition to many other authors, reports that a measure of hydrophobicity, in the form of log of the octanol/water partition coefficient (log P), is suitable to use as the sole physicochemical parameter to define aquatic toxicity. A number of log P-based QSAR for aquatic toxicity now exist. For example, Unilever have developed such QSARs for the prediction of aquatic toxicity of both anionic and non-ionic surfactants. There is now also at least one example of such a QSAR for the cationic class of surfactants known as quaternary ammonium compounds, derived by Singh *et al.*[55] It has been acknowledged that log P-based QSARs for zwitterionic surfactants do not exist.

It is important, however, to acknowledge that even though a high correlation between log P and log ($1/EC_{50}$) may exist, this does not guarantee that there are no other physicochemical properties involved. It simply indicates that log P (or hydrophobicity) is probably the overriding factor. Other factors such as size of molecule are expected to play a role in partitioning into octanol (i.e. a cavity for the solute must be created) but energy cost is presumed to be similar for octanol and the biosystem – hence non-specific size effects would cancel.[34]

1.3.5 Modes of Toxic Action

Xenobiotics act by differing modes of action. *Mode of action* may be defined as a common set of physiological and behavioural signs that characterise a type of adverse biological response.[17] Extensive research in both the U.S. and in the Netherlands in the 1980s, led to the conclusion that the majority of industrial organic chemicals (excluding pesticides and pharmaceuticals) elicit their

acute toxic effects through a *narcosis* mechanism.[56] Narcosis can be defined as a reversible state of arrested activity of protoplasmic structures resulting from exposure to an appropriate xenobiotic. In the context of the intact organism, narcosis and general anaesthesia are commonly used interchangeably. Although widely described as a non-specific mode of action, the actual mechanism of narcosis and anaesthesia remains unknown, but is an active area of research. It is expected that due to the complexity of the cell membrane, simple mechanical perturbation by non-reactive compounds could have a significant effect on the life of a cell. It is a general belief that the non-specific perturbation of membrane structure is responsible for the toxicity of simple hydrophobic chemicals, but this is perhaps an oversimplification. On the other hand, the cell membrane is not thought to be a particularly ordered system, so it is questionable if such a small perturbation alone could account for the toxic action.[34] More likely theories come in the form of the 'critical-volume hypothesis' and the 'protein-binding hypothesis'. The critical-volume hypothesis suggests that changes in the lipid component of cell membranes caused by an increase in volume, due to toxicant dissolved in this phase, substantially modifies the characteristics of the membrane bringing on narcosis. The protein-binding hypothesis postulates that narcosis arises from the toxicant binding to receptor sites of specific dimensions that are located in hydrophobic regions of proteins found in cell membranes. The latter hypothesis is supported by Franks and Lieb.[56,57]

Narcosis is reversible by stopping exposure before death and allowing the organism to recover by eliminating the chemical. This mode of toxic action is obeyed by a diverse group of chemical structures, including chemically inert gases, aliphatic and aromatic hydrocarbons, chlorinated hydrocarbons, alcohols, ethers, ketones, aldehydes, weak acids and bases, and some aliphatic nitro compounds.[56]

Following additional research, it became apparent that there were subclasses of narcotics more potent than could be predicted by what we now refer to as baseline/general narcosis (described above), that could be classified by either acute potency and/or physiological and behavioural characteristics of the narcosis response. This mode of action we now refer to as *polar* narcosis. Polar narcosis was derived from a series experiments to determine the acute toxicity of a variety of phenolic compounds to the fathead minnow.[56,58]

Furthermore, it became obvious that some industrial chemicals were significantly more toxic than either general or polar narcosis predicted, and this was because they were capable of acting by other, more specific mechanisms. Many chemicals were found to be acting as respiratory uncouplers, AchE inhibitors, membrane irritants, CNS convulsants or respiratory blockers.[17,56]

1.4 The Octanol/Water Partition Coefficient

1. 4.1 History of the Octanol/Water Partition Coefficient

Hydrophobicity can be described as a measure of the relative affinity of a solute for the hydrophobic phase of a biphasic system compared with that for the aqueous phase. It was demonstrated by Meyer and Overton, at the turn of the century, that the oil/water partition coefficient of a compound was an appropriate parameter to use as a model of the penetration of a xenobiotic through lipid membranes.[60] It was later demonstrated by Hansch *et al.*[60], that octanol would provide a more suitable lipid phase than the oily fats used by Meyer, Overton and other researchers at this time. Octanol appeared to be more suitable than oily fats in partitioning work, since as well as being a good mimic of a lipid (possessing a polar head group and a hydrophobic chain), it also had the practical advantage that it was readily available, could be prepared in a highly pure form and had the ability to solvate a wide range of organic compounds.[61,62]

The octanol/water partition coefficient is one of the principal physicochemical properties used in the development of QSAR for the prediction of biological activity. However, as a general rule, it is appreciated that biological activity should be envisaged as the result of the interplay of various parameters. In contrast, it has become apparent that the octanol/water partition coefficient, is the key and commonly sole physicochemical property used to define aquatic toxicity of organic compounds.[62,63]

1. 4. 2 Definition and Interpretation of Octanol/Water Partition Coefficient

The octanol/water partition coefficient (P) is defined as the ratio of concentrations of solute between the two immiscible phases, n-octanol and water, at equilibrium.

$$P = \frac{\text{Concentration of Solute in Octanol Phase}}{\text{Concentration of Solute in Water Phase}} \qquad \textit{(Equation 1.7)}$$

P is therefore unitless. When a partition coefficient is measured, the more lipophilic phase is, by convention, the numerator. The term 'immiscibility' does not prelude these two phases having partial miscibility. Octanol saturated with water in fact contains 27% mole of water. Furthermore, it was pointed out by Nernst, that a true partition coefficient should relate to the same molecular species in each phase.[53,61]

Organic compounds with low P values ($<10^{-2}$) are considered relatively hydrophilic and thus have greatest affinity for the aqueous phase of a partitioning system. Conversely, compounds with high P values ($>10^{4}$) are very hydrophobic with great affinity for the octanol phase. Where

octanol is used to mimic the lipid membrane, we can conclude that compounds with greatest P values are potentially most toxic, since they are expected to penetrate the lipid membrane more readily.[64]

P is most commonly expressed as its logarithm to the base ten ($\log_{10}$ P, or simply log P) for convenience, since the range of P values for organic compounds in general is extremely wide (typically 10^{-4} to 10^{8}. i.e. a range of twelve orders of magnitude).[65,66]

1. 4. 3 Determination of Log P

Log P is a widely measured property by a variety of different methods in the laboratory, but the process can often be time consuming and/or expensive and can in some cases be experimentally difficult. For example, surfactants have the tendency to reside on the interface of the biphasic system, to solubilise octanol in water and water in octanol. As the interfacial tension between the octanol and water phases is decreased, the tendency to form emulsions is greatly enhanced. Surfactants also readily form co-micelles with octanol in water and inverse micelles, with an aqueous core in the octanol phase.[67,68]

For compounds, like surfactants, where the measurement of log P is difficult or in cases where compounds are not available to test, there is also the possibility of estimating this value by calculation. The calculation of log P is particularly useful at the screening level (before commercial development), since new or 'in-development' chemicals often have no measured values. In this case estimation by a calculation method is easier, faster and cheaper than experimental determination.[69]

1.4.3.1 Measurement of Log P

Despite the fact that log P is a widely measured property, there is often considerable variation among published log P values. It is generally considered that a range of 0.3 log P units is acceptable, but even this range does not show particularly good reproducibility. Aside from random errors, there are a number of factors that can affect the measurement of a partition coefficient and these should be appreciated to ensure good practice before any attempts to determine log P by experiment.

Accuracy of partition coefficient determination can be affected by temperature, lack of mutual phase saturation, pH, buffer type and concentration, phase miscibility, solute concentration, solute and solvent purity, solute stability, phase volume ratio, solute adsorption and failure to reach equlilibrium.[61,69]

1.4.3.1.1 Stir-Flask Method

This method is the most conventional method of log P determination along with the analogous 'shake-flask' equivalent. Stirring to reach equilibrium tends to reduce the formation of emulsions, and this is why this method has proved more useful in recent years. An octanol/water system is prepared using solvents of very high purity. The biphasic system is stirred in the absence of solute to achieve mutual saturation of the phases. The solute is then added and equilibrium partition is obtained with stirring. The experiment *should be* conducted at a temperature between 20 and 25 °C, and should not de*via*te beyond ± 1°C for the duration of the experiment. At the end of the experiment an aliquot is taken from one (or both) phases and solute concentration in one (or both) phases is determined by an appropriate analytical method.[61,69]

This method is particularly reliable when precautions are taken, but it is the most time consuming. It is expected that the optimal log P range is approximately –2.5 to 4.5, but this does also depend on the detection method employed to quantify the solute.[67]

1.4.3.1.2 Filter Probe Method

The filter probe is a device that protrudes into one phase of a partitioning system, and continuously draws off that phase through a filter, passes it through an on-line analytical instrument and then returns it to the partitioning vessel. The filter is usually either a Whatman No. 4 filter paper for hydrophilic solvents, or Millipore Type LC (10μm) for hydrophobic solvents. Analysis is usually by spectrophotometry. It is claimed that a phase volume ratio of up to 1000: 1 can be used and both phases can be sampled if present in sufficient volume.

The filter probe method has been found to give almost identical results to those of the stir-flask experiment for 42 compounds tested. However, this method has a few advantages over the stir-flask method. Equilibrium is reached quickly and the on-line sampling allows one to see when equilibrium has been reached, temperature control is more easily achieved by using a jacketed vessel (the analytical instrument should also be thermostatted), partition behaviour at different temperatures and pH-partition profiles can be investigated and finally, the method can be employed for the partitioning of unstable or complexed species.[61]

1.4.3.1.3 AKUFVE Method

Reinhardt and Ryberg[70] described the use of a rapid and continuous system to measure distribution ratios in solvent extractions, which they called the AKUFVE system. This system consists of a rapid sequence of continuous unit operations: mixing, separation and on-line analysis

- like the filter probe system. Mixing of the solute with the biphasic system is achieved in a thermostatted chamber. The biphasic mixture is then passed into a flow centrifuge of special design, which permits a complete and rapid separation. Finally, the separated phases flow through measuring cells where the concentration of the solute is determined continuously using the appropriate detection mechanism.[45]

Davis and Elson later applied this technique to the measurement of partition coefficients, including those of octanol/water. They found it useful for examining the effect of temperature or pH on partitioning behaviour and they also found that equilibrium is reached rapidly using this system. However, disadvantages include the unduly large amount of organic phase consumed per determination (~500ml), the overall expense and the fact that partial stripping down of the machine and cleaning between samples is required.[61]

1.4.3.1.4 Automatic Titration Method

The titration/pH-meter technique is recommended for ionisable compounds. For these substances the partition coefficient of the neutral form of the substance may lead to the wrong assessment of the environmental fate. e.g. the ionised species of an organic acid is generally adsorbed by sediments to a much lesser degree than the neutral form because of its high water solubility and low lipophilicity. The fate of polyprotic molecules may depend on the ambient pH. Therefore, a single value of log P is an incomplete description of the lipophilicity of ionisable compounds.

The technique consists of two linked potentiometric titrations. The pK_a of the test substance is determined by titration using potentiometric measurement of pH. At the completion of the first titration, octanol is added and a second titration is performed, returning the well-mixed two-phase system to the starting pH. The second titration yields an apparent pK_a, denoted by P_oK_a. As the substance partitions into the octanol rich phase, the pK_a values derived from the two titration curves are different. Using established equations, an estimation of the partition coefficient is obtained.[71-73]

1.4.3.1.5 Generator Column Method

Generator columns are columns containing Chromosorb, or a similar material, packed onto a column, onto which the hydrophobic solvent is coated. For the measurement of an octanol/water partition coefficient, the column is loaded by pulling an unsaturated solution of the solute of interest at a known concentration in water-saturated octanol through the column. The solute is then eluted with octanol-saturated water, and the eluent is analysed by an appropriate analytical procedure. A primary advantage of this design is that there is no risk of emulsion formation, since

one phase is brought into contact with the other by slow permeation. Other advantages include the fact that it is a closed system and that a result is achieved faster than with a stir/shake-flask procedure.

It is also useful to note that there is good agreement between log P determined by this method and the stir-flask technique. However, there are a few important disadvantages associated with the method. It is rather elaborate and the column becomes stripped of solute and a new column is required for each solute. Finally, the method can only be applied to lipophilic compounds and the optimal log P range for this method is approximately 2-7.[61,62,67]

1.4.3.1.6 High Performance Liquid Chromatography (HPLC)

HPLC is an analytical technique offering a number of methods by which log P can be determined, either indirectly or directly. These methods have a number of advantages over the conventional stir/shake-flask experiment. For example, this method is far less time consuming, there is no risk of emulsions being formed, there is precision in the determination of log P from –4 to 4, for example, (corresponding to 10,000 times more analyte in octanol than water, or *vice versa*) since a retention time is being measured, impurities in the sample do not affect retention times but can affect partitioning between octanol and water, and finally extremely small quantities are required to determine retention times compared with the relatively large quantities required for stir-flask experiments.[74,75]

The most widely used *indirect* HPLC method used for determining log P involves measuring the retention time of a given analyte on a C8 or C18 reverse-phase column and thus calculating the capacity factor (k′).

$$k' = (t_r - t_o)/t_o \qquad \textit{(Equation 1.8)}$$

t_r = retention time of analyte; t_o = dead time (time for unretained compound to pass through system).

Compounds are partitioned between a mobile solvent phase and a hydrocarbon stationary phase, and hence retained in proportion to their hydrocarbon-water partition coefficient, with water-soluble compounds eluting first, followed by oil-soluble ones. This allows the relationship between the retention index on a reverse-phase column and the octanol/water partition coefficient to be established.

The method requires the establishment of a relationship between log k′ and log P for a series of similar reference compounds, usually between 5 and 10, for which log P is known. Log P

of the analytes can then be determined from their retention indices, using a calibration plot. This method allows log P in the range of 0 and 6 to be estimated, but literature reports that by modifying the mobile phase this log P range can be extended to cover the region of 6-12.

There have been a number of publications confirming excellent correlation of log k′ $_{(C8\ or\ C18)}$ with log P. The greatest errors between log P determined by this method and reported log P values (determined by calculation or by conventional methods) are observed with polar chemicals that dissociate in water, and these compounds tend to elute more rapidly than expected. This is because dissociation of the polar groups is more significant than adsorption interactions. Therefore, this method is not applicable to strong acids and bases, metal complexes, substances that react with the eluent or surface-active agents.[75-79]

The *direct* HPLC method used to determine log P involves the measurement of retention indices (log k′) on an octanol-coated HPLC column.

Early investigations by Mirelees *et al.*[80], made use of a commercial Kieselguhr support that was thoroughly silanised and slurry packed into a column. They found that it was possible to coat it with water-saturated octanol and used octanol-saturated water as the mobile phase (an *in situ* coating method). They found that by varying the column length and flow rate, they could successfully measure log P in the range of -0.3 to 3.7. For the 17 compounds investigated, an excellent correlation between log k′ and log P was observed.

Miyake *et al.*[81] developed the use of Corasil I mixed with octanol as a slurry, packed in a PTFE tube (as the stationary phase). Their method also proved successful for the analysis of a wide variety of compounds, but they considered their column to have additional benefits. As well as an easier and more efficient preparation method, the column is also expected to be more stable.

More recently, a direct method incorporating a C8 (or C18) column coated with octanol has been considered. For example, Kaune *et al.*[82] described the use of a C18 octanol-coated column to determine log P for a series of common reference compounds and a series of 44 triazines and related degradation products. Once again an extremely good correlation between log k′ and log P was reported, and the authors found that using this method a maximum log P value of ~3 could be determined.

1.4.3.1.7 Thin Layer Chromatography (TLC)

The RP-TLC technique is considered the two dimensional equivalent of the corresponding HPLC method. It uses the same basic principles, but it is inferior to HPLC with regards to precision and

is therefore less commonly used to estimate partition coefficients. However, the TLC technique actually has a wider optimal log P range (0-12 compared with 0-6 for HPLC).

For each test and reference compound, the R_f (or retention factor) is first determined, which can be defined as the ratio of the distance moved by the substance to the distance moved by the solvent front. From this, R_m, a free energy based constant, can be calculated.

$$\mathbf{R_m = \log (1/R_f - 1)} \qquad \textit{(Equation 1.9)}$$

From the constant for the system, K, and the derived R_m values, the octanol/water partition coefficients of both the reference and test compounds can be deduced.

$$\mathbf{\log P = \log K + R_m} \qquad \textit{(Equation 1.10)}$$

For reference compounds for which log P is known, R_m is plotted against log P to give the calibration plot and log P for the unknown(s) is estimated by interpolation. It is recommended that the analytes are chromatographed on the same plate as the reference compounds. Duplicate measurements are, however, performed on different plates.[67-76]

1.4.3.1.8 Centrifugal Partition Chromatography (CPC)

Most recently, Centrifugal Partition Chromatography as been explored as a novel technique for the measurement of octanol/water partition coefficients. The method uses Counter-Current Chromatography (CCC) – a technique in which the stationary and mobile phases are both liquids. These liquids are poorly miscible, hence octanol and water may be employed. A centrifugal force maintains the stationary phase, while the mobile phase is pumped through the system.[83]

Using this technique we can determine log P directly by Equation 1.11.

$$\mathbf{Log\ P_{CPC} = \log [(V_R - V_M)/V_S]} \qquad \textit{(Equation 1. 11)}$$

V_R = retention volume of analyte; V_M = retention volume of mobile phase; V_S = retention volume of stationary phase

Alternatively, we can simply determine the retention indices on a CCC system, and using carefully selected reference compounds establish a calibration plot of log k′ versus log P. For compounds of unknown log P, we can measure the retention indices in the same way and then use the established relationship to estimate log P.

1.4.3.1.9 Microemulsion Electrokinetic Chromatography (MEEKC)

Recently, MEEKC was assessed and developed as a screening tool for the indirect determination of log P. As with the RP-HPLC method, it is the capacity factor from MEEKC that is correlated with the octanol/water partition coefficient. MEEKC appears to have all the advantages of the indirect HPLC analogue to estimate lipophilicity including automation, small sample size, short analysis times and good reproducibility. However, MEEKC does not have the disadvantages of the HPLC technique such pH limitations and column degradation.

There have also been efforts to apply the technique to the analysis of small neutral organic molecules, anionic and cationic species that exist in the charged state at pH 7, and at extreme values of pH 1 and 12 allowing most compounds to be analysed in their neutral state.

The principle of the technique is based on capillary electrophoresis, where electrokinetic separations are performed along fused-silica capillaries filled with buffer solutions, across which a potential difference is applied. However, in MEEKC an oil-in-water type of microemulsion is prepared, which acts as a medium into which test solutes are dissolved before they are injected into the capillary. The most commonly used type of oil-in-water microemulsion which has proved to be a similar model to octanol/water partitioning, contains heptane as the lipophilic component, sodium dodecyl sulfate (SDS) as the surfactant and *n*-butanol as the co-surfactant, in the aqueous solution. These solutions are stable for long periods of time.[84]

Gluck *et al.*[85] showed that MEEKC was suitable for the determination of log P in the range of –1 to over 4 (covering over 5 orders of magnitude). They concluded that the technique was capable of giving accurate predictions over a log P range of 0.6 to 4.4 at pH 1 and between –1 and 4.4 at pH 12. The error was ± 0.4 log P units.

Poole *et al.*[86] found that log P in the range of 0.3 to 5.8 could be determined using this technique under the conditions investigated. They also found that the average de*via*tion in log P between literature and estimated values from MEEKC was ± 0.12 log P units.

Finally and most recently, Klotz *et al.*[87] have shown that the technique can be extended to the determination of log P in the range of –1 to 7 (six orders of magnitude). They concluded that the migration window and analysis time could be further optimised by varying the conditions of analysis to accommodate a broader or narrower range of compounds for separation. The elution window may be extended or decreased by adjusting the composition of the microemulsion, the most effective perhaps being the concentration of the surfactant component.

1.4.3.2 Calculation of Log P

In Section 1.4.3, the possibility of log P determination by calculation was discussed and situations where this is common practice were appreciated. It must be acknowledged that despite the effectiveness of current calculation methods shown by test calculations, calculated log P values may deviate quite considerably from the observed values for some compounds for many reasons.[70] We will now discuss in detail the most common approaches to log P calculation.

1.4.3.2.1 Leo and Hansch Fragment Method

Leo and Hansch's approach to the estimation of octanol/water partition coefficients is based on the assumption that log P has an additive-constitutive nature and hence can be calculated by conceptually breaking down the molecule into smaller fragments, summing the partial log P values and then applying factors to allow for the variation in how the fragments are bound together in the whole molecule. The calculations therefore use empirically derived atomic or group fragment constants (*f*) and structural factors (*F*), and are based on Equation 1.12.

$$\textbf{Log P} = \sum \textbf{fragments } (f) + \sum \textbf{factors } (F) \qquad \textit{(Equation 1.12)}$$

f = fragment values quantifying the contribution of individual molecular fragments to log P; ***F*** = factors quantifying the effects of how fragments are joined together on log P.

Since there are a large number of known *f* and *F* values this method of estimating log P is a fairly powerful one. The method only requires knowledge of the chemical structure, but for structurally complex molecules it is considered helpful to have a measured value of log P for a structurally similar compound to hand. This measured value can be modified by adding or subtracting, as required, the appropriate *f* or *F* values.

$$\textbf{log P}_{\textbf{(new chemical)}} = \textbf{log P}_{\textbf{(similar compound)}} \textbf{ +/- fragments } (f) \textbf{ +/- factors } (F) \qquad \textit{(Equation 1.13)}$$

If, for example, an estimate of log P is desired for the compound R-Br and a measured value is available for R-Cl, then

$$\textbf{log P}_{\textbf{(R-Br)}} = \textbf{log P}_{\textbf{(R-Cl)}} - f_{Cl} + f_{Br} \qquad \textit{(Equation 1.14)}$$

This approach is recommended whenever a reliable measured value of log P is available for a base compound that differs from the compound of interest by the substitution of only one or a small

number of fragments. However, if many different factors (*F* values) are likely to be involved in the two structures, this can often complicate matters and this approach should not be used.[89]

Fragment values are available for a few hundred atoms and groups. A fragment will have different f values depending on the type of structure it is bonded to. e.g. aliphatic or aromatic. As a more specific example, take the fundamental fragment NO_2, if attached to an aliphatic structure, f = -1.16; if attached to an aromatic structure, the hydrophobicity is increased and f = -0.03.[90]

To be able to define a fundamental fragment the two classes of carbon atoms must be appreciated, namely *isolated* carbon atoms (ICs) and *non-isolated* carbon atoms (NICs). ICs are carbon atoms that either have four single bonds (at least two of which are to non-heteroatoms) or else are multiply bonded to other carbon atoms. NIC are those that are multiply bonded to heteroatoms.

A *single-atom fundamental fragment* can be an IC or a hydrogen or heteroatom all of whose bonds are ICs.

Joining directly any of these types; a NIC, a hydrogen, a heteroatom, can form a multiple-atom fundamental fragment. A fundamental fragment is complete only when *all* its remaining bonds lead to ICs.

A *H-polar fragment* is one that can be expected to participate in hydrogen bonding, either as a donor or an acceptor. e.g. $-NH_2$, -OH, -O- and $-CO_2H$.

An *S-polar* (or σ polar) *fragment* is one that has strong electron withdrawing power, but little or no tendency to hydrogen bond. e.g. the halogens.

Anyone who frequently calculates log P manually will often save time if he/she prepares a list of derived fragment constants. *Multiple-atom derived fragments* are thus a combination of single-atom or multiple-atom fundamental fragments that are common or convenient to use. e. g. $-CH_3$: 3(H) + C = 3(0.23) + 0.20 = 0.89.

It is important to appreciate the reliance on *hydrophobic factors* (*F*) to maintain the integrity of the fundamental fragments that have been discussed in some detail. The most important factors affecting chains and interactions between various groups/fragments will be evaluated. Many different factors need to be considered when conducting log P calculations, but seldom are all used in the same calculation.[88,89]

The factors employed may be broadly classified into 'factors relating to chains' and 'factors accounting for interaction between polar groups'. Firstly, the *flexibility* (in terms of bond rotation and bending) of chains has been shown to have an effect on log P. For this reason a bond factor, F_b, is incorporated into the calculation. F_b has a negative sign (-0.12) since this flexibility

is thought to reduce the degree of order in the solvation shell, and is taken (n-1) times where *n* is the number of bonds. This bond applies only to C atoms - not to hydrogen or other atoms. Furthermore, we do not count bonds within a multi-atom polar fragment. The change in *F* from -0.12 for chains to –0.09 for rings fits with the concept that there is less flexibility within a ring. The change to –0.20, whenever 3 or more hydrophobic chains radiate from a tetrahedral atom, fits with the concept of hydrophobic shielding (i.e. less structured water required to enclose the same number of hydrocarbon units). *Unsaturation* has also been shown to have a noticeable effect on log P. In the calculation of log P for compounds containing a double or triple bond, log P is first calculated as if the molecule was saturated and then the appropriate $F_{=}$ (-0.55) or $F_{\equiv}$ (-1.42) constant is added. However, it is important to note that these values are subject to variation if the double or triple bond is conjugated to aryl groups. For example, if conjugated to one aryl, $F_{=}$ becomes –0.42, and if doubly conjugated to an aryl group it becomes 0.00. *Branching* is another way in which bonding can affect log P. It is advised that a branching factor should be added for each branch in a molecule. When applied to a hydrocarbon chain this branch factor is denoted by F_{c}Br and is equivalent to –0.13. However, when applied at a H-polar group a larger increase in water solubility results in a decrease in the branch factor (i. e. F_{g}Br = -0.22). For longer hydrocarbon chains on a branching (Y) carbon or nitrogen atom (such as for some tertiary alkyl amines) the reduction in log P is proportional to chain length and F_{c}Br is inadequate here. An enhanced bond factor is used (F_{bYN} = -0.20). It has also been shown that a position dependent branch factor (PDBF) may apply. Roberts stated that, PDBF = -1.44 log (1 + C_s), where C_s is the carbon number of shortest branch. It was originally derived from goldfish toxicity data for LAS homologues and isomers, and subsequently found to be applicable to aquatic toxicity of non-ionic surfactants.[23,89]

A number of factors have been derived to account for the interaction between polar groups such as halogen with halogen, halogen with H-polar fragment, H-polar to H-polar fragment and π electrons with polar fragments attached to aromatic systems. There will be a brief discussion of the factors associated with multiple halogenations and H-polar proximity.

In alkane structures, multiple halogenations on the same or adjacent carbon atoms result in a higher log P than simple additivity predicts. This seems reasonable if it is imagined that the localised dipole can be partially shielded from complimentary dipoles in the aqueous phase by the bulk of the halogen atoms themselves. For multiple halogenations on the same carbon atom (geminal substitution) the F_{mhGn} factor should be applied, whose value varies depending on the number of geminal halogens present (G_n). For G_n = 2, F_{mhGn} = 0.30; G_n = 3, F_{mhGn} = 0.53; G_n = 4, F_{mhGn} = 0.72. These 'gem-factors' must be added for each halogen. For multiple halogenations on

adjacent carbon atoms (vicinal substitution) the F_{mhVn} factor should be applied, where Vn is equivalent to the number of vicinal carbon atoms. It has been shown that $F_{mhVn} = 0.28$ (n-1), but it is important to note that this factor is thought to apply only when carbon atoms are connected by single bonds.

Likewise, there are rules for H-polar proximity. For two H-polar groups, A and B, $f_A + f_B$ should be added and this sum then multiplied by a factor which depends on the length of the spacer unit between H-polar groups. For rings this should be done for each connecting chain.[89,90]

Separation between A and B	In chains	In aliphatic rings
1C (F_{P1})	-0.42	-0.32
2C (F_{P2})	-0.26	-0.20
3C (F_{P3})	-0.10	---

Table 1.3. A table to show some H-polar proximity factors

Symbols are often used to indicate specific types of f and F values. A variety of subscripts and superscripts are now in common use in association with these values. Some rules regarding their use are outlined.[89]

- *Subscripts.*

1. A subscript may be used on f to indicate the structure (if complex).

e. g. f_{MVN} = -NH-CO-NH-

2. For *F*, the type is commonly indicated using a subscript. A number of examples are shown below:

F_b = bond factor; F_{cBr} = chain branch factor; F_{gBr} = group branch factor; $F_{=}$ = double bond factor; $F_{\equiv}$ = triple bond factor; F_{mhG} = multiple halogenation, geminal; F_{mhV} = multiple halogenation, vicinal; $F_{P\text{-}1}$ = proximity factor, H-polar fragment, one carbon separation; $F_{P\text{-}2}$ = proximity factor, H-polar fragment, two carbon separation.

3. The *underlining* of any symbol signifies its presence within a ring structure.

- *Superscripts.*

1. The absence of a superscript indicates an aliphatic structural attachment.
2. ϕ = attached to aromatic ring; if bivalent, the attachment is from left as written.

   e.g. Ar-CO_2

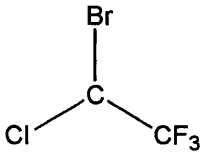
3. 1/ϕ = attached to aromatic ring; bivalent, the attachment is from right as written.
   e.g. $-CO_2-Ar$
4. ϕϕ = bivalent fragment, two aromatic attachments. e.g. Ar-NH-Ar
5. X = aromatic attachment; 1R = benzyl attachment.

The few examples that follow are taken from references 89 and 90, and are chosen to incorporate many of the interesting factors previously discussed. A comparison with observed values is made where possible.

(a)

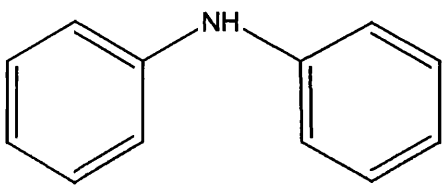
$$\log P = 4f_C + 10f_H + (3\text{-}1)F_b + F{\equiv}$$
$$= 4(0.20) + 10(0.23) + 2(-0.12) + -1.42$$
$$= \mathbf{1.44} \quad \text{(cf. observed = 1.46)}$$

(b)



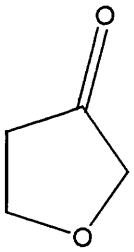
$$\log P = 2f_C + f_H + f_{Cl} + f_{Br} + 3f_F + (6\text{-}1)F_b + 2F_{mhG2}$$
$$+3F_{mhG3} + (5\text{-}1)F_{mhV5}$$
$$= 2(0.20) + 0.23 + 0.06 + 0.20 + 3(-0.38) +$$
$$5(-0.12) + 2(0.30) + 3(0.53) + 4(0.28)$$
$$= \mathbf{2.46} \quad \text{(cf. observed = 2.30)}$$

(c)



$$\log P = 2f_{C6H5} + f^{\phi\phi}{}_{NH} + (2\text{-}1)F_{bYN}$$
$$= 2(1.90) + -0.09 + -0.20$$
$$= \mathbf{3.51} \quad \text{(cf. observed = 3.22, 3.34, 3.50, 3.72)}$$

(d)



$$\begin{aligned}\log P &= 3f_{CH2} + f_{CO} + f_O + (5\text{-}1)F_b \\ &\quad + F_{P1}(f_{CO} + f_O) + F_{P2}(f_{CO} + f_O) \\ &= 3(0.66) + \text{-}1.90 + \text{-}1.82 + 4(\text{-}0.12) \\ &\quad + \text{-}0.32(\text{-}3.72) + \text{-}0.20(\text{-}3.72) \\ &= \mathbf{-0.17}\end{aligned}$$

1.4.3.2.2 ClogP Program

The Leo and Hansch method of calculating log P can be applied with computer assistance in the form of ClogP, originally devised by Chou and Jurs.[91] However, computerisation can often be less reliable since the breakdown of a molecular structure into fragments is a matter of choice. The approach is being improved with time by the incorporation of correction factors, as with the manual or 'back of envelope' approach.[34] In a comparative evaluation of commercial log P estimation software, Muller[92] concluded that the mean square error for approximately 1200 compounds indicated that ClogP generally gave good estimates of log P. ClogP was also found to give "markedly more log P estimates almost 'exact' in comparison to the observed values, than all other computerised methods investigated". This may be due to the inclusion of correction factors for more complex substructures. On the other hand, ClogP does tend to give rise to a number of discrepancies and in some cases log P cannot be calculated due to 'missing fragments'.[91]

1.4.3.2.3 KowWin Program

This program is also based on an atom/fragment contribution method requiring SMILES notation as structural input. This program was developed by the Syracause Research Cooperation (SRC), and Meylan and Howard[66] describe how it was originally developed with the prime objective being to overcome the problems associated with 'missing fragments', like those sometimes observed with the ClogP program. They were successful in making the program general enough to predict log P for almost any structure (not limited to a particular class of chemicals). Muller[92] also evaluated the use of the KowWin program for the estimation of log P for approximately 1200 compounds, as with ClogP.

He concluded, that from the mean square error, the program gave good estimates of log P. KowWin was found to display less discrepancies than ClogP, and log P for most of the compounds containing C, H, N, O, P, S and Hal could be calculated by this program. However, it was observed that ClogP was not always able to calculate log P for a lot of very similar compounds due to missing fragments.

1.4.4 The Distribution Coefficient

Nernst was responsible for putting the partition coefficient on a firmer thermodynamic basis by making it clear that the distribution ratio was indeed constant only if the molecular species were identical in the two solvent phases. This can have significance for the measurement of log P in situations where some interaction (such as dimerisation) occurs predominantly in one phase. However, it is probably of most significance if the solute to be partitioned contains an *ionisable* group. Since log P refers only to one species it will be necessary to suppress ionisation by the use of a suitable pH for the aqueous phase. The alternative is to measure the log of the distribution coefficient, log D, which involves the concentration of both ionised and non-ionised species and apply a correction factor based on the pKa values of the group(s) involved.[34] Another alternative is to use log D itself as the hydrophobic descriptor, although this may suffer from the disadvantage that it includes electronic information.[32]

1.5 References

1. Tadros, Th. F., *Surfactants*, Academic Press Inc. (1984).
2. Tolls, J. and Sijm, D. T. H. M, *Handbook of Property Estimation Methods, Environmental and Health Sciences*, edited by Robert S. Boethling and Donald Mackay, Lewis Publishers (2000), Chapter 17.
3. Karsa, D. R., *Chemistry and Industry*, **17** (1998) 685.
4. Boethling, R. S. and Lynch, D. G., Quaternary Ammonium Surfactants, *The Handbook of Environmental Chemistry*, Vol. 3 Part F, Anthropogenic Compounds – Detergents, vol. editor N. T. de Oude, Springer-Verlag (Berlin 1992).
5. Huber, L. H., *J. American Organic Chemistry Society*, **61** (1984) 377.
6. Boethling, R. S., *Water Research*, **18** (1984) 1061.
7. Attwood, D. and Florence, A. T., *Surfactant Systems. Their Chemistry, Pharmacy and Biology*, Chapman and Hall NY (1983), Chapter 3.
8. Rosen, M. J., *Surfactants and Interfacial Phenomena*, 2nd Edition, Wiley and Sons (1989).
9. Huibers, P. D. T., Lobanow, V. S., Katrisky, A. R., Shah, D. O. and Karelson, M., *Langmuir*, **12** (1996) 1462.
10. Zoeller, N. J. and Blankschtein, D., *Ind. Eng. Chem. Res.*, **34** (1995) 4150.
11. Muller, N., *Langmuir*, **9** (1993) 96.
12. Carale, T. R., Phan, Q. T. and Blankschtein, D., *Langmuir*, **10** (1994) 109.
13. Roberts, D. W., *Langmuir*, **18** (2002) 345.
14. Morall, S. W., Rosen, M. J., Zhu, Y. P., Versteeg, D. J. and Dyer, S. D., *Physical Chemical Descriptors for the Development of Aquatic Toxicity QSARs for Surfactants*, Proceedings of the 7th International Workshop on QSARs in Environmental Science (1997), SETAC Press, USA.
15. Lewis, M. A., *Water Research*, **25** (1991) 101.
16. Tolls, J., Haller, M., De Graaf, I., Thijssen, M. and Sijm, D., *Environ. Sci. Technol.*, **31** (1997) 3426.
17. Rand, G., *Fundamentals of Aquatic Toxicology: Effects, Environmental Fate and Risk Assessment*, 2nd edition, Taylor and Francis (1995).
18. Landis, W. G. and Yu, M-H., *Introduction to Environmental Toxicology. Impacts of Chemicals Upon Ecological Systems*, Lewis Publishers (CRC Press Inc.) (1995), Chapters 1-5.
19. Mason, C. F., *Biology of Fresh Water Pollution*, 3rd edition, Longman (1996).

20. Campbell, P. G. C. and Stokes, P. M., *Can. J. Fish Aquat. Sci.*, **42** (1985) 2034.
21. Holman, W. F. and Macek, K. J., *Transactions of the American Fish Society*, **109** (1980) 122.
22. Roberts, D. W., *Communicaciones presentatados a las Jornada del comite Espanol de la detergencia*, **20** (1989) 35.
23. Towell, T. W., Newsome, C. S. and Howes, D., *Water Research*, **8** (1974) 291.
24. Harshbarger, J. C. and Clarke, J. B., *Sci. Total Environ.*, 94 (1990) 1.
25. Ecotoxicology **SOP 019 11**, *Laboratory Culture of Daphnia magna* (Unilever Internal Document).
26. Doherty, F. G., *Environ. Sci. Technol.*, **17** (1983) 661.
27. Stephan, C. E., Methods for Calculating an LC50, Aquatic Toxicity and Hazard Evaluation, ASTM STP 634, F. L. Mayer and J. L. Hamelink, Eds., American Society for Testing and Materials, Philadelphia, PA (1977) 65.
28. Ecotoxicology **SOP 165 03**, *Use of Computer program Random. Bas.* (Unilever Internal Document).
29. Donkin, P., *Handbook of Ecotoxicology*, Vol. 2, edited by Peter Calow, Blackwell Scientific Publications, London (1994), Chapter 14.
30. Calow, P., *Handbook of Ecotoxicology*, Vol. 1, Blackwell Science Publications (1994).
31. Blum, D. J. W. and Speece, R. E., *Ecotoxicol. Environ. Safety*, **22** (1990) 198.
32. Livingstone, D., *Data Analysis for Chemists*, Oxford Science Publications (1995) Chapter 1.
33. Crum Brown, A. and Fraser, T., *Transactions of the Royal Society of Edingburgh*, **25** (1868-9) 151.
34. Hansch, C. and Leo, A., *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*, ACS Professional Reference Book, Amer. Chem. Soc. (1995).
35. Lindgren, F., Nouwen, J., Loonen, H., Worth, A., Hansen, B. and Karcher, W., Proceedings from Workshop 'Assessing low-dose effects of air-pollutants', Pa*via*, Italy (15-16th September, 1995).
36. Roberts, D. W. and Hodges, G., personal communication.
37. Meyer, H., *Arch. Exper. Parthol. Pharmakol.*, **42** (1899) 109.
38. Overton, E., *Z. Physik. Chem.*, **22** (1897) 189.
39. Hermens, J., *Quantitative Structure-Activity Relationships in Ecotoxicology* (March 1984).
40. *QSAR Parameter Table*: www.chem.awin.edu.au.

41. Dearden, J. C., Practical Applications in *Environmental Chemistry and Toxicology* (1990) 25.

42. Hammett, L. P., *J. Amer. Chem. Soc.*, **59** (1937) 96.

43. Roberts, D. W., Acute Lethal Toxicity Quantitative Structure-Activity Relationships for Electrophilies and ProElectrophiles: Mechanistic and Toxicokinetic Principles. In: *Aquatic Toxicology and Environmental Fate*: Eleventh volume, Eds. G. W. Suter II and M. A. Lewis, American Society for Testing and Materials, Philadelphia (1989) 490.

44. Hansch, C. and Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York (1979)

45. Saarikoski, J. and Viluksela, M., *Archives of Environmental Contamination and Toxicology,* **10** (1981) 747.

46. Greico, C., Silipo, C. and Vittoria, A., *Il Farm. Ed. Sci.*, **34** (1978) 433.

47. Abraham, M. H., Duce, P. P., Prior, D. V., Barrat, D. G., Morris, J. J. and Taylor, P. J., *J. Chem. Soc. Perkin Trans. II* (1989) 1355.

48. Taft, R. W., Abraham, M. H., Famini, G. R., Doherty, R. M., Abboud, J. L., Kamlet, M. J., *J. Pharm. Sci.*, **74** (1985) 807.

49. Kamlet, M. J., Abraham, D. J., Doherty, R. M., Taft, R. W., Abraham, M. H., *J. Pharm. Sci.*, **75** (1986) 350.

50. Dearden, J. C. and O' Hara, J. H., *Eur. J. Med. Chem.*, **13** (1978) 415.

51. Cramer, R. D., Patterson, D. E. and Bunce, J. D., J. *Amer. Chem. Soc.*, **110** (1988) 5959.

52. Berthelot, M. and Jungfleish, E., *Ann. Chim. Phys.*, **4** (1872) 26.

53. Hansch, C. and Leo, A., *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*, ACS Professional Reference Book, Amer. Chem. Soc. (1995), Chapter 4.

54. Donkin, P. and Widdows, J., *Rev. Aquat. Sci.*, **2** (1990) 375.

55. Singh, W. P., Lin, G. H. and Bockris. J. O'M., unpublished results.

56. Bradbury, S. P., *Toxicology Letters*, **79** (1995) 229.

57. Franks, N. P. and Lieb, W. R., *Environ. Health. Perspect.*, 87 (1990) 199.

58. Roberts, D. W. and Marshall, S., *SAR and QSAR in Environmental Research*, **4** (1995) 167.

59. Dearden, John C., *Environmental Health Perspectives*, **61** (1985) 203.

60. Hansch, C., Maloney, P. P., Fujita, T. and Muir. R. M., *Nature*, **194** (1962) 178.

61. Dearden, J. C. and Bresnen, G., *Quantitative Structure-Activity Relationships*, 7 (1988) 133.

62. Connell, D. W., *Handbook of Ecotoxicology*, Vol. 2, edited by Peter Calow, Blackwell Scientific Publications, London (1994), Chapter 13.

63. Rekker, R. F, *The Hydrophobic Fragment Constant*, Amsterdam, Elsevier (1977) Chapter 1.

64. Hennes Morgan, E. C. and de Oude, N. T., *Handbook of Ecotoxicology*, Vol. 2, edited by Peter Calow, Blackwell Scientific Publications, London (1994), Chapter 7.

65. Roberts, D. W., lecture on ' log P', University of Wales, Swansea (September 1999).

66. Meylan, W. M. and Howard, P., *Journal of Pharmaceutical Sciences*, **84** (1995) 83.

67. Sangster, J., *J. Phys. Chem. Ref. Data*, **18** (1989) 1111.

68. Morral, S. W., Herzog, R. R., Kloepper-Sams, P. and Rosen, M. J., *Proceedings of 4th World Surfactants Congress*, Barcelona (1996) 220.

69. *OECD Guideline for Testing of Chemicals*, **107**, Adopted 12th May 1981.

70. Reinhardt, H. and Ryberg, J., *Chem. and Ind.* (1970) 488.

71. *OECD Guideline for Testing of Chemicals – Proposal for a New Guideline*, Partition Coefficient (n-octanol/water): pH-Metric Method for Ionisable Compounds, draft document, June 1999.

72. Brandstrom, A., *Acta Chem. Scand.*, **17** (1963) 1218.

73. *OECD Guideline for Testing of Chemicals*, **112**

74. Veith, G. D., Austin, N. M. and Morris, R. T., *Water Research*, **13** (1979) 43.

75. Makovskaya, V., Dean, J. R., Tomlinson, W. R., Hitchen, S. M. and Comber, M., *Analytica Chimica Acta*, **315** (1995) 183.

76. Eadsforth, C. V. and Moser, P., *Chemosphere*, **12** (1983) 1459.

77. Carlson, Robert M., Carlson, Robert E. and Kopperman, Herbert L., *J. Chromatography*, **107** (1975) 219.

78. Konemann, R., Zelle, R., Busser, F. and Hammers, W. E., *J. Chromatography*, **178** (1979) 599.

79. *OECD Guideline for Testing of Chemicals*, Proposal for Updated Guideline **117**, Partition Coefficient (n-octanol/water), HPLC Method, draft document, July 1998.

80. Mirrlees, M. S., Moulton, S. J., Murphy, C. T. and Taylor P. J., *J. Med. Chem.*, **19** (1976) 615.

81. Miyake, K. and Terada, H., *J. Chromatography*, 157 (1978) 386.

82. Kaune, A., Bruggemann, R. and Kettrup, A., *J. Chromatography A*, 805 (1998) 119.

83. El Tayar, N., Tsai, R., Vallat, P., Altomare, C. and Testa, B., *J. Chromatography*, **556** (1991) 181.

84. Abraham, M. H., Treiner, C., Roses, M., Rafols, C. and Ishihama, T., *J. Chromatography A,* **752** (1996) 243.

85. Gluck, S. J., Benko, M. H., Hallberg, R. K. and Steele, K. P., *J. Chromatography A,* **744** (1996) 141.

86. Poole. S. K., Durham, D. and Kibbey, C., *J. Chromatography B*, **745** (2000) 117.

87. Klotz, W. L., Schure, M. R. and Foley, J. P., *J. Chromatography A*, **930** (2001) 145.

88. Lyman, Warren J., *Handbook of Chemical Property Estimation Methods for Environmental Behaviour of Organic Compounds*, edited by R. F, Reehl and D. H. Rosenblatt, Mc Graw Hill (1982), Chapter 1.

89. Hansch, C. and Leo, A., *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley and Sons NY (1979), Chapter 5.

90. Roberts, D. W., lecture on 'Calculation of log P by the Leo and Hansch Fragment Method', University of Wales, Swansea (February 2001).

91. Chou, J. T., and Jurs, P. C., *J. Med. Chem.*, **22** (1979) 476.

92. Muller, M., *Comparative Evaluation of Commercial log P_{ow} Estimation Software*, Poster Presentation, 7th International Workshop on QSARs in Environmental Sciences, Helsingor, Denmark (1996).

2.1 Introduction

Non-ionic, anionic and cationic surfactants are predominantly used in the manufacture of commercial detergents and some cosmetic formulations. In contrast, zwitterionic surfactants, renowned for their mildness, are often more suitable for use in cosmetic formulations and hair-care products. However, their utilisation is presently limited, with cocoamidopropylcarboxybetaine (a formulation containing predominantly the C12 and C14 amidopropylcarboxybetaines) being the only major commercially used zwitterionic surfactant mixture. The toxicological properties of many non-ionic, anionic and cationic surfactants have previously been investigated and QSARs for all three classes have been established to aid in the prediction of aquatic toxicity in cases where it has not been experimentally determined, and these are the main reasons why they are presently dominating the industry.

Toxicological properties for most zwitterionic surfactants remain unknown and QSARs for the prediction of aquatic toxicity remain undeveloped.

In connection with studies on the determination of log P–based QSARs for the aquatic toxicity of zwitterionic surfactants, a series of zwitterionic compounds (both short-chain surfactant and non-surfactant) have been synthesised of the quaternary alkylammonium sulfobetaine series. Here the commercial and industrial importance of this class, and similar classes, of zwitterionic surfactant will be discussed and a description the synthetic routes to the desired sulfobetaines will be given.

2.2 Betaines

The common name 'betaine' was proposed by Brühl in 1876 to cover all compounds with similar structure to the natural product, trimethylammonium acetate (Figure 2.1)[1] previously discovered by Scheibler and isolated from sugar beet juice.[2,3]

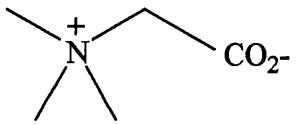
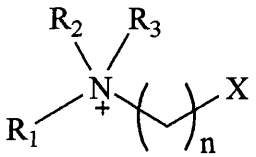


Figure 2.1 Example of a simple carboxybetaine

There are now many sub-classes of betaines, including carboxybetaines; sulfo-, sulfito- and sulfatobetaines; phosphinate-, phosphonate- phosphito- and phosphatobetaines; sulfonium

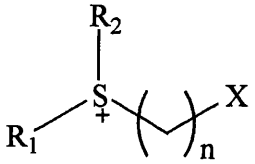
and phosphonium betaines or sulfobetaines.[4] A simple formula given by Figures 2.2, 2.3 and 2.4 may be used to describe what is meant by each of the above terms.



Carboxybetaines, X = CO_2^-; Sulfobetaines, X = SO_3^-;
Sulfitobetaines, X = OSO_2^-; Sulfatobetaines, X = OSO_3^-;
Phosphinatebetaines, X = PRO_2^-R; Phosphonatebetaines, X = PO_2RO^-;
Phosphatobetaines, X = ^-OPOHO; Phosphatobetaines, X = $^-OPO_2HO$.

R_2 and R_3 are commonly methyl groups;
R_1 and n is variable.

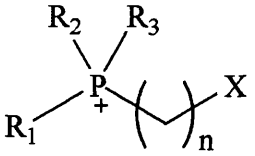
Figure 2.2 Simple formula for 8 sub-classes of betaine.



Sulfoniumcarboxybetaine, X=CO_2^-;
Sulfoniumsulfobetaine, X=SO_3^-

R_2 is commonly a methyl group;
R_1 and n is variable.

Figure 2.3 Simple formula for a Sulfoniumbetaine



Phosphoniumcarboxybetaine, X=CO_2^-;
Phosphoniumsulfobetaine, X=SO_3^-

R_2 and R_3 are commonly methyl groups;
R_1 and n is variable.

Figure 2.4 Simple formula for a Phosphoniumbetaine

2.3 Advantages, Properties and Applications of Betaines

Zwitterionic surfactants are amongst the most powerful 'detoxifiers', meaning agents that can reduce or eliminate irritation in surfactant blends, and their use in applications involving skin

contact preparations continues to increase. Not only is their use in personal care blooming, but they have also found use in hand dishwashing, both to promote mildness and improve overall performance.[4]

The fact that many physical properties of zwitterionic surfactants (such as betaines) have been investigated, reflects the commercial interest in these unique surfactants. These properties include wetting power, foaming, detergency, lime soap dispersion, CMC and antistatic properties.

The wetting power of a surfactant is generally determined by noting the time taken for a standard sample of unbleached cotton to sink. Obviously, the shorter the sinking time, the better the wetting. The best wetting is usually exhibited at an intermediate chain length since adsorption increases with increasing chain length, but mobility and solubility decrease.[4] Ernst evaluated the wetting power of a broad series of alkylammonium propanesulfonates (sulfobetaines) *via* the Draves Test-Synthron Tape Method. Criterion of a good wetting agent in this test is considered to be a sinking time of <25 seconds at 1 % concentration. He showed that the sulfobetaines, *N*-dodecyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate and *N*-tetradecyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate are particularly good wetting agents in a variety of different media (distilled and hard water, synthetic seawater, and acid and basic conditions).[4]

Another requirement of a surfactant is the capacity to produce foam. A common means of quantifying foaming power is achieved using Ross-Miles foam numbers. These have been determined for a series of ammonium propanesulfonates and these betaines were found to be good foamers in a variety of media including distilled water, hard water, acid and basic conditions and synthetic seawater. Hence foam appears to be unaffected by characteristics of the medium. However, of more interest is the synergistic property that exists between betaines and some other surfactants in mixtures. For example, Ernst noted the synergism between sodium lauryl sulfate and various ammonium propanesulfonates. Ross-Miles foam numbers greatly in excess of the individual components were achieved.[5,6]

As with foaming characteristics, it is the detergent properties of betaines in the presence of other surfactants that are of greatest interest. Sulfobetaines, however, were shown to provide improved detergency at lower wash temperatures. Ernst measured the detergency of many sulfobetaines. He did not rate them as excellent detergents, but he did appreciate the synergistic effect induced when these sulfobetaines were combined with soap flakes. Substantially higher detergent efficiencies were attained for combinations of sulfobetaines and soap flakes than for the sulfobetaines alone.[4] In 1973, Parris *et al.*[7] studied the detergency of several sulfobetaines on

cotton and cotton/polyester blends and found that the best detergent in the ammonium propanesulfonate (sulfobetaine) series appeared to be that derived from the hydrogenated tallow. When formulated with tallow soap and sodium silicate, performance approaching that of a commercial phosphate was obtained.[4]

The alkylpropanesulfobetaines were also found to exhibit the ability to disperse lime soaps formed in hard water[6] and in 1991 Ohme *et al.*[8] reported that sulfobetaines were better dispersants than alkylcarboxybetaines.[4] Parris *et al.*[7,9-11] reported lime soap dispersant requirements (LSDR) using standard methods for many sulfobetaines, sulfoamidobetaines and sulfatobetaines, and found that sulfoamidobetaines and sulfatobetaines are even better lime soap dispersants than sulfobetaines. Many betaines show excellent stability to calcium ions, as one would expect. This property was also studied using the Wilkes and Wickert method,[12] leading to the conclusion that, generally, sulfobetaines gave values greater than 1800 ppm of calcium carbonate, while the corresponding secondary amino compounds gave much lower values.[10]

Another important surfactant property requiring consideration is the antistatic property. Distler and Widder showed evidence of static charge reduction on synthetic fibres when treated with several ammonium propanesulfites. The excellent results obtained were comparable to many commercial antistatic products.[13] It is presumed that it is the zwitterionic structure of betaines that assists in dissipation of the static charge through the fibre.[5]

Any factor that increases hydrophilic repulsion relative to hydrophobic attraction will increase the CMC, whilst any factor which decreases hydrophilic repulsion relative to hydrophobic attraction will decrease the CMC. Therefore the close connection between the CMC and solubility is also appreciated. As with other classes of surfactant, increasing alkyl chain length (increasing hydrophobic interaction) of a zwitterionic surfactant also decreases the CMC. The decrease in the CMC is approximately one order of magnitude for each $(CH_2)_2$ unit added in the chain. An example is shown for the alkylcarboxybetaines in Figure 2.5.

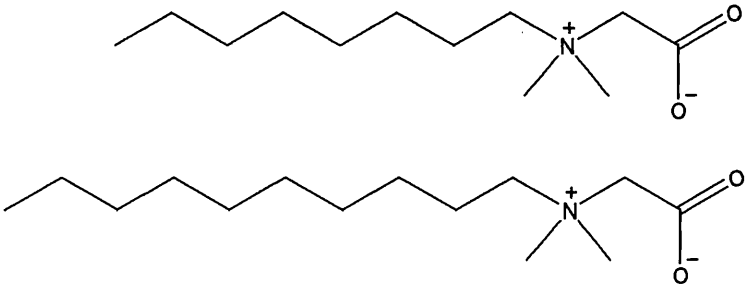


Figure 2.5 An example to show that increasing the chain length by two methylene units decreases the CMC roughly by an order of magnitude.

Similarly, introducing counterions decreases hydrophilic repulsion and leads to a decrease in the CMC. Also, the higher the charge density on the counter ion, the greater the increase in the hydrophilic repulsion. For example, inclusion of Mg^{2+} would be more effective (result in a greater decrease in CMC) than Na^{+}.[5]

The CMC is also markedly affected by the number of methylene groups between the head group charges of the zwitterionic moiety. Chevalier *et al.*[14] studied the CMC of alkylcarboxybetaines for different members of the series as a function of the number of methylene groups between the quaternary nitrogen and the carboxylate group. The CMC increases as the number of methylene groups increases from 1 to 3, reaching a maximum value and then decreases for further additions of the methylene unit. This phenomenon was similarly observed for phosphinatebetaines. The maximum value of the CMC where the number of methylene groups between the head group charges is equivalent to 3 indicates two opposing effects. As the number of methylene groups increases, the dipole moment of the head group increases making it more hydrophilic. However, an increase in the number of methylene groups between the charged moeties also increases the hydrophobic properties of the molecule. As a result there will be unequal increases in the hydrophilic repulsion and hydrophobic attraction according to the number of methylene groups present in the spacer unit. Where the number of methylene units is between 1 and 3, there is an overall increase in hydrophilic repulsion and hence an increase in the CMC, however, beyond 3, there is an overall increase in hydrophobic attraction and an accompanying decrease in the CMC.[4,14]

Another factor that can affect the CMC of a zwitterionic surfactant is pH. The relationship between pH and CMC is generally complicated. However, the CMC will be greatest when the majority of the surfactant is in the cationic form. Rosen and Zhu studied the effect of pH on the CMC of *N*-dodecyl-*N*-benzyl-*N*-methyl-ammonio-methanecarboxylate. At pH 5.85, 99.9 mole % of the surfactant was in the zwitterionic form and only below pH 2.8 did the concentration of the cationic form exceed that of the zwitterionic form. Therefore there is little change in the CMC until the solution reaches a sufficiently low pH.[4]

The six major properties of surfactants that govern their commercial, industrial, medicinal, pharmaceutical and biological applications have now been described. The wide range of applications associated with zwitterionic surfactants depending on these properties can now be appreciated. Another aspect of prime interest that drives many of their practical applications is the strong interaction of these molecules with other ionic surfactants (commonly anionic surfactants). It is important to note that betaines are commonly used in admixture with other

surfactant families thus largely improving their performance, mildness, foaming, viscosity and other surface-active properties.

Currently, the most significant commercial applications are in toiletries and personal care products, which together represent the largest area of consumption.[4,5] However, Lietz exhaustively investigated several betaines and found that, despite remarkable mildness, only a few of these (usually from the amidocarboxybetaine, sulfoamidobetaine and amidosulfohydroxybetaine series) were suitable as the sole surfactant component of shampoos for hair, and this was mainly attributed to the fact that they proved inadequate for excessively soiled hair.[15] Betaines were then found to form mixed micelles with anionic surfactants, maintaining mildness and providing many other advantageous properties in such personal care products. Hence today it is these blends have greatest commercial importance. Following Mannheimer's study on complexes of alcohol sulfates with zwitterionic surfactants, several commercial shampoos now take advantage of the low occular irritation (which can be further enhanced by the addition of a non-ionic surfactant) and high foaming power of these blends.[16-18] It appears that zwitterionic surfactants have an inhibitory effect on the irritation potential of alcohol sulfates, and it is the presence of the anionic surfactant that is responsible for the enhanced foaming power of the blend.[19]

'Fly-away' hair caused by friction that leads to a build-up of static charges on the hair's surface is commonly avoided by after-treatment with a conditioner or by simply washing with a conditioning shampoo. Cationic surfactants are used in conditioning compositions, but not so often in shampoos since they are not compatible with anionic surfactants. The presence of betaines in neutral shampoos was found to reduce the accumulation of charge on hair. Additionally, amine oxides as well as betaines, have been used to provide the desired conditioning effect in shampoos at low pH, in the isoelectric region of hair.[4]

Mixtures of alkylcarboxybetaines, amidobetaines or sulfobetaines with amine oxides were found suitable for use in shampoos to not only provide conditioning but also inhibit the growth of *pityrosporum ovale*.[20,21] Betaines have also been used as essential ingredients in high-foaming antidandruff shampoos.[22] Cocoamidopropylcarboxybetaine is commonly used in shampoos to increase the colour intensity of direct dyes deposited from shampoo. Hair treated with these formulations tends to exhibit a brighter colour and stronger shade than hair treated with betaine-absent shampoos.[23] Amidocarboxybetaines of this type are also useful in foam bath compositions because they are good thickeners and enhance foaming. Beh and James[24] found that

amidocarboxybetaines were the best stabilisers of foam, with alkylcarboxybetaines a close second.

Betaines are also used as perfume solubilisers in low content aftershave lotions[25] and many patents describe the advantages of using betaines in dentifrices for the prevention and removal of plaque. Importantly for the latter application, they were seen to improve foaming without inhibition of antibacterial activity.[26,27] Betaines are very efficient for use in deodorants for treatment of odorous air from toilets and cigarette smoke in enclosed areas[28,29] and many patents describe the use of sulfobetaines, and hydroxysulfobetaines, in particular, as components of liquid detergents, providing good skin texture after washing and having low irritation potential.[30,31] Betaines have been used in cosmetic formulations and are expected to help reduce the effects of ageing.[32] Their derivatives are often considered advantageous secondary ingredients in detergent formulations due to their lime soap dispersant properties and synergistic detergent effects with anionic surfactants.[4]

A mixture of a fatty amidopropyl-dimethylamine and an amidocarboxybetaine is a good fabric softener, which is just as effective as the quaternary salts commonly used, and has the advantage of compatibility with anionic surfactants. Blends of these two compounds with anionics are suitable for both cleaning and softening of fabrics at high or low temperatures.[33] Alkylcarboxybetaines have also been found useful for killing bilharzia-carrying cercaria, especially *schistosoma mansoni*, and can be added to laundry detergents to kill cercaria in water during washing.[34]

Finally, carboxy- and sulfobetaines are claimed to be useful dry-cleaning agents in chlorinated hydrocarbon solvents and sufficient water to remove both oily and water-soluble soil.[5]

The use of many zwitterionic surfactants in personal care and consumer products have been described in some detail. It is now convenient to briefly mention some of the most interesting industrial applications.

Many years ago, it was the textile industry that drove the development of betaine derivatives because they were found to be the most chemically stable classes of surfactants. They have since become efficient textile auxillaries. Their zwitterionic character and stability over such a wide pH range favours the use of both carboxybetaines and sulfobetaines as levelling agents, wetting agents in sulphuric acid carbonisation of wool, scouring agents, antistatic agents and softeners.[4] Riva and Cegarra[35] studied the effect of alkylcarboxybetaines and amidocarboxybetaines on the levelling of dyes on wool and Parra *et al.*[36] studied the interactions

of amidocarboxybetaines with anionic surfactants in imparting antishrinking properties to wool fibres. Betaines have been used as ingredients in printing pastes for use with polyester suede since they were seen to prevent bleeding and staining. The fact that betaines are accredited antistatic agents and softeners (as previously mentioned) is also of great importance to the textile industry. Betaines also improve the spinning properties of polyalkylene fibres and are known to improve the yield, strength and whiteness retention of acrylic textiles and rayon tyre cord. They can be used as antistatic agents in lubricant compositions than contain potassium or sodium alkylphosphate esters to avoid scum formation and charge generation. Alkylcarboxybetaines, in particular, have proved useful in leather-finishing, since they act to make leather more resistant to dry-cleaning.[4]

Betaines have also found applications as disperants and emulsifiers. For example, *N*-dodecyl-*N*,*N*-dimethyl-ammonio-methanecarboxylate has been used as a suspending agent in the aqueous polymerisation of vinyl chloride.[37] Many alkylcarboxybetaines have been used as emulsifiers in styrene polymerisation. If the latex is diluted in water and padded onto a pile of viscose rayon, acrylic or nylon carpet, there is a reduction in soiling.[38] Furthermore, betaines are useful in the paper industry. If incorporated into paper coating compositions, good printability and improved lustre are achieved.[39]

Betaines have also found applications in the paint and photographic industries. For example, in the paint industry, betaines provide paints with superior wet adhesion and freeze-thaw stability.[40] Also, they have found application in metallic pigment compositions, especially those containing aluminium. Metallic pigment compositions with intended incorporation into waterborne paints have been prepared by contacting the pigment with an organic phosphite and phosphobetaine to avoid the risk of explosion and to give the coating composition a good metallic appearance and a smooth surface, as well as good adhesion and water resistance properties.[41] In the photographic industry, betaines are useful in the preparation of photographic emulsions. *N*-octadecyl-*N*,*N*-dimethyl-ammonio-methanecarboxylate has been found suitable for use as an emulsifier for monomers (e.g. acrylonitrile) in the preparation of gelatin-compatible hydrosols. These photographic emulsions have the advantage that they are stable over a wide pH range.[42] Furthermore, sulfobetaines improve surface uniformity and reduce the coefficient of friction of gelatin-silver halide emulsions, and sulfobetaines derived from alkoxylated fatty alkylamines in silver halide emulsions improve coating properties and reduce static charge on the cellulose triacetate support.[43,44] *N*-octadecyl-*N*,*N*-dimethyl-ammonio-methanecarboxylate has been claimed to impart excellent antistatic properties in photosensitive silver halide emulsions and

substituted ethoxylated sulfobetaines added to photographic emulsions were seen to improve the fog value.[45,46]

Interestingly, betaines have been found to be useful in secondary and tertiary recovery of crude oil. It appears that alkylcarboxybetaines and sulfobetaines are the preferred compounds, but the use of other types have also been patented.[4] The possibility of enhanced petroleum recovery by flooding with a alkylcarboxybetaine or amidocarboxybetaine microemulsion has also been described.[47] Microemulsion fuel compositions have since been prepared having a high degree of phase stability, which is virtually independent of temperature, using cocoamidocarboxybetaine as the co-surfactant.[48] Sulfitobetaines with alkyl chains C12–14 are bactericides for sulfate-reducing bacteria and have applications in enhanced petroleum recovery and petrol refining.[49]

The use of oil-soluble alkylcarboxybetaines and sulfobetaines as fuel additives that inhibit the corrosion and clogging of pipelines and filters has also been appreciated.[50,51] The introduction of betaines to some liquid fertilisers was seen to almost double their storage life. Fire-fighting foams have also been developed where the formulation incorporates non-fluorinated betaines and finally, betaines have been used in the electroplating of copper, tin and zinc to obtain bright alloys.[52,54]

When the wide diversity of applications already encountered is considered, it is remarkable to learn that betaines have also found uses in medicine, pharmacy and biology.

At this stage it is important to stress that, on the whole, the antimicrobial effectiveness of betaines is not comparable to cationic surfactants.[5] However, certain betaine-derived compounds are active against specific strains or potentiate the effects of other compounds such as antibiotics. In an investigation by Distler and Widder, it was found that some sulfitobetaines were particularly efficient against *S. aureus*, *E. coli* and *A. niger*. Stearamidosulfitobetaine in particular was found to be efficient against *S. aureus* and *A. niger*, and dodecylamidosulfitobetaine against *E. coli*.[55] It was also found that the activity of chlorotetracycline on antibiotic resistant strains of *E. coli* was potentiated by dodecyl/tetradecyl carboxybetaine (as like the cationic surfactant – cetyl trimethylammonium chloride). Suling and O'Leary explained that in each case (with either the surface-active zwitterionic or cationic compound) the surfactant increases the uptake of the antibiotic, inhibiting protein synthesis in the *E. coli* cells.[56] Similarly, it has also become apparent that dodecyl/tetradecyl carboxybetaine (and cetyl trimethylammonium chloride) potentiate the activity of chlorotetracycline on antibiotic resistant strains of *Proteus mirabilis* and *Klebsiella pneumoniae*.[55]

Antimicrobial compositions containing a variety of betaine surfactants were considered useful for incorporation into odour-controlling body shampoos by Michaels and Kenny.[57,58]

Sulfobetaines were compared with non-ionic ethoxylated surfactants and sodium dodecylsulfate as emulsifiers of a triglyceride model system and extractors of proteins from 3T6 mouse fibroblast membranes. They were actually found to be superior to the non-ionic ethoxylated surfactants but inferior to sodium dodecylsulfate. However, sulfobetaines have the advantage that they do not denature either water-soluble or membrane proteins.[59]

The solubilising effects of alkylcarboxybetaines and sulfobetaines have been confirmed by several independent sources. For example, Allen and Humphries were able to dissociate milk fat globule membranes and those of red blood cells or rat liver cells with dodecyl/tetradecyl carboxybetaine, ATPase was solubilised from yeast plasma membrane with tetradecylcarboxybetaine and envelope proteins of *Bordetella pertussis* (whooping cough) were solubilised with retention of antigenic properties.[60,62]

Sulfobetaines have been found suitable for the separation and purification of antigens from biological systems for use as vaccines or diagnostic agents and mixtures of alkylcarboxybetaines and non-ionic surfactants in aqueous solutions have been used as therapeutic topical liquid compositions suitable for the treatment of burns and inflamed skin. Alkylcarboxybetaines also provide effective therapy for ulceration of the gastric mucosa, and when administered with salicylate-based drugs decrease or prevent stress–induced or salicylate-induced ulceration of the gastrointestinal tract.[63,65]

Parvin *et al.*[66] showed that during the isolation of mitochondria the loss of mitochondrial carnetine due to efflux may be prevented by the presence of 20 mM of *N*-octyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate. It was also found that sulfobetaines inhibit carnetine acetyl transferase and carnetine palmitoyl transferase at several times the concentrations required for translocase inhibition.

More recently there has been interest in zwitterionic compounds, especially sulfobetaines, as potential antihypercholesterolemic agents. These are suitable inhibitors of squalene synthase (SS), the enzyme that catalyses reductive dimerisation of two farnesyl diphosphates *via* presqualene diphosphate to form squalene. SS inhibitors, as alternatives to the statin class of HMG-CoA reductase inhibitors (which act as cholesterol lowering agents) are useful because these also inhibit of the rate-limiting step in the isoprenoid biosynthetic pathway with potential for depletion of essential non-steroid isoprenoid metabolites.

Spencer *et al.*[67] synthesised and evaluated a number of sulfobetaines as inhibitors of SS. Variation in size of the hydrophobic moiety was extensively explored and efforts were made to establish the optimum separation between the ionic centres. At best, *N*-phenylnonyl-*N,N*-dimethyl-4-ammonio-1-butanesulfonate gave a value of IC_{50} = 2 μM for SS inhibition and, despite a wide variety of modifications to this structure, SS inhibition could not be further improved. An inhibitor in the submicromolar range is required and it is expected that exploration of a different type of zwitterion is necessary if this appealing approach to SS inhibition is to provide a potential antihypercholesterolemic agent.

2.4 Sulfobetaines

Sulfobetaines are essentially trialkyl ammonium compounds similar to that shown by Figure 2.1 but they have the alkyl carboxylate group substituted by an alkyl sulfonate group (Figure 2.6). However, none of these simple compounds, would exhibit surface-active properties, due to the lack of hydrophobicity. Only when one methyl group is substituted by a hydrophobic chain in the range of 8–20 carbon atoms is a surfactant property assumed.[4]

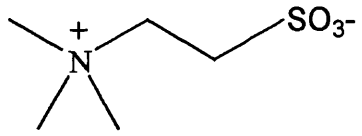
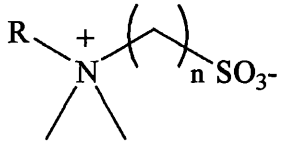


Figure 2.6 *Example of a simple sulfobetaine*

Sulfobetaines are more accurately named ammonium alkanesulfonates. The structure shown by Figure 2.7 illustrates the potential to vary the length of the spacer unit between the polar groups. The routes to twenty one sulfobetaines (both surfactant and non-surfactant) each belonging to one of three series, namely the ethanesulfonates (n = 2), the propanesulfonates (n = 3) or the butanesulfonates (n = 4) is described.



R=alkyl, aryl

Figure 2.7 *General formula to illustrate potential to vary the length of the spacer unit*

2.4.1 Literature Preparation of Sulfobetaines

2.4.1.1 Preparation of Ethanesulfonates

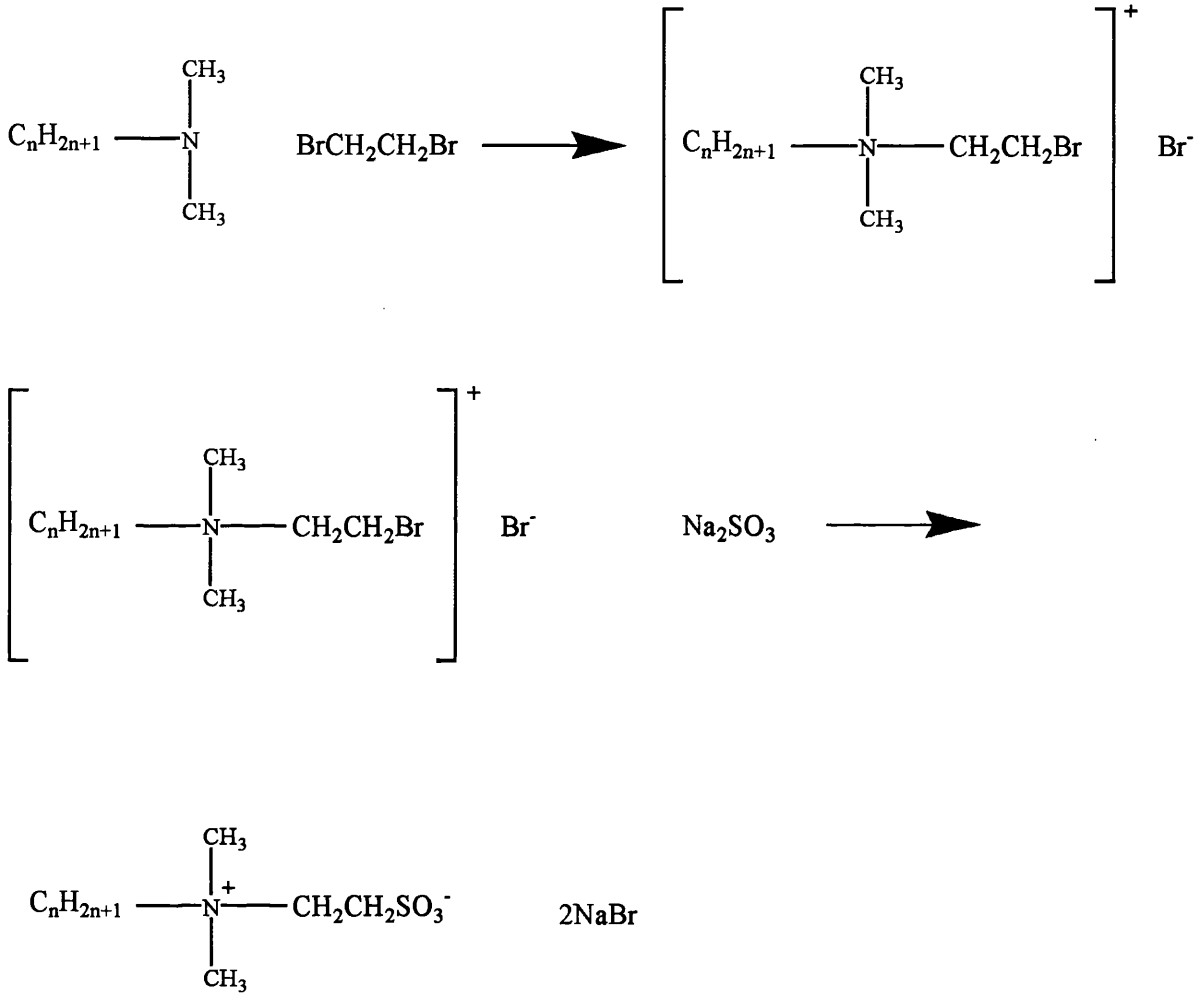
Quaternisation of trimethylamine with chloroethanesulfonic acid, or nucleophilic addition using ethenesulfonic acid, were the earliest routes to the ethanesulfonate derivatives of sulfobetaines. As early as 1885, James[68] synthesised *N,N,N*-trimethyl-2-ammonio-1-ethanesulfonate using chloroethanesulfonic acid as the alkylsulfonating agent and in 1973, Le Berre and Delacroix[69] also synthesised this compound, but *via* the nucleophilic addition route mentioned above.

Barnhurst[70] described an alternative route to ethanesulfonates as shown by Scheme 2.1. A 10-fold excess of ethylene dibromide was required for the 96 h reaction with a tertiary amine at 30 °C. The excess dibromide was removed under vacuum at 40 °C and the quaternary product was then reacted with 5 % molar excess of sodium sulfite at 85 °C for a minimum of 5 h. The sulfobetaine was seen to precipitate on cooling and it was recrystallised from water. It is perhaps useful to note that ethylene dibromide was preferred to the corresponding dichloride since it yielded crystallisable salts. In contrast, the corresponding chlorides were more hygroscopic. Le Berre *et al.*[71] used the reaction of ethenesulfonyl chloride with tertiary amines to give ethanesulfonate betaines. The reaction proceeds by nucleophilic addition of the amine to the activated double bond followed by dechlorination of the resulting compound with weak bases at moderate temperature. For example, two moles of tertiary amine, *N*-dodecyl-*N,N*-dimethyl-2-amine or *N*-octadecyl-*N,N*-dimethyl-2-amine, dissolved in acetic acid with ethenesulfonyl chloride at room temperature resulted in a 90 % yield of the corresponding ethanesulfonate. However, a drawback of this method is that ethenesulfonyl chloride has limited stability in some media and reacts explosively with pyridine and triethylamine at room temperature.[72]

King *et al.*[73] investigated alternative conditions to synthesise ethanesulfonate betaines, making use of ethenesulfonyl chloride and *trans*-1-propene-1-sulfonyl chloride (in aqueous media) for the alkylsulfonation step. Alternatively, Le Berre replaced ethenesulfonyl chloride with 2-chloroethanesulfonyl chloride for use as the alkylsulfonating agent.

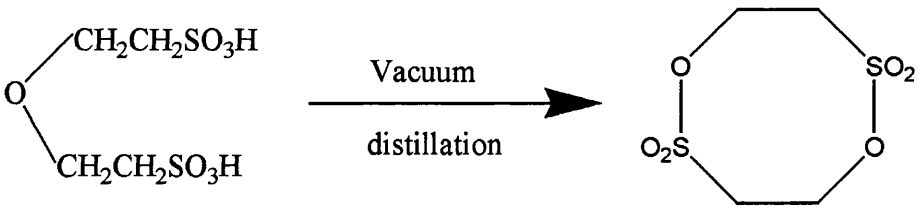
The method of Le Berre and Delacroix involving nucleophilic addition is obviously specific for the preparation of ethanesulfonates. However, the methods of James and Barnhurst are more versatile. By choice of the starting bromide/chloride, the desired number of carbon atoms can be introduced into the spacer unit between N^+ and SO_3^-, hence the equally important propane- and butanesulfonates may, in principle, be synthesised by these methods.

Okahata *et al.*[74] was responsible for the synthesis of *N,N*-didodecyl-*N*-methyl-2-ammonio-1-ethanesulfonate from sodium 2-methylamino-1-ethanesulfonate and dodecyl bromide in the presence of Na_2CO_3, followed by reaction with excess dodecyl bromide.



Scheme 2.1 Schematic Representation of route to ethanesulfonates described by Barnhurst

The direct reaction of olefins with sulfur trioxide and tertiary amines using either a complex of sulfur trioxide or a falling film reactor has also been described.[75,76] Potentially ethanesultone could also act as an alkylsulfonating agent. However, it is unstable and has never been isolated. Furthermore, a dimeric form has been made (Scheme 2.2) that could act as an alkylsulfonating agent to yield ethanesulfonates.[77]



Scheme 2.2 *Route to dimeric form of ethanesultone that can be isolated*

2.4.1.2 Preparation of Propanesulfonates

Propanesulfobetaines are commonly obtained by a reaction of tertiary amines with 1,3-propanesultone.[78,79] Sultones are internal esters of hydroxysulfonic acids and are sulfur analogues of lactones. Aliphatic saturated sultones, such as 1,3-propanesultone, are often prepared by cyclisation of hydroxysulfonic acids and their analogues or sulfonation of olefins.[80] This method of preparing propanesulfonates is widely used because the resulting sulfobetaine is substantially free from by-products. However, a major drawback is that 1,3-propanesultone is highly hazardous and is potentially carcinogenic to mammals. Industrial production *via* this route has now been abandoned, but for research purposes (smaller scale syntheses) this route still proves very popular.[4] In the literature, several authors have derived sulfobetaines from sultones.[81,82]

An alternative route to alkyldimethyl sulfobetaines bearing the propanesulfonate group, which avoids the use of 1,3-propanesultone, was developed by Parris *et al.*[83]

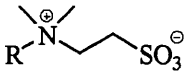
2.4.1.3 Preparation of Butanesulfonates

Similar to the method described in Section 2.4.1.2, butane sulfobetaines are commonly obtained by reaction of tertiary amines with 1,4-butanesultone.

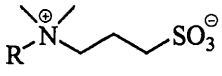
2.5 Synthesis of Quaternary Alkylammonium Sulfobetaines

A series of quaternary alkylammonium sulfobetaines of general formula $RN^{+}(CH_3)_2(CH_2)_nSO_3^{-}$, where n = 2–4, were synthesised by reacting the corresponding *N,N*-dimethylamines with either sodium 2-chloroethanesulfonate (n = 2), 1,3-propanesultone (n = 3), or 1,4-butanesultone (n = 4). In some cases the *N,N*-dimethylamines were commercially available. However, in many cases they required synthesis prior to reaction with the alkylsulfonating agent.

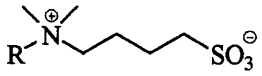
As already explained, all compounds synthesised belong to one of three series that differ in the length of the spacer unit separating the quaternary ammonium centre from the sulfonate group. Additionally, compounds **1-2**, **3-7** and **14-16** possess an alkyl chain (containing six to twelve carbon atoms), which is directly connected to the quaternary ammonium centre, and compounds **8-13** and **17-22** contain an aromatic ring, which is separated from the quaternary ammonium centre by one to four CH_2 units. In the case of **12**, **13**, **21** and **22,** the aromatic ring is separated from the quaternary ammonium centre by three CH_2 units but also carries a *para* alkyl substituent containing four or six carbon atoms. Figure 2.8 summarises the twenty one compounds synthesised, along with **6**, that was obtained commercially, giving a total of twenty two sulfobetaines for log P measurement (Chapter 3) and hence potential use in QSAR development (Chapter 4, Part II).



1 R = C_8H_{17}
2 R = $C_{12}H_{25}$



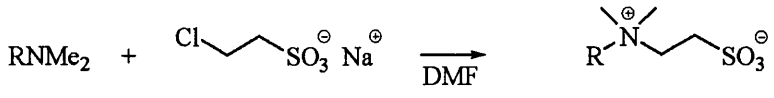
3 R = C_6H_{13}
4 R = C_7H_{15}
5 R = C_8H_{17}
6 R = $C_{10}H_{25}$
7 R = $C_{12}H_{25}$
8 R = $PhCH_2$
9 R = $Ph(CH_2)_2$
10 R = $Ph(CH_2)_3$
11 R = $Ph(CH_2)_4$
12 R = $C_4H_9C_6H_4(CH_2)_3$
13R = $C_6H_{13}C_6H_4(CH_2)_3$



14 R = C_6H_{13}
15 R = C_7H_{15}
16 R = C_8H_{17}
17 R = $PhCH_2$
18 R = $Ph(CH_2)_2$
19 R = $Ph(CH_2)_3$
20 R = $Ph(CH_2)_4$
21 R = $C_4H_9C_6H_4(CH_2)_3$
22 R = $C_6H_{13}C_6H_4(CH_2)_3$

Figure 2.8 *Structure of 22 sulfobetaines obtained*

The two ethane sulfobetaines (**1** and **2**) were synthesised by reacting commercially available *N,N*-dimethylamines with sodium 2-chloroethanesulfonate at reflux in DMF for 96 hours (Scheme 2.3). The products separated on cooling as finely powdered white solids and were isolated by filtration.

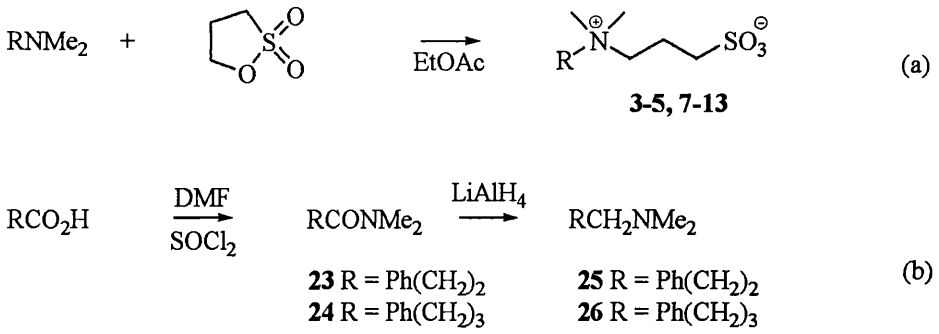


1 R = C_8H_{17}
2 R = $C_{12}H_{25}$

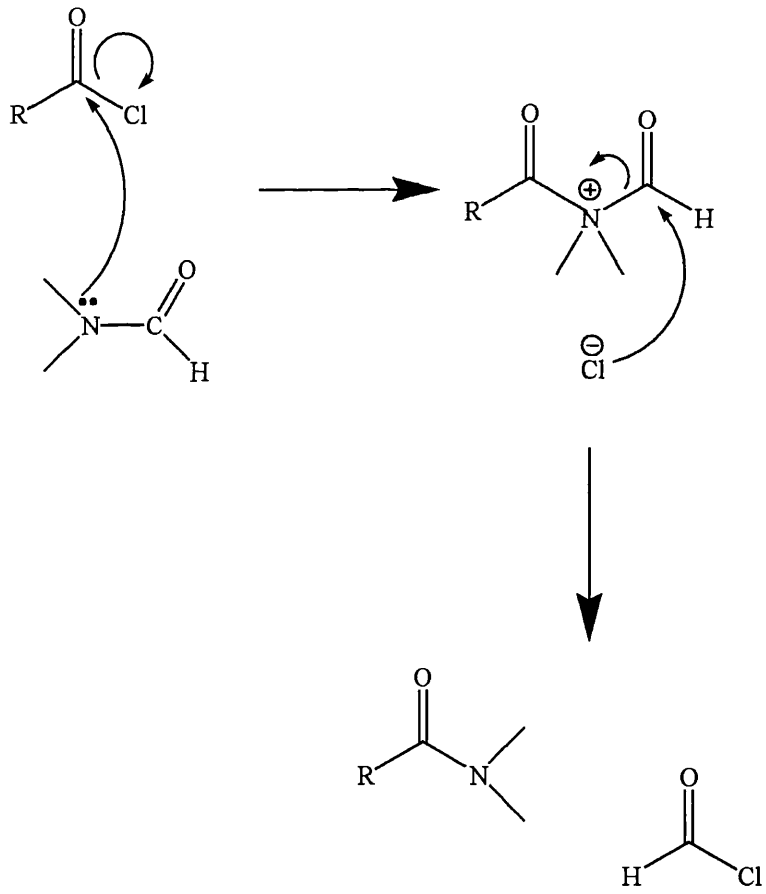
Scheme 2.3 Route to Ethane Sulfobetaines

Prior to a successful reaction involving the use of dimethylformamide as the reaction solvent, the reaction described by Scheme 2.3 was attempted using ethanol, methanol and *n*-amylalcohol as possible reaction solvents. However, with all three solvents the reaction proved unsuccessful and only mixtures of starting materials were recovered. Sodium 2-chloroethane sulfonate appeared to have only limited solubility in ethanol and it was believed that this was the main reason for reaction failure using this solvent. The salt appeared to be very soluble in methanol, but the reaction still proved unsuccessful. It was concluded that the reflux temperatures of ethanol and methanol (78 °C and 65 °C respectively) were insufficient for this reaction to proceed. *n*-amylalcohol was also investigated. It is expected that the reflux temperature (138 °C) was sufficient for the reaction to proceed; however the polarity of the solvent limited progression of the reaction. Finally, success in the use of dimethylformamide, with a slightly higher but comparable boiling point (153°C) to that of *n*-amylalcohol, indicated that a polar solvent was required. It was also beneficial that the reactants and the by-product, sodium chloride, were soluble in this solvent, while the product was at least partially insoluble and thus could be recovered by filtration followed by washing with ether. The overall yield could have been improved by separating the components of the mother liquor using ion-exchange chromatography.

The propane sulfobetaines (**3-5** and **7-13**) were prepared by reacting the corresponding *N,N*-dimethylamines with 1,3-propanesultone in ethyl acetate at room temperature (Scheme 2.4a).[67,84] The products separated out from the solution as white solids, which could in some cases be purified by recrystallisation, but were isolated following a wash with ether in excellent purity. The amines required for the preparation of **3-5** and **7-9** were commercially available, while *N,N*-dimethyl-3-phenylpropylamine (**25**) and *N,N*-dimethyl-4-phenylbutylamine (**26**) were prepared by reducing the corresponding *N,N*-dimethylamides **23** and **24** with $LiAlH_4$ (Scheme 2.4b).[85] The dimethylamides **23** and **24** in turn were prepared from hydrocinnamic acid and 3-phenylbutanoic acid respectively by reaction with DMF and thionyl chloride. The direct preparation of an *N,N*-dimethylamide by reacting a carboxylic acid with DMF and a co-reagent such as P_2O_5,[86-89] $POCl_3$,[90-91] $(COCl)_2$,[92-95] DCCI[96,97] or NaH[98] has been previously reported but, to the best of our knowledge, there are only two reports of the use of $SOCl_2$ for this purpose.[99,100] The mechanism given by Scheme 2.5 has been postulated for the reaction of an acid chloride with DMF. There are examples of mechanisms where nuclephilic attack of the lone pair of electrons on the nitrogen atom of amides have been proposed.[101]



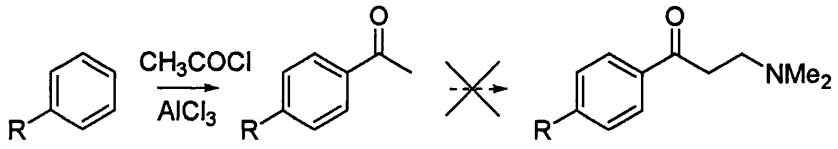
Scheme 2.4 Routes to (a) Propane Sulfobetaines and (b) N,N-dimethylamides



Scheme 2.5 Proposed Mechanism for the reaction of a carboxylic acid with DMF/$SOCl_2$ to yield an N,N-dimethylamide

For the synthesis of the amines required for the preparation of **12** and **13** the initial intention was to make use of a Mannich reaction[102] to synthesise the ketones **28** and **29** (Scheme 2.6), but all

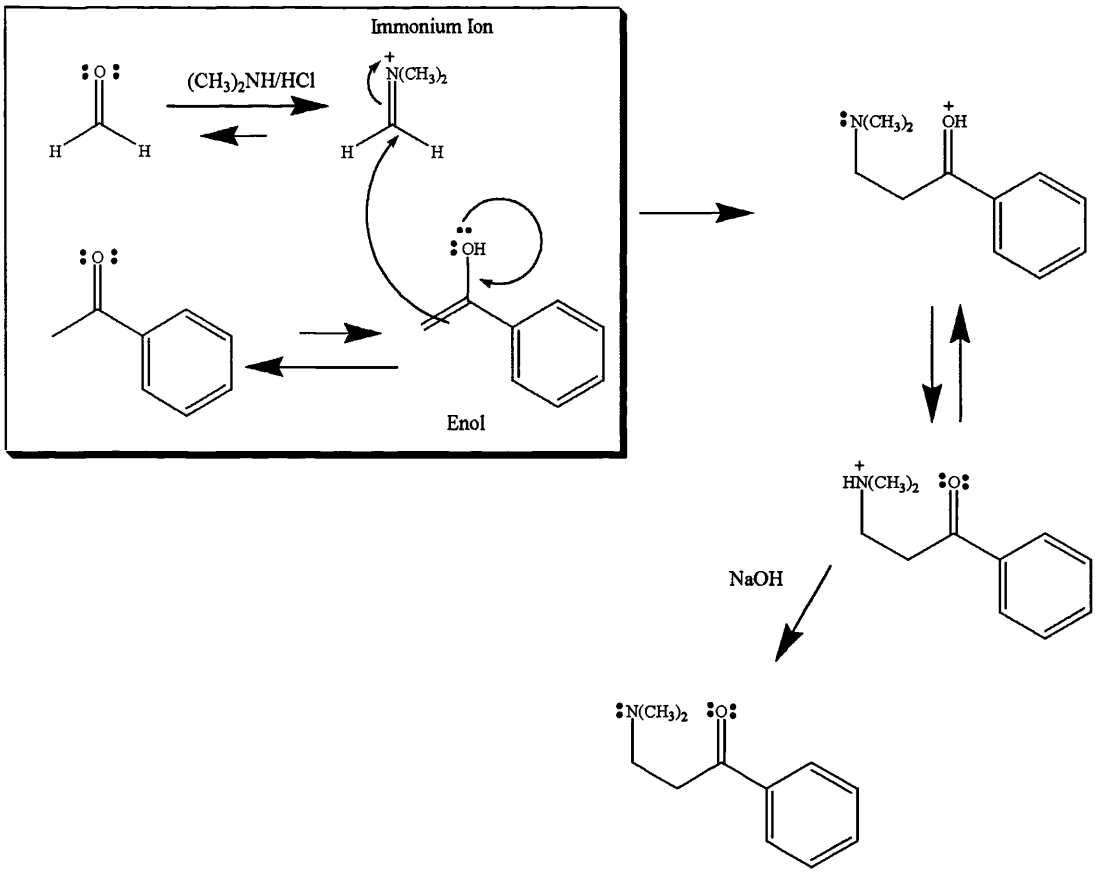
attempts to carry out a Mannich reaction on 4-hexylacetophenone (**27**) proved unsuccessful. The first attempt involved reacting compound **27** with dimethylamine hydrochloride, paraformaldehyde and concentrated hydrochloric acid in ethanol (under reflux) for 2 h, as suggested for acetophenone.[102] However, even extending the reaction time up to 24 h and/or repeating the reaction in a different solvent (e.g. *n*-amylalcohol) appeared to have no effect. After confirming that the conditions described in the literature for the analogous reaction on acetophenone resulted in a 40 % yield of the desired product, this method was abandoned as a potential route to **28** and **29**.



27 R = C_6H_{13} **28** R = C_4H_9
29 R = C_6H_{13}

*Scheme 2.6 Unsuccessful route to the amino-ketones **28** and **29**.*

As highlighted by Scheme 2.7, the mechanism of the Mannich reaction is believed to involve electrophilic attack by an immonium salt on the *enol* form of the active methylene compound.[103] It was considered possible that the inductive effect of the C6 alkyl chain could lead to stabilisation of the keto-tautomer, so that formation of the required enol would be disfavoured. However, Figure 2.9 clearly shows that the energy difference between the keto-tautomer of each compound and the enol-**A** is comparable which negates our hypothesis. Not surprisingly, the energy difference between 4-hexylacetophenone and the enol-**B** is somewhat greater. Furthermore, attempts to carry out Friedel-Crafts acylation on *N,N*-dimethyl-3-phenylpropylamine (**25**) using hexanoyl chloride also proved unsuccessful. It was decided that the possibility of synthesising the ketones **28** and **29** by means of a conjugate addition reaction[104] on the α,β-unsaturated ketones **30** and **31** (Scheme 2.8) would be investigated.



Scheme 2.7 Proposed mechanism for the Mannich reaction[103]

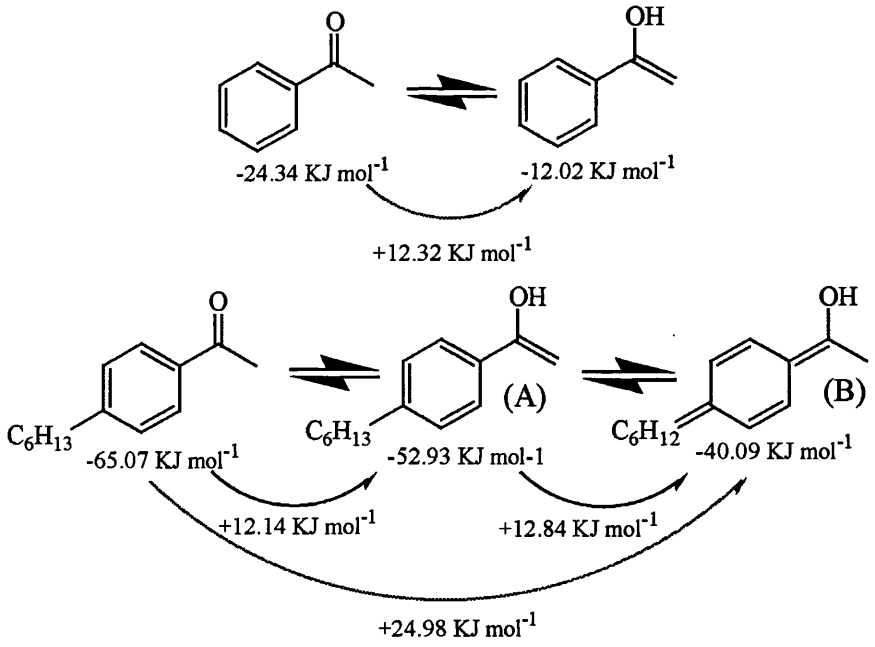
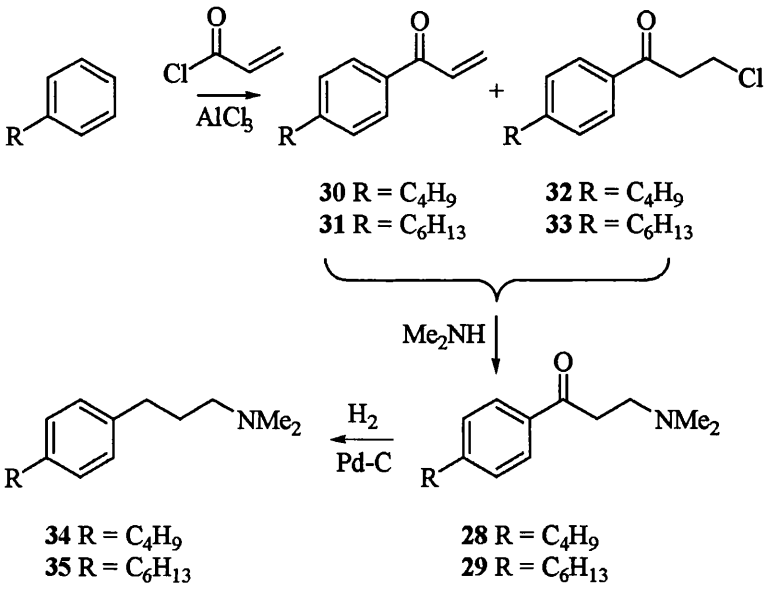


Figure 2.9 A simple AM1 calculation to compare relative heats of formation of the tautomeric forms of acetophenone derivatives

The unsaturated ketone **30** was obtained along with the 3-chloropropiophenone derivative **32** by reacting butylbenzene with acryloyl chloride.[105] Similarly, reaction of 1-phenylhexane with acryloyl chloride afforded a mixture of the unsaturated ketone **31** and the chloro ketone **33**. In each case the relative proportions of the two products varied with the reaction conditions (Table 2.1). It appeared that the relative proportions of the unsaturated ketone and the chloro ketone formed was a function of both reaction time and temperature. Acryloyl chloride was introduced to the reaction mixture at 0 °C and the temperature was allowed to slowly increase to r.t. over the course of the reaction. Table 2.1 shows that after 3 h, only the 3-chloropropiophenone derivative is formed, which suggests that this reaction time is not sufficient for r.t. to be reached which leads to the formation of compound **32** only. However, at a reaction time of 5 h 30 min, almost equal proportions of the two products (**30** and **32**) were observed. Conveniently both products reacted, without separation, with dimethylamine to give the corresponding dimethylamino ketones **28** and **29**.

Hydrogenation of **28** and **29** gave the required dimethyl amines **34** and **35** (Scheme 2.8), which reacted with 1,3-propanesultone to afford the zwitterionic compounds **12** and **13** (Scheme 2.4a).

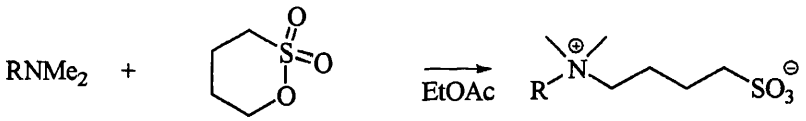


Scheme 2.8 *Schematic diagram to show route from phenylalkane to N,N-dimethylamine*

Table 2.1 Variation in relative proportions of the unsaturated ketone and its 3-chloropropiophenone derivative with reaction time.

Substrate	Reaction time	% unsaturated ketone	3-chloropropiophenone derivative
1-Phenylbutane	3 h	0	100
1-Phenylbutane	5 h 30 min	46	54
1-Phenylhexane	6 h	53	47
1-Phenylhexane	6 h 45 min	31	69

The butane sulfobetaines (**14–22**) were prepared by reacting the corresponding *N,N*-dimethylamines with 1,4-butanesultone in ethyl acetate under reflux (Scheme 2.9).[64,84] The products separated out from the solution as white solids, which were purified by washing with ether.



14 - 22

Scheme 2.9 Route to Butane Sulfobetaines

2.6 Conclusion

A series of quaternary alkylammonium sulfobetaines were synthesised by reacting the corresponding tertiary amines with either sodium 2-chloroethanesulfonate or with an appropriate sultone. The products were obtained as white hygroscopic solids, which were stored under anhydrous conditions until required. As a general rule, it appears that although hydrophilic, surfactant sulfobetaines are only slightly hygroscopic, but shorter-chain sulfobetaines tend to be the most hygroscopic. Sulfobetaines generally exhibit good water solubility and usually have high melting points due to their ionic character. Literature reports that some decompose before their melting point is reached. Most alkylsulfonate molecules are hydrophilic and only slightly soluble in hydrophobic solvents. It appears that the propanesulfonate group is the most advantageous, not only because it is the easiest group to

introduce, but also it usually offers better stability and water solubility in comparison to the analogous ethane- and butanesulfonates.[106]

N,N-dimethylamines react with sultones in equimolar quantities and the yields obtained are relatively high. The added advantage of using sultones as alkylsulfonating agents is that there is complete absence of inorganic salts during syntheses of sulfobetaines and because of the strong difference in hydrophilicity between reactants and product, the latter often precipitating during the reaction and therefore being easily isolated and in high purity. Reactions of *N,N*-dimethylamines with sodium 2-chloroethanesulfonate are not as efficient. The desired sulfobetaine may be isolated by filtration and in high purity but the yield is very low in comparison to the reactions with sultones. The yield may be greatly improved by isolation of further product from the reaction mixture; however, the desired product must be separated from other components (also soluble in the reaction solvent) by a technique such as ion-exchange chromatography. This step is obviously time consuming and is conveniently avoided in reactions involving sultones.

A whole series of these compounds were synthesised for direct experimental determination of log P (Chapter 3) for the determination of acute aquatic toxicity (Chapter 4, Part I) and for use in QSAR correlation studies (Chapter 4, Part II). Finally, they were also required for an investigation into the effectiveness of a variety of direct and indirect chromatographic methods for estimating log P (Chapter 6). It is useful to note that a selection of short-chain surfactants and non-surfactants were synthesised for the measurement of log P and for use in QSAR studies, since it is suggested in the literature by many authors that log P of commercial, long chain surfactants can not be measured with any accuracy due to interfacial phenomena.

2.7 Experimental

^1H and ^{13}C NMR spectra were recorded on a Bruker AC 400 instrument and were run in $CDCl_3$ or CD_3OD as stated. Mass spectra were recorded on a VG 12-250 low resolution quadrupole instrument or on a VG Micromass Quattro II instrument. Accurate mass measurements were made using either a ZAB-E high resolution double-focussing instrument or a Finnigan Mat 900 instrument. IR spectra were recorded on a Perkin-Elmer FT 1725X spectrometer and were measured using KBr discs. Melting points were recorded on an Electrothermal 9100 apparatus and are uncorrected.

TLC analysis was carried out on Merck 5785 Kiesegel $60F_{254}$ fluorescent plates. Purity of the sulfobetaines was assessed by reverse-phase HPLC on a Phenomenex Synergi Polar-RP or Phenomenex Synergi Max-RP column (4μ 4.6mm x 150mm) with a mixture of acetonitrile/water as the eluting solvent. Ether and dichloromethane were purified on distillation from calcium hydride. THF was passed down an alumina column and distilled from sodium/benzophenone.

Ethane Sulfobetaines 1 and 2; General Procedure

N-octyl-*N*,*N*-dimethyl-2-ammonio-1-ethanesulfonate (1)

To a stirred solution of sodium 2-chloroethanesulfonate monohydrate (8.809 g, 47.6 mmol) in DMF (75 ml), was added a solution of *N*,*N*-dimethyloctylamine (7.404 g, 47.2 mmol) in DMF (75 ml). The reaction mixture was stirred under reflux for 96 h before being allowed to cool to r.t. The resulting precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield the product **1** as a finely powdered white solid (3.313 g, 27 %); mp 242-4 °C.

^1H NMR (400MHz, CD_3OD): δ = 0.86 (t, 3H, *J*=6.8 Hz, CH_3), 1.20-1.32 (m, 10H, $(CH_2)_5$), 1.62 (m, 2H, CH_2CH_2N), 2.92 (m, 2H, CH_2SO_3) 3.02 (s, 6H, $N(CH_3)_2$), 3.26 (m, 2H, CH_2N), 3.51 (m, 2H, NCH_2). ^{13}C NMR (101MHz, CD_3OD): δ = 14.0 (CH_3), 21.7 (CH_2), 22.1 (CH_2), 25.8 (CH_2), 28.5 (CH_2), 28.5 (CH_2), 31.2 (CH_2CH_2N), 44.3 (CH_2SO_3), 50.1 (NCH_3), 60.0 (NCH_2), 62.7 (CH_2N). *m/z* (EI): 264 ($[M-H]^+$, 1%), 158 (16%), 157 (100%), 156 (10%). *m/z* (CI): 266 ($[M+H]^+$, 1%), 159 (9%), 158 (100%). *m/z* HRMS: Calcd. for $C_{12}H_{28}NSO_3$ $[M+H]^+$ 266.1790. Found 266.1786.

_N_-dodecyl-_N,N_-dimethyl-2-ammonio-1-ethanesulfonate (2)

Prepared using the procedure described above starting from _N,N_-dimethyl dodecylamine (10.77 g, 50.6 mmol). The product **2** was obtained as a finely powdered white solid (2.575 g, 16 %); mp 357-8 °C (lit.[107] mp 353 °C).

^{1}H NMR (400MHz, CD_3OD): δ = 0.90 (t, 3H, _J_=6.9 Hz, CH_3), 1.24-1.40 (m, 18H, $(CH_2)_9$), 1.80 (m, 2H, CH_2CH_2N), 3.09 (s, 6H, $N(CH_3)_2$), 3.30 (m, 4H, CH_2SO_3 and CH_2N), 3.70 (m, 2H, NCH_2). ^{13}C NMR (101MHz, CD_3OD): δ = 15.0 (CH_3), 24.0 (CH_2), 24.1 (CH_2), 27.7 (CH_2), 30.6 (CH_2), 30.8 (CH_2), 30.9 (CH_2), 31.0 (CH_2), 31.1 (CH_2), 33.4 (CH_2CH_2N), 46.0 (CH_2SO_3), 52.4 (NCH_3), 61.3 (NCH_2), 66.4 (CH_2N). *m/z* (FAB): 665 ($[2M+Na]^+$, 7%), 643 ($[2M+H]^+$, 3%), 344 ($[M+Na]^+$, 74%), 322 ($[M+H]^+$, 86%), 228 (52%), 214 (100%). *m/z* HRMS (ES^+): Calcd. for $C_{16}H_{36}NSO_3$ $[M+H]^+$ 322.2416. Found 322.2411.

Propane Sulfobetaines 3 – 5 and 7 – 9; General Procedure

_N_-hexyl-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (3)

To a stirred solution of 1,3-propanesultone (9.879 g, 81.0 mmol) in EtOAc (30 ml), was added _N,N_-dimethylhexylamine (10.476 g, 81.2 mmol) in EtOAc (30 ml). The reaction mixture was stirred at r.t. for 1.5 h. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **3** (12.79 g, 63 %) as a white crystalline solid.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.88 (t, 3H, _J_=6.9 Hz, CH_3), 1.28-1.37 (m, 6H, $(CH_2)_3$), 1.72 (m, 2H, CH_2CH_2N), 2.21 (m, 2H, $CH_2CH_2SO_3$), 2.87 (t, 2H, _J_ 6.9, CH_2SO_3), 3.19 (s, 6H, $N(CH_3)_2$), 3.30 (m, 2H, CH_2N), 3.68 (m, 2H, NCH_2). ^{13}C NMR (101MHz, $CDCl_3$): δ = 13.9 (CH_3), 19.4 (CH_2), 22.4 (CH_2), 22.6 (CH_2), 26.0 (NCH_2CH_2), 31.2 (CH_2CH_2N), 47.8 (CH_2SO_3), 50.8 (NCH_3), 63.1 (NCH_2), 64.3 (CH_2N). *m/z* (EI): 251 (M^+, 1%), 250 ($[M-H]^+$, 2%), 180 (4%), 166 (2%), 130 (19%), 129 (100%). *m/z* (CI): 252 ($[M+H]^+$, 1%), 238 (3%), 182 (1%), 140 (30%), 130 (100%). *m/z* HRMS: Calcd. for $C_{11}H_{26}NSO_3$ $[M+H]^+$ 252.1633. Found 252.1630.

_N_-heptyl-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (4)

Prepared using the procedure described above starting from _N,N_-dimethylheptylamine (9.804 g, 68.6 mmol). The product **4** was obtained as a white crystalline solid (12.30 g, 68 %); mp 188-191 °C.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.88 (t, 3H, _J_=6.9 Hz, CH_3), 1.27-1.34 (m, 8H, (CH_2)$_4$), 1.72 (m, 2H, CH_2CH$_2$N), 2.22 (m, 2H, CH_2CH$_2$SO$_3$), 2.88 (t, 2H, _J_=6.8 Hz, CH_2SO$_3$), 3.18 (s, 6H, N(CH_3)$_2$), 3.28 (m, 2H, CH_2N), 3.66 (m, 2H, NCH_2). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.1 (*C*H$_3$), 19.1 (*C*H$_2$), 22.5 (*C*H$_2$), 22.6 (*C*H$_2$), 26.3 (NCH$_2$*C*H$_2$), 28.9 (*C*H$_2$), 31.6 (*C*H$_2$CH$_2$N), 47.8 (*C*H$_2$SO$_3$), 50.8 (N*C*H$_3$), 63.0 (N*C*H$_2$), 64.3 (*C*H$_2$N). _m/z_ (EI): 265 (M^+, 1%), 264 ([M-H]$^+$, 3%), 180 (18%), 166 (6%), 144 (18%), 143 (100%). _m/z_ (CI): 266 ([M+H]$^+$, 2%), 252 (2%), 184 (2%), 182 (2%), 170 (6%), 145 (8%), 144 (100%), 140 (30%). _m/z_ HRMS: Calcd. for $C_{12}H_{28}NSO_3$ [M+H]$^+$ 266.1790. Found 266.1788.

_N_-octyl-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (5)

Prepared using the procedure above starting from _N,N_-dimethyloctylamine (10.61 g, 67.6 mmol). The product **5** was obtained as a white crystalline solid (12.69 g, 67 %); mp 203-4 °C.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.86 (t, 3H, _J_=6.7 Hz, CH_3), 1.24-1.32 (m, 10H, (CH_2)$_5$), 1.70 (m, 2H, CH_2CH$_2$N), 2.20 (m, 2H, CH_2CH$_2$SO$_3$), 2.86 (t, 2H, _J_=6.5 Hz, CH_2SO$_3$), 3.19 (s, 6H, N(CH_3)$_2$), 3.28 (m, 2H, CH_2N), 3.68 (m, 2H, NCH_2). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.1 (*C*H$_3$), 19.4 (*C*H$_2$), 22.6 (*C*H$_2$), 22.7 (*C*H$_2$), 26.3 (NCH$_2$*C*H$_2$), 29.1 (*C*H$_2$), 29.2 (*C*H$_2$), 31.7 (*C*H$_2$CH$_2$N), 47.8 (*C*H$_2$SO$_3$), 50.8 (N*C*H$_3$), 63.2 (N*C*H$_2$), 64.3 (*C*H$_2$N). _m/z_ (EI): 278 ([M-H]$^+$, 6%), 180 (72%), 166 (75%), 157 (100%), 156 (94%). _m/z_ (CI): 280 ([M+H]$^+$, 4%), 266 (3%), 184 (4%), 170 (3%), 159 (25%), 158 (100%), 140 (58%). _m/z_ HRMS: Calcd. for $C_{13}H_{30}NSO_3$ [M+H]$^+$ 280.1946. Found 280.1945.

_N_-dodecyl-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (7)

To a stirred solution of 1,3-propanesultone (2.526 g, 20.7 mmol) in EtOAc (50 ml), was added _N,N_-dimethyldodecylamine (3.815 g, 17.9 mmol) in EtOAc (50 ml). The reaction mixture was stirred at r.t. for 24 h. The resulting white precipitate was filtered, washed with diethyl ether,

and dried *in vacuo* to yield **7** (2.407 g, 68 %) as a white crystalline solid; mp 242-4 °C (lit.[108] mp 250-5 °C).

^{1}H NMR (250MHz, $CDCl_3$): δ = 0.78 (t, 2H, *J*=6.8 Hz, CH_3), 1.08-1.27 (m, 18H, $(CH_2)_9$), 1.60 (m, 2H, CH_2CH_2N) 2.10 (m, 2H, $CH_2CH_2SO_3$), 2.88 (t, 2H, *J*=6.7 Hz, CH_2SO_3), 3.10 (s, 6H, $N(CH_3)_3$), 3.15 (m, 2H, CH_2N), 3.60 (m, 2H, NCH_2); ^{13}C NMR (63MHz, $CDCl_3$): δ = 14.0 (CH_3), 19.3 (CH_2), 22.6 (CH_2), 26.3 (CH_2), 26.3 (NCH_2CH_2), 29.1 (CH_2), 29.2 (CH_2), 29.4 (CH_2), 29.4 (CH_2), 29.5 (CH_2), 31.8 (CH_2CH_2N), 47.7 (CH_2SO_3), 50.7 (NCH_3), 63.2 (NCH_2), 64.3 (CH_2N). *m/z* (EI): 335 (M^+, 27%), 213 (100%), 212 (48%). *m/z* (CI): 336 ($[M+H]^+$, 1%), 214 (100%).

***Nb*-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (8)**

To a stirred solution of 1,3-propanesultone (1.113 g, 9.12 mmol) in EtOAc (15 ml), was added *N,N*-dimethylbenzylamine (1.117 g, 8.27 mmol) in EtOAc (15ml). The reaction mixture was stirred at r.t. for 12 h. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **8** (1.416 g, 67 %) as a white crystalline solid; mp 235-7 °C.

^{1}H NMR (400MHz, CD_3OD): δ = 2.33 (m, 2H, $CH_2CH_2SO_3$), 2.90 (t, 2H, *J*=6.9 Hz, CH_2SO_3), 3.04 (s, 6H, $N(CH_3)_2$), 3.54 (m, 2H, NCH_2), 4.55 (s, 2H, $PhCH_2N$), 7.50-7.60 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 20.0 (NCH_2CH_2), 48.7 (CH_2SO_3), 50.2 (NCH_3), 64.3 (NCH_2), 68.9 (CH_2N), 128.8 (Ph, C-1), 130.3 (Ph, CH), 131.8 (Ph, CH), 134.1 (Ph, CH). *m/z* (EI): 257 (M^+, 1%), 256 ($[M-H]^+$, 1%), 182 (100%), 135 (25%), 105 (65%). *m/z* (CI): 258 ($[M+H]^+$, 3%), 185 (13%), 168 (30%), 140 (20%), 136 (100%). *m/z* (ES^+): 537 ($[2M+Na]^+$, 13%), 515 ($[2M+H]^+$, 5%), 280 ($[M+Na]^+$, 100%), 258 ($[M+H]^+$, 55%). *m/z* (ES^-): 302 (78%), 292 ($[M+Cl]^-$, 52%), 256 ($[M-H]^-$, 19%). *m/z* HRMS (ES^+): Calcd. for $C_{12}H_{20}NSO_3$ $[M+H]^+$ 258.1164. Found 258.1163.

***N*-phenylethyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (9)**

To a stirred solution of 1,3-propanesultone (4.207 g, 34.5 mmol) in EtOAc (70 ml), was added *N,N*-dimethylphenethylamine (5.138 g, 34.5 mmol) in EtOAc (70 ml). The reaction mixture

was stirred at r.t. for 112 h. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **9** (6.353 g, 68 %) as a white crystalline solid; mp 244-6 °C. ^{1}H NMR (400MHz, CD_3OD): δ = 2.24 (m, 2H, $CH_2CH_2SO_3$), 2.89 (t, 2H, *J*=6.8 Hz, CH_2SO_3), 3.10-3.16 (m, 2H, $PhCH_2$), 3.17 (s, 6H, $N(CH_3)_2$), 3.53 (m, 2H, NCH_2), 3.61 (m, 2H, CH_2N), 7.24-7.36 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 19.9 (NCH_2CH_2), 29.7 ($ArCH_2$), 48.5 (CH_2SO_3), 51.4 (NCH_3), 63.6 (NCH_2), 65.9 (CH_2N),128.3 (Ph, *C*H), 129.9 (Ph, *C*H), 130.1 (Ph, *C*H), 137.0 (Ph, C-1). *m/z* (EI): 271 (M^+, 20%), 270 ($[M-H]^+$, 100%), 231 (53%). *m/z* (CI): 272 ($[M+H]^+$, 70%), 258 (100%). *m/z* HRMS (ES^+): Calcd. for $C_{13}H_{30}NSO_3$ $[M+H]^+$ 272.1320. Found 272.1318.

N-phenylpropyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (10)

N,N-dimethyl-3-phenylpropanamide (23)

To a stirred solution of 3-phenylpropionic acid (2.000 g, 13.3 mmol) in DMF (15 ml, 0.19 mol) at 0 °C, was added $SOCl_2$ (2 ml, 27.0 mmol). The reaction mixture was stirred for 1.75 h as the temperature was raised from 40 to 90 °C. HCl evolved was trapped by bubbling into 1M aq.Na_2CO_3. Excess $SOCl_2$ and DMF were removed by reduced pressure distillation (20°C@1.5 mmHg) and the crude product was dissolved in CH_2Cl_2 (50 ml). The organic layer was washed with water (3 x 50 ml) and the CH_2Cl_2 removed *in vacuo* to yield **23** (2.199 g, 93 %) as a brown oil.

^{1}H NMR (250MHz, $CDCl_3$): δ = 2.50-2.56 (m, 2H, $PhCH_2CH_2$), 2.84 (s, 3H, NCH_3), 2.87 (s, 3H, NCH_3), 2.88-2.95 (m, 2H, $PhCH_2$), 7.14-7.25 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, $CDCl_3$): δ = 31.3 (CH_2CO), 35.2 ($ArCH_2$), 35.3 (NCH_3), 37.1 (NCH_3), 126.0 (Ph, *C*H), 128.3 (Ph, *C*H), 128.4 (Ph, *C*H), 141.4 (Ph, C-1), 172.1 (*C*O). *m/z* (EI): 177 (M^+, 56%), 105 (67%), 104 (62%), 91 (100%). *m/z* (CI): 195 ($[M+NH_4]^+$, 28%), 178 ($[M+H]^+$, 100%). *m/z* HRMS (ES^+): Calcd. for $C_{11}H_{16}NO$ $[M+H]^+$ 178.1232. Found 178.1232.

N,N-Dimethyl-3-phenylpropylamine (25)

To a stirred suspension of $LiAlH_4$ (1.550 g, 40.8 mmol) in dry THF (20 ml), under N_2, was added drop-wise a solution of **22** (2.094 g, 11.8 mmol) in dry THF (10 ml). The reaction

mixture was stirred under gentle reflux for 5 h. After cooling, the excess $LiAlH_4$ and metallic complexes were deposited by the careful addition of acetone (3 ml) to the well-stirred mixture. An equal volume of dry ether was added, followed by the careful drop-wise addition of water (2 ml) after which the mixture was filtered and the solid residue washed with ether (20 ml). The filtrate was dried over anhydrous $MgSO_4$ and the THF-ether solvent mixture removed *in vacuo* to yield **25** (1.923 g, 74 %) as a dark yellow liquid.

^{1}H NMR (250MHz, $CDCl_3$): δ = 1.71 (q^5, 2H, *J*=7.7 Hz, $PhCH_2CH_2$), 2.13 (s, 6H, $N(CH_3)_2$), 2.20(t, 2H, *J*=7.7 Hz, CH_2N), 2.54 (t, 2H, *J*=7.7 Hz, $PhCH_2$), 7.08-7.18 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, $CDCl_3$): δ = 29.4 (CH_2CH_2N), 35.6 ($ArCH_2$), 45.4 (NCH_3), 59.2 (CH_2N), 125.6 (Ph, *C*H), 128.2 (Ph, *C*H), 128.3 (Ph, *C*H), 142.2 (Ph, C-1). *m/z* (EI): 163 (M^+, 100%). *m/z* (CI): 164 ($[M+H]^+$, 100%). *m/z* HRMS (ES^+): Calcd. for $C_{11}H_{18}N$ $[M+H]^+$ 164.1439. Found 164.1439.

N-phenylpropyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (10)

To a stirred solution of 1,3-propanesultone (0.500 g, 4.1 mmol) in EtOAc (10 ml), was added a solution of **25** (0.534 g, 3.3 mmol) in EtOAc (10 ml). The reaction mixture was stirred at r.t. for 140 h. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **10** (0.531 g, 57 %) as a white crystalline solid; mp 172-7 °C.

^{1}H NMR (400MHz, CD_3OD): δ = 2.10-2.21 (m, 4H, $PhCH_2CH_2$ and $CH_2CH_2SO_3$), 2.73 (t, 2H, *J*=7.55 Hz, $PhCH_2$), 2.87 (t, 2H, *J*=7.0 Hz, CH_2SO_3), 3.10 (s, 6H, $N(CH_3)_2$), 3.35 (m, 2H, NCH_2), 3.52 (m, 2H, CH_2N), 7.20-7.35 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 19.9 (NCH_2CH_2), 25.2 (CH_2CH_2N), 33.2 ($ArCH_2$), 48.6 (CH_2SO_3), 51.4 (NCH_3), 63.8 (NCH_2), 65.0 (CH_2N), 127.6 (Ph, *C*H), 129.5 (Ph, *C*H), 129.7 (Ph, *C*H), 141.3 (Ph, C-1). *m/z* (EI): 286 ($[M+H]^+$, 44%), 285 (M^+, 57%), 284 ($[M-H]^+$, 100%), 190 (47%). *m/z* (CI): 286 ($[M+H]^+$, 58%), 272 (100%), 192 (87%), 190 (92%), 178 (42%). *m/z* HRMS (ES^+): Calcd. for $C_{14}H_{24}NSO_3$ $[M+H]^+$ 286.1477. Found 286.1475.

N-phenylbutyl-*N*,*N*-dimethyl-3-ammonio-1-propanesulfonate (11)

N,*N*-dimethyl-4-phenylbutanamide (24)

To a stirred solution of 4-phenylbutyric acid (3.285 g, 20.0 mmol) in DMF (15 ml, 0.19 mol) at 0 °C, was added $SOCl_2$ (2 ml, 27.0 mmol). The reaction mixture was stirred for 1.75 h as the temperature was raised from 40 to 90 °C. HCl evolved was trapped by bubbling into 1M aq.Na_2CO_3. Excess $SOCl_2$ and DMF were removed by reduced pressure distillation (20 °C@1.5 mmHg) and the crude product was dissolved in CH_2Cl_2 (50 ml). The organic layer was washed with water (3 x 50 ml) and the CH_2Cl_2 removed *in vacuo* to yield **24** (3.629 g, 95 %) as a brown oil.

^{1}H NMR (250MHz, $CDCl_3$): δ = 1.87 (q^5, 2H, *J*=7.5 Hz, $PhCH_2CH_2$), 2.21 (t, 2H, *J*=7.5 Hz, $PhCH_2$), 2.57 (t, 2H, *J*=7.5 Hz, CH_2CO), 2.83 (s, 6H, $N(CH_3)_2$), 7.06-7.20 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, $CDCl_3$): δ = 26.5 ($ArCH_2CH_2$), 32.4 (CH_2CO), 35.2 ($ArCH_2$), 35.3 (NCH_3), 37.2 (NCH_3), 125.8 (Ph, *C*H), 128.2 (Ph, *C*H), 128.4 (Ph, *C*H), 141.7 (Ph, C-1), 170.3 (*C*O). *m/z* (EI): 191 (M^+, 100%), 147 (17%). *m/z* (CI): 209 ($[M+NH_4]^+$, 12%), 192 ($[M+H]^+$, 100%). *m/z* HRMS (ES^+): Calcd. for $C_{12}H_{18}NO$ $[M+H]^+$ 192.1388. Found 192.1386.

N,*N*-Dimethyl-4-phenylbutylamine (26)

To a stirred suspension of $LiAlH_4$ (0.959 g, 25.2 mmol) in dry THF (20 ml), under N_2, was added dropwise a solution of **24** (2.600 g, 13.6 mmol) in dry THF (10 ml). The reaction mixture was stirred under gentle reflux for 5 h. After cooling, the excess $LiAlH_4$ and metallic complexes were deposited by the careful addition of acetone (3 ml) to the well-stirred mixture. An equal volume of dry ether was added, followed by the careful drop-wise addition of water (2 ml) after which the mixture was filtered and the solid residue washed with ether (20 ml). The filtrate was dried over anhydrous $MgSO_4$ and the THF-ether solvent mixture removed *in vacuo* to yield **26** (2.201 g, 91 %) as a dark yellow liquid.

^{1}H NMR (250MHz, $CDCl_3$): δ = 1.39 (q^5, 2H, *J*=7.6 Hz, $PhCH_2CH_2$), 1.56 (q^5, 2H, *J*=7.6 Hz, CH_2CH_2N), 2.09 (s, 6H, $N(CH_3)_2$), 2.18 (t, 2H, *J*=7.5 Hz, $PhCH_2$), 2.52 (t, 2H, *J*=7.5, CH_2N), 7.06-7.20 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, $CDCl_3$): δ = 27.3 (CH_2), 35.8 (CH_2), 39.2 (CH_2), 45.4 (NCH_3), 59.6 (CH_2N), 125.6 (Ph, *C*H), 128.2 (Ph, *C*H), 128.3 (Ph, *C*H),

142.4 (Ph, C-1). *m/z* (EI): 177 (M^+, 100%). *m/z* (CI): 178 ($[M+H]^+$, 100%). *m/z* HRMS (ES^+): Calcd. for $C_{12}H_{20}N$ $[M+H]^+$ 178.1595. Found 178.1596.

_N_-phenylbutyl-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (11)

To a stirred solution of 1,3-propanesultone (1.080 g, 8.9 mmol) in EtOAc (20 ml), was added a solution of **26** (1.126 g, 6.9 mmol) in EtOAc (20ml). The reaction mixture was stirred at r.t. for 261 h. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **11** (1.027 g, 50 %) as a white crystalline solid; mp 188-191 °C.

^1H NMR (250MHz, CD_3OD): δ = 1.62-1.88 (m, 4H, $PhCH_2(CH_2)_2$), 2.11-2.20 (m, 2H, $CH_2CH_2SO_3$), 2.71 (t, 2H, *J*=7.4 Hz, $PhCH_2$), 2.85 (t, 2H, *J*=6.8 Hz, CH_2SO_3), 3.06 (s, 6H, $N(CH_3)_2$), 3.33 (m, 2H, NCH_2), 3.45 (m, 2H, CH_2N), 7.16-7.28 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, CD_3OD): δ = 19.9 (NCH_2CH_2), 23.0 ($ArCH_2CH_2$), 29.2 (CH_2CH_2N), 36.1 ($ArCH_2$), 48.6 (CH_2SO_3), 51.2 (NCH_3), 63.8 (NCH_2), 65.2 (CH_2N), 127.1 (Ph, *C*H), 129.4 (Ph, *C*H), 129.5 (Ph, *C*H), 142.8 (Ph, C-1). *m/z* (EI): 299 (M^+, 27%), 298 ($[M-H]^+$, 69%), 202 (77%), 192 (100%). *m/z* (CI): 300 ($[M+H]^+$, 100%), 286 (97%). *m/z* HRMS (ES^+): Calcd. for $C_{15}H_{26}NSO_3$ $[M+H]^+$ 300.1633. Found 300.1629.

4-Hexylacetophenone (27)[109]

Acetyl chloride (25.92 g, 0.33 mol) in dry CH_2Cl_2 (50 ml) was added drop-wise to a suspension of $AlCl_3$ (23.92 g, 0.18 mol) in dry CH_2Cl_2 (100 ml) at r.t.. To the stirred mixture was added 1-phenylhexane (11.75 g, 72.5 mmol) in dry CH_2Cl_2 (50 ml) at r.t.. After 22 h, the mixture was poured onto a mixture of ice and conc. HCl. The aqueous solution was separated, extracted with CH_2Cl_2 (3 x 100 ml) and the combined organic layers washed with aq. Na_2CO_3 (2 x 120 ml), water (3 x 100 ml) and dried over anhydrous $MgSO_4$. The solvent was evaporated under reduced pressure to yield **27** (9.614 g, 65 %) as a dark orange liquid.

^1H NMR (250MHz, $CDCl_3$): δ = 0.78 (t, 3H, *J*=6.8 Hz, $CH_3(CH_2)_5$), 1.21 (m, 6H, $CH_3(CH_2)_3$), 1.53 (q^5, 2H, *J*=7.5 Hz, $CH_3(CH_2)_3CH_2$), 2.49 (s, 3H, $COCH_3$), 2.56 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 7.17 and 7.78 (2 x d, 4H, *J*=8.2 Hz, C_6H_4). ^{13}C NMR (63MHz, $CDCl_3$): δ = 8.86 (CH_3), 17.36 (CH_2), 21.33 ($COCH_3$), 23.71 (CH_2), 25.88 (CH_2), 26.45 (CH_2), 30.78

(CH_2), 123.25 (Ar, *C*H), 123.39 (Ar, *C*H), 129.63 (Ar, C-1/4), 143.57 (Ar, C-1/4), 192.63 (*C*O). *m/z* (EI): 204 (M^+, 14%), 190 (15%), 189 ($[M-CH_3]^+$, 100%). *m/z* (CI): 222 ($[M+NH_4]^+$, 100%), 206 (11%), 205 ($[M+H]^+$, 54%), 189 (12%). *m/z* HRMS: Calcd. for $C_{14}H_{21}O$ $[M+H]^+$ 205.1592. Found 205.1594.

N-[3-(4-butylphenyl)propyl]-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (12)

1-(4-Butylphenyl)propenone (30) and 3-chloro-1-(4-butylphenyl)propan-1-one (32)

To a suspension of $AlCl_3$ (9.922 g, 74.6 mmol) in dry 1,2-dichloroethane (30 ml) at 0 °C under N_2 was added acryloyl chloride (3.7 ml, 4.122 g, 45.3 mmol) in 1,2-dichloroethane (20 ml). A solution of 1-phenylbutane (5.8 ml, 4.988 g, 37.2 mmol) in 1,2-dichloroethane (20 ml) was added dropwise to the stirred mixture. After stirring in the dark for 5.5 h (while allowing the temperature rise from 0 °C to r.t.), the reaction mixture was poured into a mixture of ice and water (50 ml) and HCl (12N, 50 ml) and extracted with CH_2Cl_2 (3 x 100 ml). The combined organic extracts were washed with satd. aq. $NaHCO_3$ (3 x 100 ml) and brine (2 x 100 ml) and dried over anhydrous $MgSO_4$. The filtered solution was evaporated *in vacuo* to yield **30** (3.115 g, 16.6 mmol) and **32** (4.357 g, 19.5 mmol) in 97 % yield (46: 54 respectively) as a yellow powdered solid.

30 ^{1}H NMR (400MHz, $CDCl_3$): δ = 0.94 (t, 3H, *J*=7.3 Hz, C*H*$_3$), 1.37 (s^6, 2H, *J*=7.4 Hz, CH_3CH_2), 1.59-1.66 (m, 2H, $CH_3CH_2CH_2$), 2.68 (t, 2H, *J*=6.3Hz, $CH_3(CH_2)_2CH_2$), 5.90 (dd, 1H, *J*=1.5, 10.6 Hz, CH=CH*H*), 6.44 (dd, 1H, *J*=1.7, 15.4 Hz, CH=CH*H*), 7.18 (dd, 1H, *J*=10.6, 17.1 Hz, C*H*=CH_2), 7.29 and 7.89 (2 x d, 4H, *J*=8.3 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.7 (CH_3CH_2), 33.5 ($CH_3CH_2CH_2$), 36.1 (CH_2Ar), 128.7 (Ar, *C*H), 129.2 (Ar, *C*H), 130.1 (*C*H_2), 132.8 (*C*H), 135.3 (Ar, C-1/4), 149.8 (Ar, C-1/4), 190.9 (*C*O). *m/z* (EI): 188 (M^+, 5%), 161 (100%), 91 (60%). *m/z* (CI): 206 ($[M+NH_4]^+$, 100%), 189 ($[M+H]^+$, 60%), 161 (27%). *m/z* HRMS (ES^+): Calcd. for $C_{13}H_{16}O$ $[M]^+$ 189.1279. Found 189.1281.

32 ^{1}H NMR (400MHz, $CDCl_3$): δ = 0.94 (t, 3H, *J*=7.3 Hz, C*H*$_3$), 1.37 (s^6, 2H, *J*=7.4 Hz, CH_3CH_2), 1.59-1.66 (m, 2H, $CH_3CH_2CH_2$), 2.68 (t, 2H, *J*=7.7 Hz, $CH_3(CH_2)_2CH_2$), 3.44 (t, 2H, *J*=6.9 Hz, CH_2Cl), 3.93 (t, 2H, *J*=6.9 Hz, COC*H*$_2$), 7.29 and 7.89 (2 x d, 4H, *J*=8.3 Hz,

C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.7 (CH_3CH_2), 33.5 ($CH_3CH_2CH_2$), 36.1 (CH_2Ar), 39.2 (CH_2), 41.6 (CH_2), 128.6 (Ar, CH), 129.1 (Ar, CH), 134.5 (Ar, C-1/4), 149.2 (Ar, C-1/4), 196.7 (CO). *m/z* (EI) : 224 (M^+, 5%), 188 (6%), 161 (100%), 91 (40%). *m/z* (CI) : 242 ($[M+NH_4]^+$, 90%), 225 ($[M+H]^+$, 20%), 221 (30%), 182 (75%). HRMS (ES^+): Calcd. for $C_{13}H_{18}ClO$ $[M+H]^+$ 225.1046. Found 225.1045.

3-Dimethylamino-1-(4-butylphenyl)propan-1-one (27)

To a solution of **30** (4.312 g, 19.2 mmol) and **32** (3.082 g, 16.4 mmol) in THF (30 ml), a solution of dimethylamine in THF (2M, 58 ml, 3.2 equiv.) was added dropwise at 0 °C. The reaction mixture was stirred for 24 h at rt. The THF was removed *in vacuo*. HCl (2M, 60 ml) was cautiously added followed by an extraction with chloroform (3 x 75 ml) and drying over anhydrous $MgSO_4$. Following filtration, the chloroform was removed *in vacuo* to yield **28** as its hydrochloride salt (6.998 g, 73 %) as an orange powder.

^1H NMR (400MHz, $CDCl_3$): δ = 0.86 (t, 3H, *J*=7.3 Hz, CH_3), 1.28 (s^6, 2H, *J*=7.4 Hz, CH_3CH_2), 1.53 (q^5, 2H, *J*=7.6 Hz, $CH_3CH_2CH_2$), 2.61 (t, 2H, *J*=7.8 Hz, $CH_3(CH_2)_2CH_2$), 2.79 (d, 6H, *J*=4.8, $N(CH_3)_2$), 3.46 (dt, 2H, *J*=4.8, 7.0 Hz, CH_2N), 3.65 (t, 2H, *J*=7.0 Hz, $COCH_2$), 7.20 and 7.83 (2 x d, 4H, *J*=8.3 Hz, C_6H_4), 12.44 (br s, 1H, HCl). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.7 (CH_3CH_2), 33.5 ($CH_3CH_2CH_2$), 34.1 ($COCH_2$), 36.1 (CH_2Ar), 43.7 (NCH_3), 53.1 (CH_2N), 128.8 (Ar, CH), 129.2 (Ar, CH), 133.6 (Ar, C-1/4), 150.5 (Ar, C-1/4), 195.8 (CO). *m/z* (EI): 233 (M^+, 4%), 188 (11%), 161 (38%), 117 (13%), 91 (28%). *m/z* (CI): 234 ($[M+H]^+$, 45%), 208 (42%), 192 (34%), 162 (10%). *m/z* HRMS (ES^+): Calcd. for $C_{15}H_{24}NO$ $[M+H]^+$ 234.1858. Found 234.1856.

***N,N*-Dimethyl-3-(4-butylphenyl)propylamine (34)**

To a stirred solution of **28** (8.276 g, 30.8 mmol) in EtOH (100ml), was added 10% palladised charcoal (0.800g). Hydrogenation was carried out at atmospheric pressure and 50 °C with almost the theoretical number of moles of H_2 being adsorbed. The catalyst was removed by suction filtration (using celite as a filter aid) and EtOH was removed *in vacuo* to yield the hydrochloride salt of **34** (7.461g, 95 %) as an orange powder.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.85 (t, 3H, *J*=7.3 Hz, C*H*$_3$), 1.27 (s^6, 2H, *J*=7.4 Hz, CH_3CH_2), 1.45-1.54 (m, 2H, $CH_3CH_2CH_2$), 2.05-2.15 (m, 2H, CH_2CH_2N), 2.61 (t, 2H, *J*=7.3 Hz, $CH_2(CH_2)_2N$), 2.50 (t, 2H, *J*=7.7 Hz, $CH_3(CH_2)_2CH_2$), 2.69 (s, 6H, $N(CH_3)_2$), 2.87-2.91 (m, 2H, CH_2N), 6.99 and 7.03 (2 x d, 4H, *J*=8.3 Hz, C_6H_4), 12.10 (br s, 1H, *H*Cl). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.4 (CH_3), 22.7 (CH_3CH_2), 26.0 (CH_2CH_2N), 32.5 ($CH_3CH_2CH_2$), 34.0 ($ArCH_2$), 35.7 (CH_2Ar), 43.2 (NCH_3), 57.7 (CH_2N), 128.5 (Ar, *C*H), 129.2 (Ar, *C*H), 136.8 (Ar, C-1/4), 141.7 (Ar, C-1/4).

The hydrochloride (7.411 g, 29.1 mmol) was quenched with a satd. aq.Na_2CO_3 until a pH of approximately 10 was reached (confirmed by indicator paper). Extraction with EtOAc (3 x 50ml) and removal of the solvent *in vacuo* yielded **34** (6.301g, 99 %) as a brown oil.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.84 (t, 3H, *J*=7.3 Hz, C*H*$_3$), 1.28 (s^6, 2H, *J*=7.3 Hz, CH_3CH_2), 1.47-1.55 (m, 2H, $CH_3CH_2CH_2$), 1.66-1.76 (m, 2H, CH_2CH_2N), 2.14 (s, 6H, $N(CH_3)_2$), 2.19-2.24 (m, 2H, CH_2N), 2.49 (t, 2H, *J*=8.0 Hz, $ArCH_2$), 2.51 (t, 2H, *J*=8.1 Hz, $ArCH_2$), 7.07 and 7.20 (2 x d, 4H, *J*=8.0 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.4 (CH_3), 22.8 (CH_3CH_2), 29.9 (CH_2CH_2N), 33.7 ($CH_3CH_2CH_2$), 34.2 ($ArCH_2$), 35.7 (CH_2Ar), 45.9 (NCH_3), 59.8 (CH_2N), 128.6 (Ar, *C*H), 128.7 (Ar, *C*H), 139.8 (Ar, C-1/4), 142.8 (Ar, C-1/4). *m/z* (EI): 219 (M^+, 24%), 131 (33%), 117 (25%), 115 (21%), 91 (26%). *m/z* (CI): 220 ($[M+H]^+$, 100%), 206 (7%). *m/z* HRMS: Calcd. for $C_{15}H_{25}N$ (M^+) 219.1987. Found 219.1986.

_N_-[3-(4-butylphenyl)propyl]-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (12)

To a stirred solution of 1,3-propanesultone (1.961 g, 16.1 mmol) in EtOAc (30ml), was added a solution of **34** (3.136 g, 14.3 mmol) in EtOAc (30 ml). The reaction mixture was stirred for 253 h at r.t. The resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **12** (2.097g, 43 %) as an off-white crystalline solid.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.95 (t, 3H, *J*=7.4 Hz, C*H*$_3$), 1.36 (s^6, 2H, *J*=7.3 Hz, CH_3CH_2), 1.59 (q^5, 2H, *J*=7.6 Hz, $CH_3CH_2CH_2$), 2.08-2.24 (m, 4H, $CH_2CH_2SO_3$ and CH_2CH_2N), 2.59 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 2.70 (t, 2H, *J*=7.5 Hz, $ArCH_2$), 2.88 (t, 2H, *J*=6.8 Hz, CH_2SO_3), 3.09 (s, 6H, $N(CH_3)_2$), 3.31-3.36 (m, 2H, CH_2N), 3.52 (m, 2H, NCH_2), 7.14

and 7.20 (2 x d, 4H, *J*=8.1 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.7 (CH_3), 20.3 (NCH_2CH_2), 23.7 (CH_3CH_2), 25.7 (CH_2CH_2N), 33.2 ($CH_3CH_2CH_2$), 35.5 (CH_2Ar), 36.6 ($ArCH_2$), 49.0 (CH_2SO_3), 51.7 (NCH_3), 64.1 (NCH_2), 65.4 (CH_2N), 129.8 (Ar, CH), 130.2 (Ar, CH), 138.8 (Ar, C-1/4), 142.6 (Ar, C-1/4). *m/z* (ES^+): 683 ($[2M+H]^+$, 8%), 577 (10%), 358 (19%), 342 ($[M+H]^+$, 100%), 236 (93%), 220 (38%). *m/z* (ES^-): 741 ($[2M+AcO]^-$, 7%), 416 (45%), 400 ($[M+AcO]^-$, 100%), 301 (11%), 166 (24%), 130 (22%). *m/z* HRMS (ES^+): Calcd. for $C_{18}H_{32}NSO_3$ $[M+H]^+$ 342.2103. Found 342.2100.

N-[3-(4-hexylphenyl)propyl]-*N*,*N*-dimethyl-3-ammonio-1-propanesulfonate (13)

1-(4-hexylphenyl)propenone (31) and 3-chloro-1-(4-hexylphenyl)propan-1-one (33)

To a suspension of $AlCl_3$ (9.842 g, 74.0 mmol) in dry 1,2-dichloroethane (30 ml) at 0 °C under N_2 was added acryloyl chloride (3.7 ml, 4.122 g, 45.3 mmol) in dry 1,2-dichloroethane (20 ml). A solution of 1-phenylhexane (7.0 ml, 6.027 g, 37.2 mmol) in 1,2-dichloroethane (30 ml) was added dropwise to the resulting stirred mixture. After stirring in the dark for 6.75 h (while allowing the temperature to rise naturally from 0 °C to r.t), the reaction mixture was poured into a mixture of ice and water (50 ml) and hydrochloric acid (12N, 50 ml) and extracted with CH_2Cl_2 (3 x 100 ml). The combined organic extracts were washed with satd. aq.$NaHCO_3$ (3 x 100 ml) and brine (2 x 100 ml), and dried over anhydrous $MgSO_4$. The filtered solution was evaporated *in vacuo* to yield **31** (1.668g, 7.7 mmol) and **33** (4.324g, 17.2 mmol) in 67 % yield (31: 69 respectively) as a pale yellow powder.

31 ^1H NMR (400MHz, $CDCl_3$): δ = 0.81 (t, 3H, *J*=6.9 Hz, CH_3), 1.16-1.29 (m, 6H, $CH_3(CH_2)_3$), 1.55 (q^5, 2H, *J*=6.7 Hz, $CH_3(CH_2)_3CH_2$), 2.59 (t, 2H, *J*=6.3 Hz, $CH_3(CH_2)_4CH_2$), 5.83 (dd, 1H, *J*=1.8, 10.5 Hz, CH=CH*H*), 6.36 (dd, 1H, *J*=1.8, 17.1 Hz, CH=CH*H*), 7.10 (dd, 1H, *J*=10.5, 17.1 Hz, $CH=CH_2$), 7.20 and 7.80 (2 x d, 4H, *J*=8.3 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.9 (CH_3CH_2), 29.3 (CH_2), 29.7 (CH_2), 32.0 (CH_2), 36.4 (CH_2Ar), 129.0 (Ar, CH), 129.1 (Ar, CH), 129.8 (CH_2), 132.9 (CH), 135.4 (Ar, C-1/4), 149.7 (Ar, C-1/4), 190.9 (CO). *m/z* (EI): 216 (M^+, 7%), 189 (100%, 91 (22%). *m/z* (CI): 234 ($[M+NH_4]^+$, 100%), 217 ($[M+H]^+$, 73%), 189 (41%), 90 (46%). *m/z* HRMS: Calcd. for $C_{15}H_{20}O$ $[M]^+$ 216.1514. Found 216.1517.

33 ^{1}H NMR (400MHz, $CDCl_3$): δ = 0.81 (t, 3H, *J*=6.9 Hz, CH_3), 1.16-1.29 (m, 6H, $CH_3(CH_2)_3$), 1.55 (q^5, 2H, *J*=6.7 Hz, $CH_3(CH_2)_3CH_2$), 2.59 (t, 2H, *J*=6.3 Hz, $CH_3(CH_2)_4CH_2$), 3.37 (t, 2H, *J*=6.9 Hz, CH_2Cl), 3.85 (t, 2H, *J*=6.9 Hz, $COCH_2$), 7.20 and 7.80 (2 x d, 4H, *J*=8.3 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.9 (CH_3CH_2), 29.3 (CH_2), 29.7 (CH_2), 32.0 (CH_2), 36.4 (CH_2Ar), 39.2 (CH_2), 41.6 (CH_2), 128.6 (Ar, *C*H), 129.0 (Ar, *C*H), 134.6 (Ar, C-1/4), 149.1 (Ar, C-1/4), 196.7 (*C*O). *m/z* (EI): 252 (M^+, 3%), 189 (100%), 91 (22%). *m/z* (CI): 272 ($[M+NH_4]^+$, 25%), 253 ($[M+H]^+$, 18%), 189 (41%), 90 (46%). *m/z* HRMS: Calcd. for $C_{15}H_{21}ClO$ $[M]^+$ 252.1281. Found 252.1277.

3-dimethylamino-1-(4-hexylphenyl)propan-1-one (29)

To a solution of **31** (1.642 g, 7.6 mmol) and **33** (4.284 g, 17.0 mmol) in THF (30 ml), a solution of dimethylamine in THF (2M, 60 ml) was added drop-wise at 0 °C. The reaction mixture was stirred for 24 h at r.t. The THF was evaporated under reduced pressure. HCl (2M, 60 ml) was cautiously added followed by extraction with $CHCl_3$ (3 x 75 ml) and drying over anhydrous $MgSO_4$. Following filtration, the $CHCl_3$ was removed *in vacuo* to yield **29** as its hydrochloride salt (6.032g, 83 %) as a yellow/orange powder.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.81 (t, 3H, *J*=6.8 Hz, CH_3), 1.15-1.36 (m, 6H, $CH_3(CH_2)_3$), 1.55 (q^5, 2H, *J*=7.6 Hz, $CH_3(CH_2)_3CH_2$), 2.60 (t, 2H, *J*=7.7 Hz, $CH_3(CH_2)_4CH_2$), 2.78 (s, 6H, $N(CH_3)_2$), 3.45 (t, 2H, *J*=6.8 Hz, $COCH_2$), 3.65 (t, 2H, *J*=6.9 Hz, CH_2N), 7.20 and 7.83 (2 x d, 4H, *J*=8.2 Hz, C_6H_4), 12.57 (br s, 1H, *H*Cl; ^{13}C NMR (101MHz, $CDCl_3$): δ = 13.1 (CH_3), 21.5 (CH_3CH_2), 27.9 (CH_2), 30.0 (CH_2), 30.6 (CH_2), 32.7 ($COCH_2$), 35.0 (CH_2Ar), 42.3 (NCH_3), 51.7 (CH_2N), 127.4 (Ar, *C*H), 127.9 (Ar, *C*H), 132.1 (Ar, C-1/4), 149.2 (Ar, C-1/4), 194.4 (*C*O). *m/z* (EI): 261 (M^+, 100%). *m/z* (CI): 262 ($[M+H]^+$, 37%), 234 (18%), 217 (76%), 189 (20%). *m/z* HRMS: Calcd. for $C_{17}H_{27}NO$ $[M]^+$ 261.2093. Found 261.2092.

***N,N*-dimethyl-3-(4-hexylphenyl)propylamine (35)**

To a stirred solution of **29** (5.982 g, 20.1 mmol) in EtOH (100 ml), was added 10% palladised charcoal (0.600g). Hydrogenation was carried out at atmospheric pressure and 50 °C with almost the theoretical number of moles of H_2 being adsorbed. The catalyst was removed by

suction filtration (using celite as a filter aid) and EtOH was removed *in vacuo* to yield the hydrochloride salt of **35** (5.404 g, 95 %) as an orange powder.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.89 (t, 3H, *J*=6.4 Hz, CH_3), 1.26-1.40 (m, 6H, $CH_3(CH_2)_3$), 1.59 (q^5, 2H, *J*=7.1 Hz, $CH_3(CH_2)_3CH_2$), 2.18 (m, 2H, CH_2CH_2N), 2.57 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 2.70 (t, 2H, *J*=7.2 Hz, $ArCH_2$), 2.78 (s, 6H, $N(CH_3)_2$), 2.96-2.99 (m, 2H, CH_2N), 7.08 and 7.11 (2 x d, 4H, *J*=8.0 Hz, C_6H_4), 12.27 (br s, 1H, *H*Cl). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.5 (CH_3), 23.0 (CH_3CH_2), 26.0 (CH_2CH_2N), 29.4 (CH_2), 31.9 (CH_2), 32.1 (CH_2), 32.5 ($ArCH_2$), 35.9 (CH_2Ar), 43.3 (NCH_3), 57.8 (CH_2N), 128.6 (Ar, *C*H), 129.2 (Ar, *C*H), 136.7 (Ar, C-1/4), 141.7 (Ar, C-1/4). *m/z* (EI): 247 (M^+, 100%), 202 (18%), 189 (14%), 175 (16%). *m/z* (CI): 248 ($[M+H]^+$, 100%), 131 (3%). *m/z* HRMS: Calcd. for $C_{17}H_{29}N$ $[M]^+$ 247.2300. Found 247.2300.

The hydrochloride salt (5.310g, 18.8mmol) was quenched with a satd. aq.Na_2CO_3 until a pH of approximately 10 was reached (confirmed by indicator paper). Extraction with EtOAc (3 x 50 ml) and removal of the solvent *in vacuo* yielded **35** (4.495 g, 97%) as a brown oil.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.81 (t, 3H, *J*=6.7 Hz, CH_3), 1.17-1.29 (m, 6H, $CH_3(CH_2)_3$), 1.52 (q^5, 2H, *J*=7.5 Hz, $CH_3(CH_2)_3CH_2$), 1.71 (q^5, 2H, *J*=7.7 Hz, CH_2CH_2N), 2.15 (s, 6H, $N(CH_3)_2$), 2.23 (t, 2H, *J*=7.5 Hz, CH_2N), 2.49 (t, 2H, *J*=7.9 Hz, $ArCH_2$), 2.52 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 7.02 (s, 4H, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.5 (CH_3), 23.0 (CH_3CH_2), 29.9 (CH_2CH_2N), 30.0 (CH_2), 32.0 (CH_2), 33.0 (CH_2), 33.7 ($ArCH_2$), 36.0 (CH_2Ar), 45.9 (NCH_3), 59.8 (CH_2N), 128.6 (Ar, *C*H), 128.7 (Ar, *C*H), 139.8 (Ar, C-1/4), 140.7 (Ar, C-1/4).

_N_-[3-(4-hexylphenyl)propyl]-_N,N_-dimethyl-3-ammonio-1-propanesulfonate (13)

Prepared using the procedure described above starting from **35** (2.090 g, 8.5 mmol). The product **13** was obtained as an off-white crystalline solid (1.286 g, 41 %).

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.89 (t, 3H, *J*=6.7 Hz, CH_3), 1.30-1.34 (m, 6H, $CH_3(CH_2)_3$), 1.58 (q^5, 2H, *J* 7.3, $CH_3(CH_2)_3CH_2$), 2.06-2.17 (m, 4H, $CH_2CH_2SO_3$ and CH_2CH_2N), 2.56 (t, 2H, *J*=7.4 Hz, $ArCH_2$), 2.68 (t, 2H, *J*=7.5 Hz, $ArCH_2$), 2.86 (t, 2H, *J*=6.8 Hz, CH_2SO_3), 3.07

(s, 6H, $N(CH_3)_2$), 3.28-3.34 (m, 2H, CH_2N), 3.50 (m, 2H, NCH_2), 7.11 and 7.17 (2 x d, 4H, J=8.1 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.4 (CH_3), 19.8 (NCH_2CH_2), 23.7 (CH_3CH_2), 25.3 (CH_2CH_2N), 30.0 (CH_2), 32.8 (CH_2), 32.8 (CH_2), 32.9 ($ArCH_2$), 36.5 (CH_2Ar), 48.6 (CH_2SO_3), 51.3 (NCH_3), 63.7 (NCH_2), 65.0 (CH_2N), 129.4 (Ar, CH), 129.7 (Ar, CH), 138.4 (Ar, C-1/4), 142.2 (Ar, C-1/4). *m/z* (ES^+): 761 ($[2M+Na]^+$, 2%), 739 ($[2M+H]^+$, 10%), 392 ($[M+Na]^+$, 11%), 370 ($[M+H]^+$, 100%), 248 (13%). *m/z* (ES^-): 797 ($[2M+AcO]^-$, 8%), 526 (8%), 428 ($[M+AcO]^-$, 100%), 404 ($[M+Cl]^-$, 8%), 315 (15%), 157 (45%). *m/z* HRMS (ES^+): Calcd. for $C_{20}H_{36}NSO_3$ $[M+H]^+$ 370.2416. Found 370.2420.

Butane Sulfobetaines 14–22

***N*-hexyl-*N,N*-dimethyl-4-ammonio-1-butanesulfonate (14)**

To a stirred solution of 1,4-butanesultone (8.941 g, 65.7 mmol) in EtOAc (30 ml), was added *N,N*-dimethylhexylamine (9.470 g, 73.4 mmol) in EtOAc (30 ml). The reaction mixture was stirred under reflux for 44 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **14** (10.906 g, 63 %) as a white crystalline solid; mp 260-2 °C. ^1H NMR (400MHz, CD_3OD): δ = 0.90-0.96 (m, 3H, CH_3), 1.34-1.41 (m, 6H, $CH_3(CH_2)_3$), 1.79 (m, 2H, CH_2CH_2N), 1.84 (q^5, 2H, J=7.1 Hz, $CH_2CH_2SO_3$), 1.90-1.99 (m, 2H, NCH_2CH_2), 2.89 (t, 2H, J=7.1 Hz, CH_2SO_3), 3.07 (s, 6H, $N(CH_3)_2$), 3.29 (m, 2H, CH_2N), 3.34 (m, 2H, NCH_2). ^{13}C NMR (101MHz, CD_3OD): δ = 14.4 (CH_3), 22.3 (CH_2), 23.0 (CH_2), 23.5 (CH_2), 23.5 ($CH_2CH_2SO_3$), 27.1 (NCH_2CH_2), 32.4 (CH_2CH_2N), 51.2 (NCH_3), 51.3 (CH_2SO_3), 64.9 (NCH_2), 65.6 (CH_2N). *m/z* (EI): 265 (M^+, 2%), 264 ($[M-H]^+$, 15%), 194 (40%), 180 (20%), 137 (30%), 129 (100%), 128 (80%). *m/z* (CI): 266 ($[M+H]^+$, 2%), 252 (3%), 170 (7%), 168 (4%), 154 (79%), 131 (13%), 130 (100%). *m/z* HRMS: Calcd. for $C_{12}H_{28}NSO_3$ $[M+H]^+$ 266.1790. Found 266.1788.

***N*-heptyl-*N,N*-dimethyl-4-ammonio-1-butanesulfonate (15)**

Prepared using the procedure described above starting from *N,N*-dimethylheptylamine (9.545 g, 66.7 mmol). The product **15** was obtained as a white crystalline solid (13.566 g, 73 %); mp 252-3 °C.

^{1}H NMR (400MHz, CD_3OD): δ = 0.81 (t, 3H, *J*=6.8 Hz, C*H*$_3$), 1.20-1.30 (m, 8H, $CH_3(CH_2)_4$), 1.61-1.68 (m, 2H, C*H*$_2$CH$_2$N), 1.72 (q^5, 2H, *J*=7.2 Hz, C*H*$_2$CH$_2$SO$_3$), 1.79-1.86 (m, 2H, NCH$_2$C*H*$_2$), 2.76 (t, 2H, *J*=7.2 Hz, C*H*$_2$SO$_3$), 2.95 (s, 6H, N(C*H*$_3$)$_2$), 3.18 (m, 2H, C*H*$_2$N), 3.23 (m, 2H, NC*H*$_2$). ^{13}C NMR (101MHz, CD_3OD): δ = 14.3 (*C*H$_3$), 22.3 (*C*H$_2$), 23.0 (*C*H$_2$), 23.5 (*C*H$_2$), 23.5 (*C*H$_2$CH$_2$SO$_3$), 27.3 (NCH$_2$*C*H$_2$), 29.9 (*C*H$_2$), 32.7 (*C*H$_2$CH$_2$N), 51.0 (N*C*H$_3$), 51.1 (*C*H$_2$SO$_3$), 64.8 (N*C*H$_2$), 65.4 (*C*H$_2$N). *m/z* (EI): 279 (M$^+$, 1%), 278 ([M-H]$^+$, 2%), 194 (5%), 180 (5%), 144 (18%), 143 (100%), 142 (30%). *m/z* (CI): 280 ([M+H]$^+$, 2%), 266 (1%), 196 (2%), 184 (11%), 154 (59%), 144 (100%), 130 (33%). *m/z* HRMS: Calcd. for $C_{13}H_{30}NSO_3$ [M+H]$^+$ 280.1946. Found 280.1942.

_N_-octyl-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (16)

To a stirred solution of 1,4-butanesultone (7.696 g, 56.6 mmol) in EtOAc (20 ml), was added a solution of *N,N*-dimethyloctylamine (8.881 g, 56.6 mmol) in EtOAc (20 ml). The reaction mixture was stirred under reflux for 2.5 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* yielding **16** (10.465g, 63%) as a white crystalline solid; mp 251-2 °C.

^{1}H NMR (400MHz, CD_3OD): δ = 0.81 (t, 3H, *J*=6.8 Hz, C*H*$_3$), 1.18-1.32 (m, 10H, $CH_3(CH_2)_5$), 1.62-1.70 (m, 2H, C*H*$_2$CH$_2$N), 1.73 (q^5, 2H, J=7.2 Hz, C*H*$_2$CH$_2$SO$_3$), 1.80-1.88 (m, 2H, NCH$_2$C*H*$_2$), 2.78 (t, 2H, *J*=7.2 Hz, C*H*$_2$SO$_3$), 2.98 (s, 6H, N(C*H*$_3$)$_2$), 3.20 (m, 2H, C*H*$_2$N), 3.25 (m, 2H, NC*H*$_2$). ^{13}C NMR (101MHz, CD_3OD): δ = 14.4 (*C*H$_3$), 22.3 (*C*H$_2$), 23.0 (*C*H$_2$), 23.5 (*C*H$_2$), 23.6 (*C*H$_2$CH$_2$SO$_3$), 27.4 (NCH$_2$*C*H$_2$), 30.2 (*C*H$_2$), 30.2 (*C*H$_2$), 32.8 (*C*H$_2$CH$_2$N), 51.1 (N*C*H$_3$), 51.4 (*C*H$_2$SO$_3$), 64.8 (N*C*H$_2$), 65.4 (*C*H$_2$N). *m/z* (EI): 293 (M$^+$, 1%), 292 ([M-H]$^+$, 32%), 280 (8%), 194 (34%), 180 (50%), 156 (100%). *m/z* (CI): 294 ([M+H]$^+$, 3%), 280 (4%), 198 (15%), 196 (10%), 158 (100%), 154 (54%), 144 (34%). *m/z* HRMS: Calcd. for $C_{14}H_{32}NSO_3$ [M+H]$^+$ 294.2103. Found 294.2100.

_N_-benzyl-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (17)

To a stirred solution of 1,4-butanesultone (6.270 g, 46.1 mmol) in EtOAc (70 ml), was added *N,N*-dimethylbenzylamine (6.400 g, 40.8 mmol) in EtOAc (70 ml). The reaction mixture was

stirred under reflux for 48 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **17** (4.875 g, 44 %) as a white crystalline solid; mp 262-4 °C (lit.[67]mp 285-6 °C).

^{1}H NMR (400MHz, CD_3OD): δ = 1.85 (q^5, 2H, *J*=7.5 Hz, $CH_2(CH_2)_2SO_3$), 2.08 (m, 2H, $CH_2CH_2SO_3$), 2.90 (t, 2H, *J*=6.9 Hz, CH_2SO_3), 3.02 (s, 6H, $N(CH_3)_2$), 3.36 (m, 2H, NCH_2), 4.53 (s, 2H, $PhCH_2N$), 7.50-7.60 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 22.4 (NCH_2CH_2), 23.0 ($CH_2CH_2SO_3$), 50.3 (NCH_3), 51.3 (CH_2SO_3), 65.2 (NCH_2), 68.7 (CH_2N), 128.9 (Ph, C-1), 130.3 (Ph, *C*H), 131.8 (Ph, *C*H), 134.1 (Ph, *C*H). *m/z* (ES^+): 565 ($[2M+Na]^+$, 28%), 543 ($[2M+H]^+$, 8%), 294 ($[M+Na]^+$, 100%), 272 ($[M+H]^+$, 47%). *m/z* (ES^-): 346 (60%), 316 (63%), 270 ($[M-H]^-$, 100%). *m/z* HRMS (ES^+): Calcd. for $C_{13}H_{30}NSO_3$ $[M+H]^+$ 272.1320. Found 272.1322.

N-phenylethyl-N,N-dimethyl-4-ammonio-1-butanesulfonate (18)

To a stirred solution of 1,4-butanesultone (5.607 g, 41.2 mmol) in EtOAc (50 ml), was added *N,N*-dimethylphenethylamine (6.215 g, 41.7 mmol) in EtOAc (50 ml). The reaction mixture was stirred under reflux for 117 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **18** (5.951 g, 51 %) as a white crystalline solid; mp 277-8 °C (lit. mp 274-5 °C).

^{1}H NMR (400MHz, CD_3OD): δ = 1.84 (q^5, 2H, *J*=7.3 Hz, $CH_2(CH_2)_2SO_3$), 1.98 (m, 2H, $CH_2CH_2SO_3$), 2.88 (t, 2H, *J*=7.2 Hz, CH_2SO_3), 3.08-3.13 (m, 2H, $PhCH_2$), 3.16 (s, 6H, $N(CH_3)_2$), 3.42 (m, 2H, NCH_2), 3.52 (m, 2H, CH_2N), 7.24-7.36 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 22.3 (NCH_2CH_2), 23.0 ($CH_2CH_2SO_3$), 29.8 ($ArCH_2$), 51.2 (CH_2SO_3), 51.3 (NCH_3), 64.9 (NCH_2), 66.1 (CH_2N),128.3 (Ph, *C*H), 129.9 (Ph, *C*H), 130.1 (Ph, *C*H), 137.1 (Ph, C-1). *m/z* (EI): 285 (M^+, 18%), 284 ($[M-H]^+$, 100%), 231 (50%). *m/z* (CI): 286 ($[M+H]^+$, 64%), 272 (100%). *m/z* HRMS (ES^+): Calcd. for $C_{14}H_{32}NSO_3$ $[M+H]^+$ 286.1477. Found 286.1481.

_N_-phenylpropyl-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (19)

To a stirred solution of 1,4-butanesultone (2.180 g, 16.0 mmol) in EtOAc (20 ml), was added a solution of **24** (1.408 g, 8.6 mmol) in EtOAc (20 ml). The reaction mixture was stirred under reflux for 156 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **19** (1.390 g, 54 %) as a white crystalline solid; mp 261-2 °C.

^{1}H NMR (400MHz, CD_3OD): δ = 1.77-1.83 (m, 4H, $(CH_2)_2CH_2SO_3$), 2.03-2.15 (m, 2H, $PhCH_2CH_2$), 2.71 (t, 2H, *J*=7.5 Hz, CH_2SO_3), 2.84 (t, 2H, *J*=6.8 Hz, $PhCH_2$), 3.05 (s, 6H, $N(CH_3)_2$), 3.27-3.34 (m, 4H, CH_2NCH_2), 7.19-7.32 (m, 5H, C_6H_5). ^{13}C NMR (101MHz, CD_3OD): δ = 22.2 (NCH_2CH_2), 22.9 ($CH_2CH_2SO_3$), 25.3 (CH_2CH_2N), 33.1 ($ArCH_2$), 51.2 (CH_2SO_3), 51.3 (NCH_3), 64.7 (NCH_2), 64.7 (CH_2N), 127.5 (Ph, *C*H), 129.5 (Ph, *C*H), 129.7 (Ph, *C*H), 141.3 (Ph, C-1). *m/z* (EI): 299 (M^+, 40%), 298 ($[M-H]^+$, 100%), 254 (24%), 246 (28%). *m/z* (CI): 300 ($[M+H]^+$, 58%), 268 (44%), 252 (100%), 238 (55%), 222 (53%). *m/z* HRMS (ES^+): Calcd. for $C_{15}H_{26}NSO_3$ $[M+H]^+$ 300.1633. Found 300.1634.

_N_-phenylbutyl-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (20)

To a stirred solution of 1,4-butanesultone (1.000 g, 7.4 mmol) in EtOAC (25 ml), was added a solution of **26** (1.126 g, 6.9 mmol) in EtOAc (25 ml). The reaction mixture was stirred under reflux for 164 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether, and dried *in vacuo* to yield **20** (0.618 g, 29 %) as a white crystalline solid; mp 243-6 °C.

^{1}H NMR (250MHz, CD_3OD): δ = 1.67-1.94 (m, 8H, $PhCH_2(CH_2)_2$ and $(CH_2)_2CH_2SO_3$), 2.71 (t, 2H, *J*=7.2 Hz, $PhCH_2$), 2.86 (t, 2H, *J*=7.0 Hz, CH_2SO_3), 3.04 (s, 6H, $N(CH_3)_2$), 3.27-3.33 (m, 4H, CH_2NCH_2), 7.17-7.28 (m, 5H, C_6H_5). ^{13}C NMR (63MHz, CD_3OD): δ = 22.3 (NCH_2CH_2), 22.9 ($ArCH_2CH_2$), 23.0 ($CH_2CH_2SO_3$), 29.1 (CH_2CH_2N), 36.0 ($ArCH_2$), 51.2 (NCH_3), 51.3 (CH_2SO_3), 64.8 (NCH_2), 65.2 (CH_2N), 127.1 (Ph, *C*H), 129.5 (Ph, *C*H), 129.5 (Ph, *C*H), 142.8 (Ph, C-1). *m/z* (EI): 313 (M^+, 54%), 312 ($[M-H]^+$, 100%), 234 (60%), 218 (99%). *m/z* (CI): 314 ($[M+H]^+$, 66%), 300 (100%). HRMS (ES^+): Calcd. for $C_{16}H_{28}NSO_3$ $[M+H]^+$ 314.1790. Found 314.1787.

_N_-[3-(4-butylphenyl)propyl]-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (21)

To a stirred solution of 1,4-butanesultone (2.186 g, 16.1 mmol) in EtOAC (30 ml), was added a solution of **34** (3.136 g, 14.3 mmol) in EtOAc (30 ml). The reaction mixture was stirred under reflux for 253 h. After being allowed to cool the resulting white precipitate was filtered, washed with diethyl ether and dried *in vacuo* to yield **21** (1.980 g, 39 %) as an off-white crystalline solid; mp 202-4 °C.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.93 (t, 3H, *J*=7.3 Hz, CH_3), 1.34 (s^6, 2H, *J*=7.4 Hz, CH_3CH_2), 1.57 (q^5, 2H, *J*=7.6 Hz, $CH_3CH_2CH_2$), 1.78-1.90 (m, 4H, $(CH_2)_2CH_2SO_3$), 2.01-2.17 (m, 2H, CH_2CH_2N), 2.58 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 2.67 (t, 2H, *J*=7.4 Hz, $ArCH_2$), 2.85 (t, 2H, *J*=6.6 Hz, CH_2SO_3), 3.04 (s, 6H, $N(CH_3)_2$), 3.27-3.35 (m, 4H, NCH_2), 7.12 and 7.17 (2 x d, 2H, *J*=8.0 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.3 (CH_3), 22.2 (NCH_2CH_2), 23.0 ($CH_2CH_2SO_3$), 23.3 (CH_3CH_2), 25.4 (CH_2CH_2N), 32.7 ($CH_3CH_2CH_2$), 35.0 (CH_2Ar), 36.2 ($ArCH_2$), 51.3 (CH_2SO_3), 51.3 (NCH_3), 64.7 (NCH_2), 65.8 (CH_2N), 129.4 (Ar, *C*H), 129.8 (Ar, *C*H), 138.5 (Ar, C-1/4), 141.2 (Ar, C-1/4). *m/z* (ES^+): 711 ($[2M+H]^+$, 10%), 372 (7%), 356 ($[M+H]^+$, 100%). *m/z* (ES^-): 769 ($[2M+AcO]^-$, 6%), 430 (45%), 414 ($[M+AcO]^-$, 100%), 180 (20%). *m/z* HRMS (ES^+): Calcd. for $C_{19}H_{34}NSO_3$ $[M+H]^+$ 356.2259. Found 356.2263.

_N_-[3-(4-hexylphenyl)propyl]-_N,N_-dimethyl-4-ammonio-1-butanesulfonate (22)

Prepared using the procedure described above starting from **35** (2.090 g, 8.5 mmol). The product **22** was obtained as an off-white crystalline solid (1.237 g, 38 %); mp 219-221 °C.

^{1}H NMR (400MHz, $CDCl_3$): δ = 0.89 (t, 3H, *J*=6.5 Hz, CH_3), 1.27-1.38 (m, 6H, $CH_3(CH_2)_3$), 1.54-1.62 (m, 2H, $CH_3(CH_2)_3CH_2$), 1.75-1.91 (m, 4H, $(CH_2)_2CH_2SO_3$), 2.03-2.20 (m, 2H, CH_2CH_2N), 2.57 (t, 2H, *J*=7.7 Hz, $ArCH_2$), 2.67 (t, 2H, *J*=7.4 Hz, $ArCH_2$), 2.84 (t, 2H, *J*=6.8 Hz, CH_2SO_3), 3.04 (s, 6H, $N(CH_3)_2$), 3.27-3.34 (m, 4H, NCH_2), 7.12 and 7.17 (2 x d, 4H, *J*=8.0 Hz, C_6H_4). ^{13}C NMR (101MHz, $CDCl_3$): δ = 14.4 (CH_3), 22.2 (NCH_2CH_2), 23.0 ($CH_2CH_2SO_3$), 23.7 (CH_3CH_2), 25.4 (CH_2CH_2N), 30.0 (CH_2), 32.8 (CH_2), 32.8 (CH_2), 32.9 ($ArCH_2$), 36.5 (CH_2Ar), 51.3 (CH_2SO_3), 51.3 (NCH_3), 64.6 (NCH_2), 64.8 (CH_2N), 129.4 (Ar, *C*H), 129.8 (Ar, *C*H), 138.5 (Ar, C-1/4), 142.3 (Ar, C-1/4). *m/z* (ES^+): 789 ($[2M+Na]^+$, 10%),

767 ($[2M+H]^+$, 10%), 406 ($[M+Na]^+$, 11%), 384 ($[M+H]^+$, 100%). *m/z* (ES^-): 825 ($[2M+AcO]^-$, 6%), 442 ($[M+AcO]^-$, 100%), 418 ($[M+Cl]^-$, 7%), 363 (3%), 151 (13%). *m/z* HRMS (ES^+): Calcd. for $C_{21}H_{38}NSO_3$ $[M+H]^+$ 384.2572. Found 384.2576.

2.8 References

1. Brühl, J. W., *Berichte*, **9** (1976) 34.
2. Scheibler, C., *Berichte*, **2** (1869) 292.
3. Scheibler, C., *Berichte*, **3** (1870) 155.
4. Domingo, X., Chapter 3: Betaines. *In Surfactant Science Series*, Vol. **59**, *Amphoteric Surfactants*, 2nd Edition, edited by Eric G. Lomax, Marcel Dekker Inc. (1996).
5. Ernst, R. and Miller, E. J., Chapter 2: Surface-Active Betaines. *In Surfactant Science Series*, Vol. **12**, *Amphoteric Surfactants*, edited by Bluestein and Hilton, Marcel Dekker Inc. (1982).
6. Ernst, R., Patent 3, 280, 179 to Textilana corp. (1966).
7. Parris, N., Weil, J. K. and Linfield, W. M., *J. Am. Oil. Chem. Soc.*, **50** (1973) 509.
8. Ohme, R., Ballschuh, D. and Seibt, H., *Tenside Surf. Det.*, **28** (1991) 235.
9. Parris, N., Weil, J. K. and Linfield, W. M., *J. Am. Oil. Chem. Soc.*, **53** (1976) 97.
10. Parris, N., Weil, J. K. and Linfield, W. M., *J. Am. Oil. Chem. Soc.*, **53** (1976) 60.
11. Parris, N., Pierce, C. and Linfield, W. M., *J. Am. Oil. Chem. Soc.*, **54** (1977) 294.
12. Wilkes, B. G. and Wickert, J. N., *Ind. Eng. Chem.*, **29** (1937) 1234.
13. Distler, H. and Widder, R., *Tenside*, **6** (1969) 241.
14. Chevalier, Y., Storet, Y., Pourchet and Le Perchec, P., *Langmuir*, **7** (1991) 854.
15. Lietz, G., *Parfuem. Kosmet.*, **48** (1967) 159.
16. U. S. Patent 2, 781, 349 to H. S. Mannheimer (1957).
17. U. S. Patent 2, 781, 358 to H. S. Mannheimer (1957).
18. Mannheimer, H. S., *Am. Perfum.*, **76** (1961) 36.
19. Garcia Dominguez, J., Balaguer, F., Parra, J. L. and Pelejero, C., *Int. Cosmetic Sci.*, **3** (1981) 57.
20. U. S. Patent 3, 755, 559 to Colgate-Palmolive Co. (1973).
21. U. S. Patent 3, 769, 398 to Colgate-Palmolive Co. (1973).
22. U. S. Patent 4, 166, 845 to Colgate-Palmolive Co. (1979).
23. Eur. Patent 0 137 178 to Beecham Group Plc. (1985).
24. Beh, H. H. and James, K. C., *Cosmet. Toiletries*, **92** (1977) 21.

25. U. S. Patent 4, 279, 891 to American Cyanamid. Co. (1981).

26. U. S. Patent 4, 343, 785 to BASF Wyandotte Corp. (1982).

27. Eur. Patent 0 480 812 to Colgate-Palmolive Co. (1992).

28. Jap. Patent 88/171620 to Y. Yokoi and M. Kshima (1988).

29. Jap. Patent 62. 102. 760 to Aikoh Co. Ltd. (1987).

30. Brit. Patent 2 219 594 to Kao Corporation (1989).

31. Brit. Patent 2 245 281 to Kao Corporation (1992).

32. Eur. Patent 0 399 909 to Sanofi (1990).

33. U. S. Patent 4, 486, 195 to Millmaster Onyx Group Inc. (1984).

34. Ger. Patent 30 31 572 to Produits Chimiques de la Montagne Noire (1981).

35. Riva, A. and Cegarra, J., *J. Soc. Dyers Colour*, **103** (1987) 32.

36. Parra, J. L., Garcia Dominguez, J. J., de la Maza, A., Sanchez, J. and Comelles, F., *Invest. Inf. Text. Tensioactivos*, **26** (1983) 205.

37. Ger. Patent 22 44 905 to BP Chemicals International Co., Ltd. (1973).

38. U. S. Patent 3, 544, 501 to Geigy Chemical Corp. (1970).

39. Jap. Patent 59. 057. 399 to Japan Synthetic Rubber Co., Ltd (1984).

40. U. S. Patent 5, 064, 888 to National Starch and Chemical Investment holding Corp. (1991).

41. Eur. Patent 0 347 930 to Nippon Paint Co., Ltd (1989).

42. U. S. Patent 2, 835, 582 to Eastman Kodak Co. (1958).

43. Fr. Patent 1. 489. 520 to Eastman Kodak Co. (1967).

44. Jap. Patent 74/08100 to Konishiroku Photo Industry Co., Ltd (1974).

45. Jap. Patent 61. 203. 452 to Mitsubishi Paper Mills, Ltd. (1986).

46. Ger. (East) Patent 258 981 to Akademie der Wissenschaften der DDR (1988).

47. Fr. Patent 2 569 188 to Institut Français du Petrole and to Centre National de la Recherche Scientifique (1986).

48. U. S. Patent 4, 744, 796 to ARCO Chemical Co. (1988).

49. Ger. Patent 31 43 362 to BASF AG (1983).

50. U. S. Patent 2, 697, 656 to California Research Corp. (1954).

51. U. S. Patent 2, 697, 116 to California Research Corp. (1954).

52. U. S. Patent 4, 383, 846 to Occidental Research Corp. (1983).

53. Brit. Patent 2 196 846 to Mine Safety Appliances Co. (1988).

54. U. S Patent 2, 435, 967 to Westinghouse Electric Corp. (1948).

55. Distler, H. and Widder, R., *Tenside Surf. Det.*, **27** (1990) 417.

56. Suling, W. J. and O' leary, W. M., *Antimicrob. Agents Chemother.*, (1975) 334.

57. Michaels, E. B., Brit. Patent 1, 504, 837 (1978).

58. Kenney, D., *Cosmet. Toilet.*, **94** (1979) 13.

59. Gonenne, A. and Ernst, R., *Anal. Biochem.*, **87** (1978) 28.

60. Allen, J. C. and Humphries, C., *FEBS Lett.*, **57** (1975) 158.

61. Malpartida, F. and Serrano, R., *FEMS Microbiol. Lett.*, **111** (1980) 169.

62. Robinson, A. and Manchee, R. J., *FEMS Microbiol. Lett.*, **5** (1979) 131.

63. Ger. (East) Patent 0 242 716 to Akademie der Wissenschaften der DDR and Institute fur Chemische Technologie (1987).

64. Brit. Patent 2 230 444 to John Ritchie-Swanson and Joan Ritchie-Swanson (1990).

65. Eur. Patent 0 004 686 to Procter and Gamble Co. (1979).

66. Parvin, R., Goswami, T. and Pande, S. V., *Can. J. Biochem.*, **58** (1980) 822.

67. Spencer, T. A., Onofrey, T. J., Cann, R. O., Russel, J. S., Lee, L. E., Blanchard, D. E., Castro, A, Gu, P., Jiang, G. and Schechter, I., *J. Org. Chem.*, **64** (1999) 807.

68. James, J. W., *J. Chem. Soc.*, **47** (1885) 367.

69. Le Berre, A. and Delacroix, A., *Bull. Soc. Chim. Fr.*, **7/8** (1973) 2404.

70. Barnhurst, J. D., *J. Org. Chem.*, **26** (1961) 4520.

71. Le Berre, A., Etienne, A. and Dumaitre, B., *Bull. Soc. Chim. Fr.*, **3** (1970) 954.

72. Rondestvedt, C. S., *J. Am. Chem. Soc.*, **76** (1954) 1926.

73. King, J. F., Hillhouse, J. H. and Skonieczny, S., *Can. J. Chem.*, **62** (1984) 1977.

74. Okahata, Y., Tanamachi, S., Nagai, M. and Kunitake, T., *J. Colloid Interface Sci.*, **82** (1981) 401.

75. Eur. Patent 0 293 912 to Rashig AG (1988).

76. Eur. Patent 0 293 913 to Rashig AG (1988).

77. Roberts, D. W., Personal Communication.

78. Weers, J. G., Rathman, J. F., Axe, F. U., Crichlow, C. A., Foland, L. D., Scheuing, D. R., Wiersema, R. J. and Zielske, A. G., *Langmuir*, 7 (1991) 854.

79. U. S. Patent to Textilana Corp. (1966).

80. Roberts, D. W. and Williams, D. L., *Tetrahedron*, **43** (1987) 1027.

81. Essaddam, H., Pichot, C. and Guyot, A., *Colloid Polym. Sci.*, **266** (1988) 462.

82. Anton, P. and Laschewsky, A., *Makromol. Chem., Rapid Commun.*, **12** (1991) 189.

83. Parris, N., Weil, J. K. and Linfield, W. M., *J. Am. Oil. Chem. Soc.*, **53** (1976) 60.

84. Roberts, D. W., Heath, S. J. and Ward, C. J., Research Report, Unilever Port Sunlight (April 1969).

85. Saitoh, T. and Hinze, W. L., *Anal. Chem.*, **63** (1991) 2520.

86. Schindlbauer, H., *Montash. Chem.*, **99** (1968) 1799.

87. Ronald, R. C., Winkle, M. R., *Tetrahedron*, **39** (1983) 2031.

88. Schiemenz, G. P., Stein, G., *Tetrahedron*, **26** (1970) 2007.

89. Sucrow, W., *Chem. Ber.*, **105** (1972) 1621.

90. Majo, V. J., Perumal, P. T., *Tetrahedron Lett.*, **37** (1996) 5015.

91. Grinev, A. N., *Chem. Heterocycl. Compd.*, **13** (1977) 621.

92. Singer, M., Ryan, W. J., Saha, B., Martin, B. R., Razdan, R. K., *J. Med. Chem.*, **22** (1998) 4400.

93. Majumdar, G., Mal, D., *Indian J. Chem.*, Sect. B, **33** (1994) 700.

94. Baker, R. W., Hockless, D. C. R., Pocock, G. R., Sargent, M. V., Skelton, B. W., *J. Chem. Soc., Perkin Trans. 1*, (1995) 2615.

95. Chan, C., Scicinski, J. J., Srikantha, A. R., Watson, N. S., *Tetrahedron Lett.*, **37** (1996) 8925.

96. ZemLicka, J., Fernandez-Moyano, M. C., Ariatti, M., Zurenko, G. E., Grady, J. E., Ballesta, J. P. G., *J. Med. Chem.*, **36** (1993) 1239.

97. Nishioka, T., Watanabe, J., Kawabata, J., Niki, R., *Biosci. Biotechnol. Biochem.*, **61** (1997) 1138.

98. Haydock, D. B., Mulholland, T. C., Telford, B., Thorp, J. M., Wain, J. S., *Eur. J. Med. Chem.*, **19** (1984) 205.

99. Linden, A., Wright, A. D., Koenig, G. M., *Acta Crystallogr.*, Sect. C., **51** (1995) 744.

100. Prashad, M., Har, D., Repic, O., Blacklock, T. J., Chin, J. A., Shapiro, M. J., *Tetrahedron Lett.*, **38** (1997) 1313.

101. Challis, B. C. and Challis, J., *Reactions of the Carboxamide Group*, In: The Chemistry of Amides, The Chemistry of Functional Groups, edited by Zabicky, J., Interscience Publishers (1970) 731.

102. Miyano, S., Lu, L. D. L., Viti, S. M., Sharpless, K. B., *J. Org. Chem.*, **50** (1985) 4350.

103. Maitland and Jones, Jr., Organic Chemistry, W. W. Norton and Company (1997).

104. Bodendorf, K., Koralewski, G., *Arch. Pharm.*, **271** (1933) 101.

105. Bodwell, G. J., Li, J., Miller, D. O., *Tetrahedron*, **55** (1999) 12939.

106. Knoll, W., Koeberle, P., Leifield, F., Schäfer, V., *Sulfobetaines, Ethersulfonates and Polymerisable Surfactants; New Surfactant Structures Via Sulfopropylation Reactions*, Poster Presentation at 5th World Surfactants Congress, Florence, 29th May – 2nd June 2000.

107. Le Berre, A., Etienne, A., Delacroix, A., Proust, A., *Bull. Soc. Chim. Fr.*, (1975) 2531.

108. Fendler, E. J., Day, C. L., Fendler, J. H., *J. Phys. Chem.*, **76** (1972) 1460.

109. Cook, M. J., Khan, T. A., Nassi, K., *Tetrahedron*, **42** (1986) 249.

Chapter 3

Determination of Log P

3.1 Introduction

Hydrophobicity, in the form of log P, is the most common physicochemical property used for the development of QSARs for biological activity. Log P has also proved adequate for use in a single parameter approach to QSARs for aquatic toxicity, and this approach has been validated over many years through support by many authors. There are now a number of log P-based QSARs for aquatic toxicity that are commonly used in the early stages of environmental risk assessment, for screening potentially new chemicals for use in commercial products. For example, Unilever use such QSARs for predicting aquatic toxicity of anionic and non-ionic compounds at the screening level, for use as surfactants. The ultimate aim is therefore to extend this approach to cover all classes of surfactant. This chapter describes the development of a reliable method for log P determination of the 22 sulfobetaines synthesised (as described in Chapter 2) for use in QSAR correlation studies.

There are a number of methods that may be employed to determine log P. Calculation methods were found not to be applicable for zwitterionic compounds (containing both N^+ and SO_3^-) and so it was decided that log P would be determined experimentally. The aim would be to use the experimental results to amend the log P calculation method to include sulfobetaines.

From an experimental point of view, there are a number of potential methods, previously described in Chapter 1, of measuring a partition coefficient such as log P. From these the conventional *stir-flask experiment* was chosen, since it is a well-established, direct method of log P determination that is considered reliable for use with polar/ionic compounds and in situations where little is known about the compounds to be analysed.

3.2 The Stir-Flask Experiment

The stir-flask experiment is a simple partitioning experiment that is now favoured over the 'shake-flask' equivalent, due to the lack of emulsion formation when two immiscible phases are *slowly* introduced to one another.[1] If a few simple precautions are taken, inter-laboratory reproducibility for this technique should be less than ± 0.05 for the average solute and ± 0.10 for the more difficult ones with log P below –3.0 or above 6.0.[2] The method is considered applicable to the measurement of log P in the region –2 to 5 units by the OECD guideline, but there is some variation on this recommendation in the literature.[3]

3.2.1 General Procedure

A stir-flask partitioning experiment has three distinct stages: mutual saturation of the phases, partitioning of the solute in the biphasic system and quantitative analysis of the phases using an appropriate analytical technique.[3]

3.2.2 Factors Affecting Partitioning

Aside from random errors, there are many factors that affect the measurement of an octanol/water partition coefficient, P, and indeed any other partition coefficient, and this is the major reason why there is considerable variation amongst published log P values. Accuracy of partition coefficient determination is affected by temperature, lack of mutual phase saturation, pH, buffer type and concentration, phase miscibility, solute concentration, solute and solvent purity, solute stability, phase volume ratio, solute adsorption and failure to reach equilibrium.[4]

Many published methods of log P determination do not advocate *temperature* control partly due to Leo, Hansch and Elkin's statement that ''for most applications ...variations in P due to temperature are hardly comparable to those inherent in the other measurements; therefore we do not consider it a serious shortcoming that most determinations are simply 'at room temperature' without an estimation of what that may be''.[5] In contrast, Dearden and Bresnen[4] acknowledge the conceivability that inter-laboratory temperatures can vary by up to 20 °C, and for this reason they consider it essential for partition coefficients to be measured under thermostatically controlled conditions.

It has been generally accepted that for most solutes log P has low temperature dependence. i.e. 0.01 per degree[6], but a study by Quigley *et al.*[7] showed that this dependence might vary by as much as three-fold within certain homologous series if the remainder of the structure is complex.

Dearden *et al.* have measured the variation in log P with temperature for a large number of compounds in octanol and water. From the Van't Hoff isochore (Equation 3. 1) the standard enthalpy of partitioning is directly related to the difference in log P at two temperatures. For example, an increase in P by a factor of 1.5 between 15 °C and 35 °C can be related to ΔH° = 15.0 KJ mol^{-1}, and they have shown that such a variation in ΔH°, and hence log P, is not uncommon.

$$\log (P_2/P_1) = (\Delta H^\circ (T_2-T_1)) / 2.303RT_1T_2 \qquad \textit{(Equation 3.1)}$$

The OECD guideline recommends that the test temperature for the determination of an octanol/water partition coefficient, should be kept constant at ± 1 °C and lie in the region of 20-25 °C.[8] It has been shown that in cases where the two solvents used for partitioning have quite substantial mutual miscibility, temperature control is even more important. For example, in the sec-butanol/water system, it is recommended that fluctuation in temperature should not exceed 0.5 °C.[9]

Another important reason for controlling temperature when measuring an octanol/water partition coefficient is that the solubility of water in octanol increases with temperature over the normal laboratory range, while the opposite happens to the solubility of octanol in water. Unless water and octanol are equilibrated at the temperature at which the partition coefficient is to be determined, and unless measurement is made at constant temperature, turbidity will be observed in one of the phases as the temperature changes and the other phase will become unsaturated.[4]

Dearden describes it as essential that *mutual saturation* is achieved before partitioning is undertaken, otherwise the phase volume ratio is likely to change during the partitioning process, which will result in the introduction of errors. From the results of some experimental investigations, it was concluded that about three hours' continuous stirring was required for mutual saturation to be achieved, but in general the time required to achieve mutual saturation is dependent on volume ratio, stirring speed and vessel size.[4]

The time required for a solute to *equilibriate* between the two phases during the partitioning process is theoretically infinite. However, in practice a finite time is required, which is defined as the point where a further elapse of time produces no detectable change in concentration of the solute in one or both phases. Mechanical agitation, such as continuous stirring, causes an increase in the rate of partitioning by increasing the interfacial area between the phases and by reducing the thickness of the stagnant layer across which the solute diffuses.[4]

A common mistake is to assume that log P has *concentration* independence. However, in an ideal situation infinite dilutions should be used. Generally, for compounds that tend not to associate (e.g. neutral compounds) or aggregate to form micelles (e.g. surfactants) it is generally acceptable to conduct a partitioning experiment using a low concentration of solute ($\sim 10^{-1}$M). For compounds that are likely to associate, such as acids, it is recommended that measurements are made at several concentrations and log P plotted against concentration to obtain the value at infinite dilution. In the case of surfactants, one must work below the CMC in the aqueous phase, since the formation of micelles will artificially reduce a value of P.[6]

The *purity of the solute and solvents* employed is also of significance. A solute is usually considered of satisfactory purity if it gives a single peak in HPLC or GC. The purity of the solute in a partitioning experiment is particularly important, since the presence of a contaminant can potentially cause variation in the partitioning properties of the analyte. The purity of the octanol and water employed should also be as high as possible. It is convenient to use double distilled water from glass or quartz apparatus and HPLC grade octanol (99+%).

The *instability of a solute* or susceptibility to degradation can also markedly affect the measurement of a partition coefficient. Certain compounds are capable of hydrolysing slowly in aqueous solution or can be oxidised by air. These problems can be minimised by conducting the partitioning experiments as rapidly as possible in a cold room (~4 °C). The error due to temperature variation is in most cases less than that due to chemical changes at higher temperatures.[6]

It is clearly preferable that when conducting a partitioning experiment the concentration of solute in both be determined, rather than only analysing the concentration in one phase and calculating the concentration in the other by subtraction. If adsorption is expected or the solute is volatile, measurement of both phases is essential.[4] Adsorption has been found to occur extensively with ionic solutes.[5]

The *phase volume ratio* is another factor that can affect the precision of log P determination. It is common practice to use equal volumes of octanol and water, but it has been shown that variation in the phase volume ratio is very useful for the measurement of particularly high or low log P values, where otherwise there would be a low (and perhaps inaccurately measured) concentration in one of the phases.[6]

Finally, where a buffered aqueous phase is employed in a partitioning experiment, it must be appreciated that factors such as *pH, buffer type* and *concentration* come into play. The pH at which the test was carried out should always be stated for ionisable compounds, since it can indicate which form the compound was in during the partitioning experiment. Davis *et al.*[10] examined the effect of ionic strength upon the distribution of phenol between octanol and water to find that P increased with increasing ionic strength. The results of Davis also indicate that no appreciable effect on P should be observed if ionic strength is maintained below 0.1M. Wang and Lien[11] reported the log P values of various ionisable and non-ionisable compounds in the presence of various buffers. It is generally observed that log P of ionisable compounds is greatly affected by buffer type, however the effect that each buffer actually has remains undecided[7].

These authors recommend the use of a phosphate buffer in partitioning experiments, since it has caused fewest problems, and this recommendation has been supported by Taylor *et al.*[12]

3.2.3 Choice of Analytical Method

The analytical technique employed to quantify the analyte in each layer following a partitioning experiment will almost entirely depend on the nature of the analyte investigated. For example, if a chromophoric group is present one would normally use a spectrophotometric technique for analysis. GC is often applied successfully when the spectrophotometric method is unsuitable. In other cases the analyte may be derivatised to make it suitable for analysis by one of these methods. Furthermore, some compounds, such as QACs (that cannot be detected spectrophotometrically owing to the lack of suitable chromophore) can be subjected to pyrolysis and subsequent GC analysis. Another possible method of analysis is demethylation of the quaternary structure (using a mixture of sodium phenylthiolate and thiophenol in anhydrous butanone) to leave a tertiary amine.[9] Likewise, compounds from which an NH_3 molecule can be split off, such as amides and carbamates, can be analysed using Nessler's reagent (K_2HgI_4) and subsequent spectrophotometry at 410nm. These compounds are subjected to hydrolysis and the ammonia released then reacts with K_2HgI_4 to give the product $(NH_2)Hg_2I_4$, that is suitable for spectrophotometric determination .[6]

NMR spectroscopy is one of the most powerful tools available to the chemist for elucidating the structure of chemical species. However, the technique is also useful for the quantitative determination of absorbing species. The proton is undoubtedly the most commonly used nucleus in NMR spectroscopy and ^1H NMR has been applied to the structural elucidation of organic, metal-organic and biochemical molecules. ^{13}C NMR is also used for structural elucidation but is about 6000 times less sensitive than ^1H NMR. As a result, ^1H NMR is most suitable for use in quantitative analysis.

The unique aspect of NMR spectra is the direct proportionality between peak area and the number of nuclei responsible for the peak, so that pure samples are not necessarily required for calibration. Therefore, if an identifiable peak for one of the constituents of a sample does not overlap with peaks due to other constituents, peak area can be used to establish concentration of the species directly, providing the signal area per proton is known. This latter quantity can be conveniently obtained from a known concentration of an internal standard that is obviously chosen so that its characteristic peak does not overlap with any of the sample peaks.[13]

An alternative analytical method capable of analysing the layers of a stir-flask experiment is HPLC, which is scarcely mentioned in the literature for this purpose. HPLC has

the advantage that it is an extremely diverse and sensitive analytical technique, capable of interfacing with various modes of detection.

3.3 High Performance Liquid Chromatography (HPLC)

HPLC uses a stationary phase (which can be either a solid or liquid) and a liquid mobile phase that is pumped through the column at high pressure. The HPLC technique can take one of two forms, namely *normal phase* (NP) and *reverse phase* (RP). The term NP-HPLC is given to a situation where HPLC is carried out using a polar stationary phase and non-polar mobile phase. In contrast, for RP-HPLC a non-polar stationary phase and a polar mobile phase are used. The latter is by far the most popular HPLC technique used, since method development is more rugged (i.e. robust and reproducible) due to the ability to tolerate minor changes in experimental conditions. In RP-HPLC, the stationary phase is organic, commonly a long alkyl chain such as octylsilane (C8) or octadecylsilane (C18) and the mobile phase is usually water with added organic modifier.[14,15] Since the mobile phase contains water, a further advantage of the RP technique is that the aqueous phase can be buffered for the analysis of ionisable compounds.

3.3.1 Quantitative Analysis

3.3.1.1 Internal Standard Method

Quantitation using the internal standard method involves the construction of a calibration plot, which is produced by preparation and analysis of standard solutions containing different concentrations of the compound of interest with a fixed concentration of the internal standard. The peak area ratio of analyte to internal standard is determined for each calibration solution prepared, and this ratio is plotted versus the concentration ratio of analyte to internal standard. The plot can either be used directly to determine the concentration of analyte in samples or this concentration may be calculated by determining the response factor (RF), the latter making the assumption that the plot is linear with zero intercept.

$$\mathbf{RF = AR_{analyte\text{-}internal\ standard} / CR_{analyte\text{-}internal\ standard}} \qquad \textit{(Equation 3.2)}$$

(AR = Area Ratio, CR = Concentration Ratio)

RF is therefore the slope of the line, as shown by Figure 3. 1.

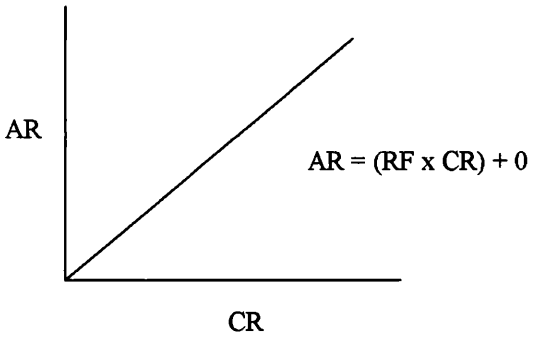


Figure 3.1 Typical form of a calibration plot

The chosen internal standard should obey a number of criteria. Most importantly it should be well resolved from the compound of interest (and any other peaks) under the conditions in use. Obviously it should not be in the original sample and must be stable and unreactive towards the sample and the mobile phase. It is also recommended that the internal standard employed should have a similar retention time to the analyte, and that it displays a comparable detector response to the analyte for the concentration used. The internal standard need not be chemically similar to the analyte, but should be available in high purity.[15]

3.3.2 Possible Detection Mechanisms

There is no single detector that can be employed for all HPLC separations, and for this reason a number of detection modes are commonly used. A suitable detector is expected to recognise when a substance is eluted from the column, and they are classified according to the property that forms the basis of the detection method. The physical parameter of the column effluent or components of the column effluent that is measured by a detector is transformed to an electrical signal. The detection modes considered in this thesis for interfacing with a HPLC system are summarised below.

3.3.2.1 UV detection

The UV detector is the most widely used detector in HPLC, partly since it is the most rugged and reliable. It is a specific detector that possesses a broad applicability range. Detection is based on the absorption of UV radiation by an appropriate functional group in the analyte. The column effluent must be UV transparent at the detection wavelength. i.e. its absorbance must be zero or at least be adjustable to zero electronically. A linear relationship exists between absorbance, A, and the molar concentration, C, as indicated by Equation 3.3.

$$A = \varepsilon C l \qquad \textit{(Equation 3.3)}$$

ε = compound and chromophore dependent molar extinction coefficient (absorptivity) at the wavelength used; l = optical path length (cm)

A is dimensionless, and absorbance is usually expressed as absorption units (AU). Since the detector response (or peak height) is a function of ε and C, compounds with greater values of ε produce a greater response as a substance passes through the detector cell, than those with a smaller value of ε, providing identical amounts of the compounds are injected. For trace analysis it is recommended that values of ε >1000 are usually required. The trace analysis of compounds with ε <100 is not normally possible.[15] Due to superior sensitivity over many of detection techniques commonly used in LC (excluding MS and fluorescence), UV detection is used routinely in quantitative analysis.

There are generally three types of UV detector that are classified according to the range of wavelengths accessible to the detector: *the fixed wavelength detector, the variable wavelength detector and the photodiode array detector.*

Fixed wavelength detectors are the simplest type and the wavelength used is principally determined by the nature of the light source. Mercury lamps are the most frequently used light sources in fixed wavelength detectors, normally used at a wavelength of 254 nm. However, a zinc lamp operated at 214 nm is also useful for detecting some types of chromophore. Most fixed wavelength detectors are designed to enable the lamp and filter to be changed very easily.

The variable wavelength detector employs a lamp that emits light over a wide range of wavelengths and by using a monochromator, light of a particular wavelength can be selected to detect different solutes. This facility enables detection sensitivity to be maximised or the choice of a detection wavelength at which interfering solutes do not absorb appreciably. A deuterium lamp is used, coupled with a suitable variable monochromator (generally the grating type) and the instrument is operated in the dual-beam mode.[16]

The photodiode array detector uses a deuterium or xenon lamp that emits light over the UV spectrum range, thus allowing simultaneous collection of chromatographs at different wavelengths during a single run. The chromatogram can be reconstructed by monitoring at a specific wavelength to depict only those compounds that absorb UV light at the chosen wavelength and may often be represented in 3-D.[16]

3.3.2.2 Refractive Index (RI) detection

The refractive index detector is a universal concentration-sensitive detector that responds to a deflection of a light beam caused by the differing RI of the contents of the sample cell compared with that of the reference cell following the elution of an analyte. It is the oldest and the most widely used universal detector and is often used for compounds with low values of ε for which UV detection is unsuccessful. This mode of detection has a number of disadvantages associated with its practical application. A major drawback is its lack of sensitivity making it unsuitable for trace analysis. This may be coupled with the unsuitability for use with gradient elution, since changes in mobile phase composition also give rise to variation in RI. Finally, the detector requires strict temperature control while in operation (possibly to a few thousandths of a degree centigrade), and is also sensitive to fluctuations in pressure and flow rate.[15, 16]

3.3.2.3 Evaporative Light Scattering Detection (ELSD)

The ELSD is a second type of universal detector but with two distinct advantages over the RI detector. It can be used in conjunction with gradient elution and it is significantly more sensitive. The column effluent is passed into a nebuliser where it is converted into a fine mist by a flow of nitrogen. The fine droplets are then carried through a temperature-controlled drift tube where evaporation of the mobile phase occurs accompanied by the formation of fine particles of the analyte. The 'cloud' of analyte particles then passes through a laser beam and the scattered radiation is detected at right angles to the flow by a silicon photodiode. Obviously the use of the ELSD is therefore restricted to the detection of non-volatile analytes carried by volatile mobile phases.[13]

3.3.2.4 Electrochemical (EC) Detection

These detectors are specific and concentration-sensitive. There are classes of EC detectors, namely direct current amperometry (DCA) and conductivity detectors.

A conductivity detector consists of two electrodes situated in the column effluent and the resistance (or more strictly, impedence) is measured by a suitable electrical circuit. The out-of-balance signal occurs when an ionic solute is present between the electrodes, but this is not linearly related to the ion concentration. As a consequence, an appropriate amplifier must be used to provide an output directly proportional to the solute concentration.

The resistance of any conductor is directly proportional to its length (l) but varies inversely with its cross sectional area (a),

$$R = \rho l / a \qquad \textit{(Equation 3. 4.)}$$

ρ = specific resistance of conductor (resistance across two opposite faces of 1cm cube of the conductor material)

The specific conductance (κ) of a solute is defined as the reciprocal of the specific resistance, and similarly the conductance (C) of a given solute is the reciprocal of the resistance.

A conductometric detector offers advantages such as simplicity, reasonably high sensitivity, convenience and wide-spread applicability, but it responds to electrolytes (e.g. buffers) in the mobile phase. The mobile phase must therefore be chosen to be non-conducting (which is often very difficult if not impossible) or the buffer electrolytes must be removed prior to detection using an ion suppression technique.[13,16]

3.3.2.5 Indirect UV Detection

From perusal of recent literature it became apparent that indirect UV detection is commonly coupled with capillary electrophoresis (CE) and reverse-phase ion-pair chromatography to detect analytes - taking advantage of the absence of a suitable chromophore for UV detection. The technique uses a mobile phase spiked with a highly UV absorbing species (in relatively low concentration) to create a large, positive background signal. When a UV inactive analyte passes through the detector, the signal decreases and a negative peak is observed. The polarity of the system may be inverted to produce positive peaks.[17] Commonly used UV absorbing species are benzyltrimethylammonium chloride (BTMAC), imidazole, aminobenzimidazole and potassium hydrogen phthalate.[18-20]

3.3.2.6 Mass Spectrometry (MS) Detection

Detection by MS is becoming more common despite its high cost. This is primarily due to its superior sensitivity compared with any other detection method. There are a number of 'hard' and 'soft' ionisation methods that may be employed in MS and the method chosen will depend mainly on the nature of the compounds to be analysed, but also on the ease of interfacing with the HPLC system.

3.3.2.6.1 Ionisation Methods

Electron Impact (EI)-MS is the most common form of MS used today, and is suitable for both qualitative and quantitative analysis. However, since it is a 'hard' ionisation method, for larger molecules (e.g. biopolymers) it is inadequate as these molecules fragment too readily. As a

result the molecular ion itself is commonly lost and a series of considerably smaller ions are observed. EI-MS allows the analysis of volatile analytes, but is generally not applicable for analytes that are polar and involatile. Other ionisation methods based on the desorption of non-volatile and thermally labile compounds directly from solutions or solid surfaces fortunately exist. These include thermospray (TSP), plasma desorption (PD), field desorption (FD), fast atom bombardment (FAB), liquid secondary-ion MS (LSIMS), laser desorption (LD), matrix-assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI). Of these, the most popular commercially available interface for HPLC is the atmospheric pressure interface of ESI.

ESI is described as a 'soft' ionisation technique since less fragmentation is observed for large, polar compounds. A spray is generated at ambient pressure and a high voltage is applied to the eluting solvent. There are now several variations on the original electrospray technique, most notably the use of a sheath or supporting gas which has often been termed ionspray (IS). As the eluent is sprayed at ambient pressure, an organic sheath liquid is mixed with the eluting aqueous solvent to reduce surface tension and enhance evaporation of the charged droplets. Analyte molecules generated from electrospray contain various charged states (varying amounts of adducted sodium ions or protons). This multicharging process produces a nearly gaussian distribution of peaks corresponding to the different *m/z* ratios of the multicharged ions, and only those analytes capable of sustaining such multiple charges are amenable to ESI or IS MS. Previous characterisation work on the zwitterionic sulfobetaines (Chapter 2) confirms that ESI-MS is suitable for their analysis by qualitative means. It should therefore be possible to make this technique quantitative in the form of HPLC/ESI-MS.[15]

3.4 Experimental

3.4.1 Chemicals

n-octanol (HPLC grade, 99+%) was purchased from Aldrich; double-distilled water was prepared by technical services at the University of Wales, Swansea, and stored in a glass vessel; urea, aniline, benzyl alcohol, formamide, acetophenone, phenol and imidazole were purchased from Aldrich; *N*-Decyl-*N*, *N*-dimethyl-3-ammonio-1-propanesulfonate was purchased from Sigma; D_2O (99.9 atom%D) was purchased from Apollo Scientific; acetonitrile (HPLC grade), methanol (HPLC grade) and water (HPLC grade) were purchased from Fisher Scientific.

3.4.2 Determination of Log P *via* the Stir-Flask Experiment

n-octanol (5 ml) and water (5 ml) were measured using separate burettes and transferred to a round-bottomed flask (50 ml). The contents were sealed and continuously stirred for 24 h to achieve mutual saturation of the phases. The test vessel was thermostatted at 22 ± 1 °C using a water bath for the duration of this process. After 24 h, the solute (55-74 mg) was added (weighed using a 4 d.p. analytical balance) and continuously stirred for a further 24 h under the same temperature control. The contents were transferred to a separating funnel (25 ml), allowed to settle for a few minutes and then separated. An aliquot of one (or both) layer(s) was taken and diluted to incorporate internal standard for quantitative analysis.

3.4.2.1 Measurement of Log P *via* the 'Stir-Flask/Evaporation and Weighing' Method

For the analysis of the reference compounds (urea, aniline, benzyl alcohol, formamide, acetophenone and phenol) the stir-flask experiment was generally conducted as described in Section 3.4.2., but with the following exceptions: the volume of the octanol and water phases employed was 100 ml and the mass of solute employed ranged from 1.00–2.22 g.

Two methods were utilised. The first involved the analysis of the water layer only and then calculation of solute concentration in the octanol layer by difference (*Method A)* and the second involved the analysis of both layers (*Method B*). When Method A was employed, the whole water layer was analysed, however when both layers were analysed a 20 ml aliquot was taken from each for analysis. Analysis involved evaporation of the solvent(s) under reduced pressure, further drying under reduced pressure for several hours and then weighing of the mass of solute remaining using a 3 d.p. balance.

For the analysis of the 10 sulfobetaines (compounds **1**, **3-5**, **8, 14-17)** that had been synthesised at this stage, the stir-flask experiment was conducted as described in Section 3.4.2., but with the following exceptions: the volume of the octanol and water phases employed was 50 ml and the mass of solute employed ranged from 0.28–0.62 g.

Method A was employed for the analysis of the sulfobetaines. Method B was attempted but proved difficult and was therefore abandoned, since it was found almost impossible to remove traces of octanol from the solute, despite drying under high vacuum for several hours.

3.4.2.2 Measurement of Log P *via* the 'Stir-Flask/Quantitative ^{1}H NMR' Method

For the analysis of the 10 sulfobetaines (compounds **1**, **3-5**, **8, 14-17)** the stir-flask experiment was generally conducted as described in Section 3.4.2., but with the following exceptions: D_2O

was used in replace of H_2O and the mass of solute used ranged from 34–62 mg. Deuterated octanol was found to be extremely expensive and so we were limited to the analysis of the aqueous layer only.

Following the partitioning experiment, a 0.5 ml aliquot was extracted from the aqueous phase using a dry syringe (1 ml) with needle. This aliquot was transferred to a sample tube containing 1–3 mg of internal standard (accurately weighed using a 4 d. p. analytical balance). The sample was shaken vigorously to ensure homogeneity and then its contents were transferred to a previously dried NMR tube for analysis. For analysis of the alkyl sulfobetaines the internal standard used was phenol, but for the phenylalkyl sulfobetaines *t*-BuOH was employed for this purpose.

The ^{1}H NMR spectra were run on a Bruker AC spectrometer at 400 MHz. The peaks were manually integrated so that any small impurities picked up during the experiment (mainly traces of octanol) could be eliminated.

3.4.2.3 Measurement of Log P *via* the 'Stir-Flask/Quantitative RP-HPLC-UV' Method' for the Phenylalkyl (and *para*-substituted phenylpropyl) Sulfobetaines

The stir-flask experiment was conducted as described by Section 3.4.2. The internal standard employed for quantitative analysis *via* this method was phenol. It is also important to note that *both* layers were analysed using this method. Before addittion of the internal standard, a sample of each layer was then analysed by HPLC to obtain an approximation of the UV response at the wavelength of detection, so that an appropriate amount of phenol could be added on dilution to obtain relative responses. The HPLC system employed was a HP1100 with quaternary pump, HP Chemstation and HP1100 variable wavelength detector. The wavelength of detection was 266 nm for analysis of the phenylalkyl sulfobetaines and 220 nm for analysis of the *para*-substituted phenylpropyl sulfobetaines. Table 3.1 shows the identities of the compounds analysed, the column employed for the investigation and the solvent composition chosen to give well resolved peaks for the analyte and internal standard, in a convenient run time (<15 min). The mobile phase in each case was eluted at 1 ml/min and the injection volume was 20 μl.

Table. 3.1 Conditions employed for the analysis of the phenylalkyl and para-substituted phenylpropyl sulfobetaines

Compound	Column	Solvent composition ($AcCN:H_2O$)
8-10, 17-19	Phenomenex Synergi Polar-RP 80 Å 4 μ 4.6 x 150 mm	15:85
11		25:75
13, 22		35:65
20	Phenomenex Synergi Max-RP 80 Å 4 μ 4.6 x 150 mm	25:75
12, 21		32:68

It was not necessary to establish a calibration plot of concentration ratio versus peak area ratio for each analyte relative to internal standard since both layers were analysed and since log P is a ratio the response factor is cancelled out in the calculation. However, for a number of sulfobetaines a calibration plot was established (using a series of standards incorporating varying concentrations of analyte and a fixed concentration of internal standard) so that that a comparison between log P calculated from the analysis of both layers and that calculated from the assumption that only one layer had been analysed, and the concentration in the remainder had been determined by subtraction, could be made.

3.4.2.4 Measurement of Log P *via* the 'Stir-Flask/Quantitative RP-HPLC-ESIMS' Method for the Alkyl Sulfobetaines

The stir-flask experiment was conducted as described by Section 3.4.2; however for a few sulfobetaines (compounds **2**, **6**, **7**) it was realised that original experiments were conducted using concentrations above the CMC.[21] For these compounds the stir-flask experiments were repeated using 6–10 mg of solute. The internal standard employed for quantitative analysis via this method was usually commercially available compound **6**. However, when compound **6** was the analyte, compound **7** was used as the internal standard. An aliquot of each layer was diluted to incorporate internal standard. A Thermofinnigan aQa, single quadropole mass spectrometer was employed for detection of these compounds. The form of ionisation was ESI, in positive ion mode, with single ion monitoring of $[M+H]^+$. The probe voltage and temperature was 4.5 kV and 300 °C respectively. The mobile phase was eluted at 1 ml/min, split to 100 μl/min to MS. The injection volume for all analyses was 20 μl. Table 3.2 shows the identities of the

compounds analysed, the column employed for the investigation and the gradient program chosen to analyse the sulfobetaines in a convenient run time (<13 min).

Table 3.2 Conditions employed for the analysis of alkylsulfobetaines

Compound	Column	Gradient Program ($AcCN:H_2O$)
1-4, 14	Phenomenex Aqua C18 125 Å 5 μ 4.6 x 250 mm	50:50 to 80:20 in 10 min
5		15:85 to 40:60 in 7 min
6		50:50 to 75:25 in 7 min. 75:25 for 3 min
15, 16		30:70 to 80:20 in 12.5 min
7	Phenomenex Synergi Max-RP 80Å 4μ 4.6 x 150mm	50:50 to 70:30 in 6 min. 70:30 for 4 min.

3.4.3 Calculation of Log P using KowWin

Log P was calculated using a computerised atom/fragment approach developed by the Syracuse Research Cooperation (SRC). The version employed was a demonstration program available for use on the internet.[22] The required input for use of this program was Simplified Molecular Input Line Entry System (SMILES) notation.

3.5 Results and Discussion

3.5.1 Stir-flask/Evaporation and Weighing Method

This method of log P determination was investigated to estimate log P of the first few sulfobetaines synthesised, prior to the investigation of a more specific analytical method. Methods A and B proved successful for the analysis of four of the six reference compounds that were chosen to validate the method (urea, formamide, benzyl alcohol and acetophenone); however, the analysis of aniline and phenol highlighted two major limitations of the evaporation method. Aniline forms an azeotrope with water and so separation on the basis of evaporation is not possible. Furthermore benzyl alcohol has a higher but similar boiling point to octanol (205 °C compared with 196 °C) and so the evaporation of solvent was not possible and separation of solute from the solvent would be extremely difficult without employing a more sophisticated separation technique. e.g. spinning band distillation. From the perusal of Tables 3.3 to 3.5, it is observed that for the reference compounds, log P values derived from Methods A and B are in

reasonable agreement. However, it should be appreciated that none of the reference compounds are very hydrophobic or hydrophilic in nature and the greatest errors are therefore expected to arise when especially small masses of analyte are recovered from one layer. This error is further amplified when only one layer is analysed and concentration in the other is determined by difference.

The series of ten sulfobetaines were analysed by Method A only. It was generally found that the method was satisfactory for the estimation of log P of compounds where the solute is reasonably evenly distributed between the lipophilic and hydrophilic phases (i.e. approx. –1.5 to 1.5; however beyond this range (where log P was expected to be somewhat lower or more negative) it was often found that the mass in the water alone was greater than that introduced to the biphasic system originally. In some cases log P values were also found not to be reproducible. As a result, log P determinations for only compounds **4**, **5** and **16** are recorded.

Table 3.3 Method (A) – Results for reference compounds

Reference compound	Mass added to system/g	Mass (g) in H_2O layer (100ml)	∴Mass (g) in octanol layer (100ml)	Log P
Urea	2.223	2.127	0.096	-1.35
Benzyl alcohol	1.009	0.051	0.958	1.27
Formamide	1.521	1.432	0.089	-1.21
Acetophenone	1.236	0.058	1.178	1.50

Table 3.4 Method (B) – Results for reference compounds

Reference compound	Mass added to system/g	Mass (g) in H_2O layer (20ml)	∴Mass (g) in octanol layer (20ml)	Log P
Urea	2.062	0.214	0.096	-1.38
Benzyl alcohol	1.012	0.051	0.958	0.94
Formamide	0.959	0.887	0.072	-1.09
Acetophenone	1.413	0.066	1.347	1.31

Table 3.5 Methods (A) and (B) – Results for reference compounds

Reference compound	Average log P from (A) and (B)	Literature log P [8,23]
Urea	-1.37	-1.21
Benzyl alcohol	1.11	1.05
Formamide	-1.15	-1.51
Acetophenone	1.41	1.32

Table 3.6 Method (A) – Results for Sulfobetaines

Sulfobetaine	Mass added to system/g	Mass (g) in H_2O layer (50ml)	∴Mass (g) in octanol layer (50ml)	Log P	Average log P
C8C3 (**5**)	0.549	0.344	0.205	-0.22	-0.24
	0.490	0.313	0.177	-0.25	
C8C4 (**16**)	0.582	0.480	0.102	-0.67	-0.74
	0.337	0.315	0.022	-0.80	
C7C3 (**4**)	0.621	0.569	0.052	-1.04	-0.92
	0.597	0.516	0.081	-0.80	

From the results shown in Table 3.6, it is concluded that a decrease in log P is observed with the addition of a methylene group into the spacer unit between the polar groups. According to Leo and Hansch, an increase in log P of f_{CH2}-F_b = 0.54 is normally seen with the addition of a hydrophobic methylene group. However, it appears that the extra hydrophobicity offered by the methylene unit is overwhelmed by the weakening of the proximity effect between charged moieties.

In contrast, the addition of a methylene group to the main alkyl chain does appear to result in an increase in log P of 0.68 (*cf.* 0.54).

3.5.2 Stir-flask/^{1}H-NMR Method

As with the evaporation and weighing method, there are also reasonably large errors associated with the quantitative NMR method, which became more obvious on determination of log P of the particularly hydrophilic sulfobetaines and where only the aqueous phase was analysed and the concentration in the octanol phase was determined by difference. By using compound **5** as an example we can do a rough calculation of the error associated with a log P value derived by this method. If a 5 % error in integration and a 0.1 mg error in any 4 d.p. balance reading is assumed, the error in the weight of internal standard becomes (0.1/1.6) x 100 = 6.3 %. The overall error for the mass of **5** in the aqueous layer therefore would be $\sqrt{(5^2 + 5^2 + 6.3^2)} = 9.5$ % and hence the value should be recorded as 39.4 mg ± 3.7 mg. This error alone is great, however, it must also be appreciated that since the mass in octanol is determined by difference, the final error is further amplified.

A total of 10 sulfobetaines were analysed; however, since it often appeared that the mass in the aqueous layer was greater than that actually introduced into the biphasic system and/or non-reproducible results were observed, log P is only recorded for the same 3 sulfobetaines as before (Table 3.7).

Table 3.7 Log P values determined by Quantitative ^{1}H-NMR

Sulfobetaine	Mass added to system/mg	Mass in aqueous phase (5ml)/mg	∴Mass in octanol (5ml)/mg	log P	Av. log P
C8C3 (**5**)	67.1	39.4	27.7	-0.15	-0.15
	66.9	38.9	28.0	-0.14	
C8C4 (**16**)	65.9	54.6	11.3	-0.68	-0.75
	62.1	53.9	8.2	-0.82	
C7C3 (**4**)	61.8	53.8	8.0	-0.83	-0.78
	66.9	56.3	10.6	-0.73	

From these results the same conclusions may be drawn as those derived from the stir-flask/evaporation and weighing method. That is, an increase in log P of almost (f_{CH2}-F_b) is observed with the addition of a methylene group to the alkyl chain, but a decrease in log P with the addition of a methylene unit to the alkyl spacer unit between polar groups.

Literature confirms that octanol is partially soluble in water. e.g. at 25 °C, the solubility of octanol in water is $8.9x10^{-3}$ mol. l^{-1} and so we could expect to see small traces of octanol in the spectrum. Certain peaks characterising the analytes were expected to overlap with octanol peaks and were hence eliminated from the calculation. An example calculation of log P using this method, employing phenol as the internal standard is given by Figure 3. 2.

Phenol (Mr = 94.0): N°. mmoles = $1.6x10^{-3}/94 = 0.0170$

5H = 2.567, ∴1H = 0.5134

Compound **5** (Mr = 279): 3H = 1.452, ∴1H = **0.484**; 10H = 7.823, ∴1H = **0.7823**; 2H = 0.888, ∴1H = **0.444***;
2H = 0.805, ∴1H = **0.4025***; 2H = 0.907, ∴1H = **0.4535***; 6H = 2.595, ∴1H = **0.4325***;
2H =0.811, ∴1H = **0.4055***; 2H = 0.839, ∴1H = **0.4195***

(*integrals used to calculate average)

Av. integration for 1H = 0.426

Mass of **5** in 0.5ml D_2O = (0.426/0.5134) x 0.0170 x 279 = 3.94mg

∴Mass of **5** in 5ml D_2O = 39.4mg

Total mass of **5** introduced to octanol/D_2O system = 67.1mg

∴Mass in octanol = 67.1 – 39.4 = 27.7mg

P = 27.7/39.4 = 0.70, ∴ **log P = -0.15**

Figure 3.2 An example calculation of log P by 'Stir-flask/^{1}H-NMR' method.

The fact that, for hydrophilic compounds, we observe a greater concentration in the aqueous layer than that originally introduced into the biphasic system, has also been appreciated by other authors – although in reverse. For hydrophobic compounds, where the octanol layer only was analysed, it was often observed that the mass determined in the octanol layer was greater than the total used in the experiment.[2] These findings help to confirm the recommendation that wherever possible (especially for particularly hydrophilic or hydrophobic compounds) both layers should be analysed for the determination of log P.

3.5.3 Stir-flask/RP-HPLC-UV

HPLC was an obvious choice of analytical method suitable for the analysis of a stir-flask experiment since it can be used effectively to quantify the solute in both layers of the experiment. Coupled with UV detection, the technique proved suitable for the analysis of the phenylalkyl- (and *para*-substituted phenylpropyl-) sulfobetaines, since they possessed a chromophore suitable for UV detection.

Method development proved time consuming since the extremely polar nature of these compounds was not at first appreciated. HPLC was first conducted on PhC1C3 (**8**), the second most polar compound in the series, using a Genesis C18 4μ 150 x 4.6mm column and employing high percentages (50-100%) of organic modifier (MeOH or AcCN). However, injection of an unretained compound, uracil, revealed that **8** was also unretained under these conditions. It was then realised that separation conditions would require a mobile phase composed of mainly water, with only a small percentage of organic modifier. Phenomenex recommends the use of a Synergi Max-RP or Polar-RP for method development of extremely polar compounds to improve peak shape and give longer retention to compounds analysed using mainly aqueous mobile phase conditions. Both these columns proved extremely successful for the analysis of the polar sulfobetaines, giving excellent peak shape without incorporation of a buffer. Phenol, the internal standard, also eluted with excellent peak shape and appeared well resolved from the sulfobetaines, under the mobile phase conditions employed for use with either column.

Log P determined by this method proved highly reproducible. Over the log P range investigated (-2.32 to 1.89), it appeared satisfactory to use a constant phase volume ratio (i.e. 5 ml octanol, 5 ml water). On dilution of the octanol and water samples to incorporate the internal standard, dilution factors were chosen carefully in an attempt to produce analyte and internal standard peaks of similar area. This is considered just as effective for the improvement of

precision of log P determinations as employing varying phase volume ratios, which is often recommended in the literature (Section 3.2.2). In both cases the precision in P is improved, and hence log P, since the concentration of solute in the layer that the solute has least affinity for can be increased, and peaks that may otherwise be undetectable, can be enlarged to be observed.

Tables A1-A28 of Appendix A show the calculation of log P for every experiment conducted, based on the analysis of both layers of the biphasic system. The compounds are displayed in order of increasing hydrophobicity. Since both layers were analysed a response factor (of analyte to internal standard) is not required for these calculations. As a result, the entries in these tables entitled 'factor related to concentration of analyte in sample' are equivalent to the '*Peak Area Ratio x Conc. of Internal Standard*' in each case. The entries entitled 'Factor related to concentration of analyte in octanol (or water) layer' take into account the dilution factors.

However, it was also decided that a comparison would be made between these results and log P calculations based on the assumption that only one layer had been analysed and the concentration of analyte in the other was determined by subtraction from the total concentration originally introduced to the stir-flask experiment. Figures A1–A7 indicate how the response factors of analyte/internal standard were derived for compounds **8**, **9**, **11**, **17-20**. Table A28 shows the calculation of log P for these compounds assuming that only *the concentration of analyte in water* was determined and the concentration of analyte in octanol was determined by subtraction. Furthermore, Table A29 shows the calculation of log P for these compounds assuming that only *the concentration of analyte in octanol* was determined and the concentration of analyte in water was determined by subtraction. Finally, Table 3.8 provides a comparison between the three methods and highlights the need for analysing both layers of the stir-flask experiment to achieve an accurate and reproducible value of log P. It is shown that the error is as low as +/-0.01 log P units in most cases rising to +/-0.13 units at the most.

It was also concluded that in cases where hydrophilic compounds are investigated, if only the aqueous phase (or phase for which solute has greatest affinity) is analysed it can sometimes appear that this layer contains more solute than introduced into the biphasic system originally. Additionally, if a value of log P is obtained using this method it cannot be considered reliable and has limited applicability. However, if the octanol phase is analysed and the concentration in the aqueous layer is determined by difference we derive a reasonable estimate of log P despite the fact that the compounds are ionic and adsorption to the glassware is a potential concern.

An attempt has been made to summarise some of the trends observed in log P for the phenylalkyl and *para*-substituted phenylpropyl sulfobetaines. As briefly described in Section 3.5.1, a surprising phenomenon is that for an increase in the length of the spacer unit (increase in hydrophobicity by the addition of a methylene group) there is an overall decrease in log P. This is thought to be attributed to the weakening of the proximity effect that exists between the oppositely charged moieties that apparently overwhelms the hydrophobicity provided by the alkyl unit. For the addition of a methylene group to the spacer unit between N^+ and SO_3^- .i.e. X=3 to X=4 (where X denotes number of methylene units), we observe a decrease in log P of approximately **0.20** units.

Secondly, we observe an increase in log P of approximately **0.34** for the addition of a methylene group to the alkyl unit between the aromatic ring and the N^+ group, providing the length of the spacer unit remains constant. The normal increment for the addition of a methylene unit to an alkyl chain is 0.54 and the fact that we observe a difference of only 0.34 suggests that there may also be a small proximity effect that exists between N^+ and Ph.

Finally, for the addition of a methylene group to the alkyl chain of the *para*-substituted phenylpropyl sulfobetaines, we see an average increase in log P of **0.66** units, which is surprisingly the exact value for f_{CH2}. However, the bond factor of –0.12 is normally applied to give the expected increment of 0.54.

Table 3.8 A comparison between log P determined from the analysis of both layers of the biphasic system and log P determined by analysis of one layer only

Compound	Av. Log P_a	Av. Log P_b	Av. Log P_c
PhC1C4 (**17**)	**-2.32** (+/-0.13)	-0.49	-2.45
PhC1C3 (**8**)	**-2.27** (+/-0.11)	-0.85	-2.28
PhC2C4 (**18**)	**-2.06** (+/-0.02)	---*	-2.01
PhC2C3 (**9**)	**-1.86** (+/-0.05)	-1.55	-1.84
PhC3C4 (**19**)	**-1.70** (+/-0.03)	---*	-1.37
PhC3C3 (**10**)	**-1.58** (+/-0.01)	---	---
PhC4C4 (**20**)	**-1.41** (+/-0.01)	---*	-1.29
PhC4C3 (**11**)	**-1.17** (+/-0.01)	-1.58	-1.1
C4PhC3C4 (**21**)	**0.30** (+/-0.01)	---	---
C4PhC3C3 (**12**)	**0.55** (+/-0.02)	---	---
C6PhC3C4 (**22**)	**1.64** (+/-0.03)	---	---
C6PhC3C3 (**13**)	**1.89** (+/-0.12)	---	---

***Av. Log P_a** = log P determined from analysis of both octanol and aqueous layer; **Av. Log P_b** = log P determined from analysis of aqueous layer only and determination of concentration of solute in octanol by difference; **Av. Log P_c** = log P determined from analysis of octanol layer only and determination of concentration of solute in aqueous by difference; * indictates compounds that were tested but for which no value of log P could be derived*

3.5.4 Stir-flask/RP-HPLC-ESIMS

It was concluded that since the alkyl sulfobetaines lacked a chromophore suitable for UV detection, they could not be detected under mobile phase conditions found suitable for the analysis of the aryl sulfobetaines, across the UV spectrum of 200 – 400 nm.

We therefore investigated RI detection as a potential means of analysing the alkyl sulfobetaines. However, it was found that a typical chromatogram contained many prominent system peaks that appeared to be broad enough to cover a significant part of the chromatogram. A few of the alkyl sulfobetaines were injected, however they were assumed to be eluting to overlap with these system peaks and hence were not detected. These system peaks could not be eliminated by ensuring that the solvent composition of mobile phase was exactly the same as the sample solvent composition, or by controlling the temperature of the system using recirculating water.

Consequently the indirect UV detection mechanism, as described in Section 3.3.2.5, was investigated as a means of detecting those sulfobetaines that lacked a chromophoric group. From

the commonly used UV active species, imidazole (UV detection @ 214 nm) was employed to create the large background signal required. Mobile phase conditions of 15 % 0.023mM imidazole in AcCN: 85 % H_2O were initially utilised and analytes **5** and **16** were injected onto a Genesis C18 4 μ 4.6 mm x 150 mm column. The analytes were dissolved in H_2O, which was expected to be UV inactive at 214 nm. The small negative peak (height = ~10 mAU; area = ~100) observed early in the chromatogram was attributed solely to H_2O and despite using highly concentrated samples no additional peaks were observed for either analyte. It became obvious that the system was very insensitive to UV inactive species passing through the detector (confirmed by the fact that the solvent peak appeared so small) and our aim was then to improve the sensitivity of the system by varying the concentration of the UV absorbing species in the mobile phase. Eight concentrations in the region of 0.023 – 400 mM were investigated; however the appearance of the chromatogram remained effectively unchanged (i.e. the area/height ratio of the solvent peak was very similar and the peak due to analyte was absent in each case).

Utilisation of a conductometric detector or an ELSD was also considered. However, we did not have access to a conductometric detector (which have been previously found applicable to the detection of cationic compounds without a chromophore, such as QACs) and purchase would have been expensive. We did however have access to an ELSD; however it was not operative due to a blocked nebuliser. Meanwhile, we had success with the investigation of a RP-HPLC-ESIMS technique.

The alkyl sulfobetaines are generally less polar than the phenylalkyl sulfobetaines, and so they could be analysed with greater proportions of organic modifier and in most cases a Phenomenex Aqua column provided satisfactory peak shape. The internal standard used was *N*-decyl-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (or *N*-dodecyl-*N, N*-dimethyl-3-ammonio-1-propanesulfonate), due to their similar molecular weight, chemistry and fragmentation pattern to the sulfobetaines analysed.

Tables A31–A50 show the calculation of log P for every experiment conducted, based on the analysis of both layers of the biphasic system. The compounds are displayed in order of increasing hydrophobicity. Since both layers were analysed a response factor (of analyte to internal standard) is not required for these calculations. As a result, the entries in these tables entitled 'factor related to concentration of analyte in sample' are equivalent to the '*Peak Area Ratio x Conc. of Internal Standard*' in each case. The entries entitled 'Factor related to concentration of analyte in octanol (or water) layer' take into account the dilution factors.

However, it was also decided that a comparison would be made between these results and log P calculations based on the assumption that only one layer had been analysed and the concentration of analyte in the other was determined by subtraction from the total concentration originally introduced to the stir-flask experiment. Figures A8 – A14 indicate how the response factors of analyte/internal standard were derived for compounds **3-6**, **14-16**. Table A51 shows the calculation of log P for these compounds assuming that only *the concentration of analyte in water* was determined and the concentration of analyte in octanol was determined by subtraction. Furthermore, Table A52 shows the calculation of log P for these compounds assuming that only *the concentration of analyte in octanol* was determined and the concentration of analyte in water was determined by subtraction. Finally, Table 3.9 provides a comparison between the three methods and highlights the need for analysing both layers of the stir-flask experiment to achieve an accurate and reproducible value of log P. The average error is +/-0.10 log P units (which is somewhat greater than that deduced for the stir-flask/RP-HPLC-UV method) with the error rising to +/-0.35 units in one extreme case.

Once again, it was concluded that in cases where hydrophilic compounds are investigated, if only the aqueous phase is analysed it can sometimes appear that this layer contains more solute than introduced to the biphasic system originally. Additionally, if a value of log P is obtained using this method it cannot be considered reliable and has limited applicability. Furthermore, if we analyse the octanol phase and determine the concentration in the aqueous layer by difference we derive a reasonable estimate of log P.

Another interesting observation is that as compounds become less hydrophilic such that the solute is more evenly distributed between the phases of a stir-flask experiment, e.g. –0.47 to 0.55 (Table. 3. 9), either the octanol or aqueous layer may be analysed (and concentration in the aqueous or octanol layer respectively may be determined by subtraction from original concentration) to give a reasonable estimate of log P.

Table 3. 9 A comparison between log P determined from the analysis of both layers of the biphasic system and log P determined by analysis of one layer only

Compound	Av. Log P_a	Av. Log P_b	Av. Log P_c
C7C3 (**4**)	**-1.97** (+/-0.35)	*---	-1.60
C7C4 (**15**)	**-1.23** (+/-0.17)	-0.79	-1.28
C6C3 (**3**)	**-1.22** (+/-0.05)	-0.57	-1.31
C6C4 (**14**)	**-1.08** (+/-0.05)	*---	-0.71
C8C3 (**5**)	**-0.47** (+/-0.06)	-0.35	-0.52
C8C4 (**16**)	**-0.36** (+/-0.03)	-0.29	-0.39
C8C2 (**1**)	**-0.06** (+/-0.17)	---	---
C10C3 (**6**)	**0.57** (+/-0.02)	0.66	0.31
C12C3 (**7**)	**1.65** (+/-0.13)	---	---
C12C2 (**2**)	**1.79** (+/-0.01)	---	---

***Av. Log P_a** = log P determined from analysis of both octanol and aqueous layer; **Av. Log P_b** = log P determined from analysis of aqueous layer only and determination of concentration of solute in octanol by difference; **Av. Log P_c** = log P determined from analysis of octanol layer only and determination of concentration of solute in aqueous layer by difference; * indictates compounds that were tested but for which no value of log P could be derived*

An attempt has been made to summarise some of the trends observed in log P for the alkyl sulfobetaines; however they appear to be more complicated than the corresponding aryl sulfobetaines. Firstly, the effect of adding a methylene group to a spacer unit of X=3 (to become X=4) is entirely opposite to that observed for the aryl sulfobetaines. For the alkyl sulfobetaines, we see an increase in log P for the addition of a methylene unit to X=3. However the extent of this increase appears to depend on the length of the main alkyl chain. For a C7 chain there is an increase in log P of 0.74 for X=3 to X=4, for a C6 chain there is an increase of log P of 0.14 and for a C8 chain the increase is equivalent to only 0.11 units. This tends to indicate that the proximity effect is weaker between the charged moieties than with the aryl sulfobetaines, perhaps due to the flexibility of the alkyl chain compared with a rigid and sterically hindered ring system. It is possible that the chains arrange themselves so to shield the proximity effect between the charged groups and thus the addition of a methylene unit simply generates an overall increase in hydrophobicity.

It is however interesting to note that where the length of the main alkyl chain remains constant, log P increases with the length of the spacer unit in the following order: -

$$X=2 > X=4 > X=3$$

These results suggest that a strong proximity effect does exist between the oppositely charged groups where X=2, which indeed does overwhelm the addition of one or even two methylene units. In this case the close proximity of the charged moieties must hold the molecule reasonably rigid and thus the flexibility of the main alkyl chain is also lost and hence cannot provide shielding of the proximity effect.

To add to the complex behaviour of this series of compounds, the effect of introducing a methylene unit to the main alkyl chain (while the length of the spacer unit remains constant) is also confusing. It appears that where an even number of carbon atoms (or methyl and methylene units) exist in the alkyl chain, for the addition of two new methylene units there is an average increase in log P of **0.45** for each new hydrophobic unit. However, alkyl sulfobetaines with an odd number of carbon atoms in the main alkyl chain do not appear to fit this pattern. i.e. C7C3 (**4**) has a lower log P than C6C3 (**3**), and similarly C7C4 (**15**) has a lower log P value than C6C4 (**14**).

3.5.5 Log P Predictions for Sulfobetaines by KowWin

Log P predicting programs investigated were poor predictors of log P for the sulfobetaines, despite providing close estimates of log P for neutral compounds such as the reference compounds mentioned briefly in Section 3.5.1. However, from all programs investigated, KowWin appeared to provide the most realistic predictions of log P for the zwitterionic sulfobetaines despite some obvious limitations of this program also. On construction of SMILES notation for each sulfobetaine, we found it essential to incorporate a sodium ion into the structure otherwise the program would make the assumption that the sulfonate group was protonated and the log P estimate would appear artificially low. By adding Na^+ and deducting its contribution at the end of the calculation log P values were generated in the expected range.

Table 3.10 Comparison between log P calculated by KowWin and experimental log P

Sub-Class	Compound	SMILES	Log $P_{(+Na)}$	Log $P_{(-Na)}$	Log P_{expt}
Alkyl ethane sulfobetaines, n=2	**2**	CCCCCCCCCCCCN(C)(C)CC S(=O)(=O)O[Na]	-3.65	**0.93**	**1.78**
	1	CCCCCCCCN(C)(C)CC S(=O)(=O)O[Na]	-5.61	**-1.03**	**-0.06**

Alkyl propane sulfobetaines, n=3	**7**	CCCCCCCCCCCCN(C)(C)CCC S(=O)(=O)O[Na]	-3.16	**1.42**	**1.65**
	6	CCCCCCCCCCN(C)(C)CCC S(=O)(=O)O[Na]	-4.14	**0.44**	**0.57**
	5	CCCCCCCCN(C)(C)CCC S(=O)(=O)O[Na]	-5.12	**-0.54**	**-0.47**
	4	CCCCCCCN(C)(C)CCC S(=O)(=O)O[Na]	-5.61	**-1.03**	**-1.97**
	3	CCCCCCN(C)(C)CCC S(=O)(=O)O[Na]	-6.11	**-1.53**	**-1.22**
Phenylalkyl-and *para*-substituted phenylpropyl sulfobetaines, n=3	**13**	CCCCCCc1ccc(CCCN(C)(C)CCC S(=O)(=O)O[Na])cc1	-2.85	**1.73**	**1.84**
	12	CCCCc1ccc(CCCN(C)(C)CCC S(=O)(=O)O[Na])cc1	-3.85	**0.73**	**0.55**
	11	c1ccccc1CCCCN(C)(C)CCC S(=O)(=O)O[Na]	-5.38	**-0.80**	**-1.17**
	10	c1ccccc1CCCN(C)(C)CCC S(=O)(=O)O[Na]	-5.87	**-1.29**	**-1.58**
	9	c1ccccc1CCN(C)(C)CCC S(=O)(=O)O[Na]	-6.36	**-1.78**	**-1.86**
	8	c1ccccc1CN(C)(C)CCC S(=O)(=O)O[Na]	-6.85	**-2.27**	**-2.27**
Alkyl butane sulfobetaines, n=4	**16**	CCCCCCCCN(C)(C)CCCC S(=O)(=O)O[Na]	-4.63	**-0.05**	**-0.36**
	15	CCCCCCCN(C)(C)CCCC S(=O)(=O)O[Na]	-5.12	**-0.54**	**-1.23**
	14	CCCCCCN(C)(C)CCCC S(=O)(=O)O[Na]	-5.61	**-1.03**	**-1.08**
Phenylalkyl- and *para*-substituted phenylpropyl butane sulfobetaines, n=4	**22**	CCCCCCc1ccc(CCCN(C)(C)CCCC S(=O)(=O)O[Na])cc1	-2.38	**2.20**	**1.64**
	21	CCCCc1ccc(CCCN(C)(C)CCCC S(=O)(=O)O[Na])cc1	-3.36	**1.22**	**0.30**
	20	c1ccccc1CCCCN(C)(C)CCCC S(=O)(=O)O[Na]	-4.89	**-0.31**	**-1.41**
	19	c1ccccc1CCCN(C)(C)CCCC S(=O)(=O)O[Na]	-5.38	**-0.80**	**-1.70**
	18	c1ccccc1CCN(C)(C)CCCC S(=O)(=O)O[Na]	-5.87	**-1.29**	**-2.06**

	17	c1ccccc1CN(C)(C)CCCCS(=O)(=O)O[Na]	-6.36	**-1.78**	**-2.32**

Before we consider comparing experimental log P determinations and calculated log P generated *via* this demonstration program in any detail it is important to appreciate a significant limitation of this program. For many compounds the program simply predicts the same log P for obviously different compounds simply because the overall number of methylene groups is the same. For example, C8C2 (**1**), C7C3 (**4**) and C6C4 (**14**) are predicted the same value of log P (-1.03) and thus the program does not account for the transition of a methylene group from a hydrophobic region of the molecule (such as the main alkyl chain) into the spacer unit between the polar groups, which is rendered more hydrophilic than perhaps expected due to a proximity effect. Experimental results show that log P values for **1**, **4** and **14** are indeed unique. i.e. –0.06, -1.97 and –1.08 respectively.

Generally the program predicts log P of the alkyl, phenylalkyl and *para*-substituted phenylpropyl propanesulfobetaines reasonably well; however, log P calculations of the alkyl ethanesulfobetaines are artificially high and in contrast log P calculations of the phenylalkyl and *para*-substituted phenylpropyl butanesulfobetaines are artificially low when compared with our experimental values.

3.6 Conclusion

Log P values (for the 22 sulfobetaines previously synthesised) measured via the conventional stir-flask experiment coupled with reliable analytical techniques (RP-HPLC-UV and RP-HPLC-ESIMS) and involving analysis of both layers of the biphasic system will be used in QSAR development in Chapter 4, Part II. Evidence is provided to confirm that for accurate and reliable determination of log P, at least 2 separate experiments should be conducted and both layers of the biphasic system should be analysed. This method has the advantage that a calibration plot need not be constructed to determine the response factor of each analyte relative to an internal standard, since this value would simply cancel out when the concentration ratio is calculated.

When analysis of both layers is for some reason difficult or unachievable we recommend that for hydrophilic compounds, where possible, the octanol layer (the layer for which the solute has least affinity) should be analysed and the concentration in the aqueous layer should be determined by subtraction to achieve a reasonable *estimate* of log P. It should be appreciated

that if the aqueous layer is analysed instead and the concentration in octanol is determined by subtraction, it may appear that the concentration of solute in the aqueous phase is greater than that introduced into the system at the start of the experiment. A further drawback of just analysing one layer is that a calibration plot must be constructed to derive the RF of analyte relative to the internal standard so that a true value of concentration may be obtained rather than a simple factor that relates directly to this concentration. This is a time-consuming process.

It was also acknowledged that in the case of solutes that are not particularly hydrophilic or hydrophobic, and thus evenly distributed between octanol and water, a reasonable estimate of log P may be achieved from the analysis of either layer.

It is therefore expected, though not investigated, that for hydrophobic compounds the one layer to be analysed should be the aqueous layer. A reasonable estimate of log P should then result if the concentration of solute in octanol is determined by subtraction. If only the octanol layer is analysed then the situation may arise in which the concentration in octanol appears to be greater than that introduced to the total system.

It may be useful to explain why it is considered appropriate to determine log P for the zwitterionic sulfobetaines in contrast to log D. The simple reason is that the sulfobetaines are effectively *ionised* as opposed to ionisable and may be thought of as overall neutral with large local maxima and minima of charge density. These compounds retain their zwitterionic character over the entire pH range (i.e. they have an extremely broad isoelectric range). Therefore, there is no doubt that these compounds were in this form in all of the experiments conducted despite the lack of pH monitoring.[24]

It has also been shown that no current log P predicting program is capable of accurately determinating log P for sulfobetaines. The major drawback is that the estimates do not account for a reasonably complicated proximity effect that has been shown to exist between the oppositely charged (N^+ and SO_3^-) groups. The zwitterionic sulfobetaines have proved to be a complex but interesting group of compounds, for which reliable values of log P could not expect to be derived without direct experimentation. However, our experimental determinations can be used to amend the methods of estimating log P for sulfobetaines and perhaps estimate log P values for similar zwitterionic species.

3.7 References

1. Brooke, D. N., Dobbs, A. J. and Williams, N., *Ecotoxicol. Environ. Safety*, **11** (1986) 251.
2. Hansch, C. and Leo, A., *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*, ACS Professional Reference Book, American Chem. Soc. (1995) Chapter 4.
3. *OECD Guideline for the Testing of Chemicals*, **117**, HPLC Method.
4. Dearden, J. C. and Bresnen, G. M., *Quantitative Structure-Activity Relationships*, 7 (1988) 133.
5. Hansch, C., Leo, A. and Elkins, D., *Chem. Reviews*, **71** (1971) 525.
6. Purcell, W. P., Bass, G. E. and Clayton, J. M., *Strategy of Drug Design: a Molecular Guide to Biological Activity*, Wiley Interscience, NY (1973)
7. Quigley, J., Fahelelbom, K., Timoney, R., Corrigan, O., *Intn. J. Pharm.*, **58** (1990) 107.
8. *OECD Guideline for the Testing of Chemicals*, **107**, "Partition Coefficient (n-octanol/water)", Flask Shaking Method, Adopted 12th May 1981.
9. Rekker, R. F., *The Hydrophobic Fragment Constant*, Amsterdam Elsevier (1997).
10. Davis, S. S, Elson, G., Tomlinson, E., Harrison, G. and Dearden, J. C., *Chem. Ind. (London)*, (1976) 677.
11. Wang, P-H and Lien, E. *J., J. Pharm. Sci.*, **69** (1980) 662.
12. Taylor, P. J., in Ramsden, C. A (Ed.), *Comprehensive Medicinal Chemistry*, Vol. **4**, Pergamon, Oxford (1989).
13. Skoog, D. A., Holler, E. J. and Nieman, T. A., *Principles of Instrumental Analysis*, Fifth Edition, Harcourt Brace and Company (1998) Chapter 19.
14. Meyer, V. R., *Practical High Performance Liquid Chromatography*, Second Edition, J. Wiley (1994).
15. Snyder, L. R., Kirkland, J. and Glajch, J., *Practical HPLC Method Development*, J. Wiley (1997).
16. Scott, R. P. W., *Techniques and Practices of Chromatography*, Marcel Dekker Inc., NY (1995) Chapter 11.
17. Kevin Altria, Capillary Electrophoresis, in *Chemistry in Britain*, **36**, November 2000
18. Heinig, K, Vogt, C. and Werner, G., *J. Chromatography A*, **781** (1997) 17.
19. Heinig, K, Vogt, C. and Werner, G., *J. Chromatography A*, **745** (1996) 281.
20. Denkert, M., Hackzell, L., Schill, G. and Sjogren, E., *J. Chromatography*, **218** (1981) 31.

21. Ernst, R. and Miller, E. J., in Bluestein, B. R. and Hilton, C. L. (Ed.), *Surfactant Science Series*, Vol. **12**, *Amphoteric Surfactants*, Marcel Dekker, NY (1982), Chapter 2.

22. http://esc.syrres.com/interkow/logkow.htm

23. Lide, David R. (Editor in Chief), *Handbook of Chem. and Phys.*, 78th Edition, CRC Press Florida, Section 16.

24. Domingo, X., in Lomax, E. G. (Ed.), *Surfactant Science Series*, Vol. **59**, *Amphoteric Surfactants* (Second Edition), Marcel Dekker (1996), Chapter 3.

Chapter 4

Part I: Acute Aquatic Toxicity Testing

Part II: QSAR Correlation Studies

4.1.1 Introduction

Aquatic toxicology is concerned with protecting aquatic organisms from the adverse effects of synthetic chemicals. To anticipate how toxicants are likely to impact on the environment a simple laboratory test may be carried out with appropriate design and control so that the result can be assigned ultimately to the presence of the test compound. The importance of understanding the test procedures used in the assessment of aquatic toxicology cannot be underestimated. In every interpretation of an EC_{50} or NOEL there should be a clear understanding of the test method used to obtain that estimate. This understanding should include knowledge of the strength and weaknesses of the test method and the vagaries of the test organism or organisms.

Obviously the test should be ecologically realistic. i.e. mimic the natural situation. However, it must be appreciated that it is often not possible to predict precise ecological effects. Normally, the operator must be content to define levels likely to have no or a minimum effect. Reproducibility is a particularly important criterion for such tests. To ensure that the same result is generated for a particular test compound, irrespective of where and when the test is carried out, tests must be standardised and hence conducted in accordance with precisely defined protocols. e.g. OECD guidelines for the testing of chemicals.

For much of the history of aquatic toxicology, the bulk of the data has been generated by acute testing, using either mortality or immobility as the response end-point. It is also true that in QSARs for aquatic toxicity the determination of toxicity is typically an acute determination – based on only a short period of an organism's life span, commonly 24-96 h, depending on the test species employed.[1,2]

The measurement of acute aquatic toxicity of the majority of the sulfobetaines synthesised (Chapter 2) will be described, whose log P values were previously derived (Chapter 3) for use in QSAR development (Part II of this Chapter).

4.1.2 Acute Aquatic Toxicity Testing

4.1.2.1 Single-Species Tests

The single-species approach (as described by Section 1.2.3.4) is commonly used for predictive tests, since it is more likely to yield an accurate and reproducible result. A single-species test is also conducted in a laboratory environment, it can be easily standardised and is the most simple to conduct, and hence these are important reasons why this approach is in common use. The most significant species used in such tests are *Daphnia*, fish and algae.[2]

4.1.2.2 *Daphnia* as Test Species

Daphnia magna and *Daphnia pulex* are the most common species used in *Daphnia* testing. They have the advantage over many other single-species for aquatic toxicity testing that they are parthenogenic, they are particularly sensitive organisms and they are easily cultured in the laboratory.[2]

D. magna and *pulex* are widely distributed in fresh water where they are important fish-food organisms. They are planktonic invertebrates that feed by filtering phytoplankton and seston in the size range of 0.9 to 18000 μm^3. Figure 4.1.1 shows a representation of the anatomy of the female *Daphnia pulex*.[3, 4]

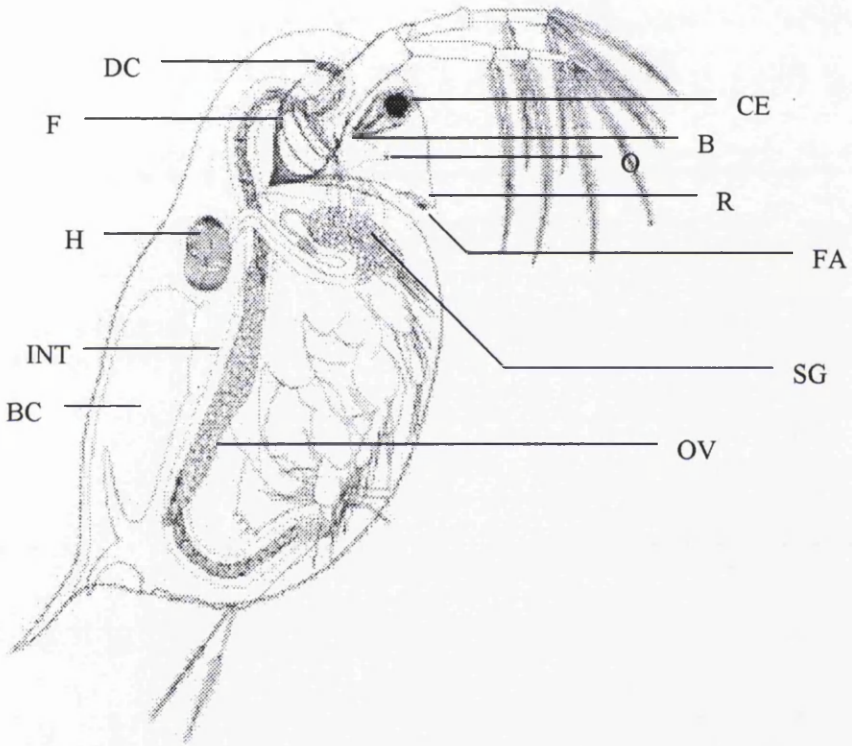


Figure 4.1.1 Anatomy of female Daphnia pulex
(B = brain; BC = brood chamber; DC = digestive caecum; CE = compound eye; F = fornix; FA = first antenna; H = heart; INT = intestine; O = ocellus; OV = ovary; R = rostrum (or beak); SG = shell gland)

D. magna are larger and tolerate harder water than *D. pulex*. In addition to this, *D. magna* are more commonly available and easier to culture, and for these reasons the use of *D. magna* in single-species toxicity tests is predominant.[2]

Whichever species is employed, it is recommended that the test organisms be derived from adults, three generations after introduction to the specific laboratory media. It is also

essential that the source and genetic history of the test organisms is known. Specifically, the original strain of the *D. magna* employed for the investigations described in latter sections of this chapter, was obtained from Shell (Tunstall Laboratories) in 1973. In 1985 the culture genotype was determined and this showed that the clone had remained genetically unchanged.[3,5]

4.1.2.2.1 Laboratory Culture of *Daphnia magna*

D. magna is held in culture around the world for use as a sensitive organism for determining the toxicity of individual chemicals, mixtures and effluents. Given suitable conditions *D. magna* will reproduce freely in laboratory cultures by an asexual process (parthenogenesis), which produces a monoculture of female *Daphnia*. Significant shifts in conditions can induce certain of the juveniles to develop as males so that sexual reproduction follows. In addition to this, the females may produce resting eggs in a chitinous case called ephippia. This biphasic life cycle is characteristic of the natural state and has survival value in the environment where it enables the species to survive severe conditions.

The culture medium used is Elendt M7, which is a complex medium that has the advantage of having a known chemical composition. It contains macro-nutrients, micro-nutrients, buffering nutrients and vitamins. Full details of the preparation of Elendt medium is given in Ecotoxicology SOP 019 11. Once prepared the water qualities of the batch should be determined and the medium must satisfy the following water quality criteria: pH in the range 7.5 – 8.5 and dissolved oxygen concentration in excess of 6 mg/l. The total water hardness should be in the range 200 – 260 mg/l as $CaCO_3$. Prepared media should be used within 1 month, and prior to use should be vigorously aerated for a minimum of 30 minutes before water qualities are determined.[3]

Cultures should be re-set when 30 days old by discarding the adults and replacing them with 10 neonates (less than 24 h old) taken from any culture containing neonates (from adults at least 12 days old). If there are no neonates released on the day when the two oldest cultures reach 30 days old, cultures should be retained until neonates become available and then re-set. When the *Daphnia* in each culture first become gravid the number of *Daphnia* is reduced to 7 and the surplus adults discarded. The cultures are staggered so that at any one moment there are a minimum of five groups of two cultures differing in age by approximately 6 days. This strategy should ensure an adequate supply of neonates for testing requirements and the continued maintenance of the cultures.[3]

Cultures should be maintained between 18 and 22 °C under a 16 h light, 8 h dark photoperiod. The temperatures of each culture should be recorded when neonates are removed for subsequent use. The cultures require daily maintenance. All neonates should be removed and discarded unless required for testing or re-setting cultures. Adult *Daphnia* are transferred with a wide bore (approx. 6 mm diameter) glass tube. Neonates may also be transferred in this way but if required for testing they can be removed from cultures by slowly pouring the medium through a fine-meshed muslin partially immersed in another beaker containing medium. If the neonates are required, the net is then swiftly and carefully inverted in another vessel of medium to release the neonates. On a daily basis, the number of live and dead adults are recorded and also the number of ephippia. It is also important to note that more than 2 ephippia in a culture is cause for investigation and if further ephippia are produced thereafter the culture should be discarded. Finally all cultures are fed daily with the alga *Chlorella vulgaris*. The algal cells have their own culturing process and *Daphnia* cultures should be given a fixed volume of algal culture on a daily basis. Occasionally the *Daphnia* cultures do not deplete their algal ration within one day. At such times it is not necessary to feed the culture until the algae have been consumed.[3]

4.1.2.3 *Daphnia* 48-h Acute Toxicity Test

This test along with the fish 96-h acute toxicity test is one of the standbys in aquatic toxicology. However, one of the major problems with conducting evaluations of fish toxicity is the reliable supply of healthy organisms. Wild organisms are often used in such tests that have to be acclimatised to the laboratory environment before carrying out the toxicity test. Furthermore, collecting the fish species may unduly stress the organism, which may lead to an overestimate of toxicity, and it is difficult to be sure that the locally collected organism is representative of the native population. Finally, vertebrates such as fish are generally considered less sensitive than invertebrates like *Daphnia*.[2]

The *Daphnia* 48-h acute toxicity test has the advantage that it has a short time frame, only small amounts of hazardous waste are generated during a typical test and the test is inexpensive to perform. Despite the advantage of sensitivity of *Daphnia* to test compounds, the sensitivity of the organisms to water quality is a disadvantage of using this test species in acute toxicity tests.[2]

4.1.2.3.1 Principle of the Test

Young daphnids or neonates, less than 24 h (at the start of the test) are exposed to the test substance at a range of concentrations, under semi-static conditions, for a period of 48 h. Obviously, neonates from a culture that has produced ephippia or has an unusually high mortality should not be used. It is also strongly recommended that first brood progeny are not used. The degree of immobilisation is recorded at 24 h and 48 h and compared with control values. The results are then analysed in order to calculate the EC_{50} at 24 and 48 h.[6,7]

4.1.2.3.2 Information on Test Substance

Useful information includes the structural formula, purity, water and photo-stability and relevant physical properties of the substance, such as log P, volatility, solubility and pK_a. It is important to note that the test is strictly designed to assess the toxicity of water-soluble, non-volatile test substances in aqueous solution. Modifications to the method may be required for substances known to be volatile, unstable or poorly soluble. The OECD guideline also recommends that a reliable analytical method for the quantification of the substance in the test solutions should be made available, so that where possible, results may be based on measured test concentrations as opposed to nominal concentrations.[6,7]

4.1.2.3.3 Validity Criteria

For the results of a toxicity test to be considered valid the following performance criteria apply: not more than 10 % of the daphnids should be immobilised in the control and the dissolved oxygen concentration at the end of the test should be ≥ 2 mg/l in control and test vessels. For the first criterion, the control daphnids should also show no signs of disease or stress, for example discolouration, unusual behaviour or trapping at the surface of the water, in addition to immobility. Additionally, for the acceptability of a test, the test vessel temperatures must be greater than 18 °C and less than 22 °C and must not vary by more than 2 °C.[6,7]

4. 2. 3. 4 Monitoring of Water Quality

Due to the sensitivity of *Daphnia magna* to the test conditions the water quality must be strictly monitored throughout a typical test. The water properties to be monitored are pH, DO and TH. The pH of all test solutions analysed should be in the region of 6–9, varying by no more than 1.5 in any given test. If when preparing the original stock solution there is evidence of marked change in the pH in the presence of the test substance (compared with pure Elendt media) the

pH of the stock may be adjusted. However, this assumes that the adjustment can be made in such a way that the concentration of the stock solution is not changed to any significant extent and that no chemical reaction or precipitation of the test substance is caused. HCl and NaOH are recommended by the OECD guideline for adjustment of pH.

Total hardness above 140 mg/l (as $CaCO_3$), but less than 250 mg/l is considered an acceptable range for any given test to determine the acute toxicity of a given test substance to *Daphnia magna*[6] and the OECD guideline recommends that the DO content should not fall below 2 mg/ml.[7]

4.1.2.3.5 Test Concentrations

Test concentrations are chosen so that ideally the lowest concentration gives no observable effect and the highest results in 100 % immobility with the intermediate concentrations providing a range of partial responses. If required a range-finding test may be conducted to determine the range of concentrations for the definitive test, unless an estimate of the toxicity of the substance is already available. For a typical definitive test at least five test concentrations are used and the concentration interval may be based on a logarithmic or geometric series. For a typical range-finding test three test concentrations are used and the concentration intervals would normally differ by a single order of magnitude. Test concentrations are prepared from a reasonably concentrated stock solution of the test substance and the dilution media used is Elendt. All tests include a control composed of Elendt media only and in all other respects treated identically to the test solutions containing varying concentrations of the test compound.[6,7]

4.1.2.3.6 Test Groups

Test vessels made from a chemically inert material, such as glass, are used to contain test media. These test vessels are filled with a constant volume of Elendt or test solution. At least 20 daphnids, preferably divided into four groups of 5 daphnids each, are used for each test concentration and for the controls in a definitive test. However, for a range-finding test, usually only 10 daphnids are employed for each test concentration and each control, and they are divided into only two groups of 5 daphnids each. It is recommended that at least 2 ml of test solution is provided for each daphnid.[7] Generally, the tests described in this chapter allow 10 ml of test solution for each daphnid (5 daphnids and a total of 50 ml test solution).

4.1.2.3.7 Incubation Conditions and Feeding

As already mentioned, the temperature should be within the range of 18 °C and 22 °C, and for each test it should be constant within ± 1 °C according to the OECD guideline. However, Ecotoxicology SOP 020 06, an internal protocol for *Daphnia* testing at Unilever, considers it acceptable for the temperature to vary by a maximum of 2 °C in any given test. It is recommended that the light cycle should be maintained as 16 h light/ 8 h dark with normal laboratory intensity, however complete darkness is also considered acceptable. The test vessels should not be aerated during the test and each test vessel should be loosely covered to minimise evaporation of water and avoid the entry of dust into the test solutions. The test is carried out without the adjustment of pH during its course, and finally, it is recommended that the daphnids are not fed for the 48 h duration of the test. [6,7]

4.1.2.3.8 Definition of Immobility

The end-point of any toxicity test involving *Daphnia* is usually immobility, since mortality is often extremely difficult to assign to small organisms like daphnids. Immobile daphnids are defined as those that are unable to swim within 15 seconds after gentle agitation of the test vessel (even if they can still move their antennae). At 24 h into the test, the numbers of immobile *Daphnia* are recorded and immobile daphnids are discarded during renewal of the test media. At 48 h the numbers of mobile versus immobile *Daphnia* are recorded.[7]

4. 2. 3. 9 Treatment of the Effect Data

The 48 h-EC_{50} is calculated using the mean of the measured concentrations where available. However, if no analytical determination of concentration is available nominal concentrations may be reported. The number of daphnids immobile at 48 h are expressed as a percentage, and then the 48 h-EC_{50} is calculated by appropriate statistical methods (Binomial, Moving Average and Probit method). The program employed for this purpose effectively plots the % immobility against the test concentrations and then the slope of the graph and the EC_{50} are calculated by each of the three statistical methods. The EC_{50} values quoted for use in QSAR development will be the figures displaying the narrowest 95 % confidence intervals.[6,7]

The Binomial method is a non-linear interpolation method that may be used to estimate the EC_{50} if there is less than two concentrations where a partial effect or no effect is observed. It makes no approximations and no assumptions and may be useful for data having a long, low tail

which defeats the Moving Average method. Estimates from the Binomial method are generally more conservative than those derived from both the Moving Average and Probit method.

The Moving Average method is an interpolation method, which has the broadest scope of all three methods, since it has been shown to be applicable to a wide range of data sets. The method assumes a monotonic response and the same number of organisms in each concentration. If the responses are not monotonic then the method will not recognise the deficiency and the output should not be used. Neither should the method be used when there are less than two partial effects or when the data has a long, low tail.

The Probit method is a parametric method that assumes a certain curve form and then determines the best fit of the actual data to that form. It is a widely used technique which is explained elswhere.[8] The method should not be used unless there are two or more partial effects, but it does have the advantage that it can take account of response data that is not monotonic.[9]

4.1.3 Experimental

4.1.3.1 Chemicals

Elendt M7 was prepared as described Ecotoxicology SOP 019 11; test compounds **1-5, 7-9, 11-13, 14-16, 18, 20-22** were synthesised as described by Chapter 2; **6** was purchased from Sigma.

4.1.3.2 General Procedure for Purity/Activity Analysis

0.0140 g of the internal standard, trioxan, was weighed in a 5 ml volumetric flask. The deuterated solvent ($CDCl_3$ or D_2O) was then added up to the mark. 0.0150 g of the analyte was weighed in a 5 ml vial and exactly 1 ml of the previously prepared trioxan solution was added using a gas tight syringe. After agitation to dissolve sample (where required), the solution was pressed through a pasteur pipette with plug of glass fibre filter paper, using a teat, into an NMR tube.

The following equation was used to calculate the % weight of the product: -

$$\text{Wt. \% of Product} = \frac{\text{Wt. of Trioxan} \times \text{No. H in Trioxan} \times \text{Integration for Product} \times \text{Mol. Wt. Product} \times 100}{\text{Integration for Trioxan} \times \text{No. H in Product} \times \text{Mol. Wt. Trioxan} \times \text{Wt. of Sample}}$$

(Equation 4.1.1)

4.1.3.3 General Procedure for a Definitive Toxicity Test

Neonates < 24 h old were removed from the cultures as described in Section 4.2.2.1 and Ecotoxicology SOP 019 11 on the day of the test, and transferred to the isolation medium (Elendt M7- whose water qualities were previously determined to ensure that conditions were similar to those of the culture medium). One litre of the stock solution was then prepared, usually at the top concentration of the proposed test. The appropriate mass of solute was weighed using a 4 d.p. analytical balance and the dilution medium used was Elendt M7. Prior to making the volume of the stock solution right up to the total volume (1000 ml mark on the graduated flask) it was considered essential to check the pH of the solution to confirm that the overall pH fell within the desired range for testing. In situations where the pH was outside this range, the pH was brought into the desired range by the dropwise addition of a concentrated solution of NaOH. The volume was then brought to the mark using Elendt, if required, following the pH rectification procedure. It is useful to note that the water qualities of the Elendt used as culture media, isolation media and dilution media must be comparable so as not to unduly stress the test organisms.

250 ml of each of the 5 test concentrations (chosen to vary according to a log scale) were then prepared using the stock solution and further dilution with Elendt.

24 glass test vessels (crystallising dishes, 120 ml) were then labelled 1 – 24. For each test concentration and the control there would be 4 replicates and to each test vessel approximately 50 ml of test media were added. At 0 h, water qualities were monitored as described by Table 4.1.1, exactly 5 neonates were added to each of the 24 test vessels and the vessels were loosely covered with glass watch glasses. The test vessels were randomly distributed according to numbered positions generated by a computerised program known as Random Bas. Random orientation of test vessels is recommended by Ecotoxicology SOP 019 11, so that the test operator is not biased with regard to observations.

At 24 h, a count of immobile *Daphnia* is completed and the test medium is renewed. Furthermore, only mobile daphnids are transferred to the new media. Water quality is determined as described in Table 4.1.1. Test solutions for the new media were prepared from the stock solution just prior to the changeover process, since these are expected to have only a 24 h expiry. In contrast, the stock solution is normally given a 1 week expiry.

At 48 h, a count of immobile versus mobile *Daphnia* is completed and once again the appropriate water qualities are determined to confirm that the results obtained are directly related to the presence of the test compound.

Finally, a light platform is used on occasion to aid with the assignment of 'immobility', *Daphnia* are not fed for the 48 h duration of the test, a 16 h photoperiod (8 h light, 8 h dark) is employed, and temperature controlled at 20 ± 2 °C.

Table 4.1.1 Water quality monitoring throughout a typical definitive test

Time (h)	New/Old Media	pH and D.O.	Total Hardness	Temperature
0	new	all concentrations	0.0 and highest concentration	Check temp. of 6 randomly chosen test vessels
24	new	0.0 and highest remaining concentration	0.0 and highest remaining concentration	Check temp. of the 6 randomly chosen test vessels
24	old	all concentrations	0.0 and highest concentration	Check temp. of the 6 randomly chosen test vessels
48	old	all remaining concentrations	0.0 and highest remaining concentration	Check temp. of the 6 randomly chosen test vessels

Table 4.1.2 defines the concentration of stock solution prepared for each test compound and the range of test concentrations chosen for each test performed. Usually, only one determination of the EC_{50} was made (Experiment A). However, there were a few situations (compounds **3**, **15** and **5**) where two determinations were made (Experiments A and B) so that an average EC_{50} could be taken.

Table 4.1.2 Table highlighting the concentration of stock solution and range of test concentrations used for each test

Test Compound	Test	Conc. of stock solution (mg/l)	Range of test concs. employed (mg/l)
C6C3 (**3**)	A	3200	0, 320, 560,1000, 1800, 3200
	B	1400	0, 100, 180, 320, 560, 1000
C6C4 (**14**)	A	5600	0, 560, 1000, 1800, 3200, 5600
C7C3 (**4**)	A	18000*	0, 1800, 3200, 5600, 10000, 18000
C7C4 (**15**)	A	10000	0, 1000, 1800, 3200, 5600, 10000
	B	18000	0, 560, 1800, 5600, 10000, 18000
C8C2 (**1**)	A	3200	0, 320, 560, 1000, 1800, 3200
C8C3 (**5**)	A	5600	0, 560, 1000, 1800, 3200, 5600
	B	5600	0, 560, 1000, 1800, 3200, 5600

C8C4 (**16**)	RF	10000	0, 100, 1000, 10000
C10C3 (**6**)	A	1000	0, 100, 180, 320, 560, 1000
C12C2 (**2**)	A	100	0, 10, 18, 32, 56, 100
C12C3 (**7**)	A	50	0, 3.2, 5.6, 10, 18, 32
PhC1C3 (**8**)	A	80000*	0, 5600, 10000, 18000, 32000, 56000
PhC2C3 (**9**)	A	40000*	0, 3200, 5600, 10000, 18000, 32000
PhC2C4 (**18**)	A	20000	0, 1800, 3200, 5600, 10000, 18000
PhC4C3 (**10**)	A	5600*	0, 560, 1000, 1800, 3200, 5600
PhC4C4 (**20**)	A	5600	0, 560, 1000, 1800, 3200, 5600
C4PhC3C3 (**12**)	A	180	0, 18, 32, 56, 100, 180
C4PhC3C4 (**21**)	A	500	0, 150, 220, 340, 500
C6PhC3C3 (**13**)	A	1000	0, 18, 32, 56, 100, 180
C6PhC3C3 (**22**)	A	320	0, 32, 56, 100, 180, 320

** indicates that pH adjustment of stock solution was required using NaOH to bring pH into the acceptable range for testing; RF = range-finder*

4.1.3.4 General Procedure for a 'Range-Finder' Toxicity Test

Procedure was as described in Section 4.1.3.3 with the following exceptions: the test vessels were labelled 1–8, test concentrations varied by an order of magnitude, a total of three concentrations and a control were employed, and for each there were only two replicates. Water qualities were determined as defined in Table 4.1.3. Finally, it is important to note that the results for the few range-finder tests conducted will not be recorded in this thesis, since they were simply employed on occasion to highlight an appropriate concentration range for definitive testing on a specific sulfobetaine. However, in the case of **16**, we only have the results of the range-finder test since time and resources did not permit us to perform a definitive test. In this instance the result is recorded (Table 4.1.2).

Table 4.1.3 Water quality monitoring throughout a typical 'range-finder' test

Time (hrs)	New/Old Media	pH and D.O.	Total Hardness	Temperature
0	new	0.0 and highest concentration	0.0 and highest concentration	Check temp. of 2 randomly chosen test vessels
24	old	0.0 and highest concentration	0.0 and highest concentration	Check temp. of the 2 randomly chosen test vessels

4.1.4 Results and Discussion

4.1.4.1 Purity/Activity Analysis

Prior to toxicity testing at SEAC, Unilever Port Sunlight/Colworth, it was decided to submit some samples of the test compounds for quantitative ^{1}H NMR spectroscopy. The NMR assay, using a suitable internal intensity standard, is extremely useful for assessing the purity (by % weight) of compounds. The internal standard employed in our investigations was trioxan, for which all protons are equivalent and thus only a single peak is observed in NMR. It was originally felt that that this type of analysis was a necessity, so that our experimental determinations of EC_{50} could be amended according to purity and depending on the nature of contaminants/impurities before incorporation into the final QSAR. Table 4.1.4 shows the results achieved *via* an NMR assay for 5 of the sulfobetaines.

Table 4.1.4

Test Compound	% weight (1)	% weight (2)	% weight (average)
8	95.7	96.4	**96.1** ± 3
7	98.4	97.2	**97.8** ± 3
5	94.2	96.0	**95.1** ± 3
16	100.1	96.3	**98.2** ± 3
3	93.9	93.4	**93.7** ± 3

In each case the assay also provided confirmation that there were no organic impurities. We could therefore conclude that these compounds were generally tested in high purity and any contamination was likely to be due to water since no inorganic materials were used in the synthesis of these sulfobetaines (Chapter 2). This was confirmed by the fact that the compound with the lowest purity, C6C3 (**3**), is the most hygroscopic and therefore may have crystallised as a hydrate.

Due to the high purity associated with the first few compounds analysed and the fact that each result is only accurate to ± 3 %, it appeared to be no need to conduct an NMR assay to assess the purity of every one of the 19 sulfobetaines to be tested. For consistency, it was also decided not to amend the EC_{50} of the 5 compounds analysed, since their purity by percentage weight was so close to 100 % (taking into account the ± 3 % error).

4.1.4.2 Monitoring of Water Quality

As emphasised in the earlier sections of this chapter, *Daphnia* are sensitive to fluctuations in the quality of their surrounding aqueous environment. Section 4.1.2.3.4 gives full details of the tolerance range for water properties such as temperature, pH, DO and TH.

Table 4.1.5 Table to show the variation in water quality (temp., pH, DO and TH) throughout each toxicity test conducted

Test Compound	Test	Temp. range (°C)	pH range	DO range (mg/l)	TH range (mg/l as $CaCO_3$)
C6C3 (**3**)	A	19.5 – 20.0	7.0 – 7.5	7.3 – 7.8	230-264
	B	21.0 – 21.5	7.2 – 7.9	8.4 – 8.9	204 - 250
C6C4 (**14**)	A	20.0	7.8 – 8.8	7.6 – 8.0	222 – 254
C7C3 (**4**)	A	20.0	7.6 – 7.8	7.3 – 8.0	220 – 256
C7C4 (**15**)	A	19.5 – 20.0	6.9 – 7.5	6.9 – 7.8	230 – 262
	B	21.0 – 21.5	7.7 – 8.3	8.3 – 8.9	188 – 250
C8C2 (**1**)	A	19.5 – 20.0	6.7 – 7.7	7.2 – 7.8	232 – 266
C8C3 (**5**)	A	21.5 – 22.5	7.0 – 7.5	7.0 – 8.0	191 – 236
	B	21.0 – 22.0	7.5 – 8.4	7.3 – 8.2	222 – 236
C8C4 (**16**)	RF	21.5 – 22.0	7.2 – 7.8	6.9 – 7.6	220 – 255
C10C3 (**6**)	A	20.5 – 21.0	7.4 – 8.0	7.9 – 8.9	184 – 250
C12C2 (**2**)	A	20.0	6.8 – 7.6	7.3 – 7.7	220 – 268
C12C3 (**7**)	A	21.0 – 21.5	7.7 – 8.2	6.7 – 7.5	205 – 270
PhC1C3 (**8**)	A	21.5 – 22.0	6.8 – 7.7	7.0 – 8.0	200 – 245
PhC2C3 (**9**)	A	21.5 – 22.0	6.4 – 7.7	6.9 – 7.4	201 – 255
PhC2C4 (**18**)	A	21.0 – 22.0	7.7 – 8.2	8.0 – 8.9	180 – 250
PhC4C3 (**10**)	A	19.5 – 20.0	7.7 – 8.8	6.9 – 7.6	218 - 254
PhC4C4 (**20**)	A	19.0 – 19.5	7.2 – 8.6	7.1 – 7.9	216 - 250
C4PhC3C3 (**12**)	A	19.5 – 20.0	7.7 – 8.8	7.0 – 8.0	218 - 256
C4PhC3C4 (**21**)	A	19.0 – 19.5	7.5 – 8.7	7.0 – 7.6	220 - 250
C6PhC3C3 (**13**)	A	18.0 – 20.0	7.0 – 8.4	7.3 – 8.0	238 - 262
C6PhC3C3 (**22**)	A	20.0	7.6 – 8.6	7.3 – 8.0	234 - 258

Since in each case the water properties remained within the permitted range, it may be concluded that for each test the result obtained is ultimately due to the presence of the test compound and not due to intolerance to the quality of the surrounding aqueous media.

4.1.4. 3 Effect Data

The raw data tables for each test conducted are given by Tables B2–B22 of Appendix B. They include the observed number of test organisms immobile at both 24 h and 48 h, and conversions to % immobility. The results of 22 separate tests are reported. In all but three cases, the raw data for a typical definitive test are reported, where there were 5 test concentrations varying over a log scale and 4 replicates for each of the chosen concentrations and the control. For compound **16** the results of a range-finder test is reported, since time and resources did not permit a definitive test to be undertaken. In this case only three test concentrations were employed varying by an order of magnitude and there were only two replicates for each of the chosen concentrations and the control. For these reasons, the result must be treated with caution and incorporation into the QSAR is questionable. For compound **6**, the number of neonates was restricted on the day of the test, and thus even though a typical definitive test was planned, for 4 of the 5 five test concentrations only 3 replicates could be employed. However, the result is still considered reliable and acceptable for use in QSAR correlation studies.

Finally, due to the shortage of test compound and hence stock solution of **6**, only 4 test concentrations were employed. However, it is useful to note that use of the full number of replicates was permitted and thus the EC_{50} result obtained is still considered reliable and acceptable for use in QSAR correlation studies.

A much more important concern is the purity of the batches of PhC4C3 (**11**) and PhC4C4 (**20**) synthesised for use in toxicity testing. These compounds were synthesised *via* a reaction involving *N,N*-dimethylphenylbutylamine and the appropriate sultone as described in Chapter 2. When the reaction was conducted on a small scale both of these sulfobetaines were isolated as white crystalline solids, however on scaling up these reactions to achieve larger quantities for toxicity testing they were isolated as pale brown solids. Attempts were made to remove the unexpected colouration but were unsuccessful. Throughout the course of the toxicity tests performed on these test compounds it was observed that orange material crystallised out of the test media and an unfavourable, distinct thiol-like smell was detected. It was expected that this insoluble impurity was inorganic, since it was not 'picked up' by any of the methods commonly used in the characterisation of organic compounds. However, the contaminating species has not been identified. As a result, it is appreciated that the EC_{50} determinations for these tests cannot be *directly* attributed to the presence of the test compound, and due to the inability to confirm the identity and level of the by-product formed in the large scale synthesis of the amine, these results were not used in QSAR development (Part II).

Table 4.1.6 shows the prediction of EC_{50} generated *via* each of the three statistical methods: the binomial, moving average and probit method, for each acute aquatic toxicity test performed. For each estimate of the EC_{50} the 95 % confidence limits (CL) are also given.

Test Compound	Test	Binomial Estimate of EC_{50} (mg/l)	95 % CL (mg/l)	Moving Average Estimate of EC_{50} (mg/l)	95 % CL (mg/l)	Probit Estimate of EC50 (mg/l)	95 % CL (mg/l)
C6C3 (**3**)	A	911	560, 1800	819	707, 957	834	703, 991
	B	496	320, 1000	506	439, 574	479	0, ∞
C6C4 (**14**)	A	5191	>3200	5191	4060, 23217	5703	3884, 13722
C7C3 (**4**)	A	4447	3200, 5600	4447	3946, 5114	4841	4094, 5728
C7C4 (**15**)	A	6757	5600, 10000	6757	5790, 7587	6054	0, ∞
	B	10396	5600, 18000	9231	7730, 10831	7768	0, ∞
C8C2 (**1**)	A	1000	560, 1800	967	789, 1180	965	790, 1178
C8C3 (**5**)	A	3533	1800, 5600	3228	2806, 3721	3214	2746, 3764
	B	2979	1800, 5600	*no estimate*	---	*no estimate*	---
C8C4 (**16**)	RF	2481	1000, 10000	*no estimate*	---	*no estimate*	---
C10C3 (**6**)	A	465	180, 1000	418	352, 495	422	349, 512
C12C2 (**2**)	A	20	10, 32	20	18, 23	20	18, 23
C12C3 (**7**)	A	30	>18	30	0, ∞	35	23, 106
PhC1C3 (**8**)	A	8810	5600, 18000	8983	7769, 10223	8898	7694, 10274
PhC2C3 (**9**)	A	3895	3200, 5600	*no estimate*	---	*no estimate*	---
PhC2C4 (**18**)	A	6752	3200, 10000	6503	5321, 7767	6008	4870, 7441
PhC4C3 (**10**)	A	1210	1000, 1800	1210	1034, 1361	1188	1037, 1370
PhC4C4 (**20**)	A	1171	1000, 1800	*no estimate*	---	*no estimate*	---
C4PhC3C3 (**12**)	A	61	32, 100	69	59, 80	67	57, 80
C4PhC3C4 (**21**)	A	460	>340	460	420, 533	468	409, 598
C6PhC3C3 (**13**)	A	56	32, 100	*no estimate*	---	*no estimate*	---
C6PhC3C3 (**22**)	A	149	100, 320	139	118, 166	135	48, 552

Comments: all concentrations (mg/l) are recorded to the nearest whole no.; ∞ = infinity.

For each test compound, the best estimate of the EC_{50} is taken as the prediction with the narrowest confidence interval associated with it. The confidence interval may be defined as the difference between the highest and lowest confidence limits. For those compounds for which only one estimate of the EC_{50} could be derived, this was the value that was taken forward for QSAR development, and for those compounds for which two separate toxicity tests were conducted, an average of the best estimate in each case was utilised in QSAR development. Table 4.1.7 summarises the best estimate of the EC_{50} for each test compound in units of mg/l and mol/l.

Table 4.1.7 Best estimate of EC_{50} for each of the 19 test compounds studied

Test Compound	Best Estimate of EC_{50}	
	mg/l	mol/l
C6C3 (**3**)	662	2.64 E-03
C6C4 (**14**)	5191	19.6 E-03
C7C3 (**4**)	4447	16.8 E-03
C7C4 (**15**)	6757	24.2 E-03
C8C2 (**1**)	965	3.64 E-03
C8C3 (**5**)	3103	11.1 E-03
C8C4 (**16**)	2481	8.47 E-03
C10C3 (**6**)	418	1.36 E-03
C12C2 (**2**)	20	6.23 E-05
C12C3 (**7**)	35	1.04 E-04
PhC1C3 (**8**)	8983	35.0 E-03
PhC2C3 (**9**)	3895	14.4 E-03
PhC2C4 (**18**)	6008	21.1 E-03
PhC4C3 (**11**)	1171	3.92 E-03
PhC4C4 (**20**)	1210	3.87 E-03
C4PhC3C3 (**12**)	69	2.02 E-04
C4PhC3C4 (**21**)	460	1.30 E-03
C6PhC3C3 (**13**)	56	1.52 E-04
C6PhC3C4 (**22**)	139	3.63 E-04

4.1.5 Conclusion

EC_{50} values have been determined for a total of 19 compounds; however only 17 were considered to be suitable for use in QSAR development (Part II of this chapter). As previously described, test compounds PhC4C3 (**11**) and PhC4C4 (**20**) were thought to be subject to contamination due to the formation of an unexpected by-product during large-scale synthesis, and thus the EC_{50} determined in these tests cannot be attributed to the sole presence of the test compound. It was unfortunate that one or two of the tests conducted deviated slightly from the typical definitive test. The results of such tests were, nevertheless, weighted with similar credibility to the other results, and with comparable accuracy for use in QSAR correlation studies.

Generally, without any attempt to be quantitative at this stage, it does appear that there is an inverse relationship between hydrophobicity and toxicity as expected. However, the strength of this relationship can only be confirmed when the log P-based QSAR is established and validated through consideration of the correlation coefficient for the data set.

The experimental portion of this chapter wascompleted in adherence to the updated Ecotoxicology Standard Operating Procedure (SOP) for the determination of the acute aquatic toxicity of test substances to *Daphnia* magna. This is an internal SOP at SEAC, Unilever Research, which has been compiled in close proximity to the well-documented OECD guideline. Advantages of conducting such tests in accordance with GLP, SOPs and/or OECD guidelines, include automatic validity, credibility and acceptability of results by several international scientific organisations and increased inter-laboratory reproducibility and accuracy.

In the evaluation of *Daphnia* magna Straus, it was found that this test species was reliable, tolerant to small fluctuations in water quality of the test media that are characteristic of such tests, and adequately sensitive to the presence of the test compounds. However, it was unfortunate from a synthetic point of view that a few of the test compounds (phenylalkylsulfobetaines) were found to display comparatively low toxicities. This meant that extremely large quantities of the test compounds were required (at the extreme 56 g, to prepare a 56 g/l stock solution). Scaling-up of reactions was not only time consuming, but more importantly expensive, since large quantities of expensive starting materials were required. In contrast, from a testing point of view, the analysis of compounds with low as well as high toxicities was beneficial, since it resulted in an increase in the range of EC_{50} values generated during this study.

An interesting extension to the toxicity studies of the sulfobetaines, would be to vary the test species in an acute toxicity test. It has been shown that algae are particularly sensitive to cationic compounds. If this was also true of zwitterionic compounds it would mean that the same range of sulfobetaines could be tested using smaller quantities of test compound in each case.

This would reduce the need for synthesis on such a large scale. However, a major drawback of the use of algae as the test species is the increase in the inter-laboratory variability when compared with *Daphnia*. It is estimated that for *Daphnia*, 20 % inter-laboratory variability is expected, however, for algae this figure is expected to increase to 50 %.[10]

Toxicity testing is extremely useful and often invaluable in environmental risk assessment, but it is time consuming and considered inhumane by many, despite attempts to abandon tests on many larger, complicated vertebrates species, including some fish species.

So far only single-species toxicity tests have been conducted, which are extremely useful in preliminary investigations into the toxicological properties of a given compound, but further testing is required in the form of multi-species testing which is expected to lead to a more realistic appreciation of the effects of sulfobetaines in the real aquatic environment that naturally includes several diverse species. Another important consideration is the effect of many suspected toxicants in the same environment. Testing using mixtures of test compounds is also very important since they often highlight synergistic or antagonistic properties of chemicals. These properties are discussed in further detail in Section 1.2.2.8.

The aim of a toxicity test is to determine an end-point that expresses the toxicity, under suitably controlled conditions so that the observation can be related directly to the presence of the test compound. The obvious difference between a test and the real situation is the control of so many parameters, but this is obviously essential if we are to achieve a result that is reproducible, and that can be also be compared with the toxicity of other test compounds. However, what we must appreciate is that the overall effect that a given test compound may have on the aquatic environment has the potential to be markedly different to that mimicked in the laboratory, due to the variability of the natural aquatic environment. Thus we could never exactly simulate the real situation, but are able to generate values relating to toxicity that are extremely useful for comparing the toxicity of one test compound with another, so that an estimate can be made as to which is likely to be more hazardous to the environment.

4.1.6 References

1. Calow, P., *Handbook of Ecotoxicology*, Vol. 1, Blackwell Science Publications (1994).
2. Landis, Wayne G. and YU, Ming-Ho, *Introduction to Environmental Toxicology*. Impacts of Chemicals Upon Ecological Systems, Lewis Publishers (CRC Press Inc.) (1995), Chapter 4.
3. Ecotoxicology **SOP 019 11**, *Laboratory Culture of Daphnia magna* (Unilever Internal Document).
4. Rand, G., *Fundamentals of Aquatic Toxicology*: Effects, Environmental Fate and Risk Assessment, 2nd Edition, Taylor and Francis (1995).
5. Calow, P., University of Sheffield. Personal Communication to Unilever in 1985. Report of a contract study for Water Research Centre, Medmenham.
6. Ecotoxicology **SOP 020 06**, *The Determination of the Acute Toxicity of Test Substances to Daphnia magna* (Unilever Internal Document).
7. OECD Guidelines for the Testing of Chemicals, *Revised Proposal for Updating Guideline 202, Daphnia sp., Acute Immobilisation Test*, Revised Draft Document, October 2000.
8. Finney, D. J., *Probit Analysis*, 3rd Edition Cambridge UP 1971, p. 333.
9. Stephan, C. E., Methods for Calculating an LC50. Aquatic Toxicology and Hazard Evaluation, ASTM STP 634, F. L. Mayer and J. L. Hamelink, Eds., ASTM (1977), 65.
10. Hodges, G., personal communication.

4.2.1 Introduction

QSARs for aquatic toxicity have proved to be extremely valuable in environmental risk assessment of surfactants. They allow the prediction of toxicological properties without the need for large-scale synthesis and biological testing in the early stages of a screening process for new surfactants. Additionally, QSARs can be used to indicate the mode of action of a given class of compounds. Log P-based QSARs have been established as reliable mathematical models to predict aquatic toxicity of various non-ionic and anionic surfactant compounds. These QSARs were derived from a series of compounds of both surfactant and non-surfactant nature. Recently, such a QSAR was developed for a series of cationic compounds, of the quaternary ammonium type. Hence, the only class of surfactant for which QSARs remain undeveloped are those of zwitterionic character, despite their wide commercial interest. Here a log P-based QSAR for the zwitterionic sulfobetaines will be derived.

4.2.2 Surfactant-Relevant Toxicity QSARs

QSARs have proved to be useful for the prediction of the physical fate of chemicals in the environment. The application of the QSAR approach to model bioaccumulation, biodegradation and aquatic toxicity (acute and chronic) has been extensive.[1] For example, Hawker and Connell[2] derived a log-P based bioaccumulation QSAR for a series of aromatic, heterocyclic and alicyclic compounds using *Daphnia* as the test species. The QSAR equation is given by Equation 4.2.1.

$$\text{Log BCF} = 0.898 \log P - 1.315 \quad (n = 22;\ r^2 = 0.93)$$

(Equation 4.2.1)

An example of a biodegradation QSAR is given by Dearden and Nicholson.[3] They derived a QSAR based on S_E (electrophilic superdelocalisability) that is given by Equation 4.2.2.

$$\text{BOD} = 0.093\ S_E - 3.163 \qquad (n = 19;\ r^2 = 0.96)$$

(Equation 4.2.2)

While there is awareness of such QSARs and their importance is appreciated, this chapter is only concerned with QSARs for the prediction of aquatic toxicity.

Since the recognition of the correlation between log P of an organic compound and aquatic toxicity, the literature is now replete with reports detailing the use of QSAR modelling to predict the aquatic toxicity of various industrial chemicals. All these references and many

other published QSARs deal with the correlation of a certain biological property with one or more physicochemical property of the organic compounds. By far the most common are correlations using log P as the structural property descriptor. A simple model is often used where,

$$\log (1/EC_{50}\ 48\ h) \text{ or } \log (1/LC_{50}\ 96\ h) = a \log P + b \qquad \textit{(Equation 4.2.3)}$$

The constants a and b will depend on the chemical structure and often on the animal species employed in the investigation. The accuracy of the model will obviously depend on the data used to build the model. QSARs can be used to predict the biological properties of 'unknowns' if they structurally resemble the group of chemicals originally used to derive the QSAR.

The mode of toxic action most easily predicted using QSARs is non-specific (or baseline) narcosis (Section 1.3.5).[1] General anaesthetics and many other compounds studied by pioneering workers in this field are thought to act in this way.[4] When toxicity is expressed in terms of concentration in the aqueous phase to which the organism is exposed, the characteristic non-specific narcotic QSAR has a single descriptor for hydrophobicity (usually log P) and is linear over the log P range of 2 to 5.5 with a slope equal to or approaching 1.[4,5] Relationships of this type have been reported for a multitude of biological systems including bacteria,[6,7] algae,[8] invertebrates (e.g. *Daphnia*),[9] fish[5,10] and isolated cells from fish and mammals.[11] The intercept of the QSAR line varies with the sensitivity of the correlated response, but both lethal and sub-lethal end-points give rise to the same general relationship. The universality of such QSARs indicates that narcosis (or anaesthesia) is a response of fundamental importance in ecotoxicity. The observation that there are no specific molecular features required to produce narcosis, and that the response is completely reversible,[12] provides a clue to the mechanism of action.

The acute aquatic toxicity of chemically unreactive non-electrolytes, such as hydrocarbons, ethers, alcohols and ketones are well-correlated with their octanol/water partition coefficients by the general narcosis equation given by Equation 4.2.4. Roberts later demonstrated the applicability of this equation to aquatic toxicity of non-ionic surfactants.[13]

$$\text{Log } (1/EC_{50}) = 0.87 \log P + 1.13 \qquad \textit{(Equation 4.2.4)}$$

The parameter EC_{50} is defined as the concentration causing a specified effect in 50 % of the population of test organisms. For Equation 4.2.4 (and Equation 4.2.5; introduced later) the specified effect was originally lethality and the original authors use LC_{50}. Equation 4.2.4 was

derived from a set of experimental guppy EC_{50} values and experimental determined log P values for fifty compounds. It has subsequently been found to apply for many other compounds and for a variety of aquatic species. For tests on *Daphnia,* EC_{50} is obviously related to 50 % immobilisation.[13]

While much progress has been made, and is being made, towards the development of QSARs for non-electrolytes, only a limited amount of work has been carried out on organic electrolytes, which are completely or partially ionised in water. However, QSARs have been reported for phenols.[14-17]

Phenols exist in equilibrium between the neutral and charged species and are more toxic than predicted by narcosis baseline toxicity models, such as Equation 4.2.4. This indicates that these compounds have specific modes of toxic action. A polar narcosis equation (Equation 4.2.5) was seen to apply to a variety of phenolic compounds and was derived from data on fathead minnow toxicity. As with Equation 4.2.4, inter-species variability tends to be small and the equation may be applied generally to fish and to aquatic invertebrates that derive oxygen from ambient water.[13]

Anionic surfactant toxicities are well-correlated by equations that are intermediate between Equations 4.2.4 and 4.2.5, but were confirmed to act as polar narcotics.

$$\text{Log } (1/EC_{50}) = 0.63 \log P + 2.52 \qquad \textit{(Equation 4.2.5)}$$

Even though many compounds have proved to be more toxic that Equation 4.2.4, there are only a few known cases of abnormally low toxicity. These are explicable in terms of low bioavailability. e.g. where the solubility in water is lower than the predicted EC_{50}.[13]

Recently, Singh, Lin and Bockris attempted to model the cationic quaternary ammonium salts for aquatic toxicity behaviour using the fathead minnow as the test species. The QSAR they presented, given by Equation 4.2.6, indicates that this class of compounds are significantly more toxic than even polar narcosis predicts.[18]

$$\text{Log } (1/EC_{50}) = 1.08 \log P + 4.19 \qquad (n = 8;\ r^2 = 0.97) \qquad \textit{(Equation 4.2.6)}$$

The log P values for the quaternary ammonium compounds used in the derivation of the above QSAR were obtained by calculation using the L and H approach. Roberts soon acknowledged that the log P calculation method for these compounds was misleading and amended this QSAR as described in Chapter 5 (Equation 4.2.7).[19]

$$\text{Log } (1/EC_{50}) = 0.72 \log P + 3.19 \qquad \textit{(Equation 4.2.7)}$$

This brings the QSAR more in line with the standard polar narcosis equation given by Equation 4.2.5. Roberts has also completed a short mixture toxicity study, the results of which suggest that cationic compounds of the quaternary ammonium type are indeed polar narcotics.[19]

Until this time the literature has lacked reports of any effort to model zwitterionic compounds (either surfactant and/or non-surfactant in nature) for aquatic toxicity behaviour. In this part, findings linked with the development of a log P-based QSAR for the prediction of acute aquatic toxicity for zwitterionic sulfobetaines are reported.

4.2.3 Experimental

For construction of an aquatic toxicity QSAR for the zwitterionic sulfobetaines two key parameters were required. One being log P values for the chosen data set and the other being the corresponding EC_{50} values.

Determination of octanol/water partition coefficients for the sulfobetaines were made experimentally using the conventional stir-flask experiment as described by Chapter 3. The EC_{50} values for aquatic toxicity to *Daphnia magna* Straus were also obtained through experimental determination. The specified effect was immobilisation (as described in Part I of this Chapter). Once the two reliable sets of data were obtained for a total of 17 sulfobetaines, simple regression analysis was performed to establish the linear relationship between log P and log $1/EC_{50}$ (48 h) for the sulfobetaines. It appeared that compound **3** did not fit the correlation and including it in the correlation significantly lowered the correlation coefficient. As a result, 16 sulfobetaines were used in the derivation of the QSAR model. The 16 sulfobetaines were **1**, **2**, **4-9**, **12-16**, **18**, **21** and **22**.

4.2.4 Results and Discussion

Good correlation was observed between log P and log $1/EC_{50}$ for the 16 compounds. The correlation coefficient for the data set was 0.87 (Figure 4.2.1) which is considered extremely reliable if we consider that *both* parameters were determined experimentally with their individual associated errors. The QSAR equation is summarised by Equation 4.2.8.

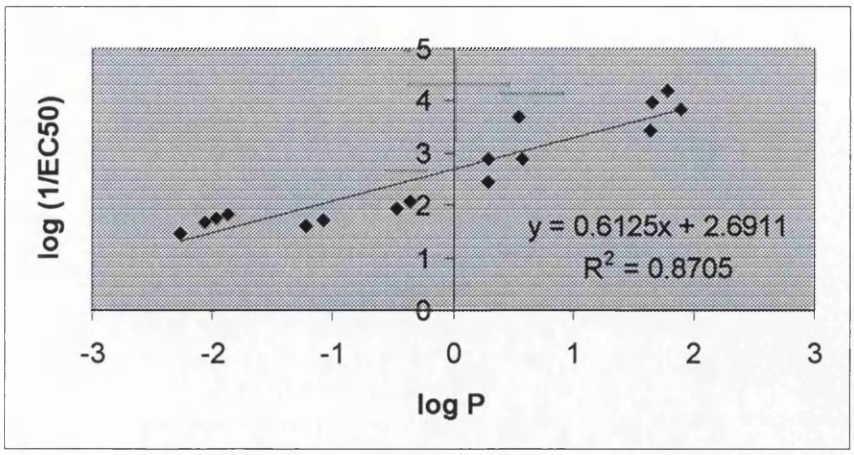


Figure 4.2.1 *Relationship between Log P and Acute Aquatic Toxicity for a Series of Zwitterionic Sulfobetaines*

The 87% correlation between aquatic toxicity and the log of the octanol/water partition coefficient suggests that the toxicity of the sulfobetaines is largely dependent on the hydrophobicity of the molecule. However, since the correlation between these parameters does deviate from 100 % we cannot rule out that other factors, aside from hydrophobicity, come into play to affect the overall toxicity.

$$\text{Log } 1/EC_{50} = 0.61 \log P + 2.69 \qquad (n = 16;\ r^2 = 0.87) \qquad \textit{(Equation 4.2.8)}$$

4.2.5 Conclusion

A linear relationship exists between log P and toxicity for sulfobetaines, that gives rise to a simple linear QSAR equation as expected. The derived QSAR equation given above, is extremely close to that of the standard polar narcosis equation (*cf.* Equation 4.2.5), which indicates that sulfobetaines are indeed polar narcotics. This is in agreement with other ionic compounds previously studied by a range of independent authors. This equation can now be used for the prediction of toxicity of other sulfobetaines, for which log P has already been deduced by calculation (as discussed in Chapter 5) or by experiment.

It is extremely interesting that compound **3** does not fit the correlation, and hence appears not to obey the same mode of action as all other members of the data set**.** Since it is the smallest molecule in the series, it is expected that size becomes more important than hydrophobicity. It is postulated that its low molecular weight allows it to penetrate the membrane irrespective of the associated log P value. It is possible that it may be small enough to be carried by an ion channel.

Toxicological information on environmental effects of new chemicals is required according to legislation. e.g. EEC directive. Currently, this directive prescribes experimental toxicity data and so the application of QSARs for new chemicals is limited to a certain extent.[20] This is expected to be linked with the difficulty in recognising to which of the many established QSARs the chemical of interest belongs.[21] Traditionally, the selection of a QSAR for use in risk assessment has been based on the implicit assumption that compounds from the same 'chemical class' should behave in a toxicologically similar manner. Although this working hypothesis seems reasonable, the delineation of 'chemical classes' is often problematic.[22] It is therefore realised that QSARs are not a panacea, but nevertheless, progress in the development and use of QSARs in predictive ecotoxicology has been rapid in recent years and it seems certain that their use will continue to increase in the future. This is explained in terms of the importance of QSARs as 'guidance information' in the systematic testing of new substances. For example, they have become very valuable in the early stages of screening processes for new surfactants in the detergent industry. They are often used to decide whether experimental evaluation of a potential pollutant is desirable as well as assisting in the improvement of our understanding of which structural characteristics actually give rise to the specific modes of toxic action. Furthermore, QSARs can help with the prediction of toxicity for 'not yet synthesised' compounds that can in turn stimulate the production of new chemicals with potentially lower environmental hazards.

While the usefulness of QSARs in environmental risk assessment is appreciated, it is acknowledged that calculated toxicities can never ultimately replace experimental data and are not considered reliable enough to be used in hazard assessments. Therefore, it is essential that calculated toxicities are always indicated as 'calculated' since as Rekker quoted, "there is a serious danger that, if such calculated data are introduced to toxicity registers, after a few through quotations they will lose the qualification 'estimate' that they once had and will be regarded as real observations or at least good equivalents".[20]

4.2.6 Future Work

It has been shown that the polar narcosis mode of toxic action is highly likely for the sulfobetaines due to such close agreement between our QSAR and the standard polar narcosis equation. However, to confirm that sulfobetaines are polar narcotics, a mixture toxicity study should be carried out. This would therefore be the aim of future work in this area. For a single mixture toxicity experiment, a reference compound (polar narcotic. e.g. LAS or phenolic compound) and a sulfobetaine is mixed in proportion to their individual EC_{50}s. Theoretically, if

they act by the same mechanism, 50 % immobility, when each component is at 100 % of its EC_{50}, is expected.[19]

4.2.7 References

1. Donkin, P., *Handbook of Ecotoxicology*, Vol. 2, edited by Peter Calow, Blackwell Scientific Publications, London (1994), Chapter 14.
2. Hawker, D. W. and Connell, D. W., *Ecotoxicol. Environ. Safety*, **11** (1986) 184.
3. Dearden, J. C. and Nicholson, R. M., In: *QSAR in Environmental Toxicology-II*, edited by K. L. E. Kaiser, Reidel, Dordrecht (1987) 83.
4. Lipnick, R. L., In: *Practical Applications of Quantitative-Structure Activity Relationships (QSAR) in Environmental Chemistry and Toxicology*, edited by W. Karcher and J. Devillers, Kluwer, Dordrecht (1990) 281.
5. Hansch, C., Kim, D., Leo, A. J., Norvellino, E., Silipo, C. and Vittoria, A., *CRC Crit. Rev. Toxicol.*, **19** (1989) 185.
6. Blum, D. J. W. and Speece, R. E., *Ecotoxicol. Environ. Safety*, **22** (1991) 198.
7. Warne, M. StJ., Connell, D. W., Hawker, D. W. and Schüürmann, G., *Ecotoxicol. Environ. Safety*, **18** (1989) 121.
8. Nendza, M. and Seydel, J. K., *Chemosphere*, **17** (1988) 1585.
9. Donkin, P. and Widdows, J., *Rev. Aquat. Sci.*, **2** (1990) 375.
10. Hermens, J. L. M., In: *Practical Applications of Quantitative-Structure Activity Relationships (QSAR) in Environmental Chemistry and Toxicology*, edited by W. Karcher and J. Devillers, Kluwer, Dordrecht (1990) 263.
11. Babich, H. and Borenfrreund, E., *Toxicol. in Vitro*, **1** (1987) 3.
12. Schultz, T. W., In: *Aquatic Toxicology and Hazard Assessment*, 12th Volume, edited by U. M. Cowgill and Williams, L. R., American Society for Testing and Materials, Philadelphia, P. A. (1989) 104.
13. Roberts, D. W. and Marshall, S. J., *SAR and QSAR in Environmental Research*, **4** (1995) 167.
14. Saarikoski, J. and Viluksela, M., *Ecotoxol. Environ. Safety*, **6** (1982) 501.
15. Moulton, M. P., *Regression Analysis in QSAR: Potential Pitfalls in Modelling Polar Narcotic Phenols*, T. Schultz (edt.)., Proceedings of 3rd International Workshop of QSAR in Environmental Toxicology.
16. Hall, L. and Kier, L. B., *Bull. Environ. Contam. Toxicol.*, **32** (1984) 354.
17. Pratic, M. and Sabljic, A., *Aquat. Toxicol.*, **14** (1989) 47.
18. Singh, W. P, Lin, G. H. and Bockris. J. O'M, unpublished results.
19. Roberts, D. W. and Costello, J. F., QSAR and Combinat. Sci., **22** (2003) 220.
20. Hermens, J., *Quantitative Structure-Activity Relationships in Ecotoxicology*, Short

Report, March (1984).

21. Turner, L., Choplin, F., Dugard, P., Hermens, J., Jaeckh, R., Marsmann, M. and Roberts, D., *Toxicol. in Vitro*, **1** (1987) 143.

22. Bradbury, S. P., *Toxicology Letters*, **79** (1995) 229.

Chapter 5

Derivation of Key Fragment Values and Interaction Factors for Use in Log P Calculations of Sulfobetaines

5.1 Introduction

The Leo and Hansch method for calculating log P is well-established and very reliable for simple, neutral organic compounds. However, log P calculations for ionic compounds are not so well-established. The method of Leo and Hansch (L and H) does not contain many ionic fragments and the interactions between ionic fragments is not accounted for. Furthermore, there are currently no rules for log P calculations of zwitterionic sulfobetaines and for this reason there was no option but to determine log P experimentally (Chapter 3) for use in QSARs (Chapter 4, Part II). However, a further aim of this work was to use our experimentally acquired log P data to help derive key fragment values (*f*) and interaction factors (*F*) for use in log P calculations of sulfobetaines. It was expected that a proximity effect would exist between the two polar groups, N^+ and SO_3^-, and only experimental results would indicate the existence and the extent of such an effect.

5.2 Log P Calculations for Quaternary Ammonium Compounds *via* the L and H Method

Rekker[1] assisted in the derivation of fragment values and interaction factors for log P calculations of quaternary ammonium compounds, *via* the L and H method. He provided values for the quaternary ammonium fragments given by Figure 5.1, but he recognised that each of the four fragments would have a different value depending on the anion it is associated with e.g. $f_{N+I^-} = -3.05$; $f_{N+Br^-} = -5.02$. As a result, we concluded that these values were inapplicable as they stand for the quaternary ammonium fragment of the alkylammonium sulfobetaines.

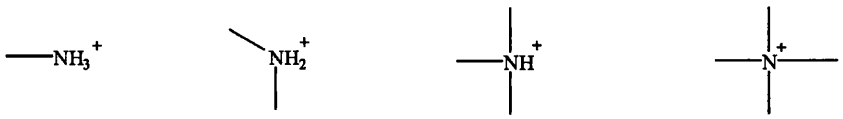


Figure 5.1 Fragments for quaternary ammonium compounds for which L and H fragment values have been derived.

For cationic compounds, of the quaternary ammonium type shown in Figure 5.2, the hydrophobicity of a carbon atom is reduced when it is directly attached or indeed in close proximity to N^+. Rekker described how the bond factor that is used to account for the distribution of charge along the carbon chains attached to N^+, can be separated into an electronic as well as a geometric component. When four chains radiate from a central N^+ the geometric portion of the bond factor is further enhanced to –0.27. The electronic portion of the bond factor

is added to the geometric component and is assumed to double in value for each unit of distance as we approach the central N^+. When plotted against the bond number counted outward from the central N^+ the total bond factors (geometric + electronic) are seen to drop sharply and level off at one of 3 levels characteristic of quaternary, tertiary and primary/secondary ammonium salts respectively. Table 5.1 shows how the bond factors F_{bx}^{+1} to F_{bx} were derived.

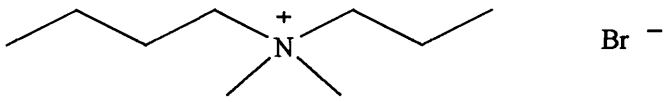


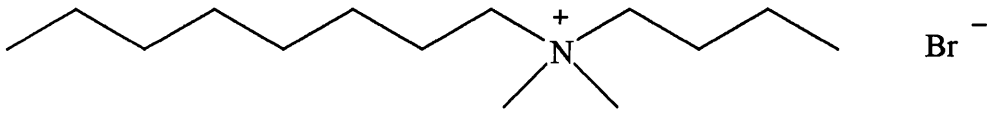
Figure 5.2 Example of an alkyldimethylammonium bromide

Table 5.1 Derivation of bond factors for quaternary ammonium compounds depending on position of C atom from N^+.

Position from N^+	Notation for bond factor	Geometric + Electronic Component	Total (bond factor)
≥ 6th bond	F_{bx}	-0.27 + 0.00	-0.27
5th bond	F_{bx}^{+5}	-0.27 + -(0.04)	-0.31 (0.30[a])
4th bond	F_{bx}^{+4}	-0.27 + -2(0.04)	-0.35
3rd bond	F_{bx}^{+3}	-0.27 + -4(0.04)	-0.45
2nd bond	F_{bx}^{+2}	-0.27 + -8(0.04)	-0.59 (0.60[a])
1st bond	F_{bx}^{+1}	-0.27 + -16(0.04)	-0.90

[a] values highlighted were rounded for convenience.

An example of a log P calculation for such a cationic compound, using the method of Rekker, is given by Figure 5.3.



$$\log P = 4f_{CH3} + 10f_{CH2} + f_{N+Br-} + 4F_{bx}^{+1} + 2F_{bx}^{+2} + 2F_{bx}^{+3} + 2F_{bx}^{+4} + F_{bx}^{+5} + 3F_{bx}$$
$$= 4(0.89) + 10(0.66) - 3.05 - 4(0.90) - 2(0.60) - 2(0.45) - 2(0.35) - 0.30 - 3(0.27)$$
$$= \mathbf{-0.40}$$

Figure 5.3 Calculation of log P for N-octyl-N, N-dimethyl-N-butylammonium bromide using the method of Rekker

Roberts[2] deduced that the inductive effect of N^+ makes the nearby carbon atoms more polar, or less hydrophobic, since a hydration sphere is created and the hydrocarbon units inside this sphere lose their hydrophobicity. This concept is illustrated by Figure 5.4. He also accepts that this effect applies to anionic fragments, such as SO_3^-, but the ability to form the anionic hydration sphere is already incorporated into the fragment value for SO_3^-. Recently, a study by Roberts led to the conclusion that the influence of N^+ does not extend such a long way down the hydrocarbon chain as the L and H method assumes. This conclusion was derived from perusal of a QSAR for acute aquatic toxicity of cationic compounds, originally developed by Singh, Lin and Bockris[3]. The log P based QSAR developed by Singh *et al.* was based on a series of quaternary ammonium compounds and used log P values deduced from calculations using the L and H approach previously described. Their QSAR (log $[1/EC_{50}]$ = 1.08 log P + 4.19; r^2 = 0.967) suggests that cationic compounds are more toxic than even the standard polar narcosis equation predicts (cf. log $[1/EC_{50}]$ = 0.63 log P + 2.52).

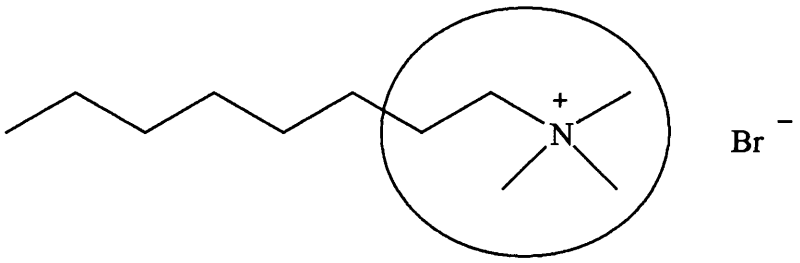


Figure 5.4 Formation of the hydration sphere that renders C atoms within it less hydrophobic.

Roberts suggested that the cationics were not as toxic as Singh *et al.* had predicted. He proposed that the log P calculation was likely to be misleading and that the cationics were actually polar narcotics, by analogy to anionics. Roberts investigated this theory by comparing log CMC values of ethoxyalcohol sulfates (anionics) with those of 'corresponding' quaternary ammonium compounds (cationics). By 'corresponding' he means that the values of both R and n are kept the same in the general formula given by Figure 5.5 for the comparison of CMC values. The CMC values were taken from a compilation produced by Barry and Wilson.

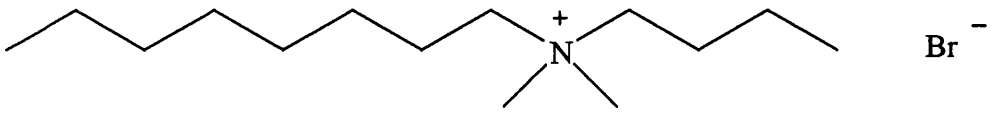
$$RO(CH_2CH_2)_n X \qquad \text{Where } X = \overset{+}{N}Me_3\overset{-}{Br} \text{ or } OSO_3^-$$

Figure 5.5 General formula used for comparison of CMCs for anionics and cationics with same value of R and n.

He concluded that 'corresponding' cationics and anionics have very similar CMC values. In the light of this result, Roberts decided to use the same approach to work out log P values. From further analysis of CMC data compiled by Rosen for cationics, he deduced that only the 2nd or 3rd CH_2 groups from the central N^+ should be rendered more hydrophilic.

Roberts then developed a log CMC based QSPR for cationics, treating Me_3N^+ and $C_5H_5N^+$ as SO_3^- for the basis of his calculations. He felt this was acceptable since he had previously observed very similar CMCs for 'corresponding' cationics and anionics. Roberts calculated log CMC for cationics using the assumption that only the first two methylene units are positioned inside the cationic hydration sphere and hence rendered more hydrophilic. Equation 1.1 describes how log CMC is calculated.

He found that a plot of log $CMC_{(obs)}$ versus log $CMC_{(calcd)}$ gave a slope very close to unity with a small intercept (log $CMC_{(obs)}$ = 0.98 log $CMC_{(calcd)}$ – 0.25; r^2 = 0.978) which indicated that the QSPR established for anionics (Equation 1.1) is capable of predicting CMC for cationics. In the light of this result, Roberts recalculated log P values for the quaternary ammonium compounds used by Singh *et al.*, using the special F_b values for the 1st and 2nd bonds from N^+ (i.e. F_{bx}^{+1} and F_{bx}^{+2}) and the normal bond factor, F_b (-0.12), for those further away. An example of his calculation method is given in Figure 5.6.



$$\log P = 4f_{CH3} + 10f_{CH2} + f_{N+Br-} + 4F_{bx}^{+1} + 2F_{bx}^{+2} + 8F_b$$
$$= 4(0.89) + 10(0.66) - 3.05 - 4(0.90) - 2(0.60) - 8(0.12)$$
$$= 1.35$$

Figure 5.6 Calculation of log P for N-octyl-N, N-dimethyl-N-butylammonium bromide using the amendments of Roberts

Roberts amended the QSAR of Singh *et al.* to include the recalculated log P values. The new aquatic toxicity QSAR for the quaternary ammonium compounds had the equation, log $(1/EC_{50})$ = 0.72 log P + 3.19, which is much closer to that of the standard polar narcosis equation. The final question arising from this work is, 'Are Cationic compounds actually polar narcotics?' To answer this question Roberts and Costello performed a short series of mixture toxicity tests, the results of which suggest that cationics are indeed polar narcotics as suspected.

However, the investigators do recognise that further tests may be required to render this result conclusive, but the result suggests that Roberts' amendments to the current log P calculation method of L and H are sound.

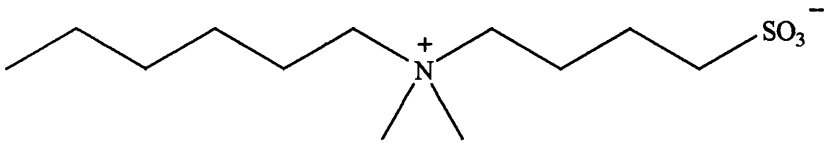
The L and H rules derived by either Rekker and/or Roberts may be applied in part to the calculation of log P for sulfobetaines. For instance, the applied bond factor (depending on the position of the C atom relative to N^+) is considered directly applicable to the main alkyl or aryl chain of the sulfobetaines. Rules for cationics can therefore be used to derive a combined fragment value and interaction factor for $N^+(CH_2)_xSO_3^-$ for use in future log P calculations *via* the L and H method.

5.3 Results and Discussion

Log P for the sum of the fragment values and interaction factors associated with $N^+(CH_2)_xSO_3^-$ (x=2,3,4) was derived from partial calculations using the L and H approach and consideration of experimental log P values determined by the stir-flask experiment for the 22 sulfobetaines. Two variations of the L and H method, one incorporating the rules of Rekker and the other incorporating the rules of Roberts were implemented.

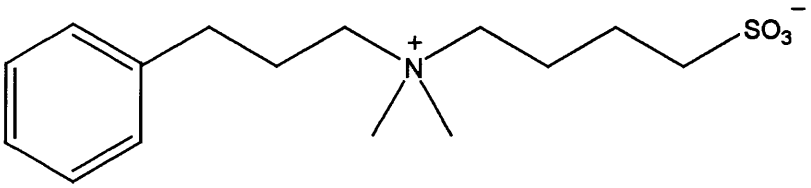
Log P predictions for sulfobetaines by KowWin (given in Section 3.5.5) were also amended. The amendments were simply based on the difference between the log P values determined experimentally *via* the stir-flask experiment and the log P values calculated using the KowWin program.

The amendments to the L and H log P calculations of Roberts (for cationics) have been used to calculate individual values of $\sum f + \Sigma F$ for $N^+(CH_2)_4SO_3^-$ ($\mathbf{X_4}$), $\sum f + \Sigma F$ for $N^+(CH_2)_3SO_3^-$ ($\mathbf{X_3}$) and $\sum f + \Sigma F$ for $N^+(CH_2)_2SO_3^-$ ($\mathbf{X_2}$). The results that were derived for each sulfobetaine using experimental log P values and the rules of Roberts are shown in Tables 5.2 to 5.4. Furthermore, Figures 5.7 to 5.9 give examples to show how X_4 for an alkyl, phenylalkyl and *para*-substituted phenylpropyl butanesulfobetaine is derived. The rules of Rekker were also used to derive values for $\mathbf{X_4}$, $\mathbf{X_3}$ and $\mathbf{X_2}$. Tables 5.5 to 5.7 show individual values of $\mathbf{X_4}$, $\mathbf{X_3}$ and $\mathbf{X_2}$ that were derived for each sulfobetaine using experimental log P values and the rules of Rekker. Figures 5.10 to 5.12 give examples to show how $\mathbf{X_4}$ for an alkyl, phenylalkyl and *para*-substituted phenylpropyl butanesulfobetaine is derived.



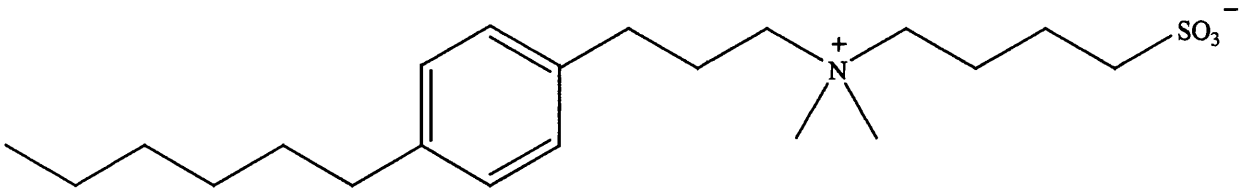
$$\log P_{expt.} = (3 f_{CH3} + 5 f_{CH2} + 3 F_{bx}^{+1} + F_{bx}^{+2} + 4 F_b) + X_4$$
$$-1.08 = ((3 \times 0.89) + (5 \times 0.66) - (3 \times 0.90) - 0.60 - (4 \times 0.12)) + X_4$$
$$X_4 = -2.07$$

Figure 5.7 An example calculation to show how X_4 is derived for an alkyl butanesulfobetaine using the rules of Roberts



$$\log P_{expt.} = (2 f_{CH3} + 3 f_{CH2} + f_{C6H5} + 3 F_{bx}^{+1} + F_{bx}^{+2} + F_b) + X_4$$
$$-1.70 = ((2 \times 0.89) + (3 \times 0.66) + 1.90 - (3 \times 0.90) - 0.60 - 0.12) + X_4$$
$$X_4 = -4.52$$

Figure 5.8 An example calculation to show how X_4 is derived for a phenyalkyl butanesulfobetaine using the rules of Roberts



$$\log P_{expt.} = (3 f_{CH3} + 8 f_{CH2} + f_{C6H4} + 3 F_{bx}^{+1} + F_{bx}^{+2} + F_b + (n-1)F_b) + X_4$$
$$1.64 = ((3 \times 0.89) + (8 \times 0.66) + 1.67 - (3 \times 0.90) - 0.60 - 0.12 - (5 \times 0.12)) + X_4$$
$$X_4 = -3.98$$

Figure 5.9 An example calculation to show how X_4 is derived for a para-substituted phenylpropyl butanesulfobetaine using the rules of Roberts

Table 5.2 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_4SO_3^-$ using the rules of Roberts

Butane sulfobetaine	$\Sigma f + \Sigma F$ for side chain and 2 methyl groups	Experimental log P	X_4
C6C4 (**14**)	2.19	-1.08	-3.27 [a]
C7C4 (**15**)	2.73	-1.23	-3.96
C8C4 (**16**)	3.27	-0.36	-3.63 [a]
PhC1C4 (**17**)	2.54	-2.32	-4.86 [b]
PhC2C4 (**18**)	2.60	-2.06	-4.66 [b]
PhC3C4 (**19**)	3.14	-1.70	-4.84 [b]
PhC4C4 (**20**)	3.68	-1.41	-5.09 [b]
C6PhC3C4 (**21**)	5.87	1.64	-4.23 [c]
C4PhC3C4 (**22**)	4.79	0.30	-4.49 [c]

Av. [a] = -3.45; Av. [b] = -4.86; Av. [c] = -4.36

Table 5.3 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_3SO_3^-$ using the rules of Roberts

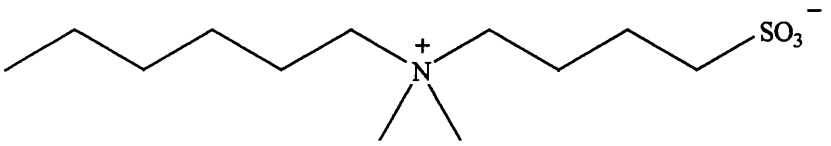
Propane sulfobetaine	$\Sigma f + \Sigma F$ for side chain and 2 methyl groups	Experimental log P	X_3
C6C3 (**3**)	2.19	-1.22	-3.41 [a]
C7C3 (**4**)	2.73	-1.97	-4.70
C8C3 (**5**)	3.27	-0.47	-3.74 [a]
C10C3 (**6**)	4.35	0.57	-3.78 [a]
C12C3 (**7**)	5.43	1.65	-3.78 [a]
PhC1C3 (**8**)	2.54	-2.27	-4.81 [b]
PhC2C3 (**9**)	2.60	-1.86	-4.46 [b]
PhC3C3 (**10**)	3.14	-1.58	-4.72 [b]
PhC4C3 (**11**)	3.68	-1.17	-4.85 [b]
C6PhC3C3 (**12**)	5.87	1.84	-4.03 [c]
C4PhC3C3 (**13**)	4.79	0.55	-4.24 [c]

Av. [a] = -3.68; Av. [b] = -4.71; Av. [c] = -4.14

Table 5.4 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_2SO_3^-$ using the rules of Roberts

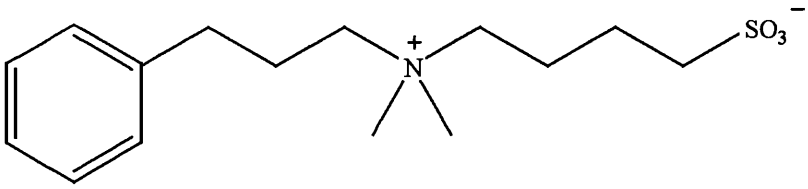
Ethane sulfobetaine	$\Sigma f + \Sigma F$ for side chain and 2 methyl groups	Experimental log P	X_2
C8C2 (**1**)	3.27	-0.06	-3.33
C12C2 (**2**)	5.43	1.78	-3.65

Av.[a] = -3.49



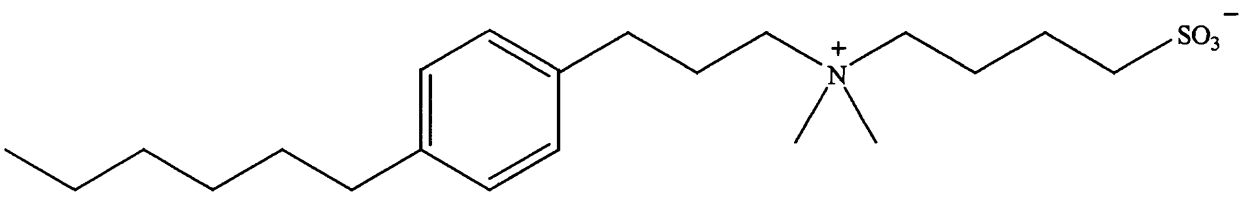
$$\log P_{expt.} = (3f_{CH3} + 5f_{CH2} + 3F_{bx}^{+1} + F_{bx}^{+2} + F_{bx}^{+3} + F_{bx}^{+4} + F_{bx}^{+5} + F_{bx}) + X_4$$
$$-1.08 = ((3 \times 0.89) + (5 \times 0.66) - (3 \times 0.90) - 0.60 - 0.45 - 0.35 - 0.30 - 0.27) + X_4$$
$$X_4 = -2.38$$

Figure 5.10 An example calculation to show how X_4 is derived for an alkyl butanesulfobetaine using the rules of Rekker



$$\log P_{expt.} = (2f_{CH3} + 3f_{CH2} + f_{C6H5} + 3F_{bx}^{+1} + F_{bx}^{+2} + F_{bx}^{+3}) + X_4$$
$$-1.70 = ((2 \times 0.89) + (3 \times 0.66) + 1.90 - (3 \times 0.90) - 0.60 - 0.45) + X_4$$
$$X_4 = -3.61$$

Figure 5.11 An example calculation to show how X_4 is derived for a phenylalkyl butanesulfobetaine using the rules of Rekker



$$\log P_{expt.} = (3f_{CH3} + 8f_{CH2} + f_{C6H4} + 3F_{bx}^{+1} + F_{bx}^{+2} + F_{bx}^{+3} + (n\text{-}1)F_b) + X_4$$

$$1.64 = ((3 \times 0.89) + (8 \times 0.66) + 1.67 - (3 \times 0.90) - 0.60 - 0.45 - (5 \times 0.12)) + X_4$$

$$X_4 = -3.63$$

Figure 5.12 An example calculation to show how X_4 is derived for a para-substituted pheylpropyl butanesulfobetaine using the rules of Rekker

Table 5.5 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_4SO_3^-$ using the rules of Rekker

Butane sulfobetaine	$\Sigma f + \Sigma F$ for side chain and 2 methyl groups	Experimental log P	X_4
C6C4 (**14**)	1.30	-1.08	-2.38 [a]
C7C4 (**15**)	1.69	-1.23	-2.92
C8C4 (**16**)	2.08	-0.36	-2.44[a]
PhC1C4 (**17**)	1.64	-2.32	-3.96[b]
PhC2C4 (**18**)	1.70	-2.06	-3.76[b]
PhC3C4 (**19**)	1.91	-1.70	3.61[b]
PhC4C4 (**20**)	2.21	-1.41	-3.62 [b]
C6PhC3C4 (**21**)	5.27	1.64	-3.63 [b]
C4PhC3C4 (**22**)	4.18	0.30	-3.88 [b]

Av.[a] = -2.41; Av.[b]= -3.75

Table 5.6 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_3SO_3^-$ using the rules of Rekker

Propane sulfobetaine	$\Sigma f + \Sigma F$ for main side chain and 2 methyl groups	Experimental log P	X_3
C6C3 (**3**)	1.30	-1.22	-2.52 [a]
C7C3 (**4**)	1.69	-1.97	-3.66
C8C3 (**5**)	2.08	-0.47	-2.55 [a]

C10C3 (**6**)	2.86	0.57	-2.29[a]
C12C3 (**7**)	3.64	1.65	-1.99[a]
PhC1C3 (**8**)	1.64	-2.27	-3.91[b]
PhC2C3 (**9**)	1.70	-1.86	3.56[b]
PhC3C3 (**10**)	1.91	-1.58	-3.49[b]
PhC4C3 (**11**)	2.21	-1.17	-3.38[b]
C6PhC3C3 (**12**)	5.27	1.84	-3.43
C4PhC3C3 (**13**)	4.18	0.55	-3.63[b]

Av.[a] = -2.34; Av.[b] = -3.57

Table 5.7 Derivation of the sum of the fragment values and interaction factors associated with $N^+(CH_2)_2SO_3^-$ using the rules of Rekker

Ethane sulfobetaine	Σf + F for side chain and 2 methyl groups	Experimental log P	X_2
C8C2 (**1**)	2.08	-0.06	-2.14
C12C2 (**2**)	3.64	1.78	-1.86

Av.[a] = -2.00

Table 5.8 Amendment to the log P calculation by KowWin for butanesulfobetaines

Butane sulfobetaine	Log P (KowWin)	Log P (expt.)	Log P (expt.) - Log P (KowWin)
C6C4 (**14**)	-1.03	-1.08	-0.05[a]
C7C4 (**15**)	-0.54	-1.23	-0.69
C8C4 (**16**)	-0.05	-0.36	-0.31[a]
PhC1C4 (**17**)	-1.78	-2.32	-0.54[b]
PhC2C4 (**18**)	-1.29	-2.06	-0.77[b]
PhC3C4 (**19**)	-0.80	-1.70	-0.90[b]
PhC4C4 (**20**)	-0.31	-1.41	-1.10[b]
C4PhC3C4 (**21**)	1.22	0.30	-0.92[c]
C6PhC3C4 (**22**)	2.20	1.64	-0.56[c]

Av.[a] = -0.18; Av.[b] = -0.82; Av.[c] = -0.72

Table 5.9 Amendment to the log P calculation by KowWin for propanesulfobetaines

Propane sulfobetaine	Log P (KowWin)	Log P (expt.)	Log P (expt.) - Log P (KowWin)
C6C3 (**3**)	-1.53	-1.22	0.31[a]
C7C3 (**4**)	-1.03	-1.97	-0.94
C8C3 (**5**)	-0.54	-0.47	0.07[a]
C10C3 (**6**)	0.44	0.57	0.13[a]
C12C3 (**7**)	1.42	1.65	0.23[a]
PhC1C3 (**8**)	-2.27	-2.27	0.00[b]
PhC2C3 (**9**)	-1.78	-1.86	-0.08[b]
PhC3C3 (**10**)	-1.29	-1.58	-0.29[b]
PhC4C3 (**11**)	-0.80	-1.17	-0.37[b]
C4PhC3C3 (**12**)	0.73	0.55	-0.18[c]
C6PhC3C3 (**13**)	1.73	1.84	0.11[c]

Av.[a] = -0.18; Av.[b] =-0.19; Av.[c] = -0.04

Table 5.10 Amendment to the log P calculation by KowWin for ethanesulfobetaines

Propane sulfobetaine	Log P (KowWin)	Log P (expt.)	Log P (expt.) - Log P (KowWin)
C8C2	-1.03	-0.06	0.97[a]
C12C2	0.93	1.78	0.85[a]

Av.[a] = 0.91

Table 5.11 Summary of $\Sigma f + \Sigma F$ for $N^+(CH_2)_xSO_3^-$ (x= 2,3,4) derived from the rules of Roberts and Rekker

$\Sigma f + \Sigma F$	Roberts' Method	Rekker's Method
$N^+(CH_2)_4SO_3^-$		
i) alkyl	-3.45	-2.41
ii) phenylalkyl	-4.86	-3.75
iii) *para*-substituted phenylpropyl	-4.36	-3.75
$N^+(CH_2)_3SO_3^-$		
i) alkyl	-3.68	-2.34
ii) phenylalkyl	-4.71	-3.57
iii) *para*-substituted phenylpropyl	-4.14	-3.57
$N^+(CH_2)_2SO_3^-$		
i) alkyl	-3.49	-2.00

Table 5.11 Summary of amendments to current log P calculations of ethane-, propane- and butanesulfobetaines using KowWin

Sulfobetaines	Amendment to log P value generated by KowWin
Butanesulfobetaines	
i) alkyl	-0.18
ii) phenylalkyl	-0.82
iii) *para*-substituted phenylpropyl	-0.72
Propanesulfobetaines	
i) alkyl	-0.18
ii) phenylalkyl	-0.19
iii) *para*-substituted phenylpropyl	-0.04
Ethanesulfobetaines	
i) alkyl	0.91

5.4 Conclusion

It has been shown that Roberts' amendments to the log P calculations for cationic compounds of the quaternary ammonium type are sound. In the light of this result, it may be concluded that calculation of log P using the method of Rekker for cationics underestimates the true log P

value. This is because the carbon atoms beyond position two from N^+ are rendered more hydrophilic than they are in reality. This is especially problematic for those who have used Rekker's method to calculate log P for use as a physicochemical descriptor for the prediction of the environmental fate of cationic quaternary ammonium compounds. Predictions of toxicity from such a log P-based QSAR would be artificially low.

However, since the method of Rekker has been established for many years it was felt necessary to also derive $\sum f + \sum F$ for $N^+(CH_2)_xSO_3^-$ using the method of Rekker, so that a comparison of the two methods can be made. It became apparent that the value of $\sum f + \sum F$ for $N^+(CH_2)_xSO_3^-$ derived using Roberts' method was less (more negative and hence more hydrophilic) than the corresponding value calculated using the Rekker method as illustrated by Table 5.11.

Originally, it was expected that a universal value of $\sum f + \sum F$ ($N^+(CH_2)_4SO_3^-$) should be derived for the alkyl, phenylalkyl and para-substituted phenylpropyl sulfobetaines, and likewise for $\sum f + \sum F$ ($N^+(CH_2)_3SO_3^-$) and $\sum f + \sum F$ ($N^+(CH_2)_4SO_3^-$). However, using the method of Roberts it became obvious that $\sum f + \sum F$ ($N^+(CH_2)_xSO_3^-$) was dependent on the nature of the hydrophobic side chain (i.e. alkyl, phenylalkyl or *para*-substituted phenylpropyl) as shown by Tables 5.2 to 5.4.

In the literature, it has been shown that the flexibility of the side chain can influence log P. Self-coiling is expected to partly shield methylene groups from solvent which prevents them from expressing their hydrophobic increment[4]. Additionally, it has been shown that the delocalisation of π-electrons in an aromatic ring system increases the polarisability of the molecule, hence reduces log P. As a result, log P of the fragment, $N^+(CH_2)_xSO_3^-$, is seen to decrease in the order obeyed below, which indeed does tie in with the potential for polarisation.

alkyl > *para*-substituted phenylpropyl > phenylalkyl

A similar trend was observed for values of $\sum f + \sum F$ for $N^+(CH_2)_xSO_3^-$ derived using the method of Rekker (Tables 5.5 to 5.7). However, log P of the fragment, $N^+(CH_2)_xSO_3^-$, could not be distinguished for the phenylalkyl and *para*-substituted phenylpropyl side chain.

alkyl > *para*-substituted phenylpropyl = phenylalkyl

On studying $\Sigma f + \Sigma F$ for $N^+(CH_2)_xSO_3^-$ (x=2,3,4) for an alkyl sulfobetaine, we see that the hydrophobicity of the fragment, $N^+(CH_2)_xSO_3^-$, decreases in the order: -

$$N^+(CH_2)_2SO_3^- > N^+(CH_2)_3SO_3^- > N^+(CH_2)_4SO_3^-$$

If there is no proximity effect between the oppositely charged moieties, N^+ and SO_3^-, the reverse order would be expected, since in this scenario the consecutive addition of a methylene unit is responsible for an incremental increase in log P. However, the order observed does in fact confirm the existence of a proximity effect between the polar groups. The fact that the hydrophobe containing the fewest number of methylene groups in the spacer unit possesses the greatest log P value suggests such a through-bond proximity effect. Additionally, the hydrophobicity of one or two extra methylene units (going from x=2 or 3, to x=4) appeared to be overwhelmed by the weakening of the proximity effect between the polar groups.

Intercharge distance–dependent hydrophobicity has been previously acknowledged in the literature. For example, Tsai *et al.*[4] studied a series of amino acids with distal NH_3^+ and CO_2^- moieties and found that hydrophobicity decreases with increasing distance between the opposite charges. They examined systematically how the distance between the charged moieties affected the hydrophobicity of homologous amino acids and found that log D values differ only minimally when the number of groups in the spacer unit varies from 1 to 6. Only the 7th methylene group shows a relatively normal increment of 0.50. It may therefore be concluded that the two opposite influences on hydrophobicity almost compensate each other at a distance of 1 to 6 carbon atoms, but when the 7th carbon atom is added, the effect of adding an extra methylene group is observed and a proximity effect ceases to exist.

It should be noted that the log P values of $N^+(CH_2)_4SO_3^-$ and $N^+(CH_2)_3SO_3^-$ for the alkyl chain, C_7H_{15}, were not included in the averages, since they appeared to artificially lower the average in each case.

Amendments to the KowWin predictions were also found to depend on the nature of the main hydrophobic group as well as on the length of the methylene spacer unit. The program totally neglects the existence of a proximity effect between N^+ and SO_3^-, hence 0.91 log P units must be added to the KowWin prediction for an alkyl ethanesulfobetaine. Other notable amendments include the subtraction of 0.82 log P units from the KowWin prediction for a phenylalkyl butanesulfobetaine and subtraction of 0.72 log P units from the KowWin prediction for a *para*-substituted phenylpropyl butanesulfobetaine, possibly arising from the fact that the

program does not consider that weakening of the proximity effect can actually provide an overall decrease in hydrophobicity when additional methylene units are introduced.

5.5 References

1. Hansch, C. and Leo, A., Substituent Constants for Correlation Analysis in Chemistry and Biology, Wiley and Sons, NY (1979), Appendix of Chapter 4.
2. Roberts, D. W. and Costello, J. F., QSAR and Combinat. Sci., **22** (2003) 220.
3. Singh, W. P., Lin, G. and Bockris, J. O'M, unpublished results.
4. Tsai, R-S., Testa, B., El Tayar, N., Carrupt, P-A., *J. Chem. Soc. Perkin Trans. 2* (1991) 1797.

Chapter 6

IAM Chromatography and Comparison of Chromatographic Methods of Estimating Log P

6.1 Introduction

Log P based QSARs are well established and are accepted as reliable mathematical models for predicting aquatic toxicity. This indicates that, although other physicochemical properties may be involved, hydrophobicity is a key factor.[1] QSARs for predicting aquatic toxicity of nonionic and anionic surfactants have been developed and, more recently, a log P based QSAR for cationics has been reported.[2] We have ourselves developed a log P based QSAR for predicting the aquatic toxicity of zwitterionic sulfobetaines which shows good correlation between log P and log ($1/EC_{50}$) (Chapter 4). However, it is recognised that octanol and water may not be the most suitable model for assessing transport phenomena across the lipid membrane of a living system[3,4]. Other models involving the use of an immobilised artificial membrane (IAM) containing phosphatidylcholine have been suggested.[5] It was therefore decided to compare the efficacy of using log k'_{IAM} and log P as parameters for predicting aquatic toxicity.

In addition other chromatographic methods for estimating log P were investigated. The conventional stir-flask method for determining log P is very reliable and the results obtained for sulfobetaines show extremely good reproducibility.[3] However conducting stir-flask experiments is very time consuming. It involves achieving mutual saturation of the octanol/water layers, partitioning of the substrate between the two layers, and quantitative analysis of aliquots from each layer. Measuring retention indices by chromatography is much more appealing in terms of efficiency. Other advantages include the fact that there is no risk of emulsions being formed, there is precision in the determination of log P values over a wide range (+4 to –4), impurities that can affect partition do not affect retention times, and only very small quantities are required compared to the relatively large amounts required for stir-flask experiments.[6]

6.2 IAM Chromatography

Since the main constituents of biomembranes are phospholipids, the use of a biologically significant phospholipid, such as phosphatidylcholine (PC), as the partitioning phase can be expected to yield *in vitro* data that better mimics the *in vivo* interactions between a solute and biomembrane. IAM chromatography phases are therefore described as solid-phase membrane mimetics that are prepared by covalently immobilising monolayers of PC to silica particles at high molecular surface densities. IAM chromatography phases prepared from PC analogues are

commercially available from Regis Technologies, Inc.[7] The IAM.PC, IAM.PC.MG, IAM.PC.DD and IAM.PC.DD2 phases are now in common use.

IAM.PC HPLC phase consists of monolayers of amphiphilic phospholipids covalently immobilised on aminopropyl silica particles through an amide linkage. The resulting IAM surface is found to be a chemically stable chromatographic material that emulates the exterior of a biological cell membrane.[8] The IAM.PC phase was first developed in 1995 by Professor Charles Pidgeon at Purdue University. It has since become subject to modification to further improve chromatography. The first modification was in the form of methylglycolate end-capping, converting residual amines to neutral amides and introducing a hydroxyl group (IAM.PC.MG). Secondly, the IAM.PC.DD material was developed by end-capping with C10 and C3 acyl chains. The IAM.PC.DD column, however, has recently been replaced by the IAM.PC.DD2 column (Figure 6.1). Generally excellent correlation is reported between the DD and DD2 columns, but the extra hydrophobicity offered by the ester bonding of the DD2 column provides longer retention times for compounds not well retained on the DD packing. Retention times are typically double for analytes run on the DD2 column compared with the DD column. Furthermore the DD2 phase has the added advantage that it is more stable at higher pH. Generally, the DD and DD2 columns are most commonly used in drug membrane permeability studies while the earlier IAM phases (PC and PC.MG) are now often used for protein purification.[7]

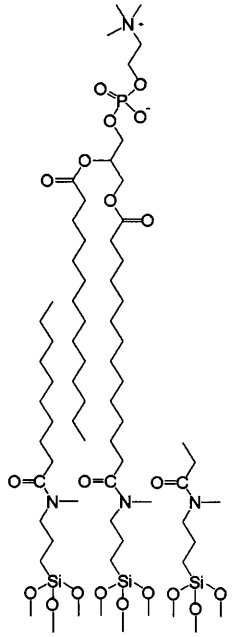


Figure 6.1 Structure of the IAM.PC.DD2 bonded phase

In IAM chromatography, a combination of hydrophobic, ion pairing and hydrogen bonding interactions may exist between the analyte and the IAM, due to the presence of both a polar and non-polar region. This combination of interactions, often referred to as phospholipophilicity, are also expected to be important in membrane transport. In contrast, where alkyl bonded phases (such as C8 and C18) are used, retention of the analyte is based solely on hydrophobicity, since only a non-polar region exists. The measurement of phospholipophilicity over hydrophobicity is expected to explain why log k'_{IAM} has been shown to provide superior correlation with experimentally determined drug permeability compared to other chromatographic parameters, such as log k'_{C8} and log k'_{C18}.[7,9]

IAM chromatography has recently gained acceptance for estimating the membrane permeability of small molecule drugs, since it appears that for these compounds, the membrane partition coefficient defines the rate-limiting step for drug absorption. The excellent correlation between drug permeability predicted by caco-2 cells and log k'_{IAM} has been appreciated in the literature. Therefore, since drug permeability on caco-2 cells is known to correlate with the oral absorption of drugs in humans, IAM is also expected to model drug absorption in humans. Significant correlations have also been reported between drug permeability predicted by inverted rat intestines and log k'_{IAM} as well as drug permeability predicted by partitioning into liposomes and k'_{IAM}. IAM chromatography has since been appreciated as a rapid and economical alternative to the more expensive and labour intensive methods of estimating drug permeability such as the use of caco-2 cell line cultures, intestinal tissue or liposome assays.[7,9] IAM chromatography has proved successful in drug discovery, where it is increasingly used in the initial screening of drug permeability of drug candidates produced by combinatorial chemistry.[10,11]

IAM bonded phases have found other applications such as prediction of solute transport across the skin,[7] predicting amino acid transport across the blood-brain barrier[12] and predicting bile salt membrane interactions,[13] since log k'_{IAM} has proved to better correlate with these membrane interactions than log P. Despite previous application in QSAR correlation studies,[14] it is surprising that the phospholipophilicity parameter, log k'_{IAM}, had not previously been used in QSARs for aquatic toxicity. Since PC is the major phospholipid found in all cell membranes, the use of log k'_{IAM} as a parameter to model the transport of zwitterionic sulfobetaines across the surface

membrane of the water flea, *Daphnia magna* was considered, and hence the development of such a QSAR for acute aquatic toxicity.

In addition to our interest in correlating log k'_{IAM} with aquatic toxicity, there was also interest in how well log k'_{IAM} correlates with log P. Amato *et al.* have reported that generally for neutral compounds, log k'_{IAM} correlates well with log P; however quite surprising chromatographic behaviour was observed for ionised compounds. They found that some hydrophilic amines could interact with phospholipids to such a surprisingly strong extent as to imply a better capability to cross biomembranes than expected on the basis of their log P values.[3]

In contradiction to the above, Escher *et al.*[5] concluded that IAM chromatography was not suitable for the prediction of membrane-water partitioning of ionic compounds, and the authors attributed this to the lack of sufficient shielding of the charges on the chromatographic support material. However, they did suggest that the technique could be useful for pre-screening of neutral species. The work of many other authors has also led to the conclusion that log k'_{IAM} does not correlate well with log P.

6.3 Indirect HPLC Methods for Estimating Log P

The most widely used *indirect* HPLC method of estimating log P involves measuring the retention time of a given analyte on a C8 or C18 RP-column and then calculating the log of the retention index (log k′), from an equation derived from a linear calibration plot of log P versus log k′ for a series of similar reference compounds for which log P is known. Log P for the reference compounds may be determined by a reliable calculation method or an experimental method such as the stir-flask experiment.[15-19] The OECD guideline recommends that at least one reference compound should have a log P value greater than that of the test substance and at least one should have log P less than that of the test substance. The OECD guideline also recommends that ionisable compounds are analysed in their non-ionised form by using a buffer at a pH below the pK_a of the free acid or above the pK_a of the free base, but it also recommends that where a log P value has potential use in environmental risk assessment the test should be performed within the pH range relevant to the natural environment (i.e. in the pH range of 5.0 and 8.5).[19] Since the sulfobetaines are in the ionised form over such a wide pH range, including the pH range of the natural environment, they were analysed in the zwitterionic form.

The transport of compounds through such columns involves partitioning between the hydrocarbon stationary phase and the mobile phase. Analytes are therefore eluted in order of hydrophobicity with the water-soluble compounds (lower log P) eluting before the lipid-soluble ones (higher log P). A form of the Collander equation states that $k'_{(C8 \text{ or } C18)}$ is proportional to the partition coefficient, $K_{(C8 \text{ or } C18)}$, and is hence directly related to P.[18]

$$\text{Log P} = a \log k'_{C8 \text{ or } C18} + b \qquad \textit{(Equation 6.1)}$$

a = gradient of graph; b = intercept of graph (therefore, a and b are constants characterising the two partitioning phases)

It is predicted by Renberg *et al.* that RP-HPLC methods are sensitive to the measurement of log P in the region of 0 to 6.[20] However, it has since been appreciated that adjustment of the mobile phase may extend this range. The indirect HPLC method has the advantage over the direct method (involving the coating of a column with octanol) that an organic modifier may be added to the eluent to shorten retention times and improve solute detectability of highly hydrophobic compounds. However, if such an organic modifier were added to the octanol-coated system the column would be immediately stripped of octanol.

It is widely accepted in the literature that an organic modifier such as methanol or acetonitrile, may be employed to shorten analysis time. This is extremely advantageous for highly hydrophobic compounds, since it is only really convenient to measure retention times of up to 2 h with accuracy. Beyond this retention time, peaks appear to be too broad, taking several minutes to fully elute, or they may be so flattened that that they cannot be observed at all at a respectable concentration. For comparison with literature values of log k', which are commonly recorded using 100 % water, $\log k'_w$, retention indices of a given analyte may be measured at different percentages of organic modifier (70 to 30 %). The linear regression equation obtained may then be used to extrapolate to conditions of 100 % water.[21-24]

It is recommended that to increase confidence in any given log P determination duplicate measurements should be made. The error in the duplicate determinations should fall within the range of ±0.1 log P units.[17] Furthermore, it has been shown that the deviation between log P determinations performed by this HPLC method and those determined by the conventional stir-flask experiment falls within ±0.5 log P units and this is perhaps the most significant disadvantage

of the technique.[19] Fujisawa and Masuhara suggest that the latter error limits may be extended if calibration plots are based only on structurally related compounds.[25]

As previously mentioned this method has a number of advantages over the conventional stir-flask experiment. The major advantages are its speed and the possibility of analysing a whole series of compounds in a single working day. In contrast, a stir-flask experiment is likely to take place over a 2-3 day period. Also a very small quantity of test substance is required for the measurement of retention data compared with that required for a stir-flask experiment.[15] Veith *et al.* acknowledged the laborious nature of the stir-flask experiment due to the possibility of emulsion formation and also recognised the need for an analytical method with the capability of *quantifying* the amount of analyte in preferably both (but at least one) of the phases. There is obviously no risk of emulsion formation when employing the HPLC method, and a detection mode is only required to monitor the elution of a given peak and so quantification is not required. This means that time is not consumed preparing standard solutions containing both analyte and internal standard to construct calibration plots. Compounds with log P of ≥4, where there is 10,000 greater affinity (or more) for octanol than water, cannot be measured with the same precision using the stir-flask experiment as chemicals that distribute themselves more evenly.[6] In environmental risk assessment, it is these hydrophobic chemicals that are of greatest concern and it is in this area that measuring retention indices on a C8 or C18 column has proved most appealing. Furthermore, partitioning between the biphasic system in a stir-flask experiment is likely to be affected by the presence of impurities, hence a reliable result would depend on using analytes of the highest purity. For example, if the hydrophobicity of a single analyte in a complex effluent was assessed by such an experiment the result would not have the desired reliability because it could not be ruled out that the bulk of the organic constituents may alter the partition behaviour. However, use of the HPLC method does not require pure analytes since impurities and degradation products do not affect retention indices. The only problem here is that the existence of such impurities can make the interpretation of the results difficult due to uncertainty in peak assignments (especially if the UV detection mode is employed since impurities may have larger absorbance that the pure test compound at the $\lambda_{detection}$).

The HPLC method is rendered unsuitable for the analysis of charged, partly charged or very polar compounds. It is also considered non-applicable to strong acids and bases, substances that are expected to react with the eluent, and surface-active agents.[19]

There have been a number of publications confirming excellent correlation between log k′ (C8 or C18) and log P for simple, neutral compounds, but correlations for ionised/ionisable compounds and more complicated molecules are less common. In an extensive review of HPLC methods for measuring log P, Braumann[26] concludes that log k'_W can successfully replace log P as a hydrophobicity parameter in most QSARs of biological interest. However, some authors disagree and their reasons are well founded. Generally, it is agreed that compared with stir-flask data, k′ is more easily obtained. However, reliable k'_w data may in some cases be just as difficult to come by, as appreciated by Garst[27] and other authors.[1]

On the whole, the greatest errors between log P determined by this method and reported log P values (determined by calculation or by conventional methods) are observed for polar compounds that dissociate in water. Here, dissociation of ionisable polar groups appears more significant than adsorption interactions. For these reasons, Veith *et al.* found that compounds such as *m*-chlorobenzoic acid, 2,4,5-trichlorophenol and diphenylamine elute more rapidly than expected from their respective log P values.[6]

A few authors have highlighted problems encountered when measuring retention indices using mobile phase conditions employing methanol as an organic modifier. Collander acknowledged that a 50 % methanol: 50 % aqueous mobile phase, and increasing proportions of methanol relative to the aqueous phase, resulted in the HPLC capacity factors becoming insensitive towards solutes with log P <1.5.[28] It has also been suggested that extrapolation of the regression line from 30 % to 0 % methanol is not actually linear.[29] The linear relationship usually observed between log k′ of analytes eluted with methanol-water and log P is also sometimes perturbed by factors such as H-bonding effects and selective solute-solvent interactions.[30] Early work by Yunger and Cramer involving the determination of octanol/water partition coefficient of 20 amino acid by the indirect HPLC method indicated anomalies due to charge and potential for H-bonding of α-amino and α-carboxylate groups.[31] Yamagami and Takao studied the relationship between log P measured by the stir-flask and the HPLC method for a series of monosubstituted pyrazines. Differing compositions of methanol/buffer (pH 9.2) were employed for the measurement of log k′.

The observed complicated behaviour of the pyrazine derivatives was ascribed to the electron-withdrawing property of the pyrazine ring, which affected the overall H-bonding effect of the molecule. This was extended to the 2-chloro-6-substituted pyrazine series and despite variation in the HPLC conditions, both series of pyrazines were found to behave similarly. Furthermore, the log k′/log P relationship was shown to become more complicated with a decrease in organic modifier concentration due to intervention of electronic interactions and the retardation effect ascribed to ester and amide substituents.[32]

Differences between HPLC determinations of hydrophobicity and stir-flask methods are often difficult to rationalise. For example, log P determinations by the stir-flask experiment for 1,3,5-trimethylbenzene (mesitylene) are lower than predicted and show a concentration dependence. This dependence indicates that mesitylene molecules may stack when completely surrounded by water. This effect is not appreciated when analysing HPLC retention times if the methanol content of the mobile phase is kept above 40 %. The lower log P values from stir-flask experiments may better model mesitylene transport in streams, for example, but the HPLC method may best model effects such as binding onto surfaces.[1]

An even greater difference between stir-flask and HPLC measurements of hydrophobicity can be seen in the study of precocene analogues studied by Camps *et al.* These precocene analogues possess ether oxygen atoms in close proximity. For most of the simple precocene analogues the HPLC method gave log P values that agreed with those derived from stir-flask experiments. However, an appreciable difference arises when oxygen atoms are placed closer to each other. For example, when two alkoxy groups are placed *ortho* to each other in the aromatic ring of precocenes, HPLC values register an approximate decrease in log P of 0.7 compared with stir-flask values.[33]

In some cases HPLC data does appear to better reflect some steric information that may model a given biological effect better than log P. For example, Wright *et al.* found that for a set of anti-inflammatory imidazolyethanols that have moieties radiating from a quaternary carbon, log P values calculated from HPLC retention times sometimes resulted in anomolous hydrophobicities for substituents. Thus, in two cases a methoxyl group was found to be much more hydrophobic than a methyl group.[34]

On analysis of hydrophobic chlorinated benzene derivatives it was found that good correlation existed between log P determined by the stir-flask method and that determined by calculation for log P values up to approximately 4, but the linear relationship begins to deteriorate for higher values of log P. The authors concluded that the steric interference of the chloro substituents could not explain this discrepancy. It seemed more likely that the deviation in log P values was attributed to the occurrence of octanol droplets in the aqueous phase that could not be removed by the usual centrifugation technique.[18] The same conclusion also emerges from data given by Tulp and Hutzinger.[35] They draw the conclusion that the HPLC method has the significant advantage of a larger application range and the capability of measuring log P ≥8. Many other authors also suggest that the HPLC method is especially useful for the determination of log P of particularly hydrophobic chemicals (log P ≥4) mainly due to the greater precision associated with such a log P value obtained by this method over the stir-flask method.

As well as attempts to measure retention indices on a C8 or C18 column the exploration of different types of HPLC bonded phases (expected to better mimic octanol) is also mentioned in the literature. Liu *et al.*[8] synthesised 12-hydroxydecyl aminopropyl silica (12-OH-silica) and evaluated the ability of this surface to predict drug partitioning into octanol/water phases and also to predict drug partitioning into fluid membranes. The 12-OH-silica is effectively an immobilised alcohol and can be considered a solid-phase model of the octanol/water partitioning system. Unlike C8- or C18-silica, 12-OH-silica contains both hydrogen donor and hydrogen acceptor capabilities at the surface. To probe the effect of H-bonding at the chromatographic interface, a surface lacking hydrogen bond donor capabilities was also prepared by immobilising 12-methoxydodecanoic acid (12-MO) on aminopropyl silica to form 12-MO-silica. Liu *et al.* chose a series of 23 drugs for analysis (7 β blockers, 6 imidazoline derivatives and 10 imidazolidine derivatives) and investigated correlations between the following parameters: -

(i) $\log k'_{C18}$ vs. solute intestinal absorption ($\log K_m$)

(ii) $\log k'_{12\text{-}OH}$ vs. $\log K_m$

(iii) $\log k'_{12\text{-}MO}$ vs. $\log K_m$

(iv) $\log k'_{IAM}$ vs. $\log K_m$

The best correlation was actually observed between $\log k'_{IAM}$ and $\log K_m$ ($r=0.864$). However, good correlations are similarly observed between $\log k'_{12\text{-}OH}$ and $\log K_m$ ($r=0.817$) and between $\log k'_{12\text{-}}$

$_{MO}$ and log K_m (r=0.812). In contrast a correlation did not exist between log k'_{C18} and log K_m (r=0.100).[8]

Additionally, a recent paper by Donovan and Pecatore[36] describes the use of a short octadecyl-polyvinyl alcohol (ODP) HPLC column to study lipophilic compounds in the log P range of 2–6 using a linear methanol-water gradient. A short column was used following Leo's suggestion that when the methanol concentration is higher than 50 %, HPLC is relatively insensitive to hydrophobicity.[37] This conclusion was supported by the work of Spencer *et al.*[30] who used 70–90 % methanol and found a poor correlation between log P and enzyme inhibition. Leo later used calculated values of log P and found a positive correlation between these values and enzyme inhibition. Lambert also recommends that the methanol content could be kept low by shortening the columns employed.[39]

The work of Donovan and Pecatore led to the conclusion that log P values could be obtained with fair accuracy and very good precision. They were not interested in the use of their acquired log P values for analytical purposes, but for developing a simple-to-use tool to better understand the biological activity and environmental fate of chemicals. They expressed confidence in the use of their method to check the validity of classically measured log P values. However, they do not suggest that their method should supersede the stir-flask experiment for registration purposes.[36]

Vallat *et al.*[40] had earlier studied the applicability of the ODP column. It was the nearly 'universal' behaviour that first drew Donovan and Pescatore's attention to these columns. Vallat *et al.* found that the column was able to predict log P with good accuracy, and this appeared true even for compounds capable of strong H-bonding. They compared the ODP column with silane-based columns and found that the ODP column demonstrated higher correlations with log P. It therefore appeared that ODP could conveniently take on the role of octanol. The manufacturers of the ODP column claim that the recommended pH range for use of the column is 2–13. Such basic conditions would destroy silane-based columns. The stability of these columns at such a high pH allows log P of the neutral form of strong bases to be measured. ODP columns have no unprotected Si-OH or Al-OH groups. In contrast, silanol-based columns have about 50 % of their solvent accessible silanol groups unprotected, even when heavily coated or end-capped. Such groups are weakly

acidic and may facilitate acid-base interactions, leading to an overestimation of log P for compounds with basic groups.

Non-silica based HPLC columns such as the polystyrene-divinylbenzene (PS-DVB) column have also been used to measure hydrophobicity. These columns have low efficiency and suffer from both shrinkage and swelling. They are also known to provide specific interactions due to their π orbitals and retention indices on such columns were found to correlate better with alkane/water partition coefficients than octanol/water partition coefficients. Interestingly, for the purpose of QSAR correlation studies, the alkane/water partition coefficient is found to better model blood/brain partitioning than the octanol/water partition coefficient.[36]

6.4 Direct HPLC Methods for Estimating Log P

The literature also describes the use of *direct* HPLC methods for determining log P of various analytes. One such method involves the measurement of log k′ on an octanol-coated column. Octanol may be entrained on the column as a result of its low solubility in water. It is now common practice to coat a RP-C8 or C18 column with water-saturated octanol and use octanol-saturated water as the mobile phase. However, the direct HPLC technique for log P determination was pioneered by Mirrlees *et al.*[41] Their column was prepared from a commercial Kieselguhr support, which was thoroughly silanised and slurry packed at high pressure into a column. The packing was then coated with water-saturated octanol, and then eluted with water-saturated octanol until no more droplets of octanol could be removed. Water-saturated octanol was then used to elute the chosen analytes. The authors called their method the *in-situ* coating method and using this method they obtained an excellent correlation between log k′ and log P for various compounds. By varying the column length (10, 20 or 30 cm) and flow rate they found that log P in the range of –0.3 to 3.7 could be measured. They also reported that for a flow rate of 2 ml/min the column does not normally require recoating in less than 50 h of use. However, the authors appreciate that the method can be unreliable when applied to compounds of very limited solubility. These compounds tend to give anomalously low log P values, and in extreme cases the peaks may accompany the solvent front. In such cases it is suspected that the partitioning process is too slow to reach equilibrium in the time taken to complete the elution process. Decreasing the flow rate sometimes helps.

Despite the success of the work of Mirrlees *et al.*, other authors in the field found that a major drawback was the tedious column preparation process. As a result, they tried to find a column that was analogous to that of Mirrlees *et al.* in terms of performance, but whose preparation was simplified.

Miyake *et al.*[42] attempted to simplify column preparation by using Corasil I as the solid support. It was heated at 110 °C overnight and then while still hot it was mixed with octanol. The slurry was packed into a PTFE tube and excess octanol removed by elution with the mobile phase (aqueous buffer saturated with octanol) until a stable baseline was obtained. The column displayed high stability and very good correlations were observed between the measured retention indices and log P for a series of compounds, including benzene, benzoic acids and phenols. They found that the use of a short column (4 cm) for compounds with high log P makes it easier to obtain accurate retention times, because peaks are sharper and operational times are shorter. Miyake *et al.* also reported in their paper that they did not obtain satisfactory results with the *in situ* coating method described by Mirrless *et al.* Octanol continually became dislodged and retention times were seen to gradually decrease as chromatography was performed.

Both Mirrlees and Miyake concluded that the value of the slope, m, of the equation defining the relationship between log k′ and log P should be unity if the retention time is governed by the partition between the mobile phase and the octanol coated on the support. Thus the value of m is a good measure of whether the stationary phase is octanol-like. Both authors derived equations where the slopes were very close to unity (Mirrlees, m = 0.960 and m = 0.965; Miyake, m = 1.0065). This indicates that the support media in both cases were almost totally inert. They also agree that the value of m = 0.72 that Henry *et al.*[43] obtained for a series of sulfonamides renders the results invalid. Mirrlees suggested that this result could indicate that the support is not totally inert and there is a direct interaction of silanol groups on the support with the solutes in the mobile phase. Henry employed a Corasil II column and coated it with 1 % octanol. If Mirrlees is correct, the coating method employed by Henry was not effective/complete or the column simply required re-coating just before these results were obtained.

Most recently, Kaune *et al.*[44] investigated a method involving the coating of octadecylsilane with octanol and an octanol-saturated aqueous mobile phase. Log k′ was determined to estimate log P of *s*-triazines and some of their degradation products. Since a maximum log P of 2.86 could be

determined using their method, they also explored the use of a gradient HPLC method to determine higher log P values. The column employed for these investigations was a LiChrisopher 100, RP-18 column. The invesigators used a flow rate of 2 ml/min for the compounds with log P ≤ 1.85 and 4 ml/min for compounds with higher log P. The equation to express the relationship between log k′ and log P for these compounds had a value of m close to unity (1.04) as expected for compounds analysed at 2 ml/min, but the value of m decreased to 0.89 for compounds run at 4 ml/min. Despite the success of the method, the authors acknowledged the high noise on the octanol-coated column compared with an uncoated column, possibly due to the displacement of small drops of octanol from the column. However, they do note that the displacement must have been small since retention times did not appear to decrease with time.

Ritter *et al.*[45] showed that the use of a longer column than Kaune *et al.* (30 mm compared with 17 mm) allowed the determination of log P of up to 3.53, applying a flow rate of 4 ml/min for the measurement of log P beyond 3.

In summary, most authors in this field would agree that the coated-column method is more difficult to perform than conventional (indirect) methods for the determination of log P. Problems involved are peak-tailing, higher noise and an increase of the baseline during a HPLC run. However, the method does appear to closely simulate the octanol/water system and is characterised by excellent reproducibility. It is restricted to log P values of about 3, which can be attributed to the practical disadvantage that an organic modifier can not be introduced to the mobile phase to reduce the retention of highly lipophilic compounds. However, for higher log P values the gradient HPLC method is very convenient and reliable.[44]

Excellent correlation has generally been reported between log k′ measurements and log P for neutral compounds and ionisable compounds for which ionisability can be suppressed by modifying the mobile phase conditions by incorporation of a suitable buffer. Whether a good correlation would be observed for the sulfobetaines (and other ionic compounds) is expected to depend on whether the octanol is effectively mobile (and hence 3-dimensional) or immobile (i.e. 2-dimensional).

6.5 Counter Current Chromatography (CCC) Methods for Measuring Log P

CCC has been described as a powerful tool that may be used to accurately measure octanol/water partition coefficients.[46-49] Generally, the technique uses stationary and mobile phases that are immiscible liquids. These can be any two liquids that can be used in a conventional stir-flask experiment. The whole volume of the liquid stationary phase is used for solute partitioning. In contrast to other chromatographic techniques, much higher concentrations can be injected without overloading. The support-free technique has the advantage that irreversible solute adsorption cannot occur, unlike the direct HPLC method for measuring log P where the packing material of the column (on which octanol is coated) may cause adsorption of very polar compounds or produce modifications of the solute at the liquid-solid interface. The simplicity of the solute retention mechanism is also a major advantage, since it depends only on one physicochemical parameter, the liquid-liquid partition coefficient.[46-51]

Where octanol and water are employed, usually octanol is used as the stationary phase and water (or a buffer) the mobile phase. This is often described as the "head-to-tail" or ascending mode. However, to extend the measurable log P range it may be convenient to switch the phase role. When octanol is the mobile phase and water the stationary phase, the technique is called "tail-to-head" or descending mode. Where water is used as the mobile phase flow rates of 1 to 4 ml/min are commonly employed. However, it is important to appreciate that due to the high viscosity associated with octanol a flow rate of 1 ml/min is recommended in cases where octanol is used as the mobile phase. In accordance with the conventional stir-flask experiment, CCC also requires pre-saturation of the phases before introduction to the apparatus.[46,47]

There is also the possibility of carrying out dual mode or back-flushing experiments. Dual mode CCC involves changing the role of the phases while chromatography is being performed with the aim of improving the retention times of the solutes. After some time in the direct mode to move the solute in the column, the role of the phases is reversed so that the stationary phase can become the mobile phase, or *vice versa*. This phase reversal will force the solute to leave the CCC apparatus. For example, Berthod *et al.*,[47] who studied a series of diuretics by CCC, described how log P for benzthiazide was derived using this technique. After some flowing time in the head-to-tail mode (130 min at 4 ml/min aqueous) the mode was switched to tail-to-head. A sharp peak was observed that corresponded to the phase change in the detector. From this point, octanol pushed the

aqueous phase. The benzthiazide peak showed up in the octanol phase at 9.65 min at 1ml/min. Hence, since P is given by V_{aq} / V_{oct}, log P = log (520/9.65) = 1.73. This method was also applied to reasonably hydrophilic compounds such as amiloride (log P = -1.25) using an aqueous phase buffered at pH 7.39. In this case the solute eluted at t_o when the direct mode was applied, however, on application of dual mode CCC, amiloride was notably retained and log P could be determined. This method is obviously extremely useful, however, it is important that the octanol and water phase volumes must not change during the switching procedure. This condition is often difficult to fulfil.

Using the methods described above, and often modifying the apparatus employed, log P has been determined accurately by CCC in the range of –2 to 2.5.

HPLC methods involving correlating log k′ and log P often give log P values that are ±0.3 log P units out, which means that P can actually be 2-fold higher or lower. In contrast, if used effectively in the direct mode, CCC can determine log P with an accuracy of ±0.02 log P units.[47]

Modern use of CCC originated with pioneering studies of Ito, Nunogaki and co-workers in Japan.[52] Ito defined two basic CCC systems. The first is called the hydrostatic equilibrium system (HSES) which consists of a stationary coiled column that is subject to a gravitational field (G), which is constant at any point in the coil. The second system, called the hydrodynamic equilibrium system (HDES), consists of a coiled column which is subjected to a fluctuating gravitational field produced by the additional rotation of the coil around its own axis, which enhances mixing of the two phases. HDES obviously provides better mixing than HSES; however to improve the efficiency of HSES the helical column of the original design has been greatly modified to improve the partition efficiency. Most HSES devices have completely lost their helical appearance. A well-known example of such a technique is droplet countercurrent chromatography (DCCC). However, literature reports that the standard octanol/water sytem has proved inadequate in DCCC due to the high viscosity of octanol.[46,53] Gago *et al.* reported that the use of octanol as the stationary phase does not allow the formation of suitable droplets to take place. Filling the columns with octanol and then pumping water or phosphate buffer (at pH 7.4) through the columns resulted in flow plugs and displacements of the stationary phase by the mobile phase. The same problem also arose when water was used as the stationary phase and octanol was made to ascend.[54] Various configurations exist for HDES depending on the relative orientation of the two rotational axes and the ratio of the

two radii to the two rotation speeds. Several permutations have been studied by Ito *et al.* and have been classified as schemes I, L, J and X. One example based on a scheme J synchronous multilayer coil planet centrifuge is called High-Speed Countercurrent Chromatography (HSCCC). The original model for semi-preparative separations has a 10 cm revolutional radius and holds a multilayer coil of 130 m long, 1.6 mm ID PTFE tubing with a total capacity of 300 ml, wound around the holder hub 10 cm in diameter. HSCCC is also suitable for analytical-scale separations. For this application, the dimensions of the column and the centrifuge radius are proportionally reduced while the revolution speed is increased to enhance CCC of the two solvent phases through the narrow lumen of the column. Analytical models now possess a coil of approximately 5 to 50 ml total capacity. HSCCC devices are very popular due to their efficiency, speedy separations, ease of operation and commercial availability at low cost.[53] More interestingly, they may be employed analytically to measure octanol/water partition coefficients.

6.6 Experimental

The compounds utilised in this study were a series of zwitterionic sulfobetaines (**1–22**) whose syntheses were described in Chapter 2. These compounds belong to three sub-series varying in the length of the spacer unit separating the quaternary ammonium centre from the sulfonate group. In addition, compounds **1–2**, **3–7** and **14–16** possess an alkyl chain (containing 6 to 12 carbon atoms), which is directly connected to the quaternary ammonium centre, and compounds **8–13** and **17–22** contain an aromatic ring, which is separated from the quaternary ammonium centre by up to four CH_2 units. In the case of **12**, **13**, **21**, and **22**, the aromatic ring is separated from the quaternary ammonium centre by three CH_2 units and carries a *para* alkyl substituent containing 4 to 6 carbon atoms.

HPLC retention times were determined using a HP1100 instrument using a UV detector set at 266 nm for analysis of the phenylalkylsulfobetaines and at 220 nm for analysis of the *para*-substituted phenylpropylsulfobetaines. Electrospray mass spectrometry (measuring total ion current) was employed for detection of the alkylsulfobetaines, since they lacked a suitable chromophore for UV detection. The Thermofinnigan aQa, single quadrupole mass spectrometer was used for these investigations which provided positive mode electrospray ionisation. Single ion monitoring of $[M+H]^+$ took place at a probe voltage and temperature of 4.5 kV and 300 °C

respectively. (Dwell time = 0.370 s, aQq max = 20 V, detector = 650 V, nitrogen nebulisation ON). All chromatographic retention data were taken as the mean of three determinations. The injection volume employed for both HPLC/UV and HPLC/ESMS was 20 µl. The log of the capacity factor, log k′, was determined by Equation 1.8.

6.6.1 Measurement of log k'_{IAM}

Samples for HPLC/UV were made up to a concentration ~1 mg/ml and for HPLC/ESMS to a concentration of ~10 µg/ml. The column employed for these investigations was an IAM.PC.DD2 column (12µ, 300Å, 4.6 mm x 100 mm) commercially available from Regis Technologies, Inc. The mobile phase was 100% HPLC grade water, which was eluted at 1 ml/min. Citric acid was used as the t_o marker and was co-injected with each analyte. The results obtained are shown in Table C1. Since *N*-[3-(4-hexylphenyl)propyl]-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (**13**) and *N*-[3-(4-hexylphenyl)propyl]-*N,N*-dimethyl-4-ammonio-1-butanesulfonate (**22**) did not elute within 1.5 h, k'_{IAM} for these compounds was determined at various percentage compositions using an organic modifier (acetonitrile) and k'_{IAM} was then determined by extrapolation (Figures C1 and C2).[55] Log P values determined by the stir-flask method and aquatic toxicity values, expressed in terms of 48 h EC_{50} to the water flea Daphnia *magna*, for the compounds are listed in Table C2.

6.6.2 Measurement of log k'_{C8} and log k'_{C18}

Retention measurements were performed on a Genesis C8 or C18 (4µ, 4.6 x 150 mm) column. Uracil was used as the t_o marker and was co-injected with each analyte. Samples were made up to a concentration of ~2 mg/ml. The mobile phase was 90 % H_2O : 10 % AcCN for the C8 column and 80 % H_2O : 20 % AcCN for the C18 column. The flow rate was 1 ml/min. The results obtained are shown in Table C3 and C4. Log k'_{C18} for *N*-[3-(4-hexylphenyl)propyl]-*N,N*-dimethyl-3-ammonio-1-propanesulfonate (**13**) and *N*-[3-(4-hexylphenyl)propyl]-*N,N*-dimethyl-4-ammonio-1-butanesulfonate (**22**) was determined at various percentage compositions of organic modifier and then extrapolated to obtain the value in 80 % H_2O : 20 % AcCN (Figures C3 and C4).

6.6.3 Measurement of log k′ (octanol-coated –C8)

Retention measurements were performed on a Genesis C8 (4μ, 4.6 x 150 mm) column, previously coated with water-saturated octanol. Water-saturated octanol and octanol-saturated water were prepared by allowing mutual saturation of water (HPLC grade) and octanol (HPLC grade). Coating was performed by passing water-saturated octanol through the column at a flow rate of 0.5 ml/min. After approximately 3 h a stable baseline was obtained. The octanol-saturated water mobile phase was then pumped through the system at 1 ml/min until no further droplets of octanol could be observed in the effluent and the baseline once again stabilised.[44] The t_o marker employed was KNO_3 and this was injected on its own prior to the injection of each analyte. Uracil, a commonly used t_o marker for HPLC, appeared to be retained under these conditions. Samples were made up to a concentration ~2 mg/ml. The mobile phase employed was octanol-saturated water, which was eluted at 0.5 ml/min and at 3.5 ml/min. The results obtained at the two different flow rates are shown in Table C5 and C6. The retention times for *N*-[3-(4-butylphenyl)propyl]-*N,N*-dimethyl-3-ammonio-1-propanesulfonate **(12)** and *N*-[3-(4-butylphenyl)propyl]-*N,N*-dimethyl-3-ammonio-1-butanesulfonate **(21)** at 0.5 ml/min were determined using higher flow rates and then extrapolated to obtain the values at 0.5 ml/min.

6.6.4 Measurement of Retention Indices in CCC

A HDES of the HSCCC-type possessing a multilayer coil, which undergoes J synchronous planetary motion around the central axis of the centrifuge was employed for these investigations. The model used was a Brunel CCC system, manufactured by the Brunel Institute for Bioengineering, Brunel University, Uxbridge, Middlesex. The model has a 110 cm revolutional radius and holds a multilayer coil 48 m long, consisting of 1.6 mm ID PTFE tubing with a total capacity of 49.9 ml. The model also uses 0.5 mm ID PTFE connection tubing and the system is temperature controlled at 30 °C (ethylene glycol refridgerated). The β-value for the system is 0.83-0.86. This corresponds to the ratio of the holder axis to each circle (r), to the distance from the holder axis to the centrifuge axis (R).

The octanol and water solvents were first mutually saturated before introduction to the chromatograph. The mobile phase of water was pumped through the CCC apparatus using a HP 1100 quaternary pump and the compounds were detected using an Applied Biosystems 783A

UV/VIS detector. A picolog ADC111 analogue to digital converter was also used. The CCC apparatus was first filled with the octanol stationary phase. The centrifuge was then turned to a stable rotational speed of 830 r.p.m. The pump was then rinsed with the aqueous mobile phase. This phase enters the apparatus in the head-to-tail (ascending) mode because it is less dense than the octanol stationary phase. So long as the apparatus is not equilibrated the octanol phase is pushed out of the apparatus and two distinct layers are seen in the collection cylinder. When the displacement of octanol ceases and only the aqueous phase is observed to elute, the instrument was ready for use. Samples were prepared in concentration of ~1 mg/ml and the injection volume employed was 20 µl. The displaced octanol volume corresponds to the aqueous phase volume (V_e) of the CCC system. Since small amounts of octanol may be carried out of the system or dissolved by the aqueous phase, V_e is taken as the time taken for the t_o marker, KNO_3, to elute. KNO_3 was detected at 210 nm.

$$P = (V_R - V_e) / V_s \quad \text{or} \quad (V_R - V_e) / (V_T - V_e) \qquad \textit{(Equation 6.2)}$$

V_R = retention volume of analyte; V_e = volume of eluent; V_s = volume of stationary phase

6.7 Results and Discussion

6.7.1 Correlation of log P and log k'_{IAM} with Aquatic Toxicity

The correlation between log P and log k'_{IAM} is shown in Figure 6.2. The correlation between log k'_{IAM} and aquatic toxicity, and between log P and aquatic toxicity, is shown in Figures 6.3 and 6.4 respectively. As can be seen from Figures 6.3 and 6.4 there is a better correlation between log k'_{IAM} and aquatic toxicity ($r^2 = 0.9258$) than between log P and aquatic toxicity ($r^2 = 0.8705$).

Compound **(3)** has been omitted from Figures 6.3 and 6.4 since in both cases it is a significant outlier. Its inclusion would change the correlation equation dramatically and lower the correlation coefficient. It therefore seems likely that this sulfobetaine displays a different mode of toxic action. Since it is the sulfobetaine with the lowest molecular weight it is possible that it is small enough to penetrate the membrane directly or be carried into the biological system *via* an ion channel, and this may override its dependence on hydrophobicity (or phospholipophilicity). Another piece of evidence that helps to support this conclusion is that compounds **(3)** and **(15)** have

almost identical log P values, but totally different aquatic toxicities; compound (**3**) is 1.7 times more toxic than compound (**15**).

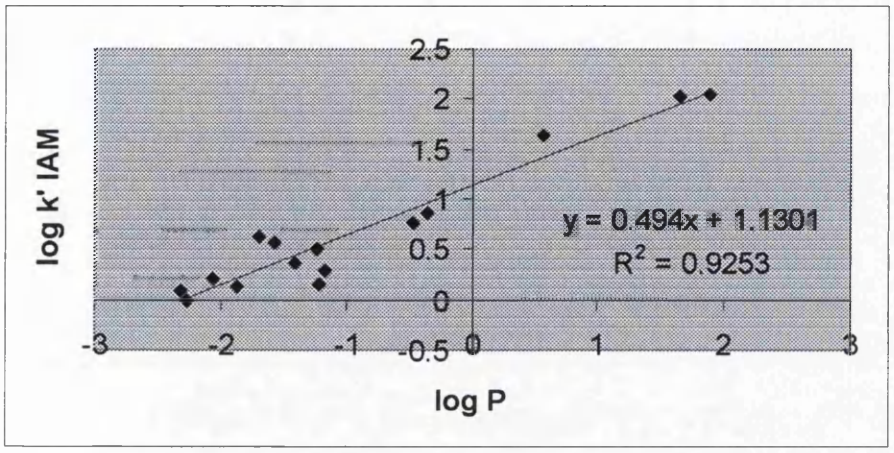


Figure 6.2 Relationship between log P determined by stir-flask method and log k'_{IAM}.

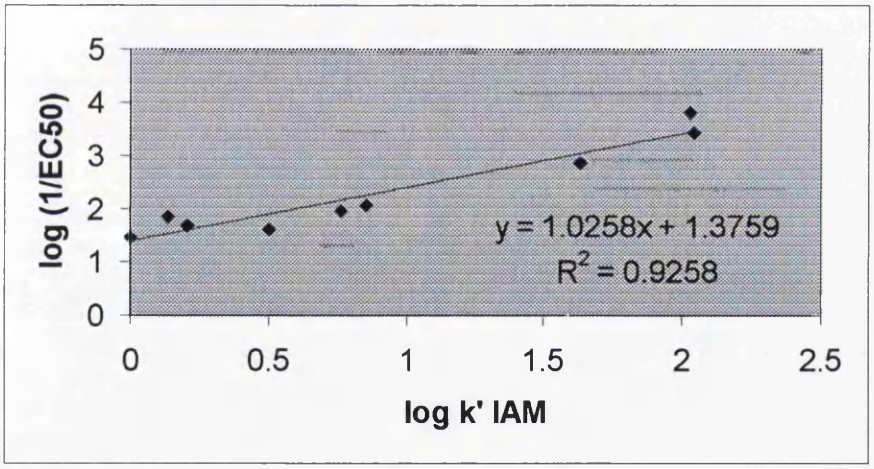


Figure 6.3 Relationship between log k'_{IAM} and Aquatic Toxicity.

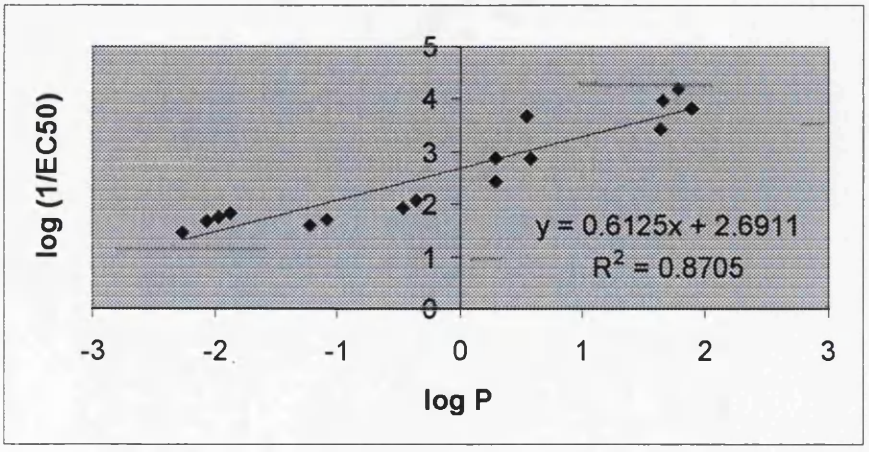


Figure 6.4 Relationship between log P and Aquatic Toxicity.

6.7.2 Correlation of log P and log k'_{IAM} with log k'_{C8} and log k'_{C18}

The correlation between log P and log k'_{C8} and between log k'_{IAM} and log k'_{C8} (Table C3) is shown in Figures 6.5 and 6.6 respectively. Surprisingly a poor correlation ($r^2 = 0.4748$) between log k'_{C8} and log P is observed, but a much better correlation ($r^2 = 0.9067$) between log k'_{C8} and log k'_{IAM}. In contrast, the results on the C18 column (Table C4) show a good correlation both between log P and log k'_{C18} ($r^2 = 0.9732$, Figure 6.7) and between log k'_{IAM} and log k'_{C18} ($r^2 = 0.9421$, Figure 6.8).

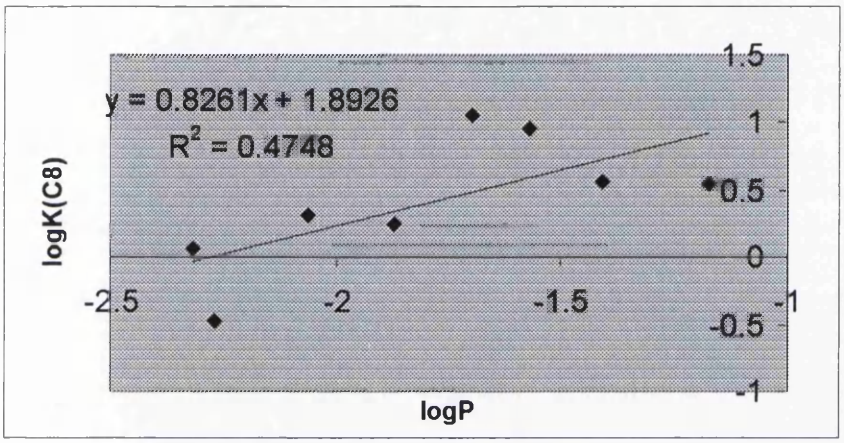


Figure 6.5 Relationship between log P and log k'_{C8}.

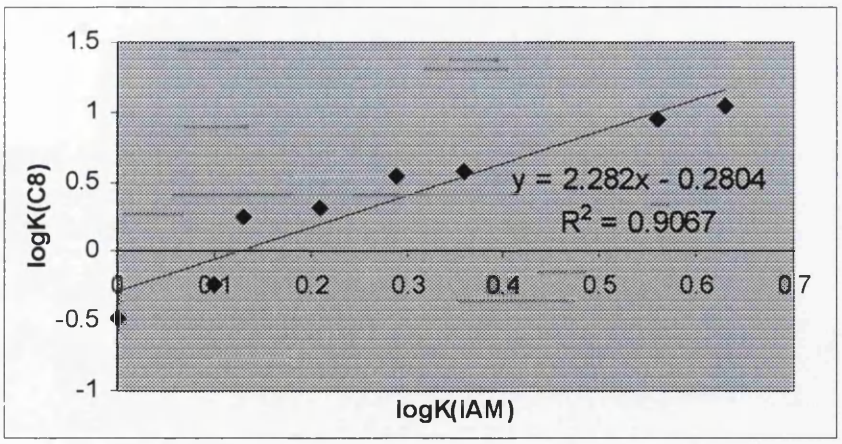


Figure 6.6 Relationship between log k'_{IAM} and log k'_{C8}.

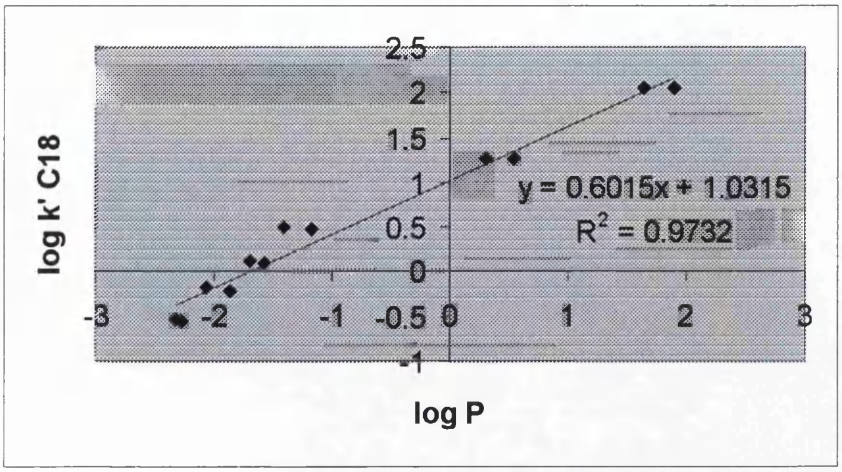


Figure 6.7 Relationship between log P and log k'$_{C18}$

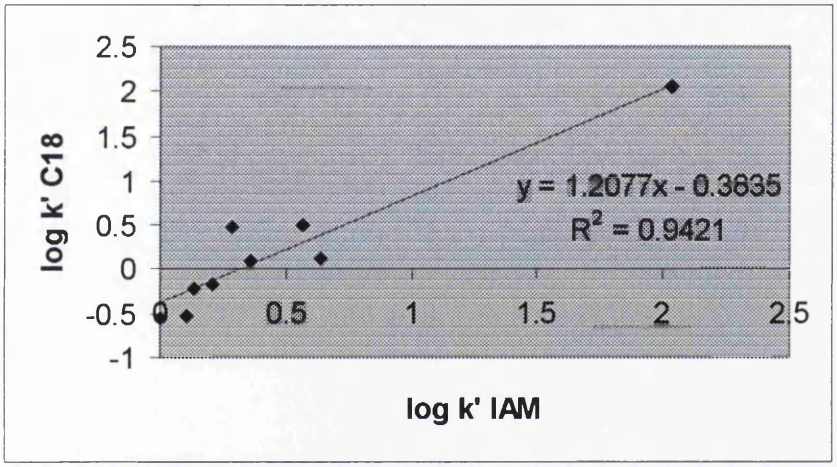


Figure 6.8 Relationship between log k'$_{IAM}$ and log k'$_{C18}$.

6.7.3 Correlation between log k' $_{(octanol\text{-}coated\ -C8)}$ and log P

The correlation between log k′ on the octanol-coated C8 column at 0.5 ml/min (Table C5) and log P is shown in Figure 6.9. An excellent correlation is observed between these parameters (r^2 = 0.9811). Log k′ was also determined at a higher flow rate for 4 compounds (Table C6) and once again, although only a small number of compounds are involved, an excellent correlation (r^2 = 0.9964) is observed (Figure 6.10). By removing the latter two points of the graph displayed in Figure 6.9 (whose log k′ values were determined by extrapolation of the flow rate – see Figures C5 and C6), Figure 6.11 shows that the correlation coefficient is reduced (r^2 = 0.9420). However, more interesting is the fact that the slope is now closer to unity (m = 0.9335).

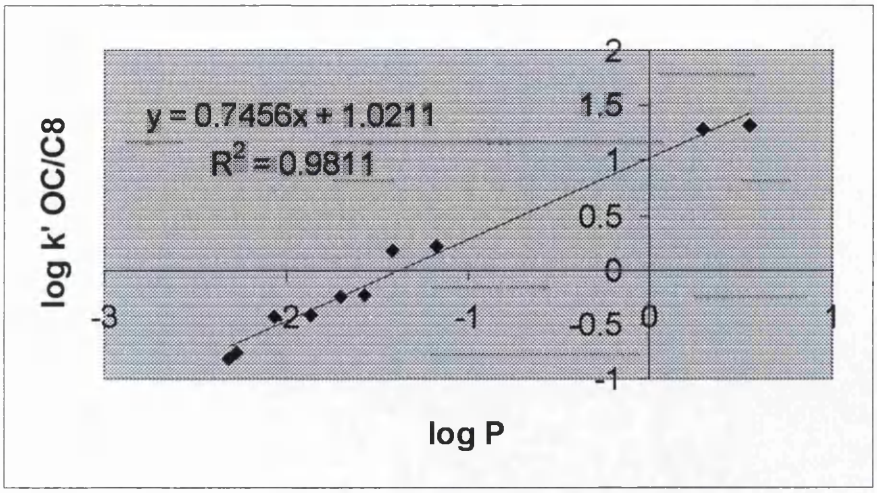


Figure 6.9 Relationship between log P and log k' on an octanol-coated C8 column (@ 0.5 ml/min).

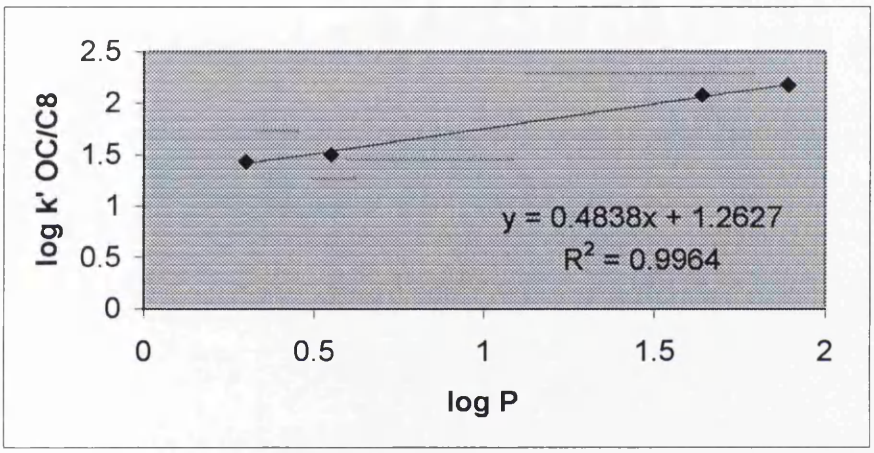


Figure 6.10 Relationship between log P and log k' on an octanol-coated C8 column (@ 3.5 ml/min).

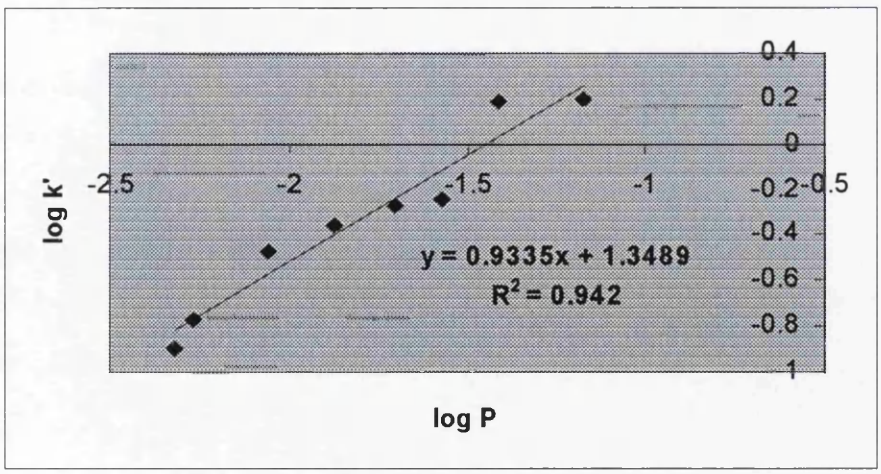


Figure 6.11 Relationship between log P and log k on an octanol-coated C8 column (@ 0.5 ml/min).

6.7.4 Correlation between Retention Indices in CCC and Log P and Direct Calculation of log P_{CCC}

Figure 6.12 shows the CCC separation of eight sulfobetaines under conditions of V_e = 27.5 ml (V_s = 22.4 ml), at 1 ml/min and 30 °C. As we can see from this Figure and from Table C7, the first six hydrophilic sulfobetaines do not show any significant difference in retention time and thus they do not elute in order of increasing hydrophobicity as we would expect. The only sulfobetaines that are notably retained are compounds (**12**) and (**21**) with log P of 0.30 and 0.55 respectively. These values were determined by the stir-flask experiment. Log P of these compounds, from the retention data of this CCC experiment, may now be calculated.

$$P_{(12)} = ((3830/60) - 27.5) / 22.4 = 1.62 \therefore \log P = \mathbf{0.21}$$

$$P_{(21)} = ((5529/60) - 27.5) / 22.4 = 2.89 \therefore \log P = \mathbf{0.46}$$

Caffeine, whose literature log P value is 0.07, was used a reference compound to check the general ability of the system to measure log P.

$$P_{(caffeine)} = ((2897/60) - 27.5) / 22.4 = 0.93 \therefore \log P = -0.03$$

Figure 6.13 shows the CCC separation of eight hydrophilic sulfobetaines under conditions of V_e = 12.5 ml (V_s = 37.4 ml), at 2 ml/min and 30 °C. As we can see from this Figure (and Table C8), it appeared that due to the hydrophilicity of the compounds analysed, equilibrium was still not being reached. This resulted in the compounds, which were members of an homologous series, eluting in a random order very close to the dead time (V_e) of the chromatographic system (as indicated by KNO_3), rather than in relation to their true hydrophobicity. In the light of this result, it seemed logical that if we could employ such a system where a similar volume of octanol could be retained and then use a reduced flow rate such as 1 ml/min, the homologues would be resolved and elute in the order of their respective log P values derived from stir-flask experiments. It is important that a similar volume of octanol is retained (if not more) to encourage interaction with the stationary phase despite the extreme hydrophilic nature of these compounds (log P of –2.32 to –1.17) before elution at a lower flow rate. A reduction in the volume of octanol retained by the CCC would otherwise outweigh the benefit of reducing the flow rate as we have seen from the previous two CCC experiments.

Figure 6.14 shows the CCC separation of six of the eight particularly hydrophilic sulfobetaines under conditions of V_e = 10.0 ml (V_s = 39.9 ml), at 1 ml/min and 30 °C. As can be seen from this Figure and from Table C9, the sulfobetaines still fail to elute in order of increasing hydrophobicity. Furthermore, if the results of the three CCC experiments are compared (Tables C7-C9) it is seen that these sulfobetaines elute in a random order from experiment to experiment and the log P values calculated from the retention data are greater than the corresponding log P values determined by stir-flask experiments.

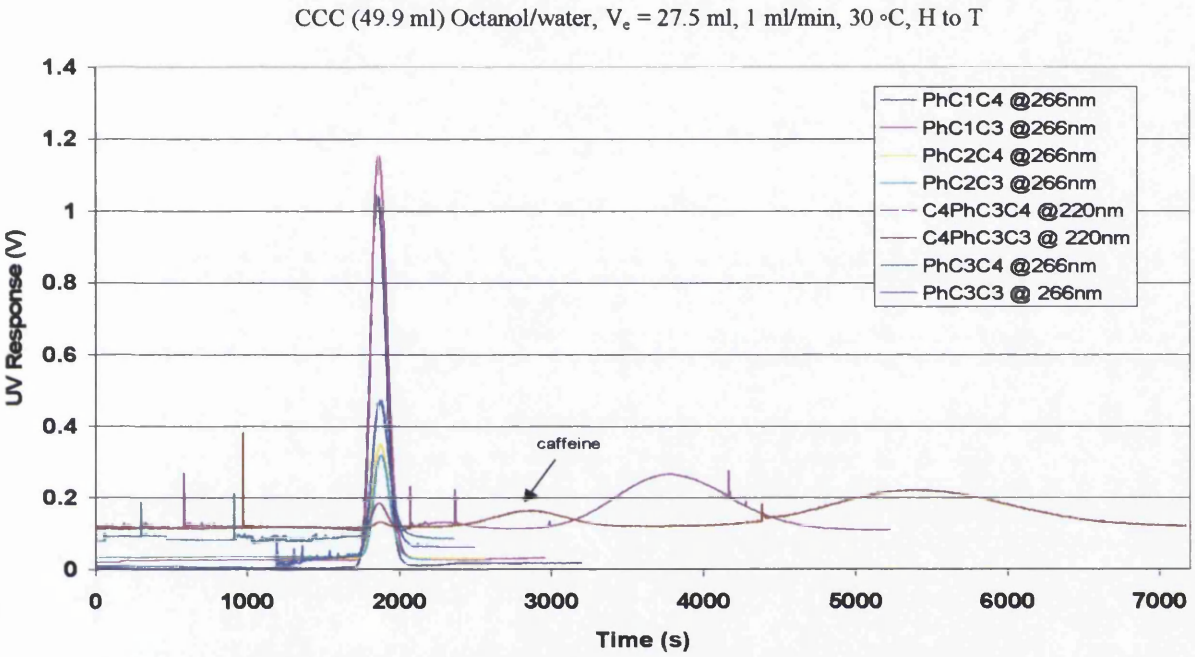


Figure 6.12 CCC Separation of 8 Sulfobetaines and a reference compound, Caffeine.

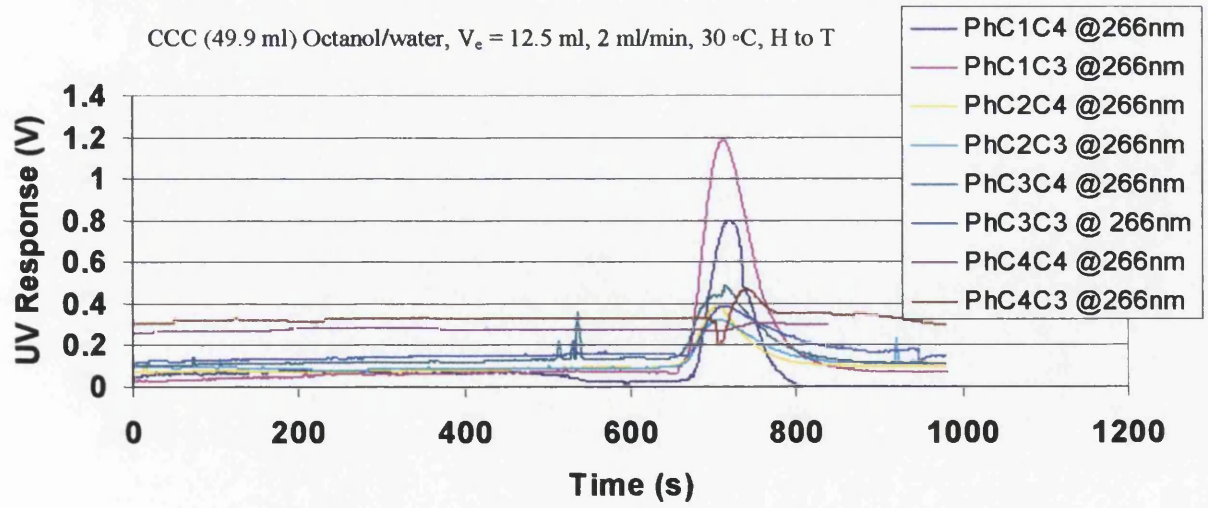


Figure 6.13 CCC Separation of 8 Hydrophilic Sulfobetaines

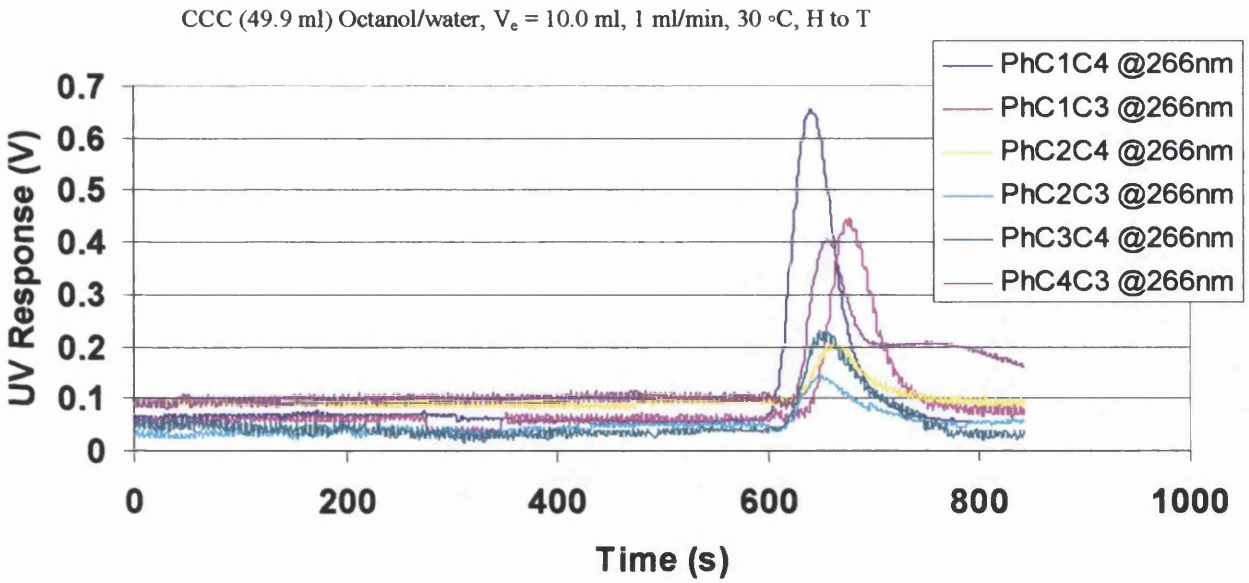


Figure 6.14 CCC Separation of 6 Hydrophilic Sulfobetaines

6.8 Conclusion

From Figure 6.2, it may be concluded that over a reasonably wide range, log P and log k'_{IAM} show good correlation, despite the fact that one parameter is a measure of hydrophobicity and the other is a measure of phospholipophilicity, in addition to the fact that one parameter can be regarded as a 3-D model while the other represents a 2-D model of the lipid membrane. However, the direct relationship between log k'_{IAM} and log P does not appear to be observed for close homologues, which would be expected to give small, incremental differences in both phospholipophilicity and hydrophobicity. In contrast, the trend in log P (measured by the stir-flask method) appears to be uniform for close members of an homologous series, e.g. compounds **(8-11)** and **(17-20)**. It is also observed that log k'_{IAM} correlates better with aquatic toxicity (Figure 6.3) than does log P (Figure 6.4). This supports the idea that the interactions involved in membrane transport and in aquatic toxicity are more complex than are revealed by the measurement of log P. It is also expected that 2-D models of the lipid membrane are more realistic, since in 2-D ionic interactions are most significant.

The fact that log P did not correlate with retention on the C8 column (Figure 6.5) was extremely surprising, especially since good correlation was observed using the C18 column (Figure

6.7). However, this can be explained by the fact that in the analysis of the applicability of the C18 column for the estimation of log P, a greater range of compounds were studied covering a wider log P range. In contrast, fewer compounds that were in fact close homologues and covered only a narrow log P range, were studied using the C8 column. As found when previously trying to correlate a 3-D (log P) with a 2-D (log k'_{IAM}) model of the lipid membrane, a direct relationship is not obeyed for homologues, although generally a good correlation does exist between log P and log k'_{C18}, since both are dependent on hydrophobicity. Furthermore, a good correlation was observed between these hydrophobicity parameters and log k'_{IAM}, the phospholipophilicity parameter (Figures 6.6 and 6.8). It seems likely that the close correlation found between log k'_{IAM}, log k'_{C8} and log k'_{C18} for close homologues is due to the fact that these parameters are all measured using immobilised HPLC-bonded phases and hence 2-D models of the lipid membrane.

Similarly, it seems logical that the close correlation between log k′ on the octanol-coated C8 column and log P for close homologues ($r^2 = 0.9811$) can be attributed to both being measures of hydrophobicity, but more importantly log P being a 3-D model and log k′ being a 'partial 3D model' of the natural membrane. The octanol-coated column is described as a 'partial 3-D model', since even though the inert C8 support is immobilised the octanol is actually effectively mobile and simply held in place by its viscosity. Despite excellent correlation between log k′ under these conditions and log P (Figure 6.9, $r^2 = 0.9811$; Figure 6.10, $r^2 = 0.9964$), according to Mirrlees and Miyake our results would be rendered invalid due to the fact that slopes of the graphs deviate from unity. Clearly, if the retention time is governed ultimately by partitioning between the mobile phase (water) and octanol, then the slope should be unity, since $k' = K (V_s/V_e)$. We can therefore appreciate that the slope is often used as a measure of whether the stationary phase is octanol-like. Despite confidence that our experimental work was not at fault, since the column was effectively coated, we are unsure why our values deviate from unity (Figure 6.9, m = **0.75** @ 0.5 ml/min; Figure 6.10, m = **0.48** @ 3.5 ml/min; Figure 6.11, m = **0.93** @ 0.5 ml/min). Since the slope appears to vary significantly with flow rate (*cf.* values of m for graphs given by Figures 6.9 and 6.10), the deviation from unity may be related to the inability to reach equilibrium due to the hydrophilic nature of most of the compounds analysed. The fact that the correct elution order (in relation to their log P) was observed suggests some retention at least. However, the interaction between the solutes and the octanol-coated column must be considerably less than that between the solutes and

octanol in a stir-flask experiment, since an increase in flow rate restricts the time for interaction. Removal of the latter two points (for which retention data were determined by extrapolation of flow rate) resulted in a marked improvement in m for the relationship (Figure 6.11). This could suggest that the relationship of log t_r to flow rate is not strictly linear. However, the value of the slope for this data set is still not unity (m = 0.93). The deviation of m from 1 for the smaller data set may be related to the ionic nature of the compounds since the model is not strictly 3-dimensional. Since it is somewhere between a 2-D and a 3-D model, ionic interactions will still potentially be more significant here than in a true 3-D model.

It was always anticipated that an excellent correlation would exist between log P_{CCC} and log P even for close members of an homologous series. This theory stems from the fact that both the CCC technique and the stir-flask system provide 3-dimensional measures of the partition coefficients. Another piece of evidence in support of our proposal was provided by Tsai *et al.* who investigated zwitterionic amino acids by CPC (using an octanol/aqueous buffer system). They were able to highlight the influence of intercharge distance on the lipophilicity of homologous piperidinyl carboxylic acids. A decrease in log D was seen with an increase in distance between the charges. This goes to show that CPC (and indeed other centrifugal chromatography techniques such as CCC) are capable of discriminating between the hydrophobicity of compounds for which log P is similar. This is due to the increase in the precision and accuracy of log P estimates provided by 3-dimensional models.

The observed trend in log D that has been described for the homologous piperidinyl carboxylic acids is similar to that observed for log P for the sulfobetaines measured by the stir-flask method. For homologous sulfobetaines, a decrease in log P is observed with increasing length of the methylene spacer unit between N^+ and SO_3^- from 3 to 4 methylene groups. Another important similarity is that both series of zwitterionic compounds (piperidinyl carboxylic acids and sulfobetaines) are very hydrophilic in nature and so this gave hope of the sulfobetaines being retained by the system if suitable conditions were sought.[56]

It was, however, unfortunate that the CCC conditions investigated for analysis of the hydrophilic sulfobetaines did not give rise to a reliable means of estimating log P. This is expected to be a result of failure to reach equilibrium. It appeared that an increase in the volume of octanol retained by the system and an accompanied reduction in flow rate did give log P predictions closer

to that derived from stir-flask experiments. However, the compounds always failed to elute in order of increasing hydrophobicity. A further decrease in flow rate may have resolved this problem; however it is more likely that a change in the role of the phases was required (where water becomes the stationary phase, and octanol the mobile phase) for analysis of compounds displaying such high hydrophilicity. If under these conditions retention times are severely extended, there is the possibility of applying dual mode CCC (as described in Section 6.5) to help solve the elution problem. Finally, the fact that log P could be determined with very good accuracy for compounds **12** and **21,** that were indeed close homologues, indicates the potential of the technique for estimating log P (Section 6.7.4), providing the retention problem for the hydrophilic sulfobetaines can be overcome.

In summary, it appears that the value of the correlation coefficient (r^2), especially for close homologues, is affected by whether the parameters to be correlated are 2-D or 3-D models of the lipid membrane. It also seems to depend on whether the correlated parameters are models of hydrophobicity or phospholipophilicity. Perhaps to a lesser extent r^2 is also affected by the size of the data set. The best correlations will therefore be obtained if the dimensions of the parameters to be compared are the same and they are both measures of the same property. Any deviation from this scenario will limit r^2 to a certain degree.

The results of this study have also suggested that for zwitterionic sulfobetaines 2-D parameters such as k'_{IAM} tend to correlate better with aquatic toxicity than the 3-D parameter, log P, that has been used as the sole physicochemical property for use in QSARs for aquatic toxicity for many years. It would appear that 3-D parameters of the lipid membrane take into consideration both ionic and non-ionic interactions; however 2-D analogues are expected to find ionic interactions of greater significance.[57] It is therefore probable that correlations between 2-D models of the lipid membrane and aquatic toxicity are more useful QSARs for zwitterionics and indeed all other ionic compounds. However, 3-D models of the lipid membrane in such QSARs are expected to be more useful for non-ionic compounds.

6.9 References

1. Hansch, C. and Leo, A., *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*, ACS Professional Reference Book, ACS (1995) Chapter 4.
2. Singh, W. P., Lin, G. H. and Bockris, J. O'M, unpublished results.
3. Amato, M., Barbato, F., Morrica, P., Quaglia, F. and La Rotonda, M. I., *Helv. Chim. Acta*, **83** (2000) 2836.
4. Nasal, A.,Sznitowska, M., Buckinski, A. and Kaliszan, R., *J. Chromatogr., A*, **692** (1995) 83.
5. Escher, B. I., Schwarzenbach, R. P. and Westall, J. C., *Environ. Sci. Technol.*, **34** (2000) 3962.
6. Veith, G. D., Austin, N. M. and Morris, R. T., *Water Research*, **13** (1979) 43.
7. Regis Technologies, Inc., Chromatography Catalogue 1998-9.
8. Liu, H., Ong, S., Glunz, L. and Pidgeon, C., *Anal. Chem.*, **67** (1995) 3550.
9. Regis Technologies, Inc., Information Booklet, July 2000.
10. Caldwell, G. W., Masucci, J. A., Evangelisto, m. and White, R., *J. Chromatogr., A*, **800** (1998) 161.
11. Pidgeon, C., Ong, S., liu, H. *et al, J. Med. Chem.*, **38** (1995) 590.
12. Duncarme, A., Neuwels, M., Goldstein, S. and Massingham, R., *Eur. J. Med. Chem.*, **33** (1998) 215.
13. Cohen, D. E., Leonard, M. R., *J. Lipid Research*, **36** (1995) 2251.
14. Salminen, T., Pulli, A., Taskinen, J., *J. Pharm. Biomed. Anal.*, **15** (1997) 469.
15. Makovskaya, V., Dean, J. R., Tomlinson, W. R., Hitchen, S. M. and Comber, M., *Anal. Chim. Acta*, **315** (1995) 183.
16. Carlson, R. M., Carlson, R. E. and Kopperman, Herbert, L., *J. Chromatogr.*, **107** (1975) 219.
17. Eadsforth, C. V. and Moser, P., *Chemosphere*, **12** (1983) 1459.
18. Könemann, R., Zelle, R., Busser, F., Hammers, W. E., *J. Chromatogr.*, **178** (1979) 599.
19. OECD Guideline for testing of Chemicals, Proposal for Updated Guideline **117**, *Partition Coefficient (n-octanol/water), HPLC Method*, Draft Document, July 1998.
20. Renberg, L., Sunström, G. and Sundh-Nygård, K., *Chemosphere*, **80** (1980) 683.
21. Belsner, K., Pfeifer, M., Willert, B. J., *J. Chromatogr.*, **629** (1993) 123.
22. Dross, K. P., Mannhold, R., Rekker, R. F., *Quant, Struct.-Act. Rel.*, **11** (1992) 36.
23. Hulshoff, A., Perrin, J. H., *J. Chromatogr.*, **120** (1976) 65.
24. Biagi, G. L., Guerra, M. C., Barbaro, A. M., Barbieri, S., Recanatini, M., Borea, P. A., Pietrogrande, M. C., *J. Chromatogr.*, **498** (1990) 179.

25. Fujisawa, S. and Mashuhara, E., *J. Biomed. Mat. Res.*, **15** (1981) 787.
26. Braumann, T., *J. Chromatogr.*, **373** (1986) 191.
27. Garst, J., Wilson, W., *J. Pharm. Sci.*, **73** (1984) 1616 and 1623.
28. Collander, R., *Acta. Chem. Scand.*, **5** (1951) 774.
29. Abraham, M. H., Chadha, H. S., Leo, A. J., *J. Chromatogr., A*, **685** (1994) 203.
30. Brugemann, W. A., Van der Steen, J. and Hutzinger, O., *J. Chromatogr.*, **238** (1982) 335.
31. Yunger, L. M. and Cramer, R. D., *Mol. Pharmacol.*, **20** (1981) 602.
32. Yamagami, C. and Takao, N., *Chem. Pharm. Bull.*, **39** (1991) 2924.
33. Camps, F., Colomina, O., Messeguer, A., Sanchez, F., *J. Liq. Chromatogr.*, **9** (1986) 23.
34. Wright, S., Harris, R., Collins, R., Corbett, R., Green, A., Wadman, E., Batt, D., *J. Med. Chem.*,**35** (1992) 3148.
35. Tulp, M. Th. M. and Hutzinger, O., *Chemosphere*, 7 (1968) 849.
36. Donovan, S. F. and Pescatore, M. C., *J. Chromatogr., A*, **952** (2002) 47.
37. Leo, A. J., *J. Pharm. Sci.*, **76** (1987) 166.
38. Spencer, R. W., Copp, I. J., Pfister, J. R., *J. Med. Chem.*, **28** (1985) 1828.
39. Lambert, W. J., *J. Chromatogr., A*, **656** (1993) 469.
40. Vallat, P., Fan, W., El Tayar, N., Carrupt, P-A., Testa, B., *J. Liq. Chromatogr.*, **15** (1992) 2133.
41. Mirrless, M. S., Moulton, S. J., Murphy, C. T., and Taylor, P. J., *J. Med. Chem.*, **19** (1976) 615.
42. Miyake, K. and Terada, H., *J. Chromatogr.*, **157** (1978) 386.
43. Henry, D., Block, J. H., Anderson, J. L., Carlson, G. R., *J. Med. Chem.*, **19** (1976) 619.
44. Kaune, A., Bruggeman, R. and Kettrup, A., *J. Chromatogr., A*, **805** (1998) 119.
45. Ritter, S., Hauthal, W., Maurer, G., *J. Chem. Eng. Data.*, **39** (1994) 414.
46. Foucault, A. P., *Anal. Chem.*, **63** (1991) 569.
47. Berthod, A., Carda-Broch, S., Garcia-Alvarez-Coque, M. C., *Anal. Chem.*, **71** (1999) 879.
48. El Tayar, N., Tsai, R-S., Vallat, P., Altomare, C., Testa, B., *J. Chromatogr.*, **556** (1991) 181.
49. Vallat, P., El Tayar, N., Testa, B., Slacanin, I., Marston, A., Hostettman, K., *J. Chromatogr.*, **504** (1990) 411.
50. Conway, W. D., *J. Chromatogr.*, **538** (1991) 27.
51. Foucault, A. P. and Chevolot, L., *J. Chromatogr., A*, 808 (1998), 3.
52. Ito, Y., Weinstein, M., Aoki, A., Horada, R., Kimura, E., Nunogaki, K., *Nature*, **212** (1966) 985.
53. Ito, Y., Counter Current Chromatography, Chapter 2. In: *J. Chromatogr. Library*, Vol. **51A**, Chromatography 5th Edition.

54. Gago, F., Alvarez-Builla, J., Elguero, J., *J. Chromatogr.*, **360** (1986) 247.

55. Taillardat-Bertschinger, A., Galland, A., Carrupt, P-A. and Testa, B., *J. Chromatography.*, *A* (2002), in press.

56. Tsai, R-S., Testa, B., El Tayar, N., Carrupt, P-A., *J. Chem. Soc. Perkin Trans. 2* (1991) 1797.

57. Roberts, D. W. and Costello, J. F., *QSAR and Combinat. Sci.*, **22** (2003) 226.

Appendix A

Chapter 3: Determination of Log P

Tables A1 & A2 *Tables to show the calculation of log P for PhC1C4 (17) from octanol/water experiments A and B*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **17 A**	54.9	2	2	4	1	Av. = 0.2795	0.2795 x 1 = 0.2795	0.2795 / (2/4) = 0.5590
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.02004	0.02004 x 0.15 = 3.006 E-03	3.006 E-03 / (3/3.25) = 3.257 E-03

$\therefore \log P = \log (3.257\ E\text{-}03/0.559) = \mathbf{-2.23}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **17 B**	55.1	2	2	4	1	Av. = 0.2768	0.2768 x 1 = 0.2768	0.2768 / (2/4) = 0.5536
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.01206	0.01206 x 0.15 = 1.809 E-03	1.809 E-03/ (3/3.25) = 1.953 E-03

$\therefore \log P = \log (1.953\ E\text{-}03/0.5536) = \mathbf{-2.45}$

Tables A3 & A4 Tables to show the calculation of log P for 17 from octanol/water experiments C and D

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **17** C	66.1	2	2	4	1	Av. = 0.4447	0.4447 x 1 = 0.4447	0.4447 / (2/4) = 0.8894
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.02273	0.02273 x 0.15 = 3.410E-03	3.410 E-03 / (3/3.25) = 3.694 E-03

$\therefore \log P = \log (3.694\ E\text{-}03/0.8894) = \mathbf{-2.38}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor Related to Concentration of analyte in water layer
Expt. **17** D	67.1	2	2	4	1	Av. = 0.4390	0.4390 x 1 = 0.4390	0.4390 / (2/4) = 0.8780
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.03305	0.03305 x 0.15 = 4.958 E-03	4.958 E-03/ (3/3.25) = 5.371 E-03

$\therefore \log P = \log (5.371\ E\text{-}03/0.8780) = \mathbf{-2.21}$

Tables A5 & A6 Tables to show the calculation of log P for PhC1C3 (8) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **8** A	54.8	2	2	4	1	Av. = 0.2915	0.2915 x 1 = 0.2915	0.2915 / (2/4) = 0.5830
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.01715	0.01715 x 0.15 = 2.573 E-03	2.573 E-03 / (3/3.25) = 2.787 E-03

$\therefore \log P = \log (2.787\text{ E-}03/0.5830) = \mathbf{-2.32}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **8** B	54.8	2	2	4	1	Av. = 0.3092	0.3092 x 1 = 0.3092	0.3092 / (2/4) = 0.6184
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.02080	0.02080 x 0.15 = 3.120 E-03	3.210 E-03/ (3/3.25) = 3.380 E-03

$\therefore \log P = \log (3.380\text{ E-}03/\ 0.6184) = \mathbf{-2.26}$

Tables A7 & A8 Tables to show the calculation of log P for **8** *from octanol/water experiments C and D*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **8** C	65.8	2	2	4	1	Av. = 0.2768	0.2768 x 1 = 0.2768	0.2768 / (2/4) = 0.5536
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.01206	0.01206 x 0.15 = 1.809 E-03	1.809 E-03/ (3/3.25) = 1.953 E-03

$\therefore \log P = \log (1.953\text{ E-}03/0.5536) = \mathbf{-2.16}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor Related to concentration of analyte in water layer
Expt **8** D	66.7	2	2	4	1	Av. = 0.4656	0.4656 x 1 = 0.4656	0.4656 / (2/4) = 0.9312
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.25	3.25	0.15	Av. = 0.02689	0.02689 x 0.15 = 4.034 E-03	4.034 E-03/ (3/3.25) = 4.370 E-03

$\therefore \log P = \log (4.370\text{ E-}03/0.9312) = \mathbf{-2.33}$

*Tables A9 & A10 Tables to show the calculation of log P for PhC2C4 (**18**) from octanol/water experiments A and B*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **18** A	64.7	3	0.5	3.5	0.286	Av. = 0.8947	0.8947 x 0.286 = 0.2559	0.2559 / (3/3.5) = 0.2985
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		4	0.2	4.2	0.019	Av. = 0.1352	0.1352 x 0.019 = 2.575 E-03	2.575E-03/ (4/4.2) = 2.704 E-03

$\therefore \log P = \log (2.704\text{ E-}03/0.2985) = \mathbf{-2.04}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **18** B	67.5	3	0.5	3.5	0.286	Av. =1.0428	1.0428 x 0.286= 0.2982	0.2982 / (3/3.5) = 0.3479
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	0.2	3.2	0.025	Av. =0.1074	0.1074 x 0.025= 2.685 E-03	2.685 E-03/ (3/3.2) = 2.864 E-03

$\therefore \log P = \log (2.864\text{ E-}03/0.3479) = \mathbf{-2.08}$

Tables A11 & A12 Tables to show the calculation of log P for PhC2C3 (9) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **9** A	59.4	3	0.5	3.5	0.286	Av. =0.8630	0.8630 x 0.286= 0.2468	0.2468 / (3/3.5)= 0.2880
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3.5	0.5	4	0.05	Av. =0.06223	0.06223 x 0.05= 3.112 E-03	3.112 E-03 / (3.5/4)= 3.556 E-03

$\therefore \log P = \log (3.556\text{ E-}03/0.2880) = \mathbf{-1.91}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor Related to concentration of analyte in water layer
Expt. **9** B	65.1	3	0.5	3.5	0.286	Av. =0.8346	0.8346 x 0.286 = 0.2387	0.2387 / (3/3.5) = 0.2785
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3.5	0.2	3.7	0.022	Av. =0.1843	0.1843 x 0.022 = 4.055 E-03	4.055 E-03/ (3.5/3.7) = 4.286 E-03

$\therefore \log P = \log (4.286\text{ E-}03/0.2785) = \mathbf{-1.81}$

Tables A13 & A14 Tables to show the calculation of log P for PhC3C4 (19) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **19** A	71.2	3	0.5	5	0.2	Av. =2.4540	2.4540 x 0.2 = 0.4908	0.4908 / (3/5) = 0.8180
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3.5	0.5	5	0.2	Av. =0.05330	0.05330 x 0.2 = 1.066 E-02	1.066 E-02 / (3.5/5) = 1.523 E-02

$\therefore \log P = \log (1.523\text{ E-}02/0.8180) = \mathbf{-1.73}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **19** B	73.7	3	0.5	5	0.2	Av. =1.9849	1.9849 x 0.2 = 0.3970	0.3970 / (3/5) = 0.6616
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3.5	0.5	5	0.2	Av. =0.04932	0.04932 x 0.2 = 9.864 E-03	9.864 E-03 / (3.5/5) = 1.409 E-02

$\therefore \log P = \log (1.409\text{ E-}02/0.6616) = \mathbf{-1.67}$

*Tables A15 & A16 Tables to show the calculation of log P for PhC3C3 (**10**) from octanol/water experiments A and B*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **10** A	62.8	3	0.5	5	0.2	Av. =1.8095	1.8095 x 0.2 = 0.3619	0.3619 / (3/5) = 0.6032
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3.5	0.5	5	0.2	Av. =0.05687	0.05687 x 0.2 = 1.137 E-02	1.137 E-02 / (3.5/5) = 1.625 E-02

$\therefore \log P = \log (1.625\text{ E-}02/0.6032) = \mathbf{-1.57}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **10** B	55.9	1.2	0.2	2	0.2	Av. = 2.1467	2.1467 x 0.2 = 0.4293	0.4293 / (1.2/2) = 0.7156
		Volume of octanol layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1.4	0.2	2	0.2	Av. = 0.06440	0.06440 x 0.2 = 1.288 E-02	0.01288 / (1.4/2) = 1.84 0 E-02

$\therefore \log P = \log (1.840\text{ E-}02/0.7156) = \mathbf{-1.59}$

*Tables A17 & A18 Tables to show the calculation of log P for PhC4C4 (**20**) from octanol/water experiments A and B*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **20** A	59.0	3	1	4	0.5	Av. =0.6258	0.6258 x 0.5 = 0.3129	0.3129 / (3/4) = 0.4172
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	1	4	0.1	Av. =0.1186	0.1186 x 0.1 = 1.186 E-02	1.186 E-02 / (3/4) = 1.581 E-02

$\therefore \log P = \log (1.581\text{ E-}02/0.4172) = \mathbf{-1.42}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **20** B	65.0	3	1	4	0.5	Av. =0.6711	0.6711 x 0.5 = 0.3356	0.3356 / (3/4) = 0.4474
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	1	4	0.1	Av. =0.1363	0.1363 x 0.1 = 1.363 E-02	1.363 E-02/ (3/4) = 1.817 E-02

$\therefore \log P = \log (1.817\text{ E-}02/0.4474) = \mathbf{-1.39}$

Tables A19 & A20 Tables to show the calculation of log P for PhC4C3 (11) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **11** A	68.7	3	1	4	0.5	Av. =0.6555	0.6555 x 0.5 = 0.3278	0.3278 / (3/4) = 0.4370
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	1	4	0.1	Av. =0.2288	0.2288 x 0.1 = 2.288E-02	2.288 E-02/ (3/4) = 3.051 E-02

$\therefore \log P = \log (3.051\ E\text{-}02/0.4370) = \mathbf{-1.16}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 2mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **11** B	67.6	3	1	4	0.5	Av. =0.6878	0.6878 x 0.5 = 0.3439	0.3439 / (3/4) = 0.4585
		Volume of octanol layer (ml)	Volume of 0.4mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		3	1	4	0.1	Av. =0.2251	0.2251 x 0.1 = 2.251 E-02	2.251 E-02/ (3/4) = 3.001 E-02

$\therefore \log P = \log (3.001\ E\text{-}02/0.4585) = \mathbf{-1.18}$

*Tables A21 & A22 Tables to show the calculation of log P for C4PhC3C4 (**21**) from octanol/water experiments C and D*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **21** C	9.8	3	1	5	0.2	Av. =0.5448	0.5448 x 0.2 = 0.1090	0.1090 / (3/5) = 0.1816
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	5	0.2	Av. =0.3543	0.3543 x 0.2 = 7.086 E-02	7.086 E-02/ (1/5) = 0.3543

$\therefore \log P = \log (0.3543/0.1816) = \mathbf{0.29}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **21** D	8.7	3	1	5	0.2	Av. =0.5239	0.5239 x 0.2 = 0.1048	0.1048/ (3/5) = 0.1746
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	5	0.2	Av. =0.3447	0.3447 x 0.2 = 6.894 E-02	6.894 E-02/ (1/5) = 0.3447

$\therefore \log P = \log (0.3447/0.1746) = \mathbf{0.30}$

Tables A23 & A24 Tables to show the calculation of log P for C4PhC3C3 (12) from octanol/water experiments C and D

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **12** C	9.9	2	0.5	2.5	0.2	Av. =0.5746	0.5746 x 0.2 = 0.1149	0.1149 / (2/2.5) = 0.1437
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	5	0.2	Av. =0.4842	0.4842 x 0.2 = 9.684 E-02	9.684 E-02 / (1/5) = 0.4842

$\therefore \log P = \log (0.4842/0.1437) = \mathbf{0.53}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. **12** D	9.0	3	1	5	0.2	Av. =0.4214	0.4214 x 0.2 = 8.428 E-02	8.428 E-02 / (3/5) = 0.1405
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	5	0.2	Av. =0.5214	0.5214 x 0.2 = 0.1043	0.1043 / (1/5) = 0.5214

$\therefore \log P = \log (0.5214/0.1405) = \mathbf{0.57}$

Tables A25 & A26 Tables to show the calculation of log P for C6PhC3C4 (22) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor Related to concentration of analyte in water layer
Expt. 22 A	54.2	3	0.5	5	0.1	Av. =0.3635	0.3635 x 0.1 = 3.635 E-02	3.635 E-02 / (3/5) = 6.058 E-02
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	10	0.1	Av. =2.7942	2.7942 x 0.1 = 0.27942	0.2742 / (1/10) = 2.7942

$\therefore \log P = \log (2.7942/6.058\ E\text{-}02) = \mathbf{1.66}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Expt. 22 B	54.2	3	0.5	5	0.1	Av. =0.4787	0.4787 x 0.1 = 4.787 E-02	4.787 E-02 / (3/5) = 7.978 E-02
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	10	0.1	Av. =3.2505	3.2505 x 0.1 = 0.3251	0.3251 / (1/10) = 3.2505

$\therefore \log P = \log (3.2505/7.978\ E\text{-}02) = \mathbf{1.61}$

*Tables A27 & A28 Tables to show the calculation of log P for C6PhC3C3 (**22**) from octanol/water experiments A and B*

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Experiment **13** A	53.1	3	0.5	5	0.1	Av. =0.2862	0.2862 x 0.1 = 2.862 E-02	2.862 E-02/ (3/5) = 4.770 E-02
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	10	0.1	Av. =2.8294	2.8294 x 0.1 = 0.2829	0.2829 / (1/10) = 2.8294

$\therefore \log P = \log (2.8294/4.770\ E\text{-}02) = \mathbf{1.77}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration of analyte in water layer
Experiment **13** B	65.9	3	0.5	5	0.1	Av. =0.2423	0.2862 x 0.1 = 2.423 E-02	2.423 E-02/ (3/5) = 4.038 E-02
		Volume of octanol layer (ml)	Volume of 1mg/ml Phenol solution (ml)	Total volume of sample (ml)	Concentration of Phenol in sample (mg/ml)	Peak Area Ratio	Factor related to concentration of analyte in sample	Factor related to concentration in octanol layer
		1	1	10	0.1	Av. =4.1611	2.8294 x 0.1 = 0.4161	0.4161 / (1/10) = 4.1611

$\therefore \log P = \log (4.1611/4.038\ E\text{-}02) = \mathbf{2.01}$

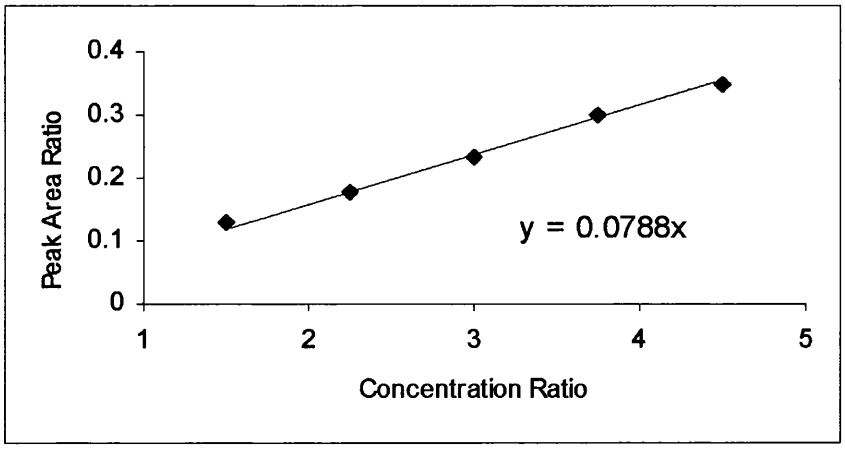


Figure A1 The Relationship between Concentration Ratio (17/Phenol) and Peak Area Ratio (17/Phenol)

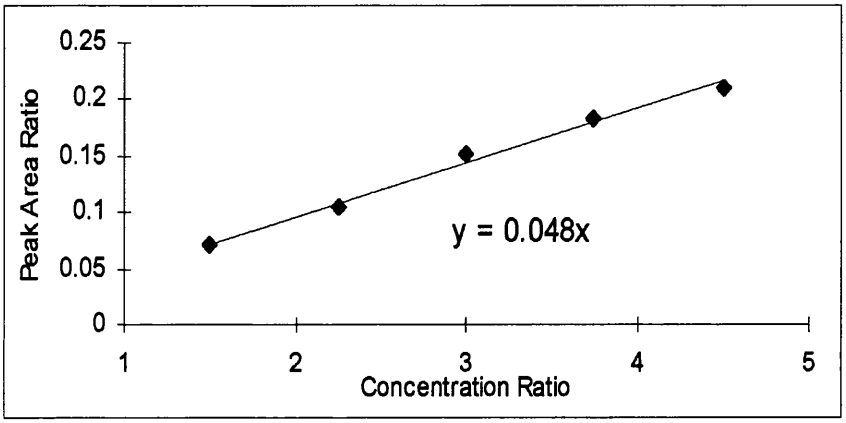


Figure A2 The Relationship between Concentration Ratio (8/Phenol) and Peak Area Ratio (8/Phenol)

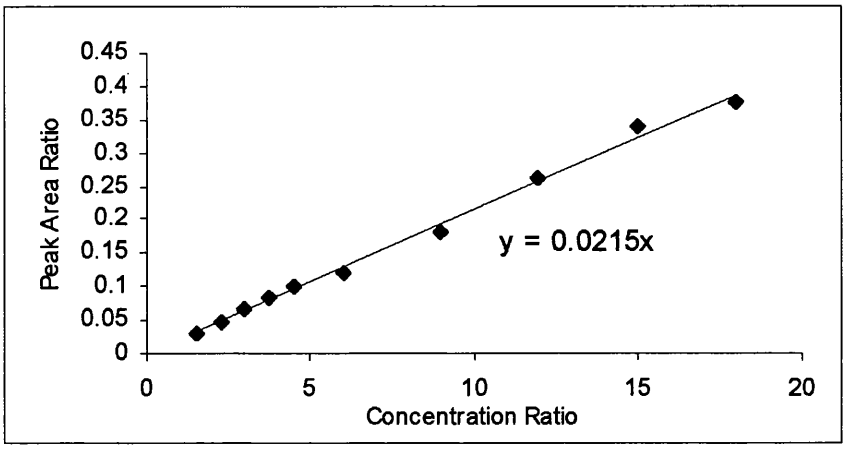


Figure A3 The Relationship between Concentration Ratio (18/Phenol) and Peak Area Ratio (18/Phenol)

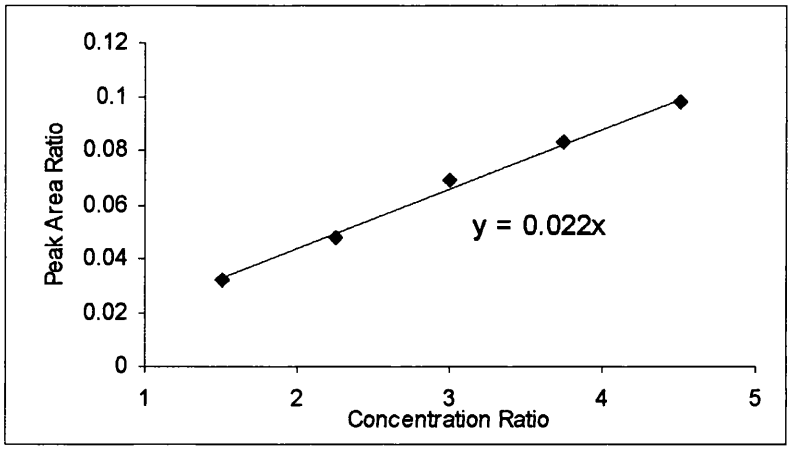


Figure A4 The Relationship between Concentration Ratio (9/Phenol) and Peak Area Ratio (9/Phenol)

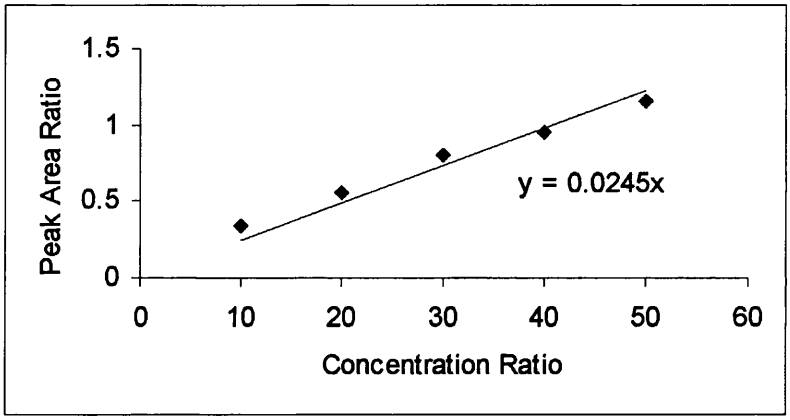


Figure A5 The Relationship between Concentration Ratio (19/Phenol) and Peak Area Ratio (19/Phenol)

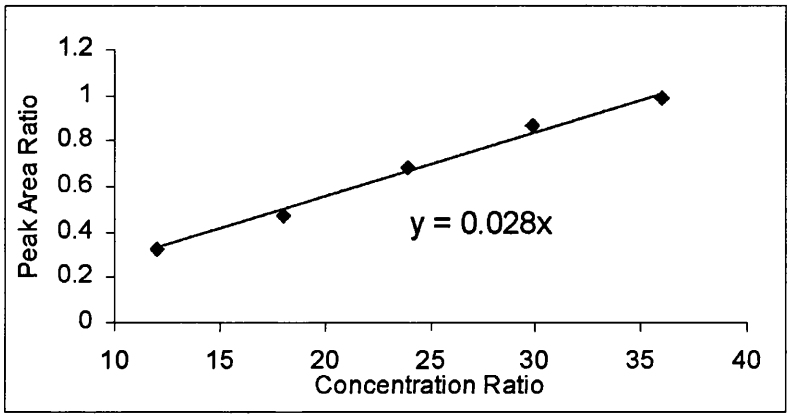


Figure A6 The Relationship between Concentration Ratio (20/Phenol) and Peak Area Ratio (20/Phenol)

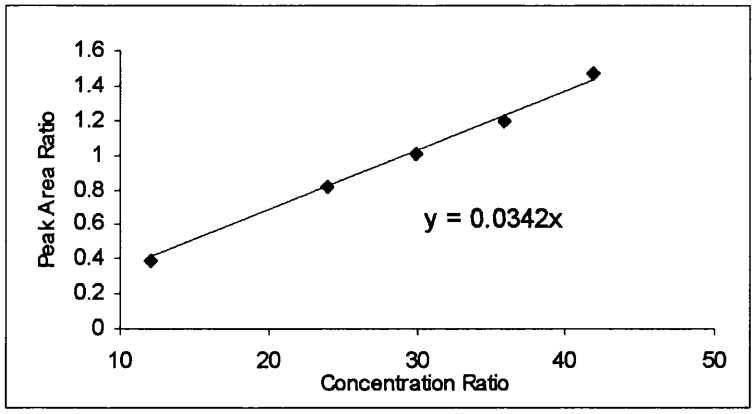


*Figure A7 The Relationship between Concentration Ratio (**11**/Phenol) and Peak Area Ratio (**11**/Phenol)*

Table A29 Table to show the calculation of log P from analysis of the aqueous layer and determination of concentration of solute in octanol by difference

Compound	Expt.	Mass introduced (mg)	Factor related to conc. of solute in H_2O layer	Response factor	Conc. in H_2O layer (mg/ml)	Conc. in H_2O layer (mg/5ml)	∴Conc. in octanol layer (mg/ml)	Log P
17	A	54.9	0.5590	0.0788	7.1	35.5	19.4	-0.26
	B	55.1	0.5536		7.0	35.1	20.0	-0.24
	C	66.1	0.8894		11.3	56.4	9.7	-0.76
	D	67.1	0.8780		11.1	55.7	11.4	-0.69
8	A	54.8	0.5830	0.0480	12.1	**60.7***	---	---
	B	54.8	0.6184		12.9	**64.4***	---	---
	C	65.8	0.5536		11.5	57.7	8.1	-0.85
	D	66.7	0.9312		19.4	**97.0***	---	---
18	A	64.7	0.2985	0.0215	13.9	**69.4***	---	---
	B	67.5	0.3479		16.2	**80.9***	---	---
9	A	59.4	0.2880	0.0220	13.1	**65.5***	---	---
	B	65.1	0.2785		12.7	63.3	1.8	-1.55
19	A	71.2	0.8180	0.0245	33.4	**167***	---	---
	B	73.7	0.6616		27.0	**135***	---	---
20	A	59.0	0.4172	0.0280	14.9	**74.5***	---	---
	B	65.0	0.4474		16.0	**79.9***	---	---
11	A	68.7	0.4370	0.0342	12.8	63.9	4.8	-1.12
	B	67.6	0.4585		13.4	67.0	0.6	-2.04

* concentration greater than that introduced to biphasic system at start of experiment; all entries quoted to a maximum of 4 s.f.

Table A30 Table to show the calculation of log P from analysis of the octanol layer and determination of concentration of solute in the aqueous layer by difference

Compound	Expt.	Mass Introduced (mg)	Factor related to conc. of solute in octanol layer	Response factor	Conc. in octanol layer (mg/ml)	Conc. in octanol layer (mg/5ml)	∴Conc. in H_2O layer (mg/ml)	Log P
17	A	54.9	3.257 E-03	0.0788	41.3 E-03	0.207	54.69	-2.42
	B	55.1	1.953 E-03		24.8 E-03	0.124	54.98	-2.65
	C	66.1	3.694 E-03		46.9 E-03	0.234	65.87	-2.45
	D	67.1	5.371 E-03		68.2 E-03	0.341	66.76	-2.29
8	A	54.8	2.787 E-03	0.0480	58.1 E-03	0.290	54.51	-2.27
	B	54.8	3.380 E-03		70.4 E-03	0.352	54.45	-2.19
	C	65.8	1.953 E-03		40.7 E-03	0.203	65.60	-2.51
	D	66.7	4.370 E-03		91.0 E-03	0.455	66.25	-2.16
18	A	64.7	2.704 E-03	0.0215	0.126	0.629	64.07	-2.01
	B	67.5	2.864 E-03		0.133	0.666	66.83	-2.00
9	A	59.4	3.556 E-03	0.0220	0.162	0.808	58.59	-1.86
	B	65.1	4.286 E-03		0.195	0.974	64.13	-1.82
19	A	71.2	1.523 E-02	0.0245	0.622	3.108	68.09	-1.34
	B	73.7	1.409 E-02		0.575	2.876	70.82	-1.39
20	A	59.0	1.581 E-02	0.0280	0.565	2.823	56.18	-1.30
	B	65.0	1.817 E-02		0.649	3.245	61.76	-1.28
11	A	68.7	3.051 E-02	0.0342	0.892	4.461	64.24	-1.16
	B	67.6	3.001 E-02		0.877	4.387	63.21	-1.16

Tables A31 & A32 Tables to show the calculation of log P for C7C3(4) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 4 A	63.4	1	100	1	2	20	10	Av. = 0.9681	0.9681 x 10 = 9.681	9.681 /(1/100) / (1/20) = 19362
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 0.1162	0.1162 x 10= 1.162	1.162 / (1/20) / (1/20) = 464.8

$\therefore \log P = \log (464.8/19362) =$ **-1.62**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 4 B	68.0	1	100	1	2	20	10	Av. = 2.4298	2.4298 x 10 = 24.298	24.298/ (1/100) /(1/20) = 48596
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 0.05767	0.05767 x 10 = 0.5767	0.5767/ (1/20) / (1/20) = 230.68

$\therefore \log P = \log (230.68/48596) =$ **-2.32**

Tables A33 & A34 Tables to show the calculation of log P for C7C4(15) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **15** A	63.2	1	100	1	2	20	10	Av. = 0.6479	0.6479 x 10 = 6.4785	6.4785 /(1/100) / (1/20) = 12957
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	—	—	2	20	10	Av. = 2.5565	2.5565 x 10= 25.5645	25.5645 / (1/20) = 511.29

$\therefore \log P = \log (511.29/12957) =$ **-1.40**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **15** B	58.2	1	100	1	2	20	10	Av. = 0.5755	0.5755 x 10 = 5.7550	5.7550/ (1/100) /(1/20) = 11510
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	—	—	2	20	10	Av. = 5.0377	5.0377 x 10 = 50.3770	50.3770/ (1/20) = 1007.54

$\therefore \log P = \log (1007.54/11510) =$ **-1.06**

Tables A35 & A36 Tables to show the calculation of log P for C6C3(3) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **3** A	55.0	1	100	1	2	20	10	Av. = 0.6846	0.6846 x 10 = 6.846	6.846 /(1/100) / (1/20) = 13692
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 0.2428	0.2428 x 10= 2.428	2.428 / (1/20) /(1/20) = 971.2

$\therefore \log P = \log (971.2/13692) =$ **-1.15**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **3** B	61.1	1	100	1	2	20	10	Av. = 0.7195	0.7195 x 10 = 7.195	7.195/ (1/100) /(1/20) = 14390
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 0.1871	0.1871 x 10 = 1.871	1.871/ (1/20) / (1/20)= 7484

$\therefore \log P = \log (7484/14390) =$ **-1.28**

Tables A37 & A38 Tables to show the calculation of log P for C6C4(14) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **14** A	67.1	1	100	1	2	20	10	Av. = 1.1045	1.1045 x 10 = 11.045	11.045 /(1/100) / (1/20) = 22090
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	1	10	10	Av. = 0.8261	0.8261 x 10= 8.261	8.261 / (1/20) /(1/10) = 1652.2

$\therefore \log P = \log (1652.2/22090) =$ **-1.13**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **14** B	63.5	1	100	1	2	20	10	Av. = 0.9626	0.9626 x 10 = 9.626	9.626/ (1/100) /(1/20) = 19252
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	1	10	10	Av. = 0.9039	0.9039 x 10 = 9.039	9.039/ (1/20) / (1/10) = 1807.8

$\therefore \log P = \log (1807.8/19252) =$ **-1.03**

Tables A39 & A40 Tables to show the calculation of log P for C8C3(5) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 5 A	61.1	1	100	1	2	20	10	Av. = 0.3511	0.3511 x 10 = 3.511	3.511 / (1/100) / (1/20) = 7022
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	2	20	10	Av. = 13.622	13.622 x 10= 136.22	136.22 / (1/20) = 2724.4

$\therefore \log P = \log (2724.4/7022) =$ **-0.41**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 5 B	58.5	1	100	1	2	20	10	Av. = 0.3694	0.3694 x 10 = 3.694	3.964/ (1/100) /(1/20) = 7388
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	2	20	10	Av. = 10.8198	10.8198 x 10 = 108.198	108.198 / (1/20) = 2163.96

$\therefore \log P = \log (2163.96/7388) =$ **-0.53**

Tables A41 & A42 Tables to show the calculation of log P for C8C4(16) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **16** A	58.2	1	100	1	2	20	10	Av. = 0.5129	0.5129 x 10 = 5.1285	5.1285 / (1/100) / (1/20) = 10257
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 1.0540	1.0540 x 10= 10.5395	10.5395 / (1/20) = 4215.8

$\therefore \log P = \log (4215.8/10257) =$ **-0.39**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **16** B	58.1	1	100	1	2	20	10	Av. = 0.5128	0.5128 x 10 = 5.1275	5.1275 / (1/100) /(1/20) = 10255
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 1.2010	1.2010 x 10 = 12.0095	12.0095 / (1/20) / (1/20) = 4803.8

$\therefore \log P = \log (4803.8/10255) =$ **-0.33**

Tables A43 & A44 Tables to show the calculation of log P for C8C2(1) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **1** A	99.8	1	100	1	2	20	10	Av. = 0.4341	0.4341 x 10 = 4.341	4.341 / (1/100) / (1/20) = 8682
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 1.2994	1.2994 x 10= 12.994	12.994 / (1/20) / (1/20) = 5197.6

$\therefore \log P = \log (5197.6/8682) =$ **-0.23**

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. **1** B	107.6	1	100	1	2	20	10	Av. = 0.3440	0.3440 x 10 = 3.440	3.440 / (1/100) /(1/20) = 6880
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	20	1	2	20	10	Av. = 2.2134	2.2134 x 10 = 22.134	22.134 / (1/20) / (1/20) = 8853.6

$\therefore \log P = \log (8853.6/6880) =$ **0.11**

Tables A45 & A46 Tables to show the calculation of log P for C10C3(6) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C12C3 solution (ml)	Total volume of sample (ml)	Conc. of C12C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 6 C	5.9	1	---	---	0.5	5	10	Av. = 4.9342	4.9342 x 10 = 49.342	49.342 / (1/5) = 246.71
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C12C3 solution (ml)	Total volume of sample (ml)	Conc. of C12C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	0.5	5	10	Av. = 17.5032	17.5032 x 10= 175.032	175.032 / (1/5) = 875.16

$\therefore \log P = \log (875.16/246.71) = \mathbf{0.55}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C12C3 solution (ml)	Total volume of sample (ml)	Conc. of C12C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 6 D	6.4	1	---	---	0.5	5	10	Av. = 4.9222	4.9222 x 10 = 49.222	49.222 / (1/5) = 246.112
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C12C3 solution (ml)	Total volume of sample (ml)	Conc. of C12C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	0.5	5	10	Av. = 19.1603	19.1603 x 10 = 191.603	191.603 / (1/5) = 958.015

$\therefore \log P = \log (958.015/246.112) = \mathbf{0.59}$

Tables A47 & A48 Tables to show the calculation of log P for C12C3(7) from octanol/water experiments A and B

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 7 A	8.8	1	---	---	0.5	5	10	Av. = 0.1868	0.1868 x 10 = 1.868	1.868 / (1/5) = 9.340
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	0.5	5	10	Av. = 11.1190	11.1190 x 10= 111.190	111.190 / (1/5) = 555.95

$\therefore \log P = \log (555.95/9.340) = \mathbf{1.77}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 7 B	9.5	1	---	---	0.5	5	10	Av. = 0.2704	0.2704 x 10 = 2.704	2.704 / (1/5) = 13.520
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	---	---	0.5	5	10	Av. = 8.9493	8.9493 x 10 = 89.493	89.493 / (1/5) = 447.465

$\therefore \log P = \log (447.465/13.520) = \mathbf{1.52}$

Tables A49 & A50 Tables to show the calculation of log P for C12C2(2) from octanol/water experiments C and D

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 2 C	8.4	1	—	—	0.5	5	10	Av. = 0.04921	0.04921 x 10 = 0.4921	0.4921 / (1/5) = 2.4603
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	—	—	0.5	5	10	Av. = 3.0572	3.0572 x 10= 30.572	30.572 / (1/5) = 152.86

$\therefore \log P = \log (152.86/2.4603) = \mathbf{1.79}$

	Mass of analyte added to octanol/water system (mg)	Volume of water layer (ml)	Dilution factor	Volume of diluted water layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (μg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. of analyte in water layer
Expt. 2 D	7.7	1	—	—	0.5	5	10	Av. = 0.05455	0.05455 x 10 = 0.5455	0.5455 / (1/5) = 2.7275
		Volume of octanol layer (ml)	Dilution factor	Volume of diluted octanol layer (ml)	Volume of 0.1mg/ml C10C3 solution (ml)	Total volume of sample (ml)	Conc. of C10C3 in sample (mg/ml)	Peak Area Ratio	Factor related to conc. of analyte in sample	Factor related to conc. in octanol layer
		1	—	—	0.5	5	10	Av. = 3.2671	3.2671 x 10 = 32.671	32.671 / (1/5) = 163.355

$\therefore \log P = \log (163.355/2.7275) = \mathbf{1.78}$

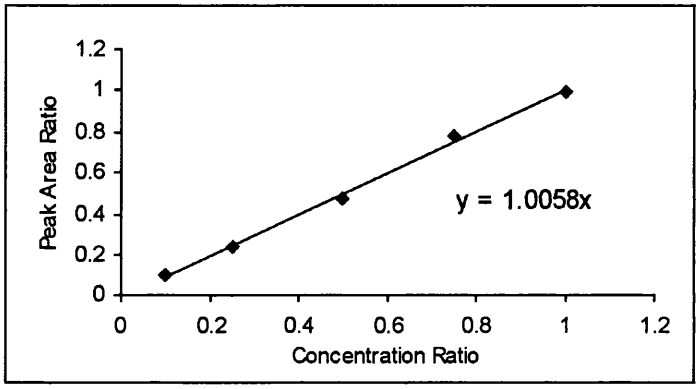


Figure A8 The Relationship between Concentration Ratio (4/C10C3) and Peak Area Ratio (4/C10C3)

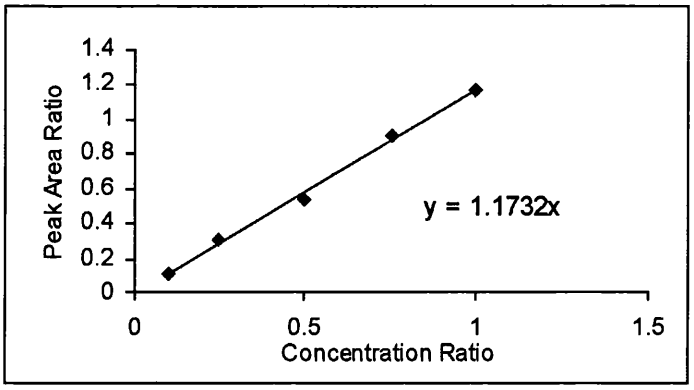


Figure A9 The Relationship between Concentration Ratio (15/C10C3) and Peak Area Ratio (15/C10C3)

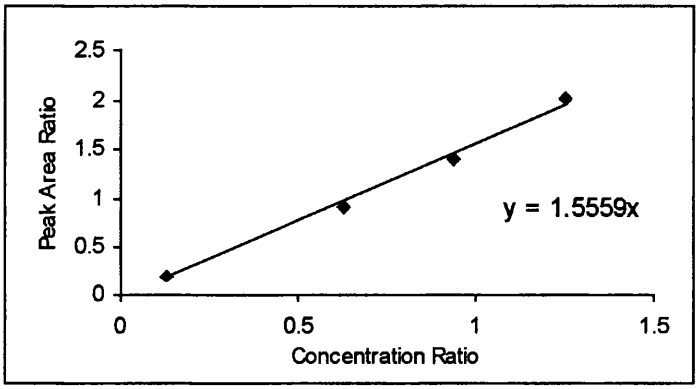


Figure A10 The Relationship between Concentration Ratio (3/C10C3) and Peak Area Ratio (3/C10C3)

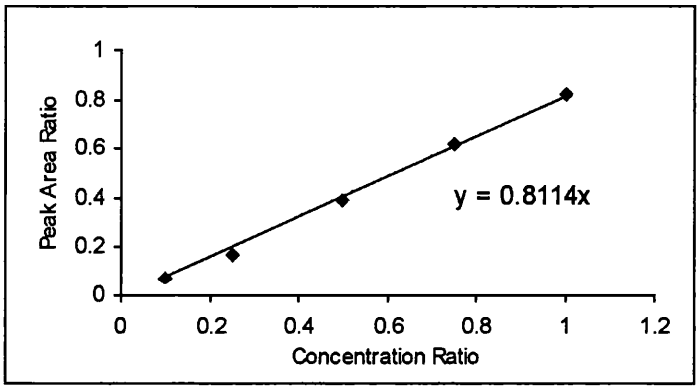


Figure A11 The Relationship between Concentration Ratio (14/C10C3) and Peak Area Ratio (14/C10C3)

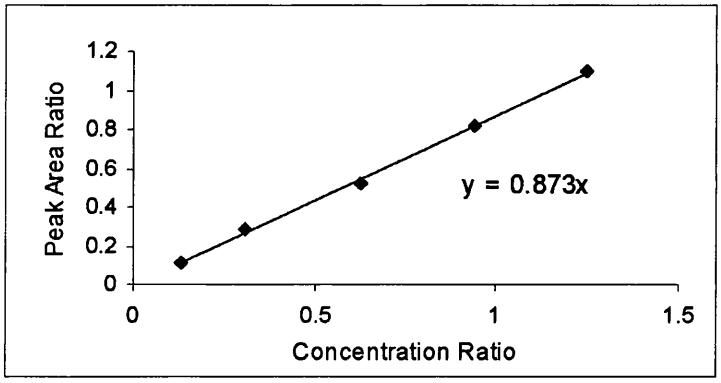


Figure A12 The Relationship between Concentration Ratio (5/C10C3) and Peak Area Ratio (5/C10C3)

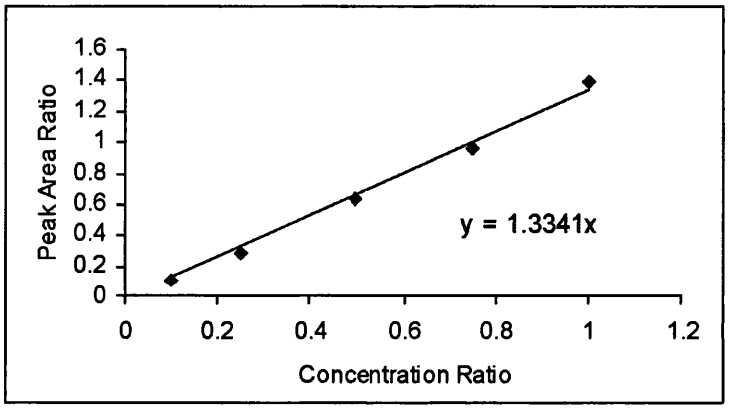


Figure A13 The Relationship between Concentration Ratio (16/C10C3) and Peak Area Ratio (16/C10C3)

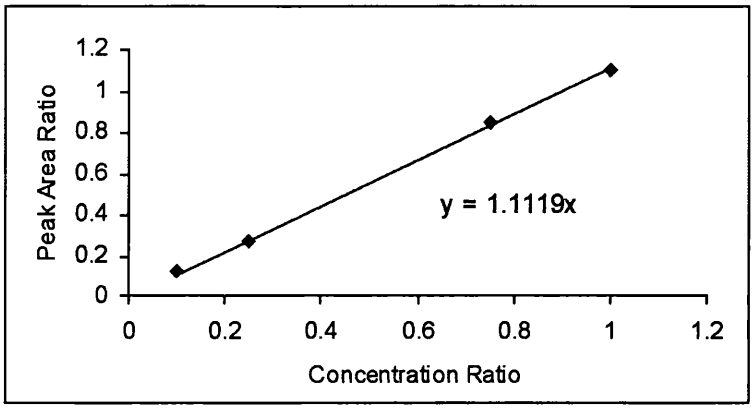


Figure A14 The Relationship between Concentration Ratio (6/C12C3) and Peak Area Ratio (6/C12C3)

Table A51 Table to show the calculation of log P from analysis of the aqueous layer and determination of concentration of solute in octanol by difference

Compound	Expt.	Mass introduced x 10^3(μg)	Factor related to conc. of solute in H_2O layer	Response factor	Conc. in H_2O layer (μg/ml)	Conc. in H_2O layer (μg/5ml)	∴ Conc. in octanol layer (μg/5ml)	Log P
4	A	63.4	19360	1.006	19250	**96250***	---	---
	B	68.0	48600		4832	**241600***	---	---
15	A	63.2	12960	1.173	11040	55220	7979	-0.84
	B	58.2	11510		9811	49050	9146	-0.73
3	A	55.0	13692	1.556	8800	44000	11000	-0.60
	B	61.1	14690		9441	47200	13900	-0.53
14	A	67.1	22090	0.8114	27230	**136100***	---	---
	B	63.5	19250		23730	**118600***	---	---
5	A	61.1	7022	0.8730	8044	40220	20880	-0.28
	B	58.5	7388		8463	42310	16190	-0.42
16	A	58.2	10260	1.334	7688	38440	19760	-0.29
	B	58.1	10260		7687	38430	19670	-0.29
6	C	5.9	246.7	1.112	222.0	1109	4791	0.64
	D	6.4	246.1		221.0	1108	5292	0.68

*concentration greater than that introduced to biphasic system at start of experiment; All entries quoted to a maximum of 4 s.f.

Table A52 Table to show the calculation of log P from analysis of the octanol layer and determination of concentration of solute in the aqueous layer by difference

Compound	Expt.	Mass introduced x 10^3(μg)	Factor related to conc. of solute in octanol layer	Response factor	Conc. in octanol layer (μg/ml)	Conc. in octanol layer (μg/5ml)	∴ Conc. in H_2O layer (μg/5ml)	Log P
4	A	63.4	464.8	1.006	462.0	2311	61090	-1.42
	B	68.0	230.7		229.0	1147	66850	-1.77
15	A	63.2	511.3	1.173	436.0	2179	61020	-1.45
	B	58.2	1008		859.0	4294	53910	-1.10
3	A	55.0	971.2	1.556	624.0	3121	51880	-1.22
	B	61.1	748.4		481.0	2405	58700	-1.39
14	A	67.1	1652	0.8114	2036	10180	56920	-0.75
	B	63.5	1808		2228	11140	52360	-0.67
5	A	61.1	2724	0.8730	3121	15600	45500	-0.46
	B	58.5	2164		2479	12390	46110	-0.57
16	A	58.2	4216	1.334	3160	15800	42400	-0.43
	B	58.1	4804		3601	18000	40100	-0.35
6	C	5.9	875.2	1.112	787.0	3935	1965	0.30
	D	6.4	958.0		862.0	4308	2092	0.31

Appendix B

Chapter 4(I): Acute Aquatic toxicity Testing

Table B1 *Table to how raw data for experiment C6C3(3) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia* Exposed	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total		24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	320	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	560	0	0	0	0	0	1	1	2	1	5	4	4	3	4	15	20	0	25
13 - 16	1000	0	0	1	1	2	3	3	2	1	11	2	2	2	3	9	20	10	55
17 - 20	1800	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100
21 - 24	3200	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Table B2 *Table to show raw data for experiment C6C3(3) B*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia* Exposed	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total		24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	100	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	180	0	0	0	0	0	1	0	0	0	1	4	5	5	5	19	20	0	5
13 - 16	320	0	0	0	0	0	0	0	1	0	1	5	5	4	5	19	20	0	5
17 - 20	560	0	2	2	0	4	3	1	3	2	13	2	2	0	3	7	20	20	65
21 - 24	1000	4	4	4	4	16	1	1	1	1	20	0	0	0	0	0	20	80	100

Table B3 *Table to how raw data for experiment C6C4(14) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	3	4	Total	Observed Number Immobile at 48 h Replicate 1	2	3	4	Total	Observed Number Mobile at 48 h Replicate 1	2	3	4	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	560	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	1000	0	0	0	0	0	1	0	1	0	2	4	5	4	5	18	20	0	10
13 - 16	1800	0	0	0	0	0	0	1	1	1	3	5	4	4	4	17	20	0	15
17 - 20	3200	0	0	0	0	0	2	1	1	0	4	3	4	4	5	16	20	0	20
21 - 24	5600	1	0	0	0	1	1	3	3	3	11	3	2	2	2	9	20	5	55

Table B4 *Table to show raw data for experiment C7C3(4) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	3	4	Total	Observed Number Immobile at 48 h Replicate 1	2	3	4	Total	Observed Number Mobile at 48 h Replicate 1	2	3	4	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	1800	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	3200	0	0	0	0	0	0	1	1	0	2	5	4	4	5	18	20	0	10
13 - 16	5600	1	1	1	1	4	3	4	2	3	16	1	0	2	1	4	20	20	80
17 - 20	10000	1	0	2	0	3	3	1	3	2	18	0	1	0	0	2	20	15	90
21 - 24	18000	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Table B5 *Table to how raw data for experiment C7C4(15) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia*	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total	Exposed	24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	1000	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	1800	0	0	0	0	0	1	0	0	0	1	4	5	5	5	19	20	0	5
13 - 16	3200	0	0	0	0	0	0	1	1	0	2	5	4	4	5	18	20	0	10
17 - 20	5600	0	0	0	0	0	2	1	1	1	5	3	4	4	4	15	20	0	25
21 - 24	10000	5	4	4	4	17	0	1	1	0	19	0	0	0	1	1	20	85	95

Table B6 *Table to show raw data for experiment C7C4(15) B*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia*	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total	Exposed	24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 – 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	560	0	0	0	1	1	0	0	0	0	0	5	5	5	4	19	20	5	0
9 – 12	1800	0	0	0	0	0	1	0	0	1	2	4	5	5	4	18	20	0	10
13 - 16	5600	0	0	0	0	0	2	0	1	1	4	3	5	4	4	16	20	0	20
17 - 20	10000	0	1	0	1	2	0	2	3	2	9	5	2	2	2	11	20	10	45
21 - 24	18000	2	1	2	1	6	3	4	3	4	14	0	0	0	0	0	20	30	70

Table B7 *Table to how raw data for experiment C8C2(1) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	3	4	Total	Observed Number Immobile at 48 h Replicate 1	2	3	4	Total	Observed Number Mobile at 48 h Replicate 1	2	3	4	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 – 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	320	0	0	0	0	0	0	1	0	0	1	5	4	5	5	19	20	0	5
9 – 12	560	1	0	0	0	1	1	1	0	0	3	3	4	5	5	17	20	5	15
13 – 16	1000	1	1	0	0	2	2	2	2	2	10	2	2	3	3	10	20	10	50
17 – 20	1800	1	2	3	4	10	2	3	2	0	17	3	4	4	4	3	20	50	85
21 – 24	3200	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Table B8 *Table to show raw data for experiment C8C3(5) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	3	4	Total	Observed Number Immobile at 48 h Replicate 1	2	3	4	Total	Observed Number Mobile at 48 h Replicate 1	2	3	4	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	560	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	1000	0	0	0	0	0	1	0	0	1	0	5	5	5	5	20	20	0	0
13 - 16	1800	0	0	1	0	1	0	0	1	0	2	5	5	3	5	18	20	5	10
17 - 20	3200	0	1	1	0	2	1	1	2	1	7	5	2	2	2	13	20	10	35
21 - 24	5600	1	1	2	2	6	4	4	3	3	20	0	0	0	0	0	20	30	100

Table B9 *Table to how raw data for experiment C8C3(5) B*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	3	4	Total	Observed Number Immobile at 48 h Replicate 1	2	3	4	Total	Observed Number Mobile at 48 h Replicate 1	2	3	4	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 – 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	560	0	0	0	1*	0	0	0	0	0	0	5	5	5	4	19	19	0	0
9 – 12	1000	0	0	0	0	1	0	0	0	0	0	5	5	5	5	20	20	5	0
13 – 16	1800	0	0	0	0	2	0	0	0	0	0	5	5	5	5	20	20	10	0
17 – 20	3200	1	0	1	0	2	1	2	4	3	12	3	3	0	2	8	20	10	60
21 – 24	5600	4	3	3	4	14	1	2	2	1	20	0	0	0	0	0	20	70	100

*crushed accidentally by test operator

Table B10 *Table to show raw data for experiment C8C4(16) RF*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate 1	2	Total	Observed Number Immobile at 48 h Replicate 1	2	Total	Observed Number Mobile at 48 h Replicate 1	2	Total	No. of *Daphnia* Exposed	Percentage Immobile 24 h	48 h
1 - 2	0.0	0	0	0	0	0	0	5	5	10	10	0	0
3 - 4	100	0	0	0	0	0	0	5	5	10	10	0	0
5 - 6	1000	0	0	0	1	0	1	4	5	9	10	0	10
7 - 8	10000	5	5	10	0	0	10	0	0	0	10	100	100

Table B11 *Table to how raw data for experiment C10C3(6) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 – 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	100	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 – 12	180	0	0	0	0	0	1	0	0	0	1	4	5	5	5	19	20	0	5
13 – 16	320	1	1	0	1	3	1	1	2	1	8	3	3	3	3	12	20	15	40
17 – 20	560	0	1	1	1	3	2	2	2	2	11	3	2	2	2	9	20	15	55
21 – 24	1000	2	3	3	1	9	3	2	2	4	20	0	0	0	0	0	20	45	100

Table B12 *Table to show raw data for experiment C12C2(2) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	10	0	2	0	0	2	0	0	0	0	2	5	3	5	5	18	20	10	10
9 - 12	18	2	1	2	1	6	0	0	0	1	7	3	4	3	3	13	20	30	35
13 - 16	32	3	4	4	4	15	2	0	1	1	19	0	1	0	0	1	20	75	95
17 - 20	56	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100
21 - 24	100	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Table B13 *Table to how raw data for experiment C12C3(7) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 – 2	0.0	0	0	--	--	0	0	0	--	--	0	5	5	--	--	10	10	0	0
3 – 6	32	0	0	0	0	0	0	0	1	0	1	5	5	4	5	19	20	0	5
7– 9	56	0	0	0	--	0	0	0	0	--	0	5	5	5	--	15	15	0	0
10 – 12	10	0	0	0	--	0	0	0	0	--	0	5	5	5	--	15	15	0	0
13 – 15	18	0	0	0	--	0	2	1	1	--	4	3	4	4	--	11	15	0	27
16 – 18	32	0	0	0	--	0	3	3	2	--	8	2	2	3	--	7	15	0	53

Table B14 *Table to show raw data for experiment PhC1C3(8) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	5600	0	0	0	0	0	0	0	1	0	1	5	5	4	5	19	20	0	5
9 - 12	10000	0	0	0	0	0	4	3	3	3	13	1	2	2	2	7	20	0	65
13 - 16	18000	1	0	0	0	1	5	5	5	5	20	0	0	0	0	0	20	5	100
17 - 20	32000	2	2	0	1	5	5	5	5	5	20	0	0	0	0	0	20	25	100
21 - 24	56000	5	5	5	5	20	5	5	5	5	20	0	0	0	0	0	20	100	100

Table B15 *Table to how raw data for experiment PhC2C3(9) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 – 2	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	3200	0	0	0	0	0	1	0	0	2	3	4	5	5	3	17	20	0	15
9 – 12	5600	0	2	2	0	4	5	3	3	5	20	0	0	0	0	20	20	20	100
13 – 16	10000	3	5	4	4	16	2	0	1	1	20	0	0	0	0	20	20	80	100
17 – 20	18000	5	5	5	5	20	0	0	0	0	20	0	0	0	0	20	20	100	100
21 – 24	32000	5	5	5	5	20	0	0	0	0	20	0	0	0	0	20	20	100	100

Table B16 *Table to show raw data for experiment PhC2C4(18) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate					Observed Number Immobile at 48 h Replicate					Observed Number Mobile at 48 h Replicate					No. of *Daphnia*	Percentage Immobile	
		1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	Total	Exposed	24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	1800	0	0	1*	0	1	0	1	0	0	2	5	4	4	5	18	20	5	10
9 - 12	3200	0	0	0	1	1	0	0	1	0	2	5	5	4	4	18	20	5	10
13 - 16	5600	0	0	0	0	0	1	3	3	0	7	4	2	2	5	13	20	0	35
17 - 20	10000	0	0	0	1	1	3	4	4	4	16	2	1	1	0	4	20	5	80
21 - 24	18000	3	3	2	3	11	2	2	3	2	20	0	0	0	0	0	20	55	100

* all other daphnids *very* active in comparison

Table B17 *Table to how raw data for experiment PhC4C3(11) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia*	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total	Exposed	24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 – 2	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	560	0	0	0	0	0	0	0	0	0	0	5	5	5	3	20	20	0	0
9 – 12	1000	0	1	0	1	2	0	1	1	1	5	5	3	4	3	15	20	10	25
13 – 16	1800	3	4	3	3	13	1	1	2	2	19	1	0	0	0	1	20	65	95
17 – 20	3200	4	5	4	4	17	1	0	1	1	20	0	0	0	0	0	20	85	100
21 – 24	5600	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Comment: At 24 h and 48 h, it was noticed that yellow/orange media had crystallised out of solution.

Table B18 *Table to show raw data for experiment PhC4C4(20) A*

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h					Observed Number Immobile at 48 h					Observed Number Mobile at 48 h					No. of *Daphnia*	Percentage Immobile	
		Replicate				Total	Replicate				Total	Replicate				Total	Exposed	24 h	48 h
		1	2	3	4		1	2	3	4		1	2	3	4				
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	560	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	1000	0	0	0	0	0	2	1	2	0	5	3	4	3	5	15	20	0	25
13 - 16	1800	5	5	4	4	18	0	0	1	1	20	0	0	0	0	0	20	90	100
17 - 20	3200	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100
21 - 24	5600	5	5	5	5	20	0	0	0	0	20	0	0	0	0	0	20	100	100

Comment: At 24 h and 48 h, it was noticed that yellow/orange media had crystallised out of solution.

Table B19. Table to how raw data for experiment C4PhC3C3(12) A

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate				Total	Observed Number Immobile at 48 h Replicate				Total	Observed Number Mobile at 48 h Replicate				Total	No. of *Daphnia* Exposed	Percentage Immobile	
		1	2	3	4		1	2	3	4		1	2	3	4			24 h	48 h
1 – 2	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	18	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 – 12	32	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	25
13 – 16	56	1	0	0	0	1	1	3	2	2	9	3	2	3	3	11	20	5	45
17 – 20	100	1	2	0	1	4	4	2	1	4	15	0	1	4	0	5	20	20	75
21 – 24	180	3	4	4	3	14	2	1	1	2	20	0	0	0	0	0	20	70	100

Table B20. Table to show raw data for experiment C4PhC3C4(21) A

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate				Total	Observed Number Immobile at 48 h Replicate				Total	Observed Number Mobile at 48 h Replicate				Total	No. of *Daphnia* Exposed	Percentage Immobile	
		1	2	3	4		1	2	3	4		1	2	3	4			24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	150	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9 - 12	220	0	0	0	1	1	0	0	0	0	1	5	5	5	4	19	20	5	5
13 - 16	340	0	0	0	0	0	0	0	1	0	1	5	5	4	5	19	20	0	65
17 - 20	500	2	1	1	1	5	1	2	2	3	13	2	2	2	1	7	20	25	100

Comment: Test concentration of 50 mg/l was abandoned due to shortage of stock solution.

Table B21. Table to how raw data for experiment C6PhC3C3(13) A

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate				Total	Observed Number Immobile at 48 h Replicate				Total	Observed Number Mobile at 48 h Replicate				Total	No. of *Daphnia* Exposed	Percentage Immobile	
		1	2	3	4		1	2	3	4		1	2	3	4			24 h	48 h
1 – 2	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 – 8	18	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
9– 12	32	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
13 – 16	56	0	0	0	0	0	3	4	1	2	10	2	1	4	3	10	20	0	50
17 – 20	100	3	3	2	1	9	2	2	3	4	20	0	0	0	0	0	20	45	100
21 – 24	180	4	3	5	4	16	1	2	0	1	20	0	0	0	0	0	20	80	100

Table B22. Table to show raw data for experiment C6PhC3C4(22) A

Test Vessel	Conc. (mg/l)	Observed Number Immobile at 24 h Replicate				Total	Observed Number Immobile at 48 h Replicate				Total	Observed Number Mobile at 48 h Replicate				Total	No. of *Daphnia* Exposed	Percentage Immobile	
		1	2	3	4		1	2	3	4		1	2	3	4			24 h	48 h
1 - 4	0.0	0	0	0	0	0	0	0	0	0	0	5	5	5	5	20	20	0	0
5 - 8	32	0	0	0	1	1	0	0	0	0	1	5	5	5	4	19	20	5	5
9 - 12	56	0	0	0	0	0	0	1	0	0	1	5	4	5	5	19	20	0	5
13 - 16	100	0	0	0	0	0	0	1	1	2	4	5	4	4	3	16	20	0	20
17 - 20	180	0	3	1	2	6	3	1	2	1	13	2	1	2	2	7	20	30	65
21 - 24	320	4	4	3	4	15	1	1	2	1	20	0	0	0	0	0	20	75	100

Appendix C

Chapter 6: IAM Chromatography and Comparison of Chromatographic Methods for Estimating log P

Table C1 Measurement of Log k'_{IAM}

Compound	Av. t_r	Av. t_o	k'_{IAM}	log k'_{IAM}
3	2.310	0.950	1.43	0.156
5	6.420	0.950	5.76	0.760
6	41.17	0.950	42.3	1.63
8	1.941	0.969	1.00	1.34×10^{-3}
9	2.267	0.963	1.35	0.132
10	3.132	0.951	2.29	0.562
11	2.801	0.951	1.95	0.289
13	---	---	---	2.04[a]
15	3.970	0.950	3.18	0.502
16	7.750	0.950	7.16	0.855
17	2.120	0.939	1.26	9.96×10^{-2}
18	2.506	0.961	1.61	0.206
19	5.076	0.956	4.31	0.634
20	4.421	0.951	3.65	0.360
22	---	---	---	2.03[a]

[a] Log k'_{IAM} for compounds **13** and **22** was determined by extrapolation (see Figures C1 and C2).

Table C2 Comparison of Log P and Log k'_{IAM} with Aquatic Toxicity

Compound	Log $P_{expt.}$	Log k'_{IAM}	Log ($1/EC_{50}$)
1	0.29	---	2.44
2	1.79	---	4.20
3	-1.22	0.156	2.71
4	-1.97	---	1.74
5	-0.47	0.760	1.97
6	0.57	1.63	2.87
7	1.65	---	3.99
8	-2.27	1.34×10^{-3}	1.46
9	-1.87	0.132	1.84
10	-1.57	0.562	---
11	-1.17	0.289	---
12	0.55	---	3.70
13	1.89	2.04	3.82
14	-1.08	---	1.71
15	-1.23	0.502	1.62
16	-0.36	0.855	2.07
17	-2.32	9.96×10^{-2}	---
18	-2.06	0.206	1.64
19	-1.70	0.634	---
20	-1.41	0.360	---
21	0.30	---	2.89
22	1.64	2.03	3.46

Table C3 Relationship of Log P and Log k'_{IAM} with Log k'_{C8}

Compound	log $P_{expt.}$	log k'_{IAM}	Av. t_r	Av. t_o	k'_{C8}	log k'_{C8}
8	-2.27	$1.34x10^{-3}$	2.562	1.924	0.33	-0.48
9	-1.87	0.132	5.820	2.106	1.76	0.25
10	-1.57	0.562	20.314	2.050	8.91	0.95
11	-1.17	0.289	9.204	2.033	3.53	0.55
17	-2.32	$9.96x10^{-2}$	3.150	1.993	0.58	-0.24
18	-2.06	0.206	6.543	2.136	2.06	0.31
19	-1.70	0.634	26.251	2.133	11.3	1.05
20	-1.41	0.360	9.670	2.068	3.68	0.57

Table C4 Relationship of Log P and Log k'_{IAM} with Log k'_{C18}

Compound	Log $P_{expt.}$	log k'_{IAM}	Av. t_r	Av. t_o	k'_{C18}	log k'_{C18}
8	-2.27	$1.34x10^{-3}$	2.196	1.711	0.283	-0.548
9	-1.87	0.132	2.765	1.715	0.612	-0.213
10	-1.57	0.562	3.859	1.717	1.25	0.0960
11	-1.17	0.289	6.773	1.711	2.96	0.471
12	0.55	---	31.57	1.635	18.3	1.26
13	1.89	2.04[a]	---	---	---	2.04[a]
17	-2.32	$9.96x10^{-2}$	2.213	1.711	0.293	-0.533
18	-2.06	0.206	2.897	1.715	0.689	-0.162
19	-1.70	0.634	3.976	1.717	1.32	0.119
20	-1.41	0.360	7.110	1.711	3.16	0.499
21	0.30	---	32.84	1.673	18.6	1.27
22	1.64	2.03[a]	---	---	---	2.06[a]

[a] Log k'_{IAM} and log k'_{C18} for compounds **13** and **22** were determined by extrapolation (see Figures C1- C4).

Table C5 Relationship between Log P and Log $k'_{OC/C8}$ @ 0.5 ml/min

Compound	log P	Av. t_r	Av. t_o	$k'_{OC/C8}$	log $k'_{OC/C8}$
8	-2.27	1.935	1.653	0.171	-0.768
9	-1.87	2.306	1.653	0.395	-0.403
10	-1.57	2.651	1.653	0.604	-0.219
11	-1.17	4.362	1.653	1.64	-0.215
12	0.55	36.18	1.653	20.9[a]	1.32
17	-2.32	1.910	1.653	0.155	-0.808
18	-2.06	2.268	1.653	0.372	-0.429
19	-1.70	2.595	1.653	0.570	-0.244
20	-1.41	4.195	1.653	1.54	0.187
21	0.30	32.97	1.653	18.9[a]	1.28

[a] t_r for compounds **12** and **21** at 0.5 ml/min was determined by extrapolation (see figures C5 and C6).

Table E6. Relationship between Log P and Log $k'_{OC/C8}$ @ 3.5 ml/min

Compound	log P	Av. t_r	Av. t_o	$k'_{OC/C8}$	log $k'_{OC/C8}$
12	0.55	11.90	0.365	31.6	1.50
13	1.89	53.74	0.365	146	2.17
21	0.30	10.18	0.365	26.9	1.43
22	1.64	42.77	0.365	116	2.07

Table C7 Relationship between Log P and Log P_{CCC} (V_e = 27.5, V_s = 22.4, 1 ml/min)

Compound	t_r (secs)	V_R	Log P	Log P_{CCC}
8	1874	31.2	-2.27	-0.77
9	1896	31.6	-1.86	-0.74
10	1888	31.5	-1.58	-0.75
12	**3797**	**62.3**	**0.30**	**0.21**
17	1868	31.1	-2.32	-0.79
18	1889	31.5	-2.06	-0.75
19	1888	31.5	-1.70	-0.75
21	**5488**	**91.5**	**0.55**	**0.46**
caffeine	**2897**	**48.3**	**0.07**	**-0.03**

Order of increasing log P: - 17 < 8 < 18 = 19 = 10 < 9 < 12 < 21

Table C8 Relationship between Log P and Log P_{CCC} (V_e = 12.5, V_s = 37.4, 2 ml/min)

Compound	t_r (secs)	V_R	Log P	Log P_{CCC}
8	714	23.8	-2.27	-0.51
9	708	23.6	-1.86	-0.53
10	716	23.9	-1.58	-0.53
11	743	24.8	-1.15	-0.48
17	719	24.0	-2.32	-0.51
18	710	23.7	-2.06	-0.52
19	716	23.9	-1.70	-0.56
20	749	25.0	-1.41	-0.48

Order of increasing log P: - 19 < 9 = 10 < 18 < 17 = 8 < 11 < 20

Table C9 Relationship between Log P and Log P_{CCC} (V_e = 10.0, V_s = 40.0, 1 ml/min)

Compound	t_r (secs)	V_R	Log P	Log P_{CCC}
8	679	11.3	-2.27	-1.48
9	653	10.9	-1.86	-1.67
11	660	11.0	-1.15	-1.60
17	642	10.7	-2.32	-1.76
18	663	11.1	-2.06	-1.58
19	656	10.9	-1.70	-1.62

Order of increasing log P: - 17 < 9 < 19 < 11 < 18 < 8

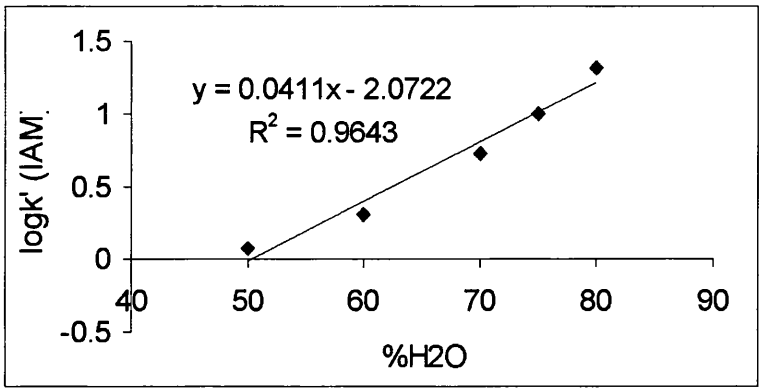


Figure C1 *Relationship between % H_2O and Log k'_{IAM} for Compound (13)*

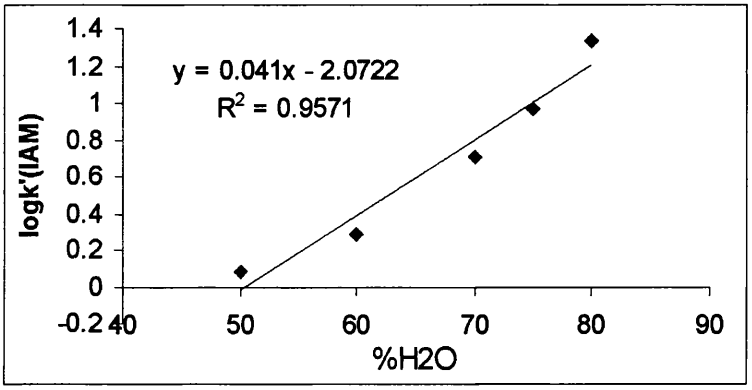


Figure C2 *Relationship between % H_2O and Log k'_{IAM} for Compound (22)*

Calculation of log k'_{IAM} for compounds **(13)** and **(22)**

(14) **log k'_{IAM} = 0.0411 (%H_2O) – 2.0722**

Therefore where %H_2O = 100, log k'_{IAM} = 2.04

(15) **log k'_{IAM} = 0.0410 (%H_2O) – 2.0722**

Therefore where %H_2O = 100, log k'_{IAM} = 2.03

Retention times can also be calculated from log k'_{IAM} using average t_o = 0.913 min. For compound **(13)**, t_r = 100.516 min (**1.68 h**); compound **(22)**, t_r = 98.248 min (**1.64 h**). These values are in line with the observation that no peaks were observed within 1.5 h at 100% aqueous.

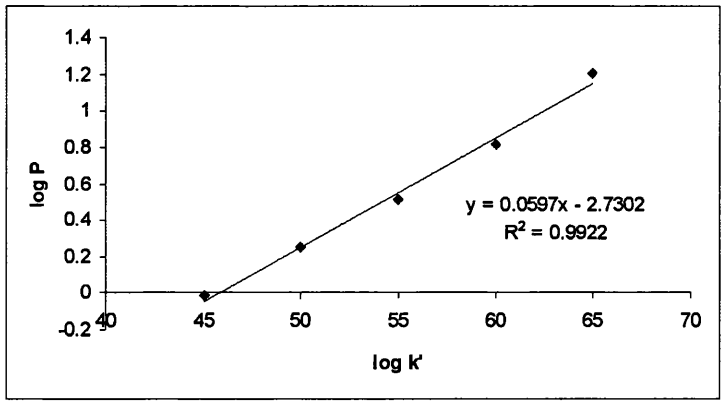


Figure C3 *Relationship between % H_2O and Log k'_{C18} for Compound (13)*

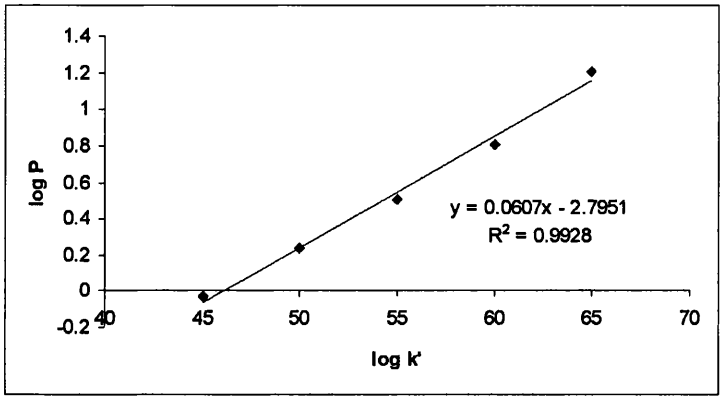


Figure C4 *Relationship between % H_2O and Log k'_{C18} for Compound (22)*

Calculation of log k'_{C18} for compounds (13) and (22)

(13) **$\log k'_{C18} = 0.0597\ (\%H_2O) - 2.7302$**

Therefore where $\%H_2O = 80$, $\log k'_{C18} = 2.05$

(22) **$\log k'_{C18} = 0.0607\ (\%H_2O) - 2.7951$**

Therefore where $\%H_2O = 80$, $\log k'_{C18} = 2.06$

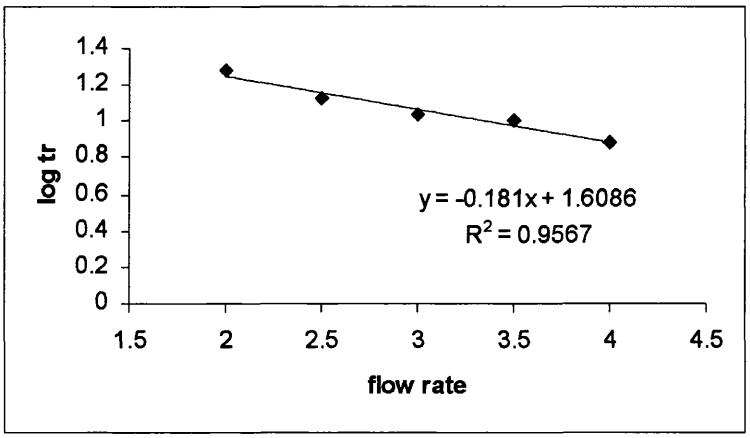


Figure C5 Relationship between Log of the Retention Time and Flow Rate for Compound (12)

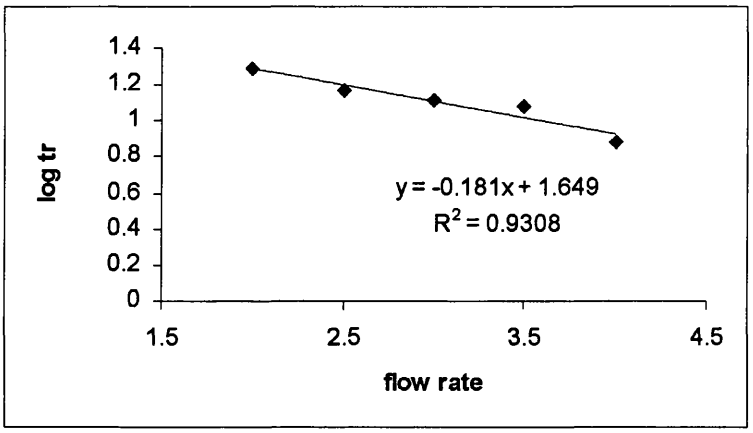


Figure C6 Relationship between Log of the Retention Time and Flow Rate for Compound (21)

Calculation of log $k'_{OC/C8}$ for compounds (12) and (21)

(12) **log $t_{r\,OC/C8}$ = -0.181 (flow rate) + 1.649**

Therefore where flow rate = 0.5 ml/min, $t_{r\,OC/C8}$ = 36.183 min

Where t_o @ 0.5 ml/min = 1.653 min, $k'_{OC/C8}$ = 1.32

(21) **log $t_{r\,OC/C8}$ = -0.181 (flow rate) + 1.6086**

Therefore where flow rate = 0.5 ml/min, $t_{r\,OC/C8}$ = 32.969 min

Where t_o @ 0.5 ml/min = 1.653 min, $k'_{OC/C8}$ = 1.08