THE PREDICTION OF CHEMOSENSORY EFFECTS OF VOLATILE ORGANIC COMPOUNDS IN HUMANS

A Thesis presented to the University of London in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science

by

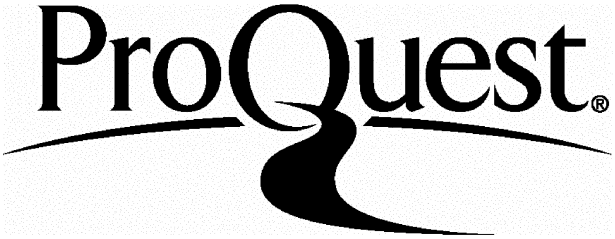
Joëlle Muriel Rachel Gola

Sir Christopher Ingold Laboratory

Chemistry Department

University College London

September 2001



To my Parents

ABSTRACT

An introduction to indoor air pollution is given, and the chemosensory effects in humans of volatile organic compounds (VOCs), singly and in binary mixtures, are described, together with the bioassays already developed to quantify the effects of VOCs. The need for predictive models that can take over the bioassays is emphasised.

Attention is drawn to the establishment of mathematical models to predict the chemosensory effects of VOCs in humans. Nasal pungency threshold (NPT), eye irritation threshold (EIT) and odour detection threshold (ODT) values are available for a series of VOCs that cover a large range of solute properties. Each of these sets of biological data are regressed against the corresponding solute descriptors, **E, S, A, B** and **L** to obtain quantitative structure activity relationships (QSARs) for log(1/NPT), log(1/ODT) and log(1/EIT) taking on the form:

$$\text{LogSP} = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L}$$

The availability of solute descriptors is investigated. It is shown that solute descriptors, **E** an excess molar refraction, **S** the solute dipolarity/polarizability, **A** the solute overall hydrogen-bond acidity, **B** the solute overall hydrogen-bond basicity and **L** the logarithmic value of the solute Ostwald solubility coefficient in hexadecane at 298K, can be obtained through the use of various thermodynamic measurements. In this way descriptors for some 300 solutes have been obtained.

A headspace gas chromatographic method is also devised to determine the 1:1 complexation constant, K, between hydrogen bond donors and hydrogen bond acceptors in octan-1-ol. The 30 complexation constants measured are then correlated with $\alpha_2^H * \beta_2^H$, a combination of the solute 1:1 hydrogen bond acidity and basicity, respectively, to give:

$$\text{Log } K_{1:1} = 2.950.\alpha_2^H * \beta_2^H - 0.741$$

Table of Contents

Chapter 7 Solvation Properties of Refrigerants

Chapter 8 Solvation Properties of Terpenes

Chapter 13 Future Work

List of Tables

List of Figures

Acknowledgements

My sincere thanks go to Dr Michael Abraham for his help over the last three years of this project. His knowledge and expertise in the area of physical chemistry has helped me to accomplish my Ph.D. I am extremely grateful for his unfailing encouragement, patience and guidance.

I would like to offer a particularly big thank you to my research group, Dr Caroline Green, Dr Joelle Le, Dr Juliette Osei-Owusu, Dr Andrea Zissimos, Dr Jamie Platts; Kei Enomoto, Rui Figueira, Vikas Gupta, who have provided me with encouragement and great friendship over the years. Thanks also to Helene Field, who carried out some of the chromatographic measurements. Finally, I would like to wish Kei and Rui good luck in the completion of their own research.

I extend my gratitude to Dr J. Enrique Cometto-Muniz and Dr William Cain for all the chemosensory threshold values included in this work, and for providing me with their expertise and advice on various issues.

My thanks go to the Center for Indoor Airborne Research, CIAR – San Diego, California – for the funding of this project and giving me the opportunity to study for this PhD.

In addition, I would like to thank all those in the chemistry department who have helped me in any way with my work or in making my time spent at University College London a happy one. In particular, Dimitra Georgeanopoulou, Camilla Forssten, and Paul Free for their tireless understanding and for always being here when I needed them. Thank you!

Thank you to all my friends in Gaillac, Toulouse, Paris and Cologne... Sophie I wish you good luck in the completion of your thesis.

Importantly, I would like to thank my family for putting up with me during this PhD. My parents, to whom this thesis is dedicated, have shown never-ending encouragement, love and support. I would not have got so far without it. A big thank you also to my sisters, Fabienne and Anne, and my two lovely nieces, Tatiana and Alexia for their love. This thesis is for you too! Finally, I would like to thank my grandma for her tireless support and love.

Merci!

Chapter 1 An Introduction to Indoor Air Pollution

1.0. Introduction

For many individuals, the perception of risks from outdoor air is substantially higher than for indoor air, although the home environment is rarely considered to be a risk in this regard. However, exposure to indoor air pollutants (IAP) is a potentially serious public health problem in a wide variety of non industrial settings, for example, residences, offices, schools and vehicles.[1] Studies from the United States and Europe show that persons in industrialised nations spend 90% or more of their time indoors.[2] Infants, elderly, and those with pre-existing respiratory diseases are virtually inside all the time. As exposure to air pollution is a function of both time and concentration, the significance of the indoor environment for the total exposure of a person to a pollutant can be high because of the time periods involved.[3]

While a good deal of public interest and concern continue to be directed at the effects of outdoor air pollution, notably traffic pollution, there is a growing tide of scientific studies that devotes sufficient attention to the indoor air environment. The understanding of the impact of indoor air pollution on human beings is of prime necessity to improve indoor air quality and, then to reduce the number of illnesses and discomfort.[1] Over the recent years, a large number of projects have been designed to gather information about the magnitude, extent, and causes of human exposure to indoor air pollutants.[1-15] Guidelines and standard levels have been proposed in an attempt to reduce exposure to IAPs.[4] Currently indoor air pollution is ranked by the United States Environmental Protection Agency (US EPA) and the Centres for Disease Control and Prevention in the top five environmental risks.[5]

Indoor air pollutants vary from heavy metals to volatile organic compounds (VOCs); it is the latter that is emphasised in this chapter. Firstly, VOCs have a wide range of physical and chemical characteristics. Chemical groups typically include hydrocarbons, halogenated hydrocarbons, aromatic hydrocarbons, alcohols, esters, ketones and aldehydes. These compounds have a wide range of boiling points and have been classed as very volatile (<0 to 50-100°C), volatile (50-100 to 240-260°C) and semi-volatile (250-260 to 380-400°C).[3] Examples in each group of concern in the indoor

air are formaldehyde, white spirit, and polyaromatic hydrocarbons, respectively. Furthermore, according to a definition given by the European Communities the expression "volatile organic compounds" means any compound having at 293.15 K a vapour pressure of 0.01 kPa or more, or having a corresponding volatility under the particular condition of use.[6] Next, VOCs contribute to a broad scale of chemicals with production levels all over the world and widespread applications in industry, trade and private households. One of the most important industrial uses of the VOCs is their supply as solvents. In Table 1.1, sources and tonnage of VOCs found in air of the United Kingdom, in 1995, are displayed. The number of identified VOCs present in ambient air has risen steadily in recent years, from 250 to more than 900 in 1989, to well more than 1,000 currently.[7] Finally, the concentrations of VOCs are typically found to be substantially higher indoors than outdoors. Because we spend the vast majority of our time indoor, the prolonged exposure to high concentrations of VOCs is all too common.[8] VOCs have been of increasing concern since the 1970s because of their potential to cause health effects similar to those reported in the Sick Building Syndrome and to contribute to respiratory problems and other diseases including cancer.[3]

The first section of this chapter deals with the various sources of air pollutants found indoors. The next section reports succinctly on the methods commonly used to assess exposure to indoor pollutant. Consequently, typical levels of indoor air pollutants are presented. Next, attention is drawn to the major health effects encountered by building occupants; effects due to exposure to VOCs are emphasised. The final section discussed the various methods put forward to control indoor air pollutant levels.

1.1. Sources of Indoor Air Pollutant

Indoor air has been shown to be a complex mixture of chemical, biological and physical agents.[10] Actually, this complexity can be illustrated by the fact that some 4000 various components have been identified in tobacco smoke alone.[11] Some of the most important IAPs currently recognised are aeroallergen (cat allergen, house dust allergen...), micro-organisms (bacteria, fungi, fungal spores...), dust and particles ($PM_{2.5}$ and PM_{10}); asbestos fibres, oxide of nitrogen (NOx), carbon monoxide, radon decay products, volatile organic compounds (VOCs) and pesticides.[12]

There are many sources of indoor air pollution at home. These include combustion sources, such as oil, gas, kerosene, coal, wood, and tobacco products (e.g. benzene), a wide variety of building materials (e.g. toluene, xylenes, decane), cleaners, office products and machines, paints, and furnishings. Further, bathing (e.g. chloroform from hot water), cooking, cosmetics, hygienic products, plants, as well as human biological processes give rise to IAP.[13] Table 1.2 gives examples of sources of VOCs known to be emitted indoor but this is far from comprehensive. Outdoor air may also contribute to indoor air contamination, particularly when air intakes are positioned near parking areas, roads, or other locations where contaminated air may be entrained into the buildings. For example, car exhaust gases and particles can infiltrate into the home, especially if windows are opened.[1]

Table 1.1. Sources of VOCs, excluding methane, in the outdoor air of UKin 1995

Source category	Estimated Emissions (thousand of tonnes)	% Total
Power station	5	-
Domestic	30	1.3
Commercial/Public service	2	-
Refineries	2	-
Iron and Steel	4	-
Other Industrial Combustion	9	0.4
Non-Combustion Sources	335	14.3
Extraction and distillation of fossil fuel	334	14.3
Solvent use	700	29.9
Road Transport	690	29.5
Off road Sources	96	4.1
Military	1	-
Railways	8	0.3
Civil Aircraft	4	-
Shipping	12	0.5
Waste treatment	26	1.1
Forest	80	3.4
Total	2338	100

Adapted from ref. 3.

Environmental tobacco smoke (ETS) is a source of particular concern because of the nuisance and irritation it can cause to users of multi-occupied buildings and the risks of disease to those inadvertently exposed to smoke. The main VOCs to be released in sidestream smoke in quantities exceeding 1 mg per cigarette are nicotine, acetaldehyde, acrolein, isoprene, and acetonitrile.[16] Very few of the constituents are unique to ETS

and this has caused difficulties in apportioning the contribution that ETS makes to concentrations of particular pollutants within buildings.[3]

Three fundamental processes control the rate of VOC emissions from building sources: (1) evaporation; (2) desorption of absorbed compounds and (3) diffusion within a material. How fast the VOCs are produced depends on the process and the source characteristics. The sources can be divided into those with continuous emissions, and discontinuous emissions.[3,10] Firstly, some sources, such as building materials and furnishings release pollutants more or less continuously. A process known as the 'Sink Effect', which is the absorption and desorption interactions between VOCs emitted and the interior sources, also prolongs the presence of pollutant in the air.[20] For instance, a forty-one day simulated chamber study indicated that an established material had absorbed about thirty VOCs, which were re-emitted to the chamber during the first thirty days of the study.[21] Only thirteen of the VOCs originally present in the first days of the study continued to be emitted in the final days, indicating that these thirteen were the only true components of the materials. Secondly, other sources, related to human activities carried out in the home, e.g. the use of solvents in cleaning, the use of pesticides in house keeping, release pollutants intermittently. The nature of emission and the variability of indoor spaces and ventilation conditions result in a dynamic behaviour of air pollutants in indoor environment.[22]

1.2. Assessment of Exposure

Assessment of exposure in humans refers to the analysis of various processes that lead to human contact with pollutants after release in the environment.[23] The term exposure refers to the length of contact with the pollutant during a specified period of time. Assessment of exposure is a science on its own. It is not the purpose of this chapter to present the various processes involved in exposure assessment. Here, attention is mainly drawn on methods used to determine pollutant concentration in indoor environments. Personal monitoring and biological monitoring are also defined. The issue of exposure assessment has been recently reviewed recently by Wallace[24] and by Ozkaynak[25].

Table 1.2. Specific Indoor Sources of VOCs (Adapted from Ref.3)

VOCs	Source Material
p-Dichlorobenzene	Moth crystals, room deodorants
Styrene	Insulation, Textiles, disinfectants, plastics, paints
Benzyl chloride	Vinyl tiles
Benzene	Smoking
Tetrachloroethylene	Dry cleaned clothes
Chloroform	Chlorinated water
1,1,1-Trichloroethane	Dry cleaned clothes, aerosol sprays, fabric protectors
Carbon tetrachloride	Industrial strength cleaners
Aromatic hydrocarbons (toluene, xylenes, ethylbenzene, trimethylbenzenes)	Paints, adhesives, gasoline, combustion products
Aliphatic hydrocarbons	Paints, adhesives, gasoline, combustion products
Terpenes	Scented deodorisers, polishes, fabrics, fabric softeners, cigarettes, food, beverages
Polyaromatic hydrocarbons (PAHs)	Combustion products (smoking, woodburning, kerosene heaters)
Alcohols	Aerosols, window-cleaners, paints, paint thinning, cosmetics, and adhesives
Ketones	Lacquers, varnishes, polish removers, adhesives
Ethers	Resin, paint, varnishes, lacquers, dyes, soaps, cosmetics
Esters	Plastics, resins, plasticizers, lacquer solvents, flavours, perfumes

1.2.1. Measurement of Volatile Organic Compounds in Indoor Air

A wide range of sampling and analytical methods has been applied to determine the nature and concentration of VOCs.[3,24-26] The most common methods are based on collection using absorbents, e.g. Tenax TA, contained in a sampling tube or badge. The absorbent can be thermally desorbed and the VOCs can be determined by gas chromatography coupled with various detection systems. Other absorbents are available and are chosen according to the polarity of the VOCs investigated. The sampler can be used in either an active or a passive mode. To obtain sufficient sensitivity for analysis

the exposure period of the passive sampler is typically between one to four weeks whereas for active sampling this can be achieved with sampling times of one hour. Other sampling methods are grab sampling, condensation, or liquid or solid-phase extraction. Solid-phase or liquid extraction is commonly used for less volatile compounds. A modified solid-phase extraction technique was put forward by Pawliszyn and co-workers.[27] The solid-phase microextraction technique (SPME) has now found a number of applications in environmental fate studies.[27,28] More recently, Elke and co-workers[29] proposed an improved SPME technique to analyse benzene, toluene, ethylbenzene and xylene (BTEX) in indoor air. Wallace has published a comprehensive review gathering methods available to sample and analyse IAPs.[25] In 1999, Clement et al. reviewed the latest techniques developed to examine IAPs.[26] Theory behind air sampling and analysis is explained in a book.[30]

Discomfort experienced in poor indoor air quality environment, is caused by simultaneous presence of individuals and air pollutants or in other words, individual's exposure to air pollutants. The individual or personal exposure level is best assessed by measuring an individual's contact with pollutants using personal monitoring or biological monitoring.[31]

1.2.2. Personal Monitoring

Measurements from personal monitoring indicate the level of external exposure.[24] They are carried out by means of small devices, which sample IAPs, placed on the individuals. For instance, bubblers, vapour adsorption tubes and passive samplers are widely used in personal monitoring to measure the concentration of airborne volatile chemical in the region of the mouth.[32] These measurements can be carried out in real environment or in simulated exposure chamber as explained later.

1.2.3. Biological Monitoring

According to the International Union of Pure and Applied Chemistry[23] (IUPAC), biological monitoring (BM) is a 'systematic continuous or repeated measurement and assessment of workplace agents or their metabolites either in tissues secreta, excreta or any combination of these to evaluate exposure or health risk

compared to appropriate reference'. In other words, BM allows one to assess the integrated exposure by different routes, including ingestion, inhalation, dermal absorption, blood, exhaled air and urine. Biological marker or biomarkers for exposure is an endogeneous substance or its metabolite or the product of interaction between a xenobiotic and some target molecule or cell that is measured in a compartment with an organism.[23] Biomarkers are focused on the amount of the pollutant penetrating to the organism. Several biomarkers are relevant to indoor air pollution: e.g. urinary excreted nicotine is used for exposure to ETS, the carboxyhaemoglobin level in blood is used to characterise exposure to CO, and the presence of VOCs in exhaled air breath is used to mark these compounds.[32] Examples of the use of BM are available in the literature. Imbriani and co-workers[33] developed a method for the BM of exposure to enflurane in operating room personnel based on the measurement of the unchanged anaesthetic in urine. Andreoli[34] used a SPME method to determine level of hydrocarbons in blood and urine. Jo and Pack[35] employed a breath analysis for exposure to benzene associated with active smoking. Mathews et al.[36] studied endogenous VOCs in breath. Further examples of use of the BM technique can be found in a chapter dealing with 'methods in human inhalation toxicology'[32]. Biological monitoring is extensively applied in practical occupational medicine in many developed countries. As a result, biological threshold values have been evaluated to control the worker's exposure.

1.2.4. Controlled Exposure Chamber

Assessment of exposure is mainly carried out according to the above methods. These techniques can be used in real environments e.g. offices, homes, vehicles, or in controlled or simulated exposure chambers. Controlled exposure chambers have been devised to study single or multiple pollutants in relative 'pure' form, without potential interference of other materials. These exposure systems vary from small volume, personalised mouth-piece or face mask configurations to large chambers, which are often used to characterise pollutants emitted from wallpaper, carpet, and other materials. Beside, small devices are more appropriate for controlled human exposure studies, which are carried out to evaluate human response to pollutants. The main drawback is that true simulation of the 'real' environment is impractical.[31]

1.3. Level of Indoor Air Pollutants

Levels of indoor air pollutants depend on several factors. From continuous sources, the magnitude of emissions often depends on temperature, relative humidity, and sometimes air velocity, and varies within a time scale of months.[10] On the other hand, discontinuous emissions are much more time dependent and may change within hours or minutes.[3] Levels of indoor air pollutant are also dependent on human's activity.[1-16] Over recent years, the combination of reduced ventilation rates, warmer and more humid conditions indoors, together with the greater use and diversity of materials, furnishing and consumer products, has resulted in accumulation of a wide range of pollutants occurring indoors at level often exceeding those outdoors.[1,3,10,24,37-45] Age of the building plays an important factor in concentration of pollutants. The US EPA studies of new buildings indicated that eight of thirty-two target chemicals measured within days after completion of the buildings were evaluated 100 fold higher compared to outdoor levels: xylenes, ethylbenzene, ethyltoluene, trimethylbenzene, decane and undecane.[37] VOCs are commonly present as mixtures, with mean concentration below 50 microgram per cubic meter ($\mu g.m^{-3}$) in established buildings, but much higher in new buildings.[3]

The United State Environmental Protection Agency (EPA) has carried out a number of studies to determine levels and also exposure to IAPs in several urban, non-urban and industrial and non-industrial areas in various places in the US. One of the most comprehensive studies designed to determine the exposure of individuals to IAPs within their homes was the Total Exposure Assessment Measurement (TEAM) study.[25,38-42] This study pointed out that personal exposures to VOCs were two to five times higher than outdoor concentration, even though the outdoor concentration were measured in heavily polluted areas such northern New Jersey and Los Angeles. Further, much of the difference is attributable to exposure to indoor sources, such as environmental tobacco smoke.[42] Comparison between personal and outdoor air concentration is shown in Figure 1.1. The values taken from reference 24 were measured during fall 1981 in New Jersey. More recently, The EPA conducted a study about indoor and outdoor air concentrations of twenty-seven hazardous air pollutants.[42] Indoor and outdoor exposures (i.e. concentrations breathed multiplied by duration of time breathed) were estimated from data of the literature. The results emerging from this

study showed that the indoor air concentration of these hazardous compounds are generally one to five times outdoor concentration, and indoor exposures are ten to fifty times outdoor exposures as displayed in Table 1.3. In Europe, the UK Building Research Establishment (BRE) has intensively studied the indoor air environment of 174 family homes in the Avon.[43] Attention was drawn to nitrogen dioxide, formaldehyde and other VOCs, house dust mites, bacteria and fungi. Examples of typical indoor and outdoor levels measured in the BRE study are shown in Table 1.4. In this study, the researchers considered the total of volatile organic compounds[44] (TVOC) encountered in the various houses. Finally, researches are also conducted in developing nations such as Brazil. Brickus et al. [45] measured indoor and outdoor concentration of several air pollutants on the first, ninth, thirtieth and twenty-fifth floor of an office building in Rio de Janeiro, see Table 1.5. The results emerging from this study agreed well with those obtained in developed countries; pollutant concentration is generally higher indoors than outdoors.

A series of studies on personal exposure to environmental tobacco smoke (ETS) has been carried out in several countries.[1,3,24] Some results emerging from these studies showed that persons with a smoking partner had mean exposure of 219 $g.m^{-3}$, compared to about 170 $\mu g.m^{-3}$ for persons without a smoking partner, a difference of about 49 $\mu g.m^{-3}$ for those without.[24] Studies have also shown that where smoking rates are high and ventilation minimal there is a clear contribution to formaldehyde concentrations from ETS of the order of a few 10s of $\mu g.m^{-3}$. Concentrations of aliphatic and monocyclic aromatic hydrocarbons can rise from 2-20 $\mu g.m^{-3}$ to 50-200 $\mu g.m^{-3}$. Detailed studies in the USA of residential exposures suggest that non-smoking households experience an exposure to approximately 7 $\mu g.m^{-3}$ of benzene, while households with a smoker are exposed to approximately 11 $\mu g.m^{-3}$. Excess exposure to styrene and xylenes range from 0.5 to 1.5 $\mu g.m^{-3}$. [3]

Table 1.3. Summary of Indoor and Outdoor Air concentrations and Exposures.(ref. 42)

HAP	Typical concentration [a]			Typical Daily Exposure [b]		
	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio
Acetaldehyde [c]	<10	<3	*[5]*	<216	<7.2	*[50]*
Benzene [c]	5	5	*[2]*	108	12	*[20]*
Captan	<0.001	Na	10	<0.02	<0.0002	~100
Carbon tetrachloride[c]	<5	1	*[2]*	<108	2.4	*[20]*
Chlordane	0.2	0.01	20	4.32	0.024	200
Chloroform[c]	1	0.2	*[5]*	21.6	0.48	*[50]*
Cumene	1	0.2	5	21.6	0.48	50
2,4-D (salts, esters)	0.001	0.00003	30	0.0216	0.000072	300
DDE	0.0005	<0.002	>0.2	0.0108	<0.005	>2
Dichlorvos	0.05	0.001	50	1.08	0.0024	400
Ethylbenzene	5	1	*[3]*	108	2.4	*[30]*
Formaldehyde	50	4	10	1080	9.6	100
Formaldehyde[c]	0.1	0.002	50	2.16	0.0048	400
Hexachlorobenzene	0.0005	0.0001	5	0.0108	0.00024	50
Hexane	5	4	*[10]*	108	9.6	*[90]*
Methoxychlor	0.0001	0.00003	3	0.00216	0.000072	30
Methyl Ethyl Ketone	10	<1	*[4]*	216	<2.4	*[40]*
Methylene chloride	10	1	*[10]*	216	2.4	*[90]*
Naphthalene	1	1	*[4]*	21	2.4	*[40]*
Paradichlorobenzene	1	<0.05	*[5]*	21.6	<0.12	*[50]*
Propoxur	0.1	0.01	10	2.16	0.024	90
Radionuclides[d]	2	0.1	20	43.2	0.24	200
Styrene [c]	2	0.6	*[5]*	43.2	1.44	*[50]*
Tetrachloroethylene[c]	5	2	*[3]*	108	4.8	*[30]*
Toluene	20	5	*[5]*	432	12	*[50]*
Trichloroethylene[c]	5	0.5	*[5]*	108	1.2	*[50]*
Xylenes (o+m+p)	15	10	*[2]*	324	24	*[20]*

[a] Indoor and outdoor concentrations in $\mu g.m^{-3}$ Typical values from the literature for U.S. locations. I/O ratios based on typical concentrations *[or reported ratios as indicated by values in Italics and brackets].*
[b]Exposure, in $\mu g.m^{-3}.h$, based on assumption that the typical person spends 90% of the typical day indoors.
[c]Urban Air Toxic substances, [d]Radionuclides/Rn in $pCi.dm^{-3}$

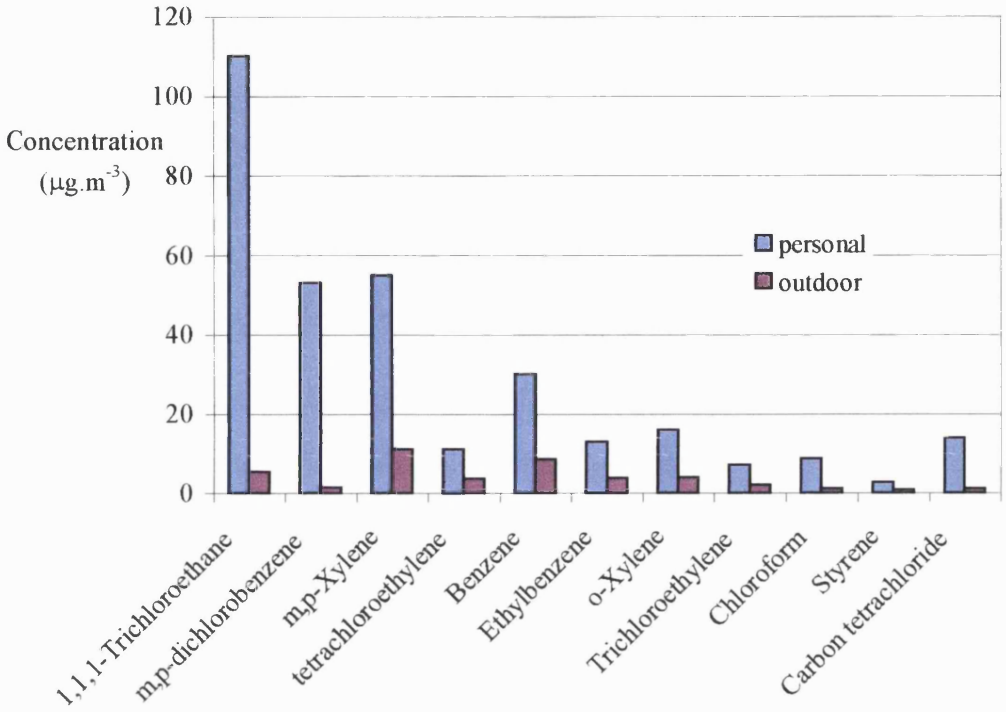


Figure 1.1. Personal exposure compared to outdoor air concentrations ($\mu g.m^{-3}$) of VOCs (New Jersey, Fall 1981) Adapted from ref. 24.

Table 1.4. Concentration of VOCs in indoor (main bedroom) and outdoor air in Avon, England.[(a)]

	Indoor Concentration ($\mu g.m^{-3}$)		Outdoor concentration ($\mu g.m^{-3}$)	
VOC	Min.	Max.	Min.	Max.
Formaldehyde	1	205	ND	11
Benzene	ND	78	ND	16
Toluene	2	1793	2	254
Undecane	ND	797	ND	3
TVOC	21	8392	4	317

Source: adapted from ref. 3

[(a)]Abbreviations: TVOC, total volatile organic compounds; ND, not detected; NA, not applied

Table 1.5. Average concentration of indoor and outdoor pollutants ($\mu g.m^{-3}$) in the offices. [a]

Floors	1st			9th			13th			25th		
Pollutants	I	O	I/O	I	O	I/O	I	O	I/O	I	O	I/O
TSP	91.4	141.4	0.7	28.7	32.8	0.9	53.5	58.8	0.9	66.6	43.5	1.5
Nicotine	1.5	ND	NA	ND	ND	NA	0.7	ND	NA	0.5	ND	NA
UV-RSP	10.4	5.8	1.8	7.7	5.5	1.4	6.6	5.0	1.3	8.1	1.8	4.5
Formaldehyde	42.0	19.7	4.7	93.8	19.8	4.7	19.5	12.1	1.6	22.6	7.9	2.8
Acetaldehyde	23.7	23.0	1.0	4.1	18.6	0.2	30.1	16.9	1.8	18.6	10.0	1.8
TVOC	709.2	346.1	2.0	88.9.2	215.9	4.1	1196.0	271.3	4.4	450.1	38.3	11.8

Adapted from Ref. 45

[a]Abbreviations: I, indoor air; O, outdoor air; I/O, indoor/outdoor ratio; TSP, total suspended particles; UV-RSP, ultraviolet respirable particle matter; TVOC, total volatile organic compounds; ND, not detected; NA, not applied

Level of particles and particulate matters in buildings and homes have also been investigated.[46,47] As for gaseous compounds, personal exposure was found higher than outdoors concentration.

Office buildings almost always contain a complex mixture of VOCs encompassing many different classes of organic compounds. Therefore, Molhave and Nielsen have proposed the concept of TVOC as an overall index to represent the entire mixture, in an attempt to overcome the task of identifying and quantifying hundreds of VOCs.[44] The levels observed are useful as indicators of background levels of VOCs. For example and European study reported the average TVOC concentrations in 54 buildings in nine countries to be below 500 $\mu g.m^{-3}$. When the USEPA studied the TVOC levels in 16 randomly selected buildings, the levels for the most part do not exceed 500 $\mu g.m^{-3}$. A study of the literature from 1983-1993 revealed a TVOC range from 20 to 5300 $\mu g.m^{-3}$. It should be noted that even the highest level observed is only about 1ppm.

Sources, levels of indoor air pollutants together with the assessment of exposure have been presented in the previous sections. Attention is now drawn on the effects of indoor air pollutants, notably VOCs, on humans.

1.4. Health Effects of Indoor Air Pollutants

Health effects from IAPs may be experienced soon after exposure or possibly years later. Firstly, immediate effects may show up after a single exposure or repeated exposures. These include irritation of the eyes, nose and throat, headaches, dizziness, and fatigue. Such immediate effects are usually short term and treatable. The likelihood of immediate reactions to IAP depends on several factors. Age and pre-existing medical conditions are two important influences. In other cases, whether a person reacts to a pollutant depends on individual sensitivity, which varies from person to person. Secondly, other health effects may show up only after long or repeated periods of exposure. These effects which include some respiratory diseases, heart disease, and cancer, can be fatal. [1,3,48-50]

Immediate effects reported by building occupants have been classed in various categories whose most widely referred to are:

- Multiple Chemical Sensitivity
- Building Related Illness
- Sick Building Syndrome

These categories are now presented in the above order; the sick building syndrome (SBS) being emphasised.

1.4.1. Multiple Chemical Sensitivity

Multiple Chemical Sensitivity (MCS) is postulated to be the development of responsiveness, including manifestation of often disabling symptoms, to extremely low concentrations of chemicals following sensitisation.[50]

1.4.2. Buildings Related Illness

The Building Related Illness (BRI) term is used when symptoms of diagnosable illness are identified and can be attributed directly to airborne contaminants. BRI complaints include cough, chest tightness, and fever, chill and muscles aches. The BRI symptoms can be defined and have clearly identified causes. Complainants may require prolonged recovery times after leaving the building.[50]

1.4.3. Sick Building Syndrome

The term Sick Building Syndrome (SBS) is used to describe situations in which building occupants experience acute health and comfort effects that appear to be linked to time spend in a building, but no cause or specific illness can be identified. The complaints may be localised in a particular room or zone, or may be widespread throughout the building. The complaints may decline short time after leaving the building.[48]

Inadequate ventilation, biological contaminants such as bacteria, moulds, pollen and viruses, and chemical contaminants are thought to play important roles in SBS. However, the main cause is believed to be common VOCs found in indoor environment. These elements may act in combination, and may supplement other complaints such as inadequate temperature, humidity, or lightning. It is also generally recognised that

psychological factors can influence an individual's perception of indoor air health effects as they can influence other disease processes.[48]

The SBS is characterised by a range of symptoms including, but not limited to, central nervous system complaints such as headache and fatigue, eye, nose, and throat irritation, and dry skin. Contrary to BRI, the SBS is by no means a well-defined medical entity, and often the symptoms vary widely from person to person within a given building.[8]

Among the various symptoms evoked, sensory responses, especially irritative response figure prominently. As Molhave noted,[51] "SBS, in brief, is an unexplainable sensory irritation appearing in a large fraction of the occupants of the affected building". Similarly, Cain pointed out that when air smells bad, or when it irritates the nose or eyes, people commonly feel threatened.[52] For these reasons, in the non-industrial environment, nasal irritation (or pungency) together with odour perception is believed to be appropriate indicators of indoor air quality.[52]

In studies on these chemosensory issues, VOCs have deserved particular attention.[52] Actually rarely does a VOC lack the potential to cause irritation and experiments could be developed to set standard for VOCs.

1.4.4. Volatile Organic Compounds and Building-Related Complaints

Several factors have led specialists to think that VOCs could adversely contribute to indoor air quality. First, both in terms of number of compounds and concentration, VOCs predominate in indoor air.[53] VOCs are found at higher concentrations in indoor air compared to outdoor concentration. Secondly, complaints about indoor air are typically more prevalent in new or refurbished buildings when concentrations of VOCs are highest. Furthermore, many VOCs can cause or contribute to a wide range of health effects from non-specific sensory responses to specific toxicity to target organs. In Table 1.6 are listed some selected VOCs and their various impacts on human beings whose irritative response figured prominently. However, although mixtures of VOCs have the potential to affect indoor air quality, studies of the relation between exposure to indoor air VOCs and symptoms reported by building occupants have shown only a sparse or inconsistent association between observed VOCs levels and health effects.[48]

Probably, the most informative studies on health effects of VOCs based on the TVOC method are obtained from simulated chamber studies using defined concentrations of mixtures and defined endpoints. Human subjects were exposed to a typical mixture of twenty-two common VOCs (the "Molhave mixture") at different concentrations.[54-58] They follow a gradient from sensory effects (for example odor at almost 3 $mg.m^{-3}$) to indications of subacute stress reactions at about 25 $mg.m^{-3}$. Sensory responses and behavioural impairment were examined through the use of questionnaires and objectives tests. The major findings of these studies are summarised:

- Complaints of poor air quality as a result of perceived odour intensity were highly correlated with total VOC concentrations and were significant at the lowest tested concentration, $3mg.m^{-3}$.
- Irritation of nose, throat, and eyes was significant at or above 5 $mg.m^{-3}$.
- Other symptoms such as headache and general discomfort appeared at 25 $mg.m^{-3}$.
- There was some limited but still controversial evidence to suggest that behavioural impairment such as reduced short-term memory may appear at 25 $mg.m^{-3}$.

While the above are the concentrations associated with SBS symptoms in controlled experiments, Molhave[55] has found that SBS complaints are likely to arise when total VOC concentration exceed 1.7 $mg.m^{-3}$. This suggests that building occupants may respond to VOCs in the real-life environment at a lower concentration than in a controlled environment. Furthermore, these real-life level are below those at which toxicological or sensory effects would be expected in humans, see Tables 1.7 and 1.8 for formaldehyde as an example. [32,54] These findings show how it is difficult to set guidelines and standards for VOCs that are needed to reduce exposure to VOCs and, by extent, to improve indoor air quality and to reduce discomfort and illnesses.

Moreover, concerns about low-level exposures usually arise around the perception of odor and irritation, and VOCs are considered to be only one of a combination of factors that cause such complaints. Studies of low levels of VOCs and sensory responses that involve odor, nasal pungency, and eye irritation show that mixtures of VOCs cause responses at concentration far below what would be expected for each component of the mixture. It seems that increasing the number of VOCs in a complex mixture can lower the thresholds for odor as well as for eye and nasal irritation.

Table1.6. Health Effects of selected Volatile Organic Compounds

VOC	Health Effects
Benzene	Carcinogen; respiratory tract irritant
Xylenes	Narcotic; irritant; affects heart, liver, kidney, and nervous system
Toluene	Narcotic; possible cause of anemia
Styrene	Narcotic; affects control of nervous system; probable human carcinogen
Toluene diisocyanate	Sensitizer; probable human carcinogen
Trichloroethane	Affects central nervous system
Ethyl benzene	Severe irritation of eyes and respiratory tract; affects central nervous system
Dichloromethane	Narcotic; affects control of nervous system; probable human carcinogen
1,4-Dichlorobenzene	Narcotic; eyes and respiratory tract irritant; affects heart, liver, kidney, and nervous system
Benzyl chloride	Central nervous system irritant depressant, affects liver and kidney; eye and respiratory tract irritant
2-Butanone	Irritant; central nervous system depressant
Petroleum distillate	Affects central nervous system, liver and kidneys
4-Phenycyclohexene	Eye and respiratory tract irritant; central nervous system effects

Sources adapted from ref. US EPA 'introduction to IAQ' report no. EPA/400/3-91/003, Washington, DC, 1991.

Table 1.7. Indoor sources for formaldehyde exposure.

Sources	Concentration
Cigarette smoke	40 ppm in 40 cm^{-3} per puff
Dose per pack for smoke	0.38 μg per pack
Environmental tobacco smoke	0.25 ppm
Clothing made with synthetic fibres	
Men's polyester-cotton blend	2.7 $\mu g.g^{-1}$ per day
Women's dress	3.7 $\mu g.g^{-1}$ per day
Furnishings	
Particle board	0.4-0.8 $\mu g.g^{-1}$
Plywood	1.5-5.3 $\mu g.g^{-1}$
Panelling	0.9-21 $\mu g.g^{-1}$
Draperies	0.8-3 $\mu g.g^{-1}$
Carpet / upholstery fabric	= 0.1 ppm

Adapted from ref. 32

Table 1.8. Effects of formaldehyde on humans after short-term exposure

Effects	Formaldehyde Concentration ($mg.m^{-3}$)
Odour detection threshold	0.06-1.2
Eye irritation threshold	0.01-1.9
Throat irritation threshold	0.1-3.1
Biting sensation in nose and eye	2.5-3.7
Tolerable for 30 minutes (lacrymation)	5.0-6.2
Strong lachrymation, lasting for 1 hour	12-25
Danger to life, oedema, inflammation, pneumonia	37-60
Death	60-125

Adapted from Ref. 32

1.5. Guidelines and Standards

The purpose of the guidelines is to provide a basis for protecting public health from the adverse effects of air pollution and for eliminating, or reducing to a minimum, those air contaminants that are known to be, or likely to be, hazardous to health and well being. For many of the classic air pollutants, the guidelines were based on controlled exposure studies, or on epidemiological studies, which demonstrated a threshold of effect. These guidelines were statements of levels of exposure at which, or below which, no adverse effects can be expected. On the other hand, air quality standard is a description of a level of air quality that is adopted by a regulation authority as enforceable.[59]

The effects of many contaminants in the industrial workplace have been well established by groups such as the American Conference of Governmental Industrial Hygienists (ACGIH). However, the industrial standards apply to situations where the concentrations are much higher than in offices or homes. In addition, industrial standards are intended to protect a healthy worker from individual compounds in an industrial environment.

The ACGIH has recommended Threshold Limit Values (TLV) as guideline of exposure to a number of substances recognised as harmful to humans, and in the United States, the Occupational safety and Health Administration (OSHA) has set workplace standards in this area. For example, OSHA presently sets the current permissible benzene exposure level at one part per million (ppm) for an 8-h average with a short-

term exposure limit of five ppm. Currently, there is no standard or regulation for acceptable benzene levels / exposures within the home. Although the ACGIH threshold limit values are much higher than the usual indoor levels, see Table 1.9, the exposures in non-industrial indoor settings may be chronic, 20 hour or more per day, seven day per week, 52 weeks per year. Occupational threshold limit values are normally based on evidence from exposures to concentrations that are much higher than those found in the indoor environment. Furthermore, such exposures often focus on health effects that are much more severe than the relatively harmless irritation. More information is needed to be able to set guidelines and standards more appropriate to indoor air pollution.[59]

Table 1.9. Typical concentration of Indoor VOCs ($\mu g.m^{-3}$) compared to Threshold Limit Values (TLV) ($\mu g\,m^{-3}$)

VOC	Indoor Concentration	TLV
Chloroform	0.00088	49
1,1,1-Trichlorobenzene	0.38000	1910
Benzene	0.00400	32
Carbon tetrachloride	0.00044	31
Chlorobenzene	0.00011	46
n-Decane	0.38000	525
m-p-Dichlorobenzene	0.00140	60
Ethylbenzene	0.08400	434
Styrene	0.00830	213
Tetrachloroethylene	0.00670	170
Trichloroethylene	0.00120	269
n-Undecane	0.17000	525
m ,p-Xylenes	0.14000	434
o-Xylene	0.07400	434

Adapted from Ref. 3.

Irritation and odour annoyance are the two main indicators for indoor air quality.[52] However, only few studies have been carried out on setting guidelines for these nonspecific effects, including the use of total volatile organic compounds (TVOC).[44] Molhave and Nielsen[44] have presented the prospect of setting a nasal irritation based standard for mixtures. Exposures to VOCs below 0.2 $mg.m^{-3}$ are unlikely to have irritation effects. At concentration higher than 3 $mg.m^{-3}$, complaints often occur. However, associations between TVOC concentration and health effects are unclear. One major problem is that using a TVOC level assumes that each VOC of that mixture is equally important in relation to health, but in fact one VOC may be more hazardous than others.

Cain and Cometto-Muniz have proposed the determination of odour detection threshold (ODT) together with nasal pungency (NPT) and eye irritation thesholds (EIT) for singly VOC to develop standard values for VOCs.[60] ODT, NPT and EIT values correspond the minimal concentration of a VOC at which the corresponding sensory effect is detected. The field of sensory impact of VOCs is presented in more detail in chapter 2. In the other hand most studies on setting standards relate to specific toxicity such as cancer development or individual organ toxicity.[61]

1.6. References

1. P.T.C. Harrison, Chemistry and Industry (1997) 677.
2. J. Robbinson, W.C. Nelson, National Human Activity Pattern Survey Database, EPA/600/R-96/074, USEPA, Research Triangle Park, NC, (1995).
3. L. Sheldon, H. Zeton, J. Sickes, C. Eaton, T. Hartwell, 'Indoor Air Quality in Public Buildings, Vol.II', Research Triangle NC: EPA/600/6-88/009b, 1988b.
4. J.Q. Koenig, in Health Effects of Ambient Air Pollution, Kluwer Academic Publisher, (2000) 195.
5. K. R. Spaeth, Prevent. Med., 31 (2000) 631.
6. Commission of the European communities, Proposal for a council Directive on limitation of emissions of volatile organic compounds due to the use of organic solvents in certain industrial activities. 96/0276 (SYN) Article 2, (1996).
7. E.W. Miller, R.M. Miller, 'Indoor Air Pollution, Contemporary World Issues' ABC Clio, Santa-Barbara, (1998).
8. US EPA, 'What Causes Indoor Air Quality Problem', http://epa.gov, (2000).
9. L. Rushton, K. Cameron, in ' Selected Organic Chemical', (1999),
10. World Health Organisation (WHO) 'Air Quality Guidelines', WOH Regional Publication Europe, Copenhagen, (1999).
11. E.L. Anderson, R.E. Albert, 'Risk Assessment and Indoor Air Quality', Lewis Publisher, CRC Press LLC, (1998).
12. P.T.C Harrison, Issue Env. Sc. Techn., Air Pollution, 10 (1998) .
12. E.J. Bordana, Jr, A. Montanaro, 'Indoor Air Pollution and Health', Marcel Decker, Inc, New York, (1997).
13. W.E. Lambert, J.M. Samet, 'Occupational and Environmental Disease', Eds. P.Harber, M.B.Schenker, J.R.Balmes, Mosby, St Louis, (1996) 784.
14. H.J Moriske, M. Drews, G. Ebert, G. Menk, M. Scheller, m. Schondube, L. Konieczny, Toxicol Lett., 88 (1996) 349.
15. V.M. Brown, D.R. Crump and M. Gavin, 'Indoor Air Quality in homes: Part 1, The Building Research Establishment Indoor Environment Study', Eds. R.W. Berry, S.K. D. Coward, D.R. Crump, M. Gavin, C.P. Grimes, D.F. Higham, A.V. Hull, C.A. Hunter, I.G. Jeffrey, R.G. Lea, J.W. Llewelly and G.J. Raw, Construction Research Communications, London, (1996) 18.
16. L.Rushton, K. Cameron, 'Air Pollution and Health', Eds. S.T. Holgate, J.M. Samet, H.S. Koren, R.L. Maynard, Acad. Press, London, (1999) 813.
17. National Research Council, Environmental Tobacco Smoke, Measuring and Assessing

Health Effects, National Academy Press, Washington D.C., (1986).

18. US EPA seltzer, book (1987).
19. M. Guerin, R. Jenkins and B. Thomkins, 'The Chemistry of Environmental Tobacco Smoke', Lewis Publisher, Michigan, (1992).
20. H. Gustafsson, 'Building materials Identified as Major Sources of Indoor Air Pollutants' Swedish Council for Building Research, Document D10, Stockholm, 1992.
21. B. Berglund, I. Johansson, Proceeding of the 4th international conference on Indoor Air Quality and Climate, 1 (1987) 16.
22. B. Siefert, H. Knoppel, R. Lanting, A. Person, P. Siskos and P.Wolkoff, ' Strategy for Sampling Chemical Substances in Indoor Air', European Concerted Action, Report 6, EUR 12617 EN, Commission of the European Communities, Luxembourg, (1989).
23. J.H. Dufus, Pure Appl. Chem, 65 (1993) 2003.
24. L. Wallace, 'Total Exposure Assessment Methodology (TEAM) Study: Summary and Analysis', vol.1, US EPA, (1987).
25. H. Ozkaynak, 'Air Pollution and Health', Eds. S.T. Holgate, J.W. Samet, H.S. Koren and R.L. Maynard, Academic Press, (1999) 148.
26. R.E. Clement, P.W. Yang, C.J. Koester, Anal. Chem, 71 (1999) 257R.
27. J. Pawliszyn, in 'Solid Phase Microextraction: Theory and Practice', Wiley-VCH, New-York, (1997).
28. P.A. Martos, J. Pawliszyn, Anal. Chem., 69 (1997) 206.
29. K. Elke, E. Jermann, J. Begerow, L. Dunemann, J. Chromatogr. A, 826 (1998), 191.
30. C.F. Poole and S.K. Poole, 'Chromatography Today', Elsevier Sci. Pub. B.V., Amsterdam, (1991).
31. H. Janturen, J.J.K. Jaakola, M. Krzyzanowski, 'Assessment of Exposure to Indoor Air Pollutant', WHO regional Publication European Series, Copenhagen, 78 (1997).
32. L.J. Folinsbee, C.S. Kim, H.R. Kerl, J.D. Prah, R.B. Devlin, in 'Handbook of Toxicology
33. M. Imbriani, S. Ghittori, G. Pezzagno, E. Capodaglio, Arch. Environ. Health, 49 (1994) 135.
34. R. Andreoli, P. Manini, E. Bergamaschi, A. Brustolin, A. Mutti, Chromatographia, 50 (1999) 167.
35. W.-J. Jo, K.-W. Pack, Environ. Res. Sec. A, 83 (2000) 180.
36. J.M. Mathews, J.H. Raymer, A.S. Etheridge, G. R. Velez and J.R. Bucher, Toxicol. App. Pharmacol., 146 (1997) 255.
37. US EPA, 'indoor Air Quality in Public Building vol II', EPA600/6-88/009b, (1988).
38. L. Wallace, Risk Anal., 13 (1993) 135.
39. L. Wallace, 'Risk Assessment and Indoor Air Quality', Edited by E.L. Andersonand R.E. Albert, (1998) 161.

40. L. Wallace, E.D. Pellizari, T.D. Hartwell, Environ. Res., 43 (1987) 290.
41. L. Wallace, 'IEH assessment on Indoor Air Quality in the Home', (assessment A1), Leicester: Institute for Environment and Health, (1996) 338.
42. US EPA, 'Inside IAQ', EPA/600/N-98/002, (1998) 1.
43. R.W. Berry, in 'Indoor Air Quality in Homes. The Building Research Establishment Indoor Environment Study', London: Construction Research Communications, (1996).
44. L. Molhave and G.D. Nielsen, Indoor Air, 2 (1992) 65.
45. L.S. Brickus, J. Cardoso and F.R. De Aquino Neto, Environ. Sci. Technol., 32 (1998) 3485.
46. C.A Clayton, R.L. Perritt, E.D. Pellizzari, Analy Environ. Epidemiol., 3 (1993) 227.
47. E. Abt, H.H. Suh, P. Catalano and P. Koutrakis, Environ. Sci. Tech., 34 (2000) 3579.
48. K. Sexton, R.S. Dyer, 'Indoor Air Quality and Human Health, 2nd Ed.', Eds. R. B. Gammage, B.A. Berven, CRC Press, Boca Raton, (1996) 1.
49. C. Rose, 'Indoor Air Quality and Human Health, 2nd Ed.', Eds. R. B. Gammage, B.A. Berven, CRC Press, Boca Raton, (1996) 127.
50. US EPA, 'A Guidebook to Comparing Risks and Setting Environmental Priorities', EPA 230-8-93-003, Washington D.C., (1993).
51. L. Molhave, Ann. N.Y. Acad. Sci., 641 (1992) 46.
52. W.S Cain, J.E. Cometto-Muniz, Occupational Med., 10 (1995) 133.
53. P. Wolkoff, P.A. Clausen, P.A. Nielsen, L. Molhave, Indoor Air, 4 (1991) 478.
54. K.M. Hau, D.W. Connell, B.J. Richardson, Regul. Toxicol. Pharmacol., 32 (2000) 36.
55. L. Molhave, ASHRAE Trans. 92 (1A) (1986) 306.
56. D.A. OttoL. Molhave, G. Rose, H.K. Hudnell, D. House, Neurotoxicol. Teratol., 12 (1990) 649.
57. S. Kjaergaard, L. Molhave, O.F. Pedersen, Atmos. Environ., 25A (1991) 1417.
58. H.K. Hudnell, D.A. Otto, D.E. House, L. Molhave, Arch. Environ. Heal., 47 (1992) 31.
59. W.S. Cain, J.E. Cometto-Muniz, Occup. Med., 10 (1995) 133.
60. World Health Organisation, 'Air Quality Guidelines for Europe', WHO Regional Pub., European Series No 23, Reg. Offices for Europe, WHO, Copenhagen, (1987).
61. H.R. Pohl, H. Hansen, C.-H.S.J. Chou, Regul. Toxicol. Pharmacol., 26 (1997) 322.

Chapter 2 Chemoreception
Biological facts, Measurements and Predictions

2.0. Introduction

Detection of airborne chemicals by humans relies on two sensory channels: olfaction and the so-called chemical senses (CCS) or chemical irritation sense.[1] Odour sensations are carried out by the olfactory nerve (cranial nerve I). They originate in specialised structures: the olfactory neurons located in the olfactory epithelium that covers a relatively small patch in the upper back portion of the nasal cavity. Unlike smell, the CCS lacks specific receptor structures and is widely distributed since there is general chemical sensitivity in all body mucosae. Of special relevance to sensory irritation generated by indoor air are two face mucosae: conjunctival and nasal. Common chemical sensations from the eyes and the nose are mediated by free nerve endings from the trigeminal nerve (cranial nerve V). An array of responses elicited in the nose by CCS stimulation are generally grouped together under the term nasal pungency. These pungent sensations include: stinging, prickliness, tingling, irritation, burning, piquancy, and freshness, among others. CCS responses in the conjunctivae are referred to eye irritation, although most of the terms just mentioned can be applied to the eye.[1-3]

Measurements of the dose-response characteristics of sensory stimuli, including odorants and irritants, constitute a branch of science known as psychophysics. Psychophysical studies describe either the perceived intensity as a function of exposure level, psychophysical scaling, or the minimum exposure necessary for conscious perception to occur, threshold determination. A number of psychochemical techniques has been developed.[1-3] However, data from one study rarely seem trustworthy enough for another investigator to use them.[4]

Over the past ten years, Cometto-Muniz and co-workers have carried out a systematic investigation into thresholds for sensory irritation and odor, using panels of human subjects under carefully controlled conditions. [5-11] Of course, in order to obtain thresholds for nasal pungency, Cometto-Muniz and co-workers have had to study anosmics – persons with no sense of smell. As a result of their efforts, they now have a

database of some 50 VOCs for which chemosensory thresholds are known. However, the number of VOCs that can be encountered at home or in the work place is orders of magnitude larger. There are more than 100,000 industrial chemicals, and even if only a third could be classed as VOCs or semi-volatile organic compounds, it is obvious that experimental determination of potency towards humans cannot possibly be extended to more than a very small proportion. The use of animal experiments does allow the study of VOCs that are too toxic to be tested on humans, but does not help very much as regards the sheer number of compounds examined. The comprehensive survey of Schaper on upper respiratory tract irritation on mice includes data on only 244 VOCs.[12]

There is therefore a very definite need for some type of prediction of the potency of VOCs towards humans. Even if restricted to VOCs that act through 'physical' mechanisms, rather than through 'chemical' or 'reactive' mechanisms, such predictions would considerably help to fill the gap between the relatively small number of VOCs studied to date, and the very large number of chemicals that could be encountered.

First a summary of the biochemical study of olfaction is presented. Biological information on nasal and eye irritation is also reported. Some methods available to measure odor and sensory irritation are listed. Attention is mainly focused on standardised methods developed by Cometto-Muniz and Cain to measure odor detection thresholds, nasal pungency thresholds and eye irritation thresholds. Finally, quantitative structure activity relationships, QSARs, developed to correlate chemosensory thresholds, are introduced. QSARs based on the Abraham solvation method are emphasised; they are quite simple algorithms that can be used for prediction of a very large number of additional threshold values.

2.1. Sensory Channels - Generality

2.1.1. Olfaction

2.1.1.1. Biological facts

In mammals, odorants are detected by the olfactory epithelium. Two others organs are thought to be involved in olfaction, they are the trigeminal nerves and the vomeronasal organ, see Figure 2.1. The role of the vomeronasal organ is however still unknown.[13]

The olfactory epithelium is a 5 cm^2 patch of yellow tissue located at the top of the nasal cavity approximately on a level with eye.[13] The mucosa is roughly 75 μm thick and is host to between 10^7 to 10^8 receptor cells distributed across the tissue Olfactory cells are bipolar neurons unique in the central neuron system due to their ability to regenerate every 30-45 days.[14] From the dendrite end of each neuron, there are several hairlike appendages called cilia, which extend outward into the mucus that covers and protects the olfactory epithelium.[15] The primary events of odor discrimination are thought to involve the interaction of odorous molecules with specific receptors on the cilia. The olfactory neurones also give rise to axons that are bundled to traverse the cruciform plate reaching the olfactory bulb in the brain. In the bulb these axons synapse onto secondary neurons known as mitra cells. This synapse, which is known as a glomerulus, is complicated and consists of a single mitral cell upon which 500 olfactory axons converge; from these signals are relayed to higher regions of the brain. The peripheral location of olfactory neurons, their remarkable capacity for post-natal regeneration and their direct axon link to the brain sets olfactory neurons apart from other neurons of the central nervous system. Figure 2.2. is a schematic of the anatomy of the human olfactory. A detailed account of the anatomy of the human olfactory neuroepithelium is provided by Morrison and Moran.[16]

In the mid-1980s, the biochemistry of olfaction was stimulated by the discovery that the level of cyclic adenosine monophosphate, cAMP, in isolated olfactory cilia increased rapidly when the cilia were exposed to certain odorants. It was subsequently demonstrated that odorants that have no effects on the level of cAMP induced a rapid change in the levels of inositol triphosphate, IP_3.

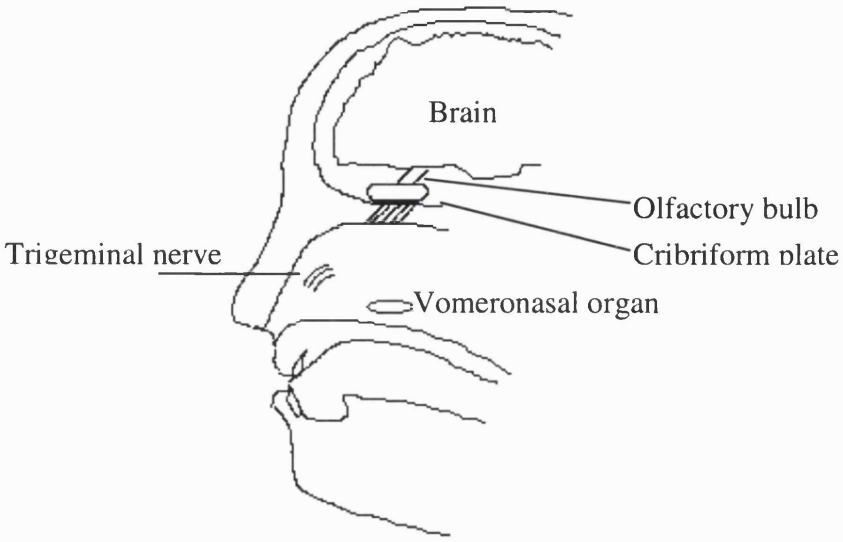


Figure 2.1. Keys organs thought to be involved in the perception of odour

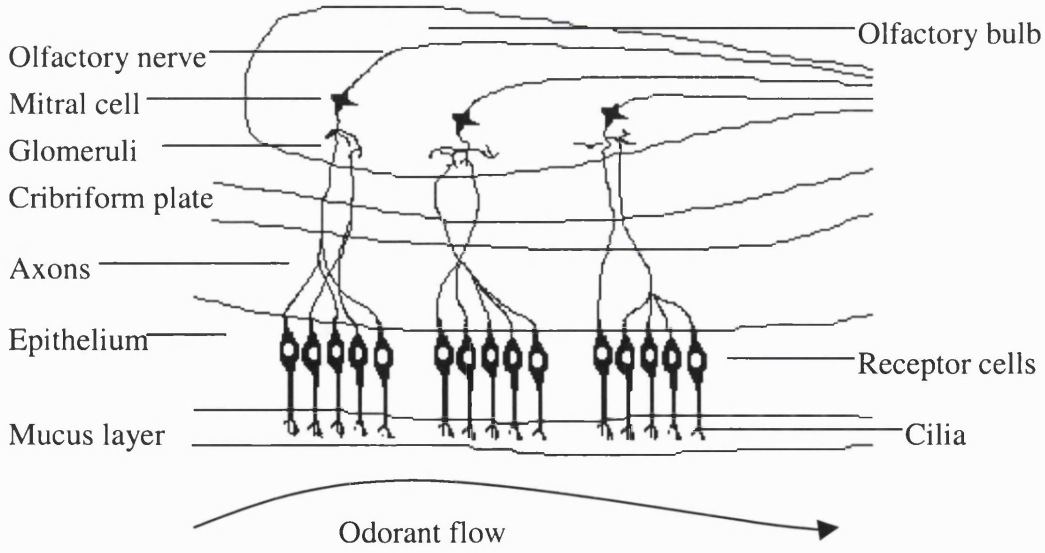


Figure 2.2. Convergence of receptor cells signals on to glomeruli on the olfactory bulb.

It is well known that these two secondary messengers cause the opening of ion channels. Influx of positive ions into the cells initiates a decrease in voltage across the cell membrane, which ultimately results in the generation of a nerve impulse. The signal then runs to the olfactory bulb, which in turn, sends impulses to the primary olfactory cortex. From there, information is relayed both to higher cortical areas and to the limbic system, which controls emotion. The former processes the odor sensations so that subjects can recognize them, for a review see reference 17.

The above findings on the odor –induced cAMP and IP_3 responses fit well with the discovery that guanine nucleotide binding proteins, G-proteins, were also involved in olfactory transduction. G-proteins mediate guanine triphosphate, GTP, dependent responses which act as intermediaries between ligand receptor and targets such as adenyl cyclase and phospholipase C, the intracellular enzymes responsible for the production of cAMP and IP_3 respectively. Similar types of transduction process are linked to receptor proteins that are inserted into the cell membrane and that cross the membrane in seven places, see Figure 2.3. Hence, it was postulated that odor receptors might also belong to this seven-helix family of G-protein-coupled receptors.[18] G-proteins are cell membrane proteins, which act as intermediaries between receptor activation and the subsequent activation of it second messenger. They have three different subunits. The portions that pass through the membrane exist as α-helices. These α-helices come together to form a cylinder and it is thought that the binding sites lie inside the cylinder. This is consistent with hormone receptors and with optical receptors in which the retinal-derived pigment is held in such a way. [18]

In 1991, Buck and Axel[19] discovered the family of transmembrane proteins that were believed to be the odor receptors and some of the genes that encode them. They cloned and characterised eighteen different members of an extremely large multigene family that encodes the seven transmembrane proteins whose expression were restricted to the olfactory epithelium. The proteins found all contained the seven helical transmembrane structure and contained sequence similarity to other members of the 'G-protein' linked receptor family. It is now estimated that there are between 500-750 odorant receptor genes in humans.

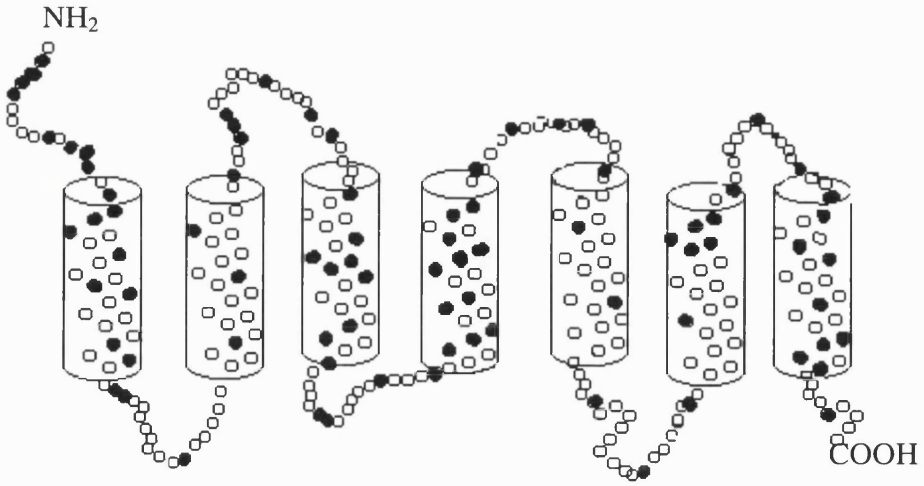


Figure 2.3. Schematic representation of a putative odorant receptor. The N-terminus is located extracellularly and the C-terminus intracellularly. The vertical cylinders delineate the seven putative α helices spanning the membrane.

The olfactory proteins are still poorly characterised and their binding affinity for odorants has yet to be demonstrated, although there is suggestive evidence that at least one of these receptors can actually respond to odor molecules.[20,21] Studies based on molecular modelling have been carried out on the OR5 and OR17 olfactory receptor genes.[22,23] First, it was pointed out that the OR5 gene product mediates the activation of G-protein-coupled IP_3 production by odorants such as lilial and lyral, methoxypyrazines and carboxylic acids.[22] Furthermore, Singer and Shepherd tested docking of the odor molecule lyral in simulated docking experiments using molecular modeling.[23] The results pointed to specific ligand-binding residues distributed in helices 3 through to 7 that form a binding pocket of approximately 12 Å from the extracellular surface of the receptor. This work supports the work of Axel and Buck[19] that helices 3, 4 and 5, which exhibit wide sequence diversity from one receptor to another, are involved in odorant binding. Generalisation of this finding to other olfactory receptors is supported by the recent observation by Lancet and Ben-Arie[24] that the OR5 receptor shares homology with a wide range of olfactory receptors found in mammalian species, including rat, dog and human. Using a similar approach, Singer constructed a molecular model of the OR17 receptor. Octanal was automatically docked in the model.[25] These results

predicted an odor-binding pocket of around 10Å from the extracellular surface. However, the modelling of this important class of proteins is highly speculative because of insufficient knowledge of the olfactory mechanisms.[15]

Currently the number of receptor types is estimated to be 500-750. This raises the question of how are these receptor are distributed in the olfactory epithelium. Are individual members of the receptor family expressed by every olfactory neuron or by only a subset of neurons? The latter is supported by evidence. Each sensory neuron expresses only one receptor and is therefore functionally distinct. Recent studies have shown that the olfactory epithelium is divided into broad zones according the types of receptors found in each zone, but that within each zone there is a random distribution of neurons expressing the same receptors and that members of each subset send their axons to only one or a few of some 2000 glomeruli.[26]

The olfactory mucus is roughly 10-30 μm thick.[13] In addition to protect the mucosa against damage by xenobiotics or microorganisms, the mucus has to provide a favourable milieu into which odorants can partition and gain access to and interact with receptors. The discovery in the mucus of small acidic proteins that accommodate hydrophobic molecules in aqueous environment provided new hopes in odorant transport process understanding. Odorant binding proteins, OBPs, belong to one of the most abundant class of proteins found in the olfactory apparatus. By their amino acid sequence, they belong to the family of lipocalins, generally involved in the transport of hydrophobic ligands.[13] Their concentration in mucus and respiratory epithelium of mammals is estimated to be high: 1% of soluble nasal proteins. OBP is a homodimer that has a central pocket lined with hydrophobic residues of which high proportion are aromatics. The central pocket has dimensions of 11*10*7 Å, i.e. 770 $Å^3$with an opening size of 6*7 Å, although a much larger cavity of 1100*1300 Å has been suggested.[27]

Assuming their involvement in the olfaction process, different physiological roles for OBPs have been proposed. Due to the hydrophobic nature of the odorant molecules and of the lipocalin pocket, it has been proposed that OBPs could carry the odorant molecules from the air to the olfactory receptors through the aqueous barrier of the mucus. It is also proposed that OBP-odorant complexes interact with the receptors. Subsequent to activation of the receptors, OBP may also be involved in removing odorants from the receptor sites to allow termination of the olfactory signal.[13] Furthermore, based on the relatively weak affinity of OBPs for odorant molecules, K_{diss}

value in the micromolar range, it has been proposed that OBPs would avoid the saturation of the olfactory receptor when odorant molecule are present at high concentration. It has also been demonstrated that OBPs bind with good affinity to 6-11 carbon alkylic aldehydes.[28]

2.1.1.2. Theories of olfaction

Over the recent years a number of theories have been proposed in an attempt to explain how humans perceive odors. Here are reviewed the most prominent theories.

The Stereochemical Theory of Odor: In 1946 Pauling indicated that a specific odor quality is due to the molecular shape and size of the chemical.[29] Similarly, he extended the idea of a 'Steric Theory of Odor' originally proposed by Moncrieff in 1949 that stated airborne chemical molecules are smelled when they fit into certain complimentary receptor sites on the olfactory receptor system.[30] In 1952, Amoore postulated there is a limited number of receptor types, each of these recognizes a particular molecular shape and, when triggered, generates the signal for a primary odor.[31] He proposed primary odors, viz. ethereal, camphoraceous, musky, floral, minty, pungent and putrid. The molecular volume and shape similarity of typical odorants of each class and proposed shaped receptors for the first five and the generation of charged species for the last two, positive for pungent and negative for putrid. Amoore subsequently carried out 'specific anosmia' experiments in an attempt to prove the existence of primary odors and identify all them.[32] This work was inspired by Guillot's suggestions that specific anosmia, the inability to detect one particular odor, was due to the affected person lacking one of the primary odor receptors.[33] One of the main objections to the stereochemical theory is that there are many examples of substances that have a similar shape but different odors because of a difference in functional group.

The Shape Theory: In 1957, Beets proposed the concept of modality instead of primary odor and used a more statistical approach. He considered that a pure odorant is composed of identical molecules arriving in the vicinity of the receptor sites with different orientations and different conformations. These differently oriented molecules interact with a variety of sites to produce a set of interactions, finally generating a

pattern of information containing several modalities at various levels of intensity and corresponding to a given odor.[34]

The Penetrating and Puncturing Theory: In 1967, Theimer and Davies proposed that odorant molecules must be absorbed into a thin-walled site on the olfactory nerve ending, penetrating it and leading to the formation of a small hole in the membrane. Accordingly, the theory stresses the primary importance of rates of desorption and molecular cross-sectional areas.[35]

The Vibration Theory: In 1938, Dyson suggested that the infrared resonance, which is a measurement of a molecules vibration, might be associated with odor.[36] This idea was popularised by Wright in the mid 1950's as infrared spectrophotometers became generally available for such spectral measurements, which Wright was able to correlate with certain odorants.[37] By the mid-70's it appeared that the theory of Wright had failed a critical test. The optical enantiomers of menthol and of carvone smelled distinctly different, although the corresponding infrared spectra were identical. And so this theory fell from favour.

Vibrational Induced Electron Tunelling Spectroscope Theory: In 1996, Turin postulated the presence of an electric potential gap in a protein, with nicotinamide adenine dinucleotide and zinc ions providing the electrodes. Electrons cannot cross the gap unless an odorant molecule is placed between the so-called electrodes. To cross the gap the electron must lose energy and this it does by tunnelling through the orbitals of the odorant molecule and exciting vibrational modes in it as a result. Thus Turin has moved the search for correlations from the infrared to inelastic electron tunnelling spectra.[38]

2.1.1.3. A combinatorial Process for odor interpretation

In 1999, Buck and co-workers appeared to have unravelled the mystery of how the nose can interpret abundance of different odors.[39] They showed that the sense of smell in mammals is based on a combinatorial approach to recognising and processing odors. Instead of dedicating an individual odor receptor to a specific odor, the olfactory system uses an alphabet of receptors to create a specific smell response within the neurons of the brain. The olfactory system seems to use combinations of receptors to

greatly reduce the number of receptor types actually required to convey a broad range of odors. In the reported study, individual mouse neurons were exposed to a range of odorants. Using a technique called calcium imaging, the researchers detected which nerve cells were stimulated by a particular odor. Using this approach, it was shown that:

(i). Single receptors can recognise multiple odorants,

(ii). A single odorant is typically recognised by multiple receptors,

(iii). Different odorants are recognised by different combinations of receptors, thus indicating that the olfactory system uses a combinatorial coding scheme to encode the identities of odors.

This new combinatorial approach is illustrated in Table 2.1.

This explains how 500-750 receptors can describe many thousands of different odors. Buck and her colleagues also demonstrated that even slight changes in chemical structure activate different combinations of receptors. Thus, octanol smells like oranges, but the similar compound octanoic acid smells like sweat. Similarly, it was found that large amounts of a chemical bind to a wider variety of receptors than do small amounts of the same chemical. This explains why a large quantity of the chemical indole smells putrid, while a trace of the same chemical smells flowery.[39]

2.1.1.4. Odor characterisation

Odor can be described using a number of different dimensions, viz. quality (floral woody), intensity (strong, moderate, weak) and threshold. Scientists studying human olfaction have developed physical, physiological and sensory techniques for two purposes: to study the human olfactory process and to quantify the sensory activity of chemicals.[15,40]

Physical methods

Physical instruments have been designed to measure both odor quality and intensity. Electronic noses are the most prominent example. They are capable of identifying the aroma from different brands of coffee. This modern technique consists of an array of gas-sensitive semi-conductors, which are connected to a neural network. The signals recorded by the sensors produce a characteristic pattern for a given odor.

Table 2.1. Comparison of the Receptor Codes for Odorants that have Similar Structures but Different Odors[a]

	S1	S3	S6	S18	S19	S25	S41	S46	S50	S51	S79	S83	S85	S86	
Hexanoic acid					■										Rancid, sweaty,sour,goat-like, fatty
Hexanol		■				■									Sweet, herbal, woody, cognac, Scotch, whiskey
Heptanoic acid	■			■	■		■			■	■				Rancid, sweaty sour, fatty
Heptanol		■			■	■									Violet, sweet, woody, herbal, fresh, fatty
Octanoic acid	■			■	■		■	■		■	■	■			Rancid, sour, repulsive, sweaty, fatty
Octanol				■	■		■			■					Sweet, orange, rose, fatty, fresh, powerful, waxy
Nonanoic acid	■			■	■		■	■		■		■		■	Waxy, cheese, nut-like, fatty
Nonanol				■	■		■			■		■			Fresh, rose, oily floral, odor of citronella oil, fatty

[a]Aliphatic acids and alcohols with the same same carbon chains were recognised by different combinations of Ors, thus providing a potential explanation for why they are perceived as having strikingly different odors. Taken from Ref.39. (Ors Oder Receptors)

The electronic nose can be used not only for odor characterisation, but also for the quantitative determination of the concentrations of individual molecules in a complex environment. Although the concept of electronic nose has been developed to imitate the signal processing in the biological noses, in practice the elements of signal transduction, signal processing, identification of chemical patterns currently differ from those believed to be present in the biological system. Comparisons have been made between the human nose and electronic noses. In all cases, the human nose was more sensitive than its electronic counterpart. The response to concentration changes was also different for the two "noses". The electronic nose was shown to respond linearly, whereas the human nose responds logarithmically. Neuner-Jehle and Etzweiler have reviewed theory and the prospect of development of electronic noses. [41]

Physiological method

Measurable physiological responses to odor stimuli include changes of electrical potential at the olfactory bulb or olfactory receptor. Attempts have been made to correlate such successful electrical activity to odor perception. It has been shown that the intensity of an odorant stimulus is related to the amplitude of a DC-recorder electrical potential and to the frequencies of an AC-recorded electrical impulse.[15] Most of the studies have been carried out on animals. Electrical activity in the human brain has been developed to study the unpleasantness of odors. Results have been reviewed.[15]

Sensory methods

Another way to obtain a human's perceptual response to an odor is to ask him for a verbal expression of the odor intensity and quality.

Quality

Can you measure the difference between one kind of smell and another ? The tests used to describe odor quality are known as odor profiling tests. These are the most complex of the sensory tests and, to ensure good quality, accurate and reproducible data are only carried out by highly trained and experienced sensory panellists.[42]

Intensity

How strong is an odor? A number of different types of sensory scales are used to measure perceived intensity of odor. The most widely used technique of psychophysical

scaling is the method of magnitude estimation. With this method, subjects are presented with randomly varying concentrations of an odorant. The respondent ranks the first stimulus with an arbitrarily chosen number; subsequent stimuli are then assigned numbers on the respondent of their own scale that reflect the perceived intensity compared with the first stimulus. After normalisation to a common scale, the data tend to reflect a similar dose-response slope, or psychophysical function, across subject. Another method, that of intensity matching, involves trained observers who match the perceived odor intensity of a given concentration of a test compound with known concentrations of an index compound, usually n-butanol. Intensity estimates so generated are said to be normalised to the butanol scale.[42]

Two laws, Fechner's law and Stevens' law, describing sensorial perception development as a function of stimulus intensity are important when measuring odor intensity. Let C stand for concentration, R for sensorial response intensity; then the two laws are respectively expressed by the following equations:[42]

Fechner's Law

$$R - Ro = n (\log C - \log Co) \qquad (2.1)$$

Stevens' Law

$$\log R - \log Ro = n (\log C - \log Co) \qquad (2.2)$$

The first function is said to be logarithmic and the second a power function, since one can also write

$$R/Ro = (C/Co)^n \qquad (2.3)$$

This law states that equal changes in stimulus magnitude, C, produce the corresponding change in perceived intensity, R. The increase in perceived intensity with concentration can be represented by a straight line, as shown in Figure 2.3, for odorants A and B. The slope indicates how fast odor intensity rises with concentration and the intercept defines the detection threshold. There are important characteristics of this type of data:

- At higher concentration a larger increase in concentration is necessary to give the same change in perceived intensity.

- At different concentrations of two odorants the rank order of their perceived odor intensity can change. This is true for odorants A and B in Figure 2.4. At log Cy, odorant A is stronger that odorant B but at log Cx odorant B is perceived as stronger than odorant A.

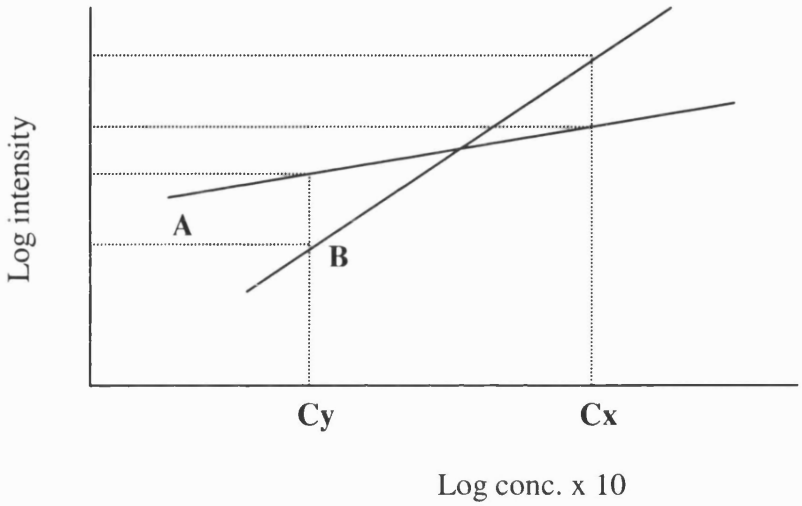


Figure 2.4. Perceived intensity vs. log perfume concentration for two odorants.

Thresholds

Threshold itself can be described at three different levels:

- Detection threshold. Can odor be detected?
- Recognition threshold. Can an odor be identified?
- Difference threshold. Is there a difference between two odors?

The detection threshold is the lowest stimulus intensity (odor concentration) that the subject can distinguish from an odor-free situation. The subject's response indicates whether the presence of an odor has been perceived or not. Correspondingly, the recognition threshold is the minimum concentration at which an odor can be identified. Such a threshold typically occurs at a concentration about threefold above the point of detection.[42]

Over the years, data have been accumulated by many investigators on the human detection thresholds of several hundred organic compounds. Laffort tabulated the results obtained by 24 earlier authors on the thresholds for 192 substances, usually by means of air-dilution olfactometers. Authors at this laboratory have generally employed water dilutions of odorants in so-called sniff-tests. The method of Guadagni et al. using pairs of Teflon squeeze-bottles has been applied to more than 400 compounds. The method of

Amoore et al., which uses sets of five glass-stoppered flasks has been tested with 146 compounds.[4]

Olfactory thresholds present an extraordinary person-to-person variability. The reported values for the same compound and different subjects may vary within three to five orders of magnitude. Recently, Guichard et al.[43] showed that even with a pair of enantiomers the sensibility may vary from person to person and moreover be completely different for the two enantiomers.[43] Hence, a direct comparison among thresholds obtained with different panels is not correct since the agreement between measures obtained using different techniques is often quite poor. For example Amoore and Hautala noted that the 29 then published olfactory thresholds for n-butanol vary from $1.45.10^{-7}$ to $1.88.10^{-4}$ $g.dm^{-3}$.[44] In an attempt to increase the reliability of published thresholds Laffort, Patte et al. and finally Devos et al established from several compilations a list of standardised values for a series of 529 compounds.[45-47]

The combination of gas chromatography, GC, with olfactometry, O, is an important tool in the analytical chemistry of flavour and fragrance materials. This technique consists of a GC with a chemical detector followed by or in parallel with a 'sniff port' and a trained human sniffer. Provided that the GC conditions adequately separate the components of a mixture, each component can be smelt in an olfactory pure state at the exit port of a GC column. It is not uncommon for an odor to be perceived at a position in the gas chromatogram where there is no peak. An area in which this technique has found wide applicability is the analysis of complex natural products. GCO dilution analyses, such as CharmAnalysis and Aroma Extraction Dilution Analysis, AEDA, can also be used to determine the odor threshold values of an odorant. Acree has recently reviewed developments in gas chromatography / olfactometry.[48]

2.1.2. Chemical Senses

The trigeminal system mediates a variety of protective reflex responses to potentially life-threatening chemical irritants. Sensory irritation is related to chemical reactivity. For an irritant gas or vapour, one speaks of the substance's warning properties, or its ability to produce immediate upper respiratory irritation. Such irritation triggers protective physiologic reflexes, alerts the exposed individual to danger, and initiates escape behaviour. For a given degree of chemical reactivity, the warning

properties of a gas or vapor tend to correlate with its water solubility. Examples of irritants that have good warning properties, i.e. high water solubility, are ammonia and sulfur dioxide.[1-3] Alarie uses the term sensory irritant or upper-respiratory tract irritant to denote irritant gases or vapors that have good warning properties.[49]

The irritation effect of gases and vapors is believed to be caused by their direct interaction with one or more receptors on trigeminal nerve endings in the cornea and nasal mucosa. The chemosensitive trigeminal nerves are C-fibres and possibly also δA-fibres. They are part of the somatic sensory system, which conveys peripheral impulses to the central nervous system. Stimulation of the trigeminal nerve endings results in a stinging sensation, which can increase to a burning and painful sensation.[1-3]

Unlike olfaction, which relies upon stimulation of hundreds of types of receptors to give the spectrum of odor quality, chemesthesis for nonreactive VOCs appears to rely upon stimulation of a very small variety. Indeed, one type of broadly tuned receptor may transduce the majority of VOCs, except for some special molecules, such as nicotine. One VOC might accordingly behave like another with respect to its dose-response ,i.e. psychometric, function.[1-3]

2.1.2.1. Bio-Assays for Nasal Pungency

Two main bio-assays have been developed for nasal pungency. First, there is the mouse bio-assays developed by Alarie, which is now a ASTM standard method.[50] In this assay a stream of air carrying a constant concentration of VOC is passed over the head of a mouse for 10-15 min, and the reduction in respiratory rate is obtained using a plethysmograph.[50] The VOC concentration in ppm that causes a 50% reduction rate is taken as the bio-assay endpoint, and denoted as RD_{50}. Values for some 150 VOCs have been listed in recent reviews. Secondly, there is the measurement of nasal pungency thresholds in humans in which the VOC is administered as a short 'sniff' lasting a few seconds.[5-11] The threshold is taken as the lowest concentration in ppm the subject can detect; these thresholds are denoted as NPT. Cain and Cometto-Muniz[5-11] have measured nasal pungency thresholds in persons lacking a functional sense of smell, viz. anosmics, for whom odor does not interfere. Values of NPT for 50 VOCs have been recently listed.[11]

Some observations have indicated that normosmics could localise the nostril stimulated by a chemical only when it triggered chemesthesis. Fortunately, the concentration at which a person can localise the nostril coincides almost exactly with the concentration for chemesthetic detection. This approach produces sensory data for chemesthetic detection free of olfactory sensations.

The negative mucosal potential seems to offer an objective measure of sensory irritation. Pungency stimulation can elicit a surface potential from the nasal respiratory mucosa. Because of its predominant negative peak, with amplitude up to hundreds of microvolts, the response has become known as the negative mucosal potential. Recorded from a limited region of the septum, the signal represents most likely the aggregate receptor potential of many thousands of free nerve endings of the stimulation. Kobal and colleagues ruled out various epiphenomenal sources, such as blood flow, olfaction, and activity from sympathetic fibers. The finding that the NMP correlates closely with feelings of irritation, expressed in ratings of magnitude, argues for a trigeminal source.[51]

2.1.2.2. Bio-assays for Eye Irritation

Two approaches have been used to determine eye irritation thresholds. First, the rabbit eye irritation test developed by Draize, Woodard and Calvery[52] is still regarded as the accepted standard for assessing eye irritation hazard. In the Draize test the substance under study is applied to the eye of a living rabbit. The effects of the substances on the cornea, iris and conjunctivae are graded on individual scales and given weighted scores. The final eye irritation score is the sum of the weighted scores for the cornea, iris and conjunctivae. Data from the rabbit eye irritation test have been acceptable for assessing the hazard of chemical irritation to the eye because of two assumptions: (1) substances that are irritant to the rabbit eye are also irritant to the human eye because the mechanisms which operate in rabbit eye irritation also generally operate in human irritation; (2) the rabbit eye test is thought to overestimate the human hazard and so provide a margin of safety to protect human health. Eye irritation in humans, EIT in ppm, have been gathered for 17 VOCs. Noteworthy, the human and animal differ in an important way.[1-11] The human data came from vapor phase stimulation with various VOCs, whereas the animal data came from direct application of VOCs as bulk liquid.

2.2. Standardised method for odor detection, nasal pungency and eye irritation thresholds

Cain and Cometto-Muniz have developed a standardised method for odor detection, nasal pungency and eye irritation threshold in humans of individual VOC as well as mixtures of VOCs.[1-3,5,11]

Cain and Cometto-Muniz measured nasal pungency thresholds in persons lacking a functional sense of smell, viz. anosmics, for whom odor did not interfere. Odor thresholds, on the other hand, were measured on subjects with normal olfaction, viz. normosmics, matched to the anosmics by age, gender, and smoking status, which are variables known to affect chemosensory sensitivity.

Cometto-Muniz and Cain have designed a two-alternative forced-choice procedure with presentation of increasing concentrations served to measure all chemosensory thresholds. The method requires participants to use the assigned nostril to select on each trial, the stronger of the two stimuli: one is always a blank, solvent, and the other a dilution step of the tested substance, starting with a concentration clearly below threshold. Selection of the blank triggers, on the following trial, presentation of the next step, i.e. a higher concentration, paired with a blank. Selection of the stimulus triggers, on the following trial, presentation of the same dilution step from a duplicate set, also paired with a blank. The procedure continues until the subject selects the substance over the blank five times in a row. The dilution step where this first occurs is taken as the threshold. Once the thresholds is measured for one nostril or eye, the other nostril or eye is tested. After this, testing begins with another stimulus. If the experimenter were testing localisation on the right nostril, he/she would be asked which presentation led to a stronger sensation in the right nostril. If the experimenter were testing localisation on the left nostril, he/she would be asked which presentation led to a stronger sensation in the left nostril. Next for all types of sensory endpoints, steps of increasing concentration will continue to be presented four times, paired with blanks, until the subject achieves 100 % detection at two successive concentrations. This is taken as the detection threshold.

Stimuli are delivered from cylindrical, squeezable, high-density polyethylene bottles (270-cm^3 capacity) containing 30 cm^3 of solution, see Figure 2.5. For measurement of odor and nasal pungency, the bottle closure has a pop-up spout that

fitted into the nostril being tested. Each nostril is tested separately. For measurements of eye irritation, the bottle caps held a tube that led a 25 cm^3 measuring chamber, the rim of which is placed around the eye. Each eye is tested separately. A squeeze of the bottle delivered a puff of vapor into the measuring chamber where the eye was exposed. A polyethelene dust cover closed the open end of the measuring chamber when the bottle is not in use. Cometto-Muniz and co-workers have recently devised a new stimulus delivery system, or glass vessel, and used it to measure nasal pungency threshold in anosmics. The design of the 270-cm^3 glass vessels aimed at producing environmentally realistic thresholds through the following improvement of the squeeze bottles: (i) the avoidance of dilution of the stimulus by providing a tight nose-piece-nostril connection; (ii) increase in the volume of stimulus vapour available to accommodate a human sniff completely; and (iii) elimination of any low-odor background inherent in the plastic squeeze bottles. The results showed that nasal pungency thresholds measured via the glass vessels were significantly lower than those measured via the squeeze bottles while showing the same trend within the members of each series.[53]

Quantification of all vapour stimuli, whether in the squeeze bottles or in glass vessels, is achieved through gas chromatography (FID detector) by direct sampling via a gas sampling valve or a gas-tight syringe. Concentration measurements are made off-line right after preparation of the stimuli, concomitantly with testing and after all subjects had been tested, to confirm stability. All readings are then referred to those of the undiluted chemical, assumed to represent saturated vapour at room temperature, 296K.

In addition to measurement of threshold values for single VOC, Cometto-Muniz determined dose-response, i.e. detectability, functions for binary mixtures via the two-alternative forced-choice procedure with an ascending concentration order of presentation.[54,55] To obtain the stimulus-response, i.e. psychometric, functions for individual compounds, a series of twofold dilution steps of the individual compounds, a series of twofold dilution steps of the undiluted chemical (100% v/v) is prepared, i.e. 50,25,12.5,6.25, etc. % v/v. If two chemicals A and B have different chemosensory potency, their detectability functions will be displaced along the concentration axis with the function for the more potent chemical displaced to the left.

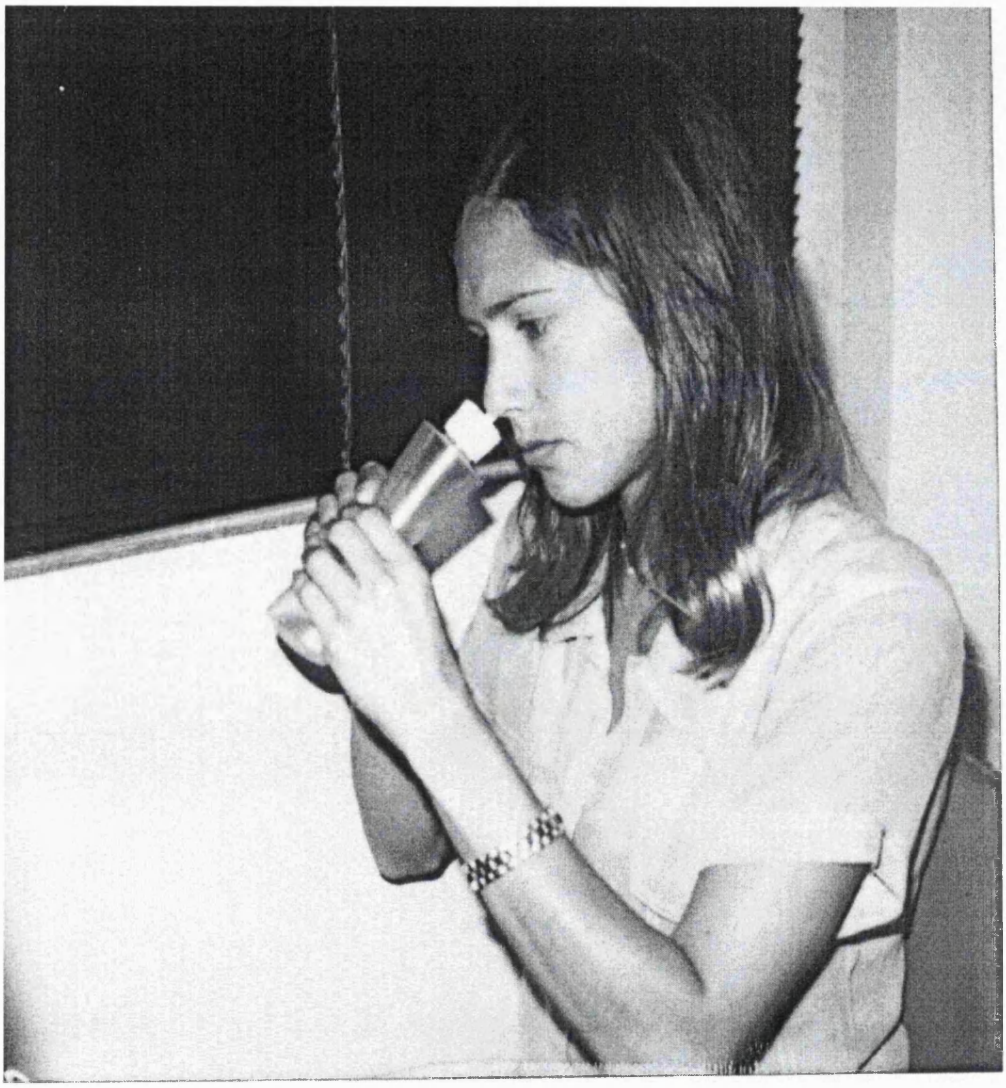


Figure 2.5. Monorhinic nasal testing via pop-up spout on the cap of the squeeze bottle.

Having these functions, for any concentration of one chemical, one can interpolate the corresponding probability of detection (p) for that concentration. With this value of "p" an interpolation in the function for the other chemical can be done in order to obtain the sensory equivalent concentration of each chemical. In summary, Cometto-muniz and co-workers found the concentrations of this second chemical, e.g. p = 0.60. These two concentrations were called sensory equivalent. This procedure allows the scientist to express the mixture of a certain concentration of chemical A with a certain concentration of chemical B in terms of concentration of only one chemical, A or B. For instance, the concentration of chemical B into can be transformed into the sensory equivalent concentration of chemical A. Then, the two concentrations are added for

each of the various mixtures and compare their trends in detectability with the corresponding trend for detectability of the single chemical.

2.2.1. Chemosensory detectability of single chemical

Detectability functions for single compounds across a wide and orderly array of substances, i.e. homologous series can provide insight into the functional properties of each sensory modality investigated. They can also provide an additional way to probe the role of physicochemical properties, arranged as a continuum along homologous series, on the sensory potency of chemicals. Odor thresholds are always below, most times well below, nasal pungency and eye irritation thresholds. The investigations have also shown that all sensory thresholds tend to decline with carbon chain length in homologous series. Quite often, though, and particularly for the three to five members, odor thresholds decline faster than nasal pungency thresholds with carbon chain length. Therefore the gap between odor and pungency grows larger across those first few members of each series. Across all the series studied, the size of the gap has varied between one and five orders of magnitude. The decline in odor thresholds has tended to approach a plateau whereas the decline in nasal pungency thresholds has tended to show an abrupt cut-off. From certain homologue on, e.g. octan-1-ol, octyl acetate and propyl benzene, a threshold for nasal pungency failed to be reached, even with presentation of undiluted chemical. As stressed by Cometto-Muniz and Cain, substances not usually regarded as irritants, e.g. heptan-1-ol, nonan-2-one, not only can be detected by anosmics, i.e. they have pungency, but their pungency threshold is lower than that of more typical irritants, i.e. methanol, acetone.[5-11]

Eye irritation thresholds, as a rule, fell into the register with nasal pungency thresholds. Nevertheless, for some of the homologous series studied, the highest member, i.e. longer chain-length, could not evoke nasal pungency but did evoke eye irritation. This might indicate broader chemesthetic responsiveness to airborne chemicals in the ocular mucosa than in the nasal mucosa.[5-11]

Nasal localisation, i.e. determination of whether a vapor has entered the right nostril or the left, has provided a way to measure chemesthetic potency in normosmic persons. Nasal localisation thresholds for homologous n-alcohols and for selected terpenes tended to agree between normosmic and anosmic. In turn, these thresholds were either the same or only slightly higher, typically less than half an order of

magnitude, than the respective nasal pungency thresholds, but were as expected substantially higher than the respective odor thresholds.[5-11]

Both the nasal thresholds (pungency and odor) for the initial set of 42 VOCs were plotted as a function of a simple physicochemical property, e.g. saturated vapor concentration at room temperature, 296K. Pungency thresholds taken as a whole exhibit a linear relationship with saturated vapour, r = 0.97, having a slope of 1.02. Pungency for individual homologous series conforms to the general picture (slopes: 0.90 for the alcohols, 1.07 for the acetates, 1.06 for the ketones, 1.17 for the alkylbenzenes, and 0.95 for miscellaneous compounds. The linear correlation with slope close to 1.00 suggests that when a certain uniform percentage of vapour saturation is achieved, nasal pungency would occur in the anosmics if it were to be evoked at all.[1-3,5-11]

On the other hand, other thresholds as a whole depicted a more substance-to-substance scatter than pungency thresholds. They generally failed to show a linear relationship with saturated vapour. The odour for no individual homologous series exhibited a strong a correlation with vapour saturation as did the pungency thresholds for all series and miscellaneous substances grouped together. For odor, the best correlation occurred with the ketones, where r = 0.95, and the worst for acetates, r = 0.87. The slopes of the relationships for odor thresholds commonly departed from unity.[1,2,3,5-11]

2.2.2. Chemosensory detectability of mixtures

The chemosensory discomfort experienced in any place may in principle come from a single chemical compound but more commonly comes from a mixture. Hence, in order to understand, and predict sensory impact, both single compounds and mixtures require study. The level of understanding of the impact of single compounds has reached a state of maturity that quite naturally invites attention to mixtures, though at this point only simple ones.

Odor responses, at suprathreshold level, to very simple mixtures of volatile compounds have been the subject of a number of investigations on human olfaction.[56-59] The typical outcome has shown that the intensity of a mixture is significantly less than the sum of the intensities of the components. The particular combinatorial rules to predict odor intensity of mixtures have proven straightforward in the binary case. Hence from knowledge of the perceived intensities of unmixed components, the perceived

intensity of a mixture can be calculated with reasonable precision. At suprathreshold level, odors seem to antagonise one another, apparently more as they become stronger. One might imagine that antagonism would therefore diminish toward threshold and this seems roughly to be so.

Comparatively few studies have explored odor perception of chemical mixtures at threshold level. The data generally suggest that a mixture reaches its odor threshold when each of its components lies below its individual threshold, a result that indicates some degree of agonism among components.[60-64] In fact, such odor agonism at threshold is usually larger than seen at suprathreshold level, a result in line with the observation that weak odors add more potently than stronger odors.[57] Cain and co-workers have also seen partial agonism among odorants for the perception of odor thresholds in mixtures of up to nine components.[65]

Chemesthesic responses, at suprathreshold level, to mixtures of substances, have received attention from studies done principally on rodents[66-68] and to a much lesser extent, on humans.[69-70] The studies on rodents have employed the classical respiratory depression technique developed by Alarie that measures RD_{50}.[71,72] Results have shown cases of competitive agonism[66-68] or antagonism[67] depending on the particular mixtures studied. The studies on humans have indicated, that in contrast to the outcome for odor, the suprathreshold nasal pungency evoked by formaldehyde and ammonia adds, or even, potentiates to produce the pungency of the binary mixture.

Regarding chemesthetic responses at threshold level for the respiratory depression technique, the use of an RD_0 value, an extrapolated threshold concentration for the respiration effect has been suggested.[71,72] This term was introduced for compounds with a low slope of the log concentration-response curve, and which, even at high concentrations, were not able to produce a 50% decrease in respiratory rate.[73] Pharmacologically, such compounds are considered partial agonists of the sensory irritation receptor, having a low intrinsic activity. It is interesting to speculate whether the decrease in intrinsic activity seen along some homologous[73] series reveals itself as a change on the quality of the threshold nasal pungency response of anosmics, that changes from a sharp, crisp sensation for lower homologous to a 'pastel' sensation for higher homologs.[7]

Cometto-Muniz and Cain measured sensory thresholds: odor, nasal pungency and eye irritation for individual VOCs (propan-1-ol, hexan-1-ol, ethyl acetate, heptyl acetate, pentan-2-one, heptan-2-one, toluene, ethyl benzene and propyl benzene) and

mixtures of them.[65] Various degree of stimulus agonism effects were observed for each of the three sensory channels when testing mixtures. As the number of components and the lipophilicity of such components in the mixtures increased, so did the degree of agonism. Synergistic stimulus agonism characterised the eye-irritation response for the most complex and the most lipophilic mixtures.

Studies of the functional properties of the olfactory and trigeminal chemosensory systems in humans have often focused on threshold measurements on a single point on a dose-response, i.e. detectability, function or on suprathreshold functions over a range of concentration. Odor and chemesthetic detectability functions provide a means to compare olfactory and trigeminal functionality as both chemosensory systems cross the boundary between threshold and suprathreshold responses. Using butan-1-ol and heptan-2-one and more recently butyl acetate and toluene, Cometto-Muniz et al, measured detectability functions for the odor, nasal pungency, and eye irritation of these four substances alone and in binary mixtures (butan-1-ol / heptan-2-one or butyl acetate / toluene).[54,55] When all stimuli, single and mixtures, were transformed into concentration units of one, or the other chemical, a single function could fit at all data from the same sensory end point with a correlation of 0.91 or higher. The outcome lends supports to the notion of chemosensory agonism, in the sense of dose additivity, between the members of binary mixtures presented at perithreshold levels.

2.3. Quantitative structure activity relationships for chemosensory functions

2.3.1. Background

The first studies of biological potency of gases and vapors were carried out from 1900 onward by two workers. H. Meyer investigated anaesthesia on mice and salamanders,[74] and was interested in the relationship between the potency of gases and vapors and the solubility of the anesthetic gases and vapors in organic liquids such as olive oil. Overton[75] showed for numerous series of compounds that the narcotic concentration, C', of aqueous solute toward the tadpole was related to the water/olive partition coefficient, P', with the latter defined as:

P' = (concentration in organic phase) / (concentration in water) (2.1)

In 1920, K.H. Meyer and H. Gottlieb-Billroth[76] expressed the relationship in the quantitative form,

C'.P' = constant (2.2)

They showed that the data of Overton led to a reasonably constant value of 0.05 for the product C'.P'. This work of K. H. Meyer and Gottlieb-Billroth represents the first quantitative structure-activity relationship ever reported.

Some years later, Ferguson advanced what has been the most widely used principle in gaseous narcosis, namely that the gaseous narcotic concentration, P^{nar}, is inversely proportional to the saturated vapor pressure, P^{o}, where P^{nar} is the partial pressure of a series of compounds giving rise to a particular effect.[77] Ferguson suggested that the existence of a direct correlation between gaseous toxic concentration and solubility or vapor pressure could be used as a criterion of a physical mechanism of toxicity, rather than a chemical mechanism. There are cases where the Ferguson rule holds quite well, especially for homologous or related series. Later on, Brink and Posternack generalised Ferguson's conclusions; they refer to the ratio P^{nar}/P^{o} as the thermodynamic activity of the narcotic and put forward the rule of 'equal narcotic effect

at equal thermodynamic activity'.[78] However, Abraham et al. [79] showed that the Ferguson rule has no thermodynamic basis, contrary to the position of Brink and Posternak. The authors set out conditions under which P^{nar}/P^{o} might be expected to be roughly constant and their investigations showed that such an observation is consistent with a receptor area in which the liquid narcosis solubilities are roughly constant. It was shown that the biological potency of nonreactive gases and vapors can be controlled either by an equilibrium between the agonist in a receptor or by an equilibrium between the agonist in the gas phase and the agonist in the receptor phase.[79]

Hydrophobic binding of chemicals to proteins, receptors and membranes can be regarded as specific cases of partition processes. This enabled Franke to develop a simple physicochemical method, which for homologous substances could distinguish between the two extreme possibilities: (i) the receptor site is totally within an hydrophobic environment, or (ii) the hydrophobic binding site is covered with water so that a molecule will adsorb to the receptor at one side of the molecule whereas the opposite side of the molecule will be in contact with the water layer. In the first case, it was shown from theoretical considerations and numerous examples that the slope of the relationship between the negative logarithm of the equipotent concentration and the logarithm of the octan-1-ol / water partition coefficient was approximately 1. In the latter case the slope was expected to be approximately equal to 0.5.[80]

In the next sections, will be presented an overview of various models and techniques developed in an attempt to predict odor thresholds, sensory potency in mice, nasal pungency and eye irritation.

2.3.2. Models for odor detection thresholds

As already pointed out, odor can be defined according to its quality, intensity and threshold. Here, attention is mainly drawn to odor detection thresholds as one of the tasks of this thesis will be to propose a predictive model based on such data. Rossiter[16] and Chastrette[40] have reviewed advances in structure odor relationships.

Numerous attempts have been made to understand the trends in odor detection thresholds that are displayed by the human olfactory sense. High odor detection thresholds are observed for most odorants that are gases under standard pressure and temperature conditions, whereas odorants with low vapor pressure generally have low odor detection thresholds. QSARs have been formulated in an attempt to correlate

trends in olfactory odor intensity with specific microscopic and macroscopic properties of various odorants.

Numerous studies have underlined the importance of particular functional groups and structural fragments in the fragrant activity of the organic molecules. Thus for examples, the influence of double bonds and some functional groups on ODT of some molecules has been put forward. Dravniek[81] showed that ODT values can be linearly correlated with the number of atoms in the odorants, the number of hydroxyl group, the number of nitrogen or sulphuric atoms, etc.. More recently, Devillers et al. described a set of 173 chemicals by means of 18 structural fragments known to have an influence on the ODT. The authors used a neural network to estimate the ODT of organic molecules. Other workers have empirically correlated trends in human ODT with boiling point of the liquid phase of the odorant species. Some workers have noted the correlation between ODT and the vapor pressure of the odorant.

Christoph and Drawert used threshold values obtained from GCO dilution analysis to study the relationship between structure and odor intensity for a series of homologous saturated aliphatic aldehydes, alcohols, acetates and ethyl esters, monoterpenes and alcohols. They found that the compounds with the lowest thresholds are molecules with 8,9 and 10 heavy atoms in the straight chain. The thresholds of acyclic monoterpenes were shown to be significantly lower than those of monocyclic terpenes and these in turn still lower than those of bicyclic compounds. An additional prerequisite for low threshold appeared to be branching of alkyl chains.

Branching was also considered to be associated with odour strength by Jurs and Edwards in their study of 58 structurally diverse compounds. This group of workers quantified odor intensity by determining the concentration (C) of these compounds required to produce an odor strength equivalent to that of butan-1-ol at 87ppm. The following criteria were also considered to be important in determining odor intensity:
(1) In the molecular weight range of their data set (41-168), increasing molecular weight led to a lower value of log (C).
(2) Odor intensity increased as the partial charge on the most negative atom became more negative
(3) Unsaturation in a molecule led to an increase in odor strength,

Jurs also focused on the strength of one individual character, sweetness of odor. No direct structure-sweetness relationship could be uncovered using this diverse set of 73 industrially important fragrance compounds. A logarithmic transformation of the

sweetness did provide substantially improved statistics. The difficulties encountered in this study prompted Jurs to criticise the procedures used to obtain the data: 120-150 panellists characterised the odor profiles of the fragrances using a glossary of 146 standard semantic descriptors. The model's limited applicability may also be a function of the immense structural diversity represented in the data set. By restricting his studies to a homologous series and by using odour threshold data from a single source, Jurs developed QSARs for a set of 53 alcohols and a set of 74 mono- and disubstituted pyrazines.[89-90]

Very recently, Turner and Willett put forward a PLS model for ODT using the EVA 3-D descriptors.[91] The EVA (**E**igen **VA**lues) descriptor is derived from fundamental infra-red and Raman range molecular vibrational frequencies. EVA is sensitive to 3-D structure, but has an advantage over field-based 3-D QSAR methods as it is invariant to both translation and rotation of the structures concerned and thus structural superposition is not required. EVA was originally developed as a descriptor for QSAR and it has been shown to perform well with a wide range of datasets. The EVA spectral descriptor has been applied to a selection of ODT values measured by Cometto-Muniz and Cain.[91] No details were given other than for 52 ODT values, the coefficient of cross-correlation, q^2, was 0.57 and for 44 ODT values q^2 was 0.71; unfortunately EVA's main limitation is that PLS regression models are very difficult to interpret in terms of molecular features and so cannot lead to any mechanistic conclusions.

The hypothesis of Laffort and co-workers is that the intermolecular forces involved in olfactory process are similar to those involved in solution.[92,93] In other words, these forces are only the van der Waals and hydrogen bonding types and not the highly specific lock-and –key type as generally encountered in pharmacology. Laffort and co-workers set out a gas-liquid chromatographic (GLC) method for the determination of solute descriptors. Using retention data on five stationary phases, they defined five descriptors:

α is an apolar factor, proportional to the molar volume at the boiling point.

ω is an orientation factor, proportional to the square of the dipole moment for simple molecules.

ε is an electron factor which for molecules with a regular distribution of electrons is proportional to the ratio between molar refraction and the molal volume at the boiling point.

π is a proton donor, or acidity factor.

β is a basicity factor.

These solubility factors were determined for 240 solutes and were then used as descriptors in a number of linear free energy relationships, LFERs. Laffort and Patte then examined the relative retention times as log $t_{R'}$, of solutes across the frog olfactory mucosa, together with the corresponding solubility factors.

$$\text{Log } t_{R'} = 0.71 + 4.80\,\beta - 0.63\,\pi - 0.99\,\omega \qquad (2.3)$$

$N = 12$, $r^2 = 0.902$, sd = 0.270, F = 25.4

Here and elsewhere, n is the number of data points, r is the correlation coefficient, sd is the standard deviation in the dependent variable, and F is the Fisher F-statistic. Interpretation of equation (2.3) is by no means easy. It is not clear what the origin is of the negative coefficients of p and w, because any solute-mucosa interaction should lead to a positive coefficient. It is also not reasonable that the coefficient of the solute basicity factor, b = 4.80, is so large, since this implies that the solvent acidity in equation (2.3) is larger than in the equation for gas / water partition, b = 3.40.[93] Yet water is an extremely strong hydrogen bond acid, and it is doubtful if any biological system is more acidic.

Yamanaka showed that odor thresholds of Davos et al.[96] for several homologous series could be correlated with the odorant activity coefficient in water, γ_w, through a set of equations of the type,

$$\text{Log } (1/\text{ODT}) = a \log \gamma_w + b \qquad (2.4)$$

where a and b differ for each homologous series.

A much more detailed analysis was carried out by Hau and Connell.[97] The authors proposed the following representation of a possible mechanism,

$$[\text{VOC}]\text{air} \xrightarrow{K^{AM}} [\text{VOC}]\text{mucus} \xrightarrow{K_{MB}} [\text{VOC}]\text{bio} \xrightarrow{K_R} \text{VOC-R}$$

Figure 2.5. Olfactory mechanism according to Hau and Connell

Here, [VOC]air is the concentration of an odorant, or VOC in air, [VOC]mucus is the concentration in the mucus, [VOC]bio is the VOC concentration in the biophase that contains the olfactory receptor, and VOC-R is the concentration of the VOC-receptor complex. The equilibrium constants for the three stages are denoted here as K^{AM}, K_{MB} and K_R. The partition coefficient between octan-1-ol and water, P^{oct}, and that between water and air, K^W, were used to model the partition process of VOCs into the biophase where the olfactory signal is transformed. QSARs were developed for four homologous series. The ODT values were from the AIHA compilation.[7] The model proposed is as follows,

$$\text{Log } [\{ODT\}K_W] = b + a \log P^{oct} \qquad (2.5)$$

Where a and b are coefficients depending on the compounds analysed. Hau and Connell[94] applied their model to four homologous series and a combination of compounds, see Table 2.2.

Table 2.2. Model for odor detection thresholds proposed by Hau and Connel (a,b)

Series	a	b	n	r
Alkyl acetate	-1.65	4.44	8	0.92
Alcohols	-1.65	5.19	9	0.84
Amines	-0.91	2.74	7	0.91
Ketones	-1.88	4.58	6	0.91
Various	-1.93	4.95	23	0.91

(a) adapted from Ref. [94]

(b) where n is the number of data point and r is the correlation coefficient.

Hau and Connell based their study on the theory for the binding of molecules to biological receptors developed by Franke.[80] Extrapolating this theory to slopes larger than unity, they postulated that slopes larger than one can occur if P^{oct} influences the partition process as well as the binding equilibrium K^A. Consequently, they proposed that the slope of the regression line is a measure of the hydrophobicity of the environment

surrounding the receptor site plus hydrophobic interaction between the odorant molecule and the actual receptor molecule. The larger-than-unity slopes for the acetates, alcohols and ketones are most likely due to the highly hydrophobic matrix surrounding the receptor site and the enhanced binding effect arising from the hydrophobic interaction between the receptor molecule and the odorant molecules.

2.3.3. Models for sensory irritating potency in mice

Alarie and co-workers have proposed equations to estimate the sensory irritating potency in mice of nonreactive VOCs by using several physicochemical variables, either singly or in combinations.[48-100] Earlier studies also attempted to estimate log RD_{50} from various physicochemical variables. Inclusion of a chemical as 'nonreactive' was based on previous findings that within homologous series such as alkylbenzenes and saturated alcohols, ratios of the chemical RD_{50} concentration (ppm) over the chemical saturated vapor pressure ate 296-298K, P^{RD50}/P^{o}, are approximately constant and usually above 0.1; thus indicating a physical rather than a chemical interaction mechanism for activation of the sensory irritant receptor.

The best log RD_{50} estimates were obtained by combining several physicochemical variables, which appeared to be reasonable in terms of defining logical attributes for nonreactive VOC of diverse chemical classes to fit the sensory irritant receptor. These are the **S** and **A** Abraham solvation parameters in combination with the Ostwald partition coefficient on hexadecane, L^{16}, or olive oil L^{oil}. In effect, the sensory irritant receptor was defined as a hydrophobic pocket, into which diverse nonreactive VOC would fit because of specific physicochemical attributes their diverse chemical classes being of secondary importance. Logarithmic value of VOC saturated vapor pressure at 298K, P^{o}, was also found to be excellent in the estimation of log RD_{50}. Even more convenient would be molecular weight or molecular volume, Vx to estimate log RD_{50} since these variables can be easily calculated.

$$\text{Log } RD_{50} = 6.603 - 0.592\ S - 1.030\ A - 0.851 \log L^{oil} \qquad (2.6)$$

$n = 50,\ r^2 = 0.91$

Abraham put forward a solvation model that is the combination of five parameters, for sensory potency in mice. This model will be presented below.

2.3.4. Models for nasal pungency thresholds

QSAR for nasal pungency thresholds have been scarcely reported. Cometto-Muniz and Cain have investigated the relationships between the human data and a number of physicochemical factors for a range of VOCs.[1-11] Abraham et al.[105] have developed a QSAR using a solvation equation. Recently, Hau, Connell and Richardson[106] developed a model for describing the triggering of nasal pungency in humans, based on the partition of VOCs between the air phase and the biophase. The authors used the same approach as for the odor detection model presented above, see equation (2.5). Regression coefficients for three homologous series and for the combination of 33 VOCs are given in Table 2.3.

Table 2.3. Model for nasal detection thresholds proposed by Hau and Connel (a,b)

Series	a	b	N	r
Alkyl acetate	-1.29	7.72	12	0.996
Alcohols	-0.86	7.63	11	0.971
Ketones	-1.00	7.34	4	0.972
Various	-1.16	7.69	33	0.971

(a) adapted from Ref. [106]

2.3.4. Models for eye irritation thresholds

In 1993, Klopman et al.[107] used the multi-CASE methodology to develop a predictive model of ocular irritation based on an evaluation of the significance of chemical structure in the induction of eye irritation. Multi-CASE is an artificial intelligence system capable of analysing a set of chemicals with defined biological activity. In the Multi-CASE approach, a generation of structural descriptors and the selection of those pertinent to eye irritation activity proceed automatically. In the course of analysis, an informational network is created that can be used to predict the activity

of VOCs not in the database. The Multi-CASE methodology was applied to the semi-quantitative analysis of eye irritation potential of 186 organic chemicals.[107]

In 1994, Cronin carried out a QSAR on a set of 38 organic liquids.[108] The DES was first altered by converting it into a molar eye score through MDES = DES .MW/1000d, where the molecular weight and density of the liquid are MW and d, respectively. The descriptors used were ClogP, the energy of the lowest unoccupied molecular orbit LUMO, and the zero order kappa alpha index $\kappa\alpha 0$. However, only very poor correlations were obtained, with values of the Fisher statistic between 7.6 and 8.2 and with r^2 0.35 at best. It was argued that the descriptors might not be the appropriate ones.

Eye irritation scores on a scale of 1-10 for a set of aliphatic alcohols have been shown to correlate well with four parameters, viz. log P_{oct}, minor principal inertial axes (Ry and Rz) and dipole moment, by means of neural network. Analysis of the dataset by both PCA and neural network analysis, showed a clear discrimination between irritants (those with eye irritation scores of 8 and 9) or non-irritants (scores of 1) using the four parameters. These analyses support the validity of the original four-parameter eye irritation QSAR model for neutral organic chemical. Within the dataset, irritant chemical are found in the region of parameter space defined by intermediate hydrophobicity, relatively small cross-sectional area and intermediate dipole moment; to be an eye irritant, a chemical must partition easily into the membrane and have a certain dipole moment.

Abraham et al.[109] developed a solvation model for eye irritation in humans, as shall now be outlined.

2.3.5. Abraham Solvation Model

The Abraham general approach is to regard the phenomena of sensory irritation and odor as mainly physical processes in which a VOC is transported to a receptor or receptor area. Abraham et al.[43,110] have constructed a rather simple equation that has already been applied to the partition of compounds between the gas phase and water, alcohols and other solvents and a number of biological phases as well as to numerous set of gas chromatographic data.

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (2.7)$$

In equation (2.7), the dependent variable SP is some property of a series of VOCs in a given system. The independent variable in equation (2.7) are properties of the VOCs as follows: **E** is an excess molar refraction, **S** is the dipolar / polarisability, **A** and **B** are the overall or effective hydrogen-bond acidity and basicity, and **L** is defined through log L^{16}, where L^{16} is the solute Ostwald solubility coefficient in hexadecane at 298K. The coefficients, e, s, a, b and l, are found by multiple linear regression analysis. There are not simply fitting coefficients, because they reflect the complementary properties of the solvent phase or biophase. The e-coefficient gives the tendency of the phase to interact with VOCs through polarisability – type interactions, mostly via electron pairs. The s-coefficient is a measure of the phase dipolarity / polarisability. The a –coefficient represents the complementary property to VOC hydrogen-bond acidity and so is a measure of the phase hydrogen – bond basicity. Likewise, the b-coefficient is a measure of the phase hydrogen bond acidity. Finally, the l-coefficient is a measure of the hydrophobicity of the phase. The theory behind the Abraham solvation equation was presented in detail in chapter 3.

Equation (2.7) has been applied to RD_{50}, NPT, EIT and ODT values, as shall be described now. Reciprocals of RD_{50}, NPT, EIT and ODT have been used so that the more potent the VOC the larger the values of log ($1/RD_{50}$), log (1/NPT), log (1/EIT) and log (1/ODT).

Application of equation (2.7) to the log ($1/RD_{50}$) values is not straightforward, because reactive VOCs will always appear as outliers to any equation for nonreactive VOCs. However, these outliers can be identified, because they will always appear more potent than calculated. Of course, there is an arbitrary element in deciding if a VOC is reactive or not. In this case, Abraham and co-workers applied equation (2.7) to all the VOCs and then gradually remove the VOCs which have observed values of log ($1/RD_{50}$) larger than calculated, until obtaining a set of VOCs that scatter randomly about a regression line of identity. The latter is given by equation (2.8):

$$\text{Log}\,(1/RD_{50}) = -7.136 + 0.338.\mathbf{E} + 1.056.\mathbf{S} + 2.333.\mathbf{A} + 0.706.\mathbf{B} + 0.786.\mathbf{L} \qquad (2.8)$$

$n = 65$, $r^2 = 0.965$, $sd = 0.221$

Abraham and co-workers have used equation (2.7) to correlate NPT values in ppm for 43 varied compounds, leading to equation (2.9),

$$\text{Log (1/NPT)} = -8.519 + 3.522.\mathbf{S} + 1.397.\mathbf{A} + 1.397.\mathbf{B} + 0.860.\mathbf{L} \qquad (2.9)$$

$n = 43$, $r^2 = 0.955$, $sd = 0.270$

Note that the e-coefficient of the independent variable, **E**, was statistically not significant.

A comparison of equation (2.8) and equation (2.9) indicates that the solute factors that govern log (1/NPT) and log ($1/RD_{50}$) are qualitatively the same, in that solute dipolarity / polarisability, hydrogen bond acidity, hydrogen bond basicity and size, as **L**, all increase VOC potency in both bio-assays, although quantitatively the factors are somewhat different. This is important because it shows that the bio-assays for nonreactive VOCs have much in common, even though the two bio-phases are not exactly the same. Finally, the alkyl acetates, the aliphatic aldehydes and the carboxylic acids were found to be of reactive nature in the mouse bio-assay, as compared to their perceived nonreactive nature in the human bio-assay. In the mouse experiment, mice have their whole head exposed to the stimulus for ten or more minutes, but in the NPT test, humans are presented with a brief stimulus puff (1-3 seconds) to one nostril only. A simple explanation of the reactive nature of the VOCs in the mouse bio-assay is that VOCs have a much longer time to react. However, this is not a convincing explanation, because a given molecule of a VOC will spend only the time of one breath in the mouse (0.24 seconds) or in human (4.0seconds) and hence has much less time to react in one breath in the mouse. But since the mouse bio-assay is a continuous test, each time the mouse breathes it will inhale a fresh dose of the VOC at a rate of some 250 breaths per minutes. Thus even if the amount of reaction is very small in each breath, there can be an accumulation of reactive VOCs to chemically react with the receptors leading to an enhanced potency.

In the case of eye irritation thresholds, not enough VOCs had been studied to construct a similar equation. However, Abraham[109] showed that scores for the Draize rabbit irritation test for VOCs administered as the pure liquids could be transformed into equivalent EIT values through equation (2.10),

$$\text{Log (DES/Po)} - 0.63 = \log (1/\text{EIT}) \qquad (2.10)$$

In equation (2.10), DES is the Draize score, Po is the saturated vapor pressure of the liquid VOC in ppm at 298K, and serves to correct the DES scores from the state of pure liquid to the gaseous state. The constant 0.63 is purely empirical and simply connects the DES and EIT scales. 54 EIT values were then obtained, of which 17 directly measured on human subjects, and 37 of which were obtained via equation (2.10). Application of equation (2.7) to the EIT values, as log (1/EIT) then yielded equation (2.11),

$$\text{Log (1/EIT)} = -7.918 - 0.482\,\mathbf{E} + 1.422.\mathbf{S} + 4.025.\mathbf{A} + 1.219.\mathbf{B} + 0.853.\mathbf{L} \qquad (2.11)$$

$n = 54$, $r^2 = 0.928$, $sd = 0.36$

Just as for the NPT equation, equation (2.11) covers a wide enough spread of log (1/EIT) values (from 0.20 to –5.46) and type of VOC to be useful as general predictive equation.

Comparison of coefficients of equation (2.8), equation (2.9) and equation (2.10) with those for gas / solvent phase transport properties shows that those equations are very comparable to simple transport equations for partition from the gas phase to organic solvents or to biological phases. Thus for irritation effects, the transport step must be the main process involved.

In contrast with the robustness of the Abraham model to describe, and even predict, chemesthetic responses in the nose and eyes, description of olfactory responses, i.e. odor thresholds, was less robust. These results suggest that, compared to the chemesthetic response, the olfactory response is more finely tuned to the chemical features of the odorant, an explanation supported by the estimated existence of up to 1,000 different olfactory receptors.

2.4 References

1. J.E. Cometto-Muniz, W.S. Cain, in Indoor Air and Human Health, 2nd Ed., Eds. R.B. Gammage, B.A. Berven, CRC Press, Inc,(1996) 53.
2. J.E. Cometto-Muniz, W.S. Cain, in Indoor Air and Human Health, 2nd Ed., Eds. R.B. Gammage, B.A. Berven, CRC Press, Inc, (1996) 23.
3. W.S. Cain, J.E. Cometto-Muniz, in Occupational Medecine: State of the Art Reviews, Hanley & Belfus, Philadelphia, 10 (1995) 133.
4. P.M. Wise, M.J. Olsson, W.S. Cain, Chem. Senses, 25 (2000) 429.
5. J.E. Cometto-Muniz, W.S. Cain, Physiol. Behav., 48 (1990) 719.
6. J.E. Cometto-Muniz, W.S. Cain, Pharmacol. Biochem. Behav, 39 (1991) 983.
7. J.E. Cometto-Muniz, W.S. Cain, Arch. Environ. Health., 48 (1993) 309.
8. J.E. Cometto-Muniz, W.S. Cain, Am. Ind. Hyg. Ass., 55 (1994) 811.
9. J.E. Cometto-Muniz, W.S. Cain, Chem. Senses, 20 (1995) 191.
10. J.E. Cometto-Muniz, W.S. Cain, Exp. Brain Res., 118 (1998) 180.
11. J.E. Cometto-muniz, W.S. Cain, Pharmacol Biochem Behav, 60 (1998) 765.
12. M. Schaper, Am. Ind. Hyg. Assoc. J. 54 (1993) 488.
13. C. Sell, in The Chemistry of Fragrances, Eds. D.H. Pybus, C. Sell, (1999) 216.
14. T.C. Pearce, J.W. Gardner, W. Gopel, in Sensors Update, Vol. 3, Ed. H. Baltes, W. Gopel, J.Hesse, Wiley-VCH, (1998) 61.
15. K. Rossiter, Chem. Senses, 96 (1996) 3201.
16. E.E. Morrison, D.T. Moran, in Handbook of Olfaction and Gustation, Ed. R.L. Doty, Marcel Deker, Inc, New-York, (1995) 75.
17. D. Lancet, Ann. Rev. Neurosci., 9 (1986) 329.
18. R.R.H. Anholt, Crit. Rev. Neurobiology, 7 (1993) 1.
19. L. Buck, R. Axel, Cell, 65 (1991) 175.
20. P. Mombaerts, Science, 286 (1999) 707.
21. D. Krautwurst, K-W Yau, R.R. Reed, Cell, 95 (1998) 917.
22. K. Raming, J. kreiger, J. Strotmann, I. Boekhoff, S. Kubik, C. Baumstark, H. Breer, Nature, 361 (1993) 353.
23. M. Singer, G.M. Shepherd, NeuroReport, 5 (1994) 1297.
24. D. Lancet, N. Ben-Arie, Curr. Biol., 3 (1993) 668.
25. M.S. Singer, Chem. Senses, 25 (2000) 155.
26. H. Koshimoto, K. Katoh, Y. Yoshihara, K. Mori, NeuroReport, 3 (1992) 521.
27. M.A. Bianchet, G. Bains, P. Pelosi, J. Pevsner, S.H. Snyder, H.L. Monaco, L.M. Amzel, Nat. Struct. Biol., 3 (1996) 934.

28. F. Vincent, S. Spinelli, R.Ramoni, S. Grolli, P. Pelosi, C. Cambillau, M. Tegoni, J. Mol. Biol., 300 (2000) 127.
29. L. Pauling, Chem. Eng. News, 24 (1946) 1375.
30. R.W. Moncrieff, Am. Perfumer, 54 (1949) 453.
31. J.E. Amoore, Perfurm. Essent. Oil Rec., 43 (1952) 321.
32. J.E. Amoore, Proc. Sci. Sect. Toilet Goods Assoc., 37 (1962) 13.
33. M. Guillot, Compt. Rend. Acad. Sci., 26 (1948) 143.
34. M.G.J. Beets, Structure Activity Relationships in Humans Chemoreception, Applied Science Publisher Ltd, London, (1978) 408.
35. E.T. Theimer, J.T. Davies, J. Agric. Food Chem., 15 (1967) 6.
36. G.M. Dyson, Chem. Ind., 57 (1938) 647.
37. R.H. Wright, The Sense of Smell, CRC Press, Boca Raton, FL (1982).
38. L. Turin, Chem. Senses, 21 (1996) 791.
39. B. Manic, J. Hirono, T. Sato, L.B. Buck, Cell, 96 (1999) 713.
40. M. Chastrette, SAR QSAR Environ. Res., 6 (1997) 215.
41. N. Neuner-Jehle, F. Etzweiler, in Perfumes – Art, Science, Technology, Eds. P.M. Muller, D. Lamparsky, Elsevier Applied Science, New-York, (1991) 153.
42. A. Richardson, in The Chemistry of Fragrances, Eds. D.H. Pybus, C. Sell, (1999) 216.
43. E. Guichard, I. Lesschaeve, N. Rournier, S. Issanchou, Riv. Ital. EPPOS, 4 (1993) 8.
44. J.E. Amoore, E. Hautala, J.Appl. Toxicology, 3 (1983) 272.
45. P. Laffort, Arch. Sci. Physiol., 17 (1963) 75.
46. F. Patte, M. Echeto, P. Laffort, Chem. Senses Flavor, 1 (1975) 283.
47. M. Devos, F. Patte, J. Rouault, P. Laffort, L.J. Van Gemert, Standardised Human Olfactory thresholds, IRL Press, Oxford, (1990) 165.
48. T.E. Acree, Anal. Chem., (1997) 175.
49. Y. Alarie, Environ. Health Perspect., 13 (1966) 433.
50. Y. Alarie, Environ. Health Perspect., 42 (1981) 9.
51. W.S. Cain, N-S Lee, P.M. Wise, R. Schmidt, B-H Ahn, J.E.Cometto-Muniz, M.H. Abraham, submitted.
52. J.H. Draize, G. Woodard, H.O. Calvery, J. Pharmacol. Experim Therapeutics, 82 (1944) 377.
53. J.E. Cometto-Muniz, W.S. Cain, T. Hiraishi, M.H. Abraham, J.M.R. Gola, Chem. Senses, 25 (2000) 285.
54. J.E. Cometto-Muniz, W.S. Cain, M.H.Abraham, J.M.R. Gola, Physiology & Behavior, 67 (1999) 269.
55. J.E. Cometto-Muniz, W.S. Cain, M.H.Abraham, J.M.R. Gola, submitted.
56. B. Berglund, M.J. Olsson, Percept. Psychophys., 53 (1993) 475.

57. W.S. Cain, F.T. Schiet, M.J. Olsson, R.A. de Wijk, Chem. Senses, 20 (1995) 625.
58. P. Laffort, A. Dravnieks, Chem. Senses, 7 (1982) 153.
59. D.G. Laing, H. Panhuber, M.E. Willcox, Physiol. Behav., 33 (1984) 309.
60. R.A. Baker, Ann. N.Y. Acad. Sci., 116 (1964) 495.
61. D.G. Guadagni, R.G. Buttery, S. Okano, H.K. Burr, Nature, 200 (1963) 1288.
62. M. Laska, R. Hudson, Chem. Senses, 16 (1991) 651.
63. M.Q. Patterson, J.C. Stevens, W.S. Cain, J.E. Cometto-Muniz, Chem. Senses, 18 (1993) 723.
64. A.A. Rosen, J.B. Peter, F.M. Middleton, J. Water Pollut. Control Fed., 35 (1962) 7.
65. J.E. Cometto-Muniz, W.S. Cain, H.K. Hudnell, Percept & Psychophys., 59 (1997) 665.
66. F.R. Cassee, J.H. Arts, J.P. Groten, V.J. Feron, Arch. Toxicol., 70 (1996) 329.
67. L.E. Kane, Y. Alarie, Am. Ind. Hyg. Assoc., 39 (1978) 270.
68. G.D. Nielsen, U. Kristiansen, L. Hansen, Y. Alarie, Arch. Toxicol. 62 (1988) 209.
69. J.E. Cometto-Muniz, M.R. Garcia-Medina, A.M. Calvino, Chem. Senses, 14 (1989).
70. J.E. Cometto-Muniz, S.M. Hernandez, Percept. Psychophys., 47 (1990) 391.
71. L.F. Hansen, G.D. Nielsen, Arch. Toxicol., 68 (1994) 193.
72. L.F. Hansen, G.D. Nielsen, Toxicology, 88 (1994) 81.
73. U. Kristiansen, G.D. Nielsen, Arch. Toxicol., 61 (1988) 419.
74. H. Meyer, Arch. Exp. Pathol. Pharmakol., 46 (1901) 109.
75. E.Overton, Studien uber die Narkose, Fisher, Jena, Germany, (1901).
76. K.H. Meyer, H. Gottlieb-Billroth, Z. Physiol.Chem., 112 (1920) 55.
77. J. Ferguson, J.Proc.Roy.Soc. (London) B127 (1939) 387.
78. F. Brink, J.M. Posternak, J.Cell.Comp.Physiol., 32 (1994) 211.
79. M.H.Abraham, G.D. Nielsen and Y. Alarie,J. Pharmaceutic. Sci., 83 (1994) 680.
80. R. Franke, Theoretical Drug Design Methods, Elsevier, Amsterdam, (1984).
81. A. Dravnieks, Ann. N. Y. Acad. Sci., 237 (1974) 144.
82. D.G. Guadigni, R.G. Buttery, S.J. Okano, J. Sci. Food Agric., 14 (1963) 761.
83. P. Laffort, C.R. Acad. Sci. Paris, 256 (1963) 5618.
84. L.H. Beck, Ann. N.Y. Acad. Sci., 116 (1964) 448.
85. M. Osako, K. Nishida, Nippon Eiseigaku Zasshi, 46 (1991) 913.
86. J. Devillers, C. Guillon, D. Domine, in Neural Networks in QSAR and Drug Design, Ed. D. Devillers, Academic Press, London, (1996) 97.
87. N. Christoph, F. Drawert, in Topics in Flavor Research, Eds. R.G. Berger, S. Nitz, P. Schreier, Eichhorn, H.: Marzling-Hangenham, Fed. Rep. Ger., (1985) 59.
88. P.C. Jurs, P.A. Edwards, Chem.Senses, 14 (1989) 281.
89. L.M. Egolf, P.C. Jurs, Anal. Chem., 65 (1993) 3119.
90. P.A. Edwards, L.S. Anker, P.C. Jurs, Chem. Senses, 16 (1991) 447.

91. D.B. Turner, P. Willett, Eur. J. Med. Chem., 35 (2000) 367.
92. P. Laffort, F. Patte, J. Chromatogr., 126 (1976) 625.
93. F. Patte, M. Echeto, P. Laffort, Anal. Chem., 54 (1982) 2239.
94. P. Laffort, F. Patte, J. Chromatogr., 406 (1987) 51.
95. T. Yamanaka, Chem. Senses, 29 (1995) 471.
96. M. Devos, F. Patte, J. Roualt, P. Laffort and L.J. Van Gemert, in Standardized Human Olfactory Thresholds, IRL Press, Oxford (1990).
97. K. M. Hau, D. W. Connell and B.J. Richardson, Regulatory Toxicology and Pharmacology, 31 (2000) 22.
98. M.H. Abraham, G.S. Whiting, Y. Alarie, J.J. Morris, P.J. Taylor, R.M. Doherty, R.W. Taft, G.D. Nielsen, Quant. Struct. Act. Rela., 9 (1990) 6.
99. M.H. Abraham, G.D. Nielsen, Y. Alarie, J. Pharm. Sci., 83 (1994) 680.
100. Y. Alarie, G.D. Nielsen, J. Andonian-Haftvan, M.H. Abraham, 134 (1995) 92.
101. G.D. Nielsen, Y. Alarie, Toxicol. Appl. Pharmacol., 65 (1982) 429.
102. J. Muller, G. Greff, Food Chem. Toxicol. 22 (1984) 661.
103. D.W. Roberts, Chem. Biol. Interact., 57 (1986) 325.
104. Y. Alarie, M. Schaper, G.D. Nielsen, M.H. Abraham, SAR and QSAR in Environmental Research, 5 (1996) 151.
105. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, Arch. Toxicol., 72 (1998) 227.
106. K.M. Hau, D.W. Connell, B.J. Richardson, Toxicological Sciences, 47 (1999) 93.
107. G. Klopman, D. Ptchelintsev, M. Frierson, S. Pennisi, K. Renkers, M. Dickens, ATLA, 21 (1993) 14.
108. M.T.D. Cronin, D.A. Basketter, M. York, Toxicology in Vitro, 8 (1994) 21.
109. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, Toxicol. in Vitro, 12 (1998) 201.
110. M.H. Abraham, Chem. Soc. Revs, 22 (1993) 73.

Chapter 3 An Introduction to the Abraham Solvation Equation

3.0. Introduction

The purpose of structure activity models such as linear free energy relationship, LFER, and quantitative structure activity relationship,QSAR, is to quantify the effects of changes in chemical structures on physicochemical activity and biological activity, respectively. First, LFERs derive from the major use of these relationships as mathematical tools for correlating changes in free energy in different reaction series. The Hammett equation[1,2] is the best known example of a LFER and this work has formed a sound basis for many other studies to be carried out involving LFERs. On the other hand, QSARs, have been used to correlate molecular structural features of compounds with their known biological properties. Attempts to establish relationships between chemical structure and biological effects may be traced back[3] as far as the work of Crum Brown and Fraser[4,5], but Meyer and coworkers[6,7] may have been the first to use a quantitative form showing that the product between narcotic concentration and oil/water or solvent/gas concentration ratio is remarkably constant.

The solvation model of Abraham, Kamlet and Taft[8] has offered a good base for the understanding, description and prediction of the ways in which solutes and solvents interact with each other. One important approach to the problem of solute / solvent interactions has been based on solvatochromism. The solvatochromic comparison approach developed by Kamlet and Taft has provided quantitative scales of solvent hydrogen bond abilities and solvent polarity. A number of quantitative relationships between a variety of physicochemical solute properties and certain solvent parameters have been put forward. Particularly successful and very extensive applications of the solvation model have evolved from the Abraham solvation equation. The Abraham model contains five solute descriptors that are linearly correlated with a given solute property. This model is not simply numerical in nature but offers a physicochemical interpretation of a large number of processes. This approach is a well-known and well-used equation for the description of relationships between structure and both physicochemical and biochemical property.

The first section of this chapter deals with the solvation model and the solvatochromic approach. Next, the Abraham solvation equation is presented in detail. In the final section the multiple linear regression analysis is discussed.

3.1. Linear Free Energy Relationships

3.1.1. Solvation Model

Abraham, Kamlet and Taft developed a simple solvation model[8-10] which describes the solvation of a gaseous solute into a solvent. The model can be broken down into the following stages:

1) A cavity of suitable size is made in the solvent, this involves the endoergic breaking of solvent / solvent bonds.
2) The solvent molecules are reorganised into their equilibrium position around the cavity.
3) The solute is inserted into the cavity and various solute / solvent interactions are set up.

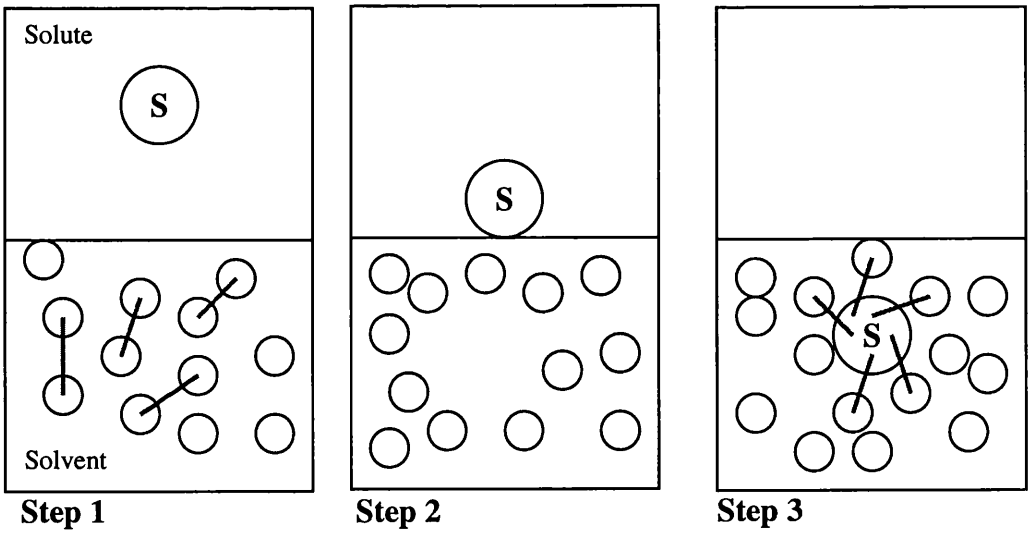


Figure 3.1. The three steps of the cavity model of solvation; — interactions.

The term solvation refers to the surrounding of each dissolved molecule by a shell of more or less tightly bound solvent molecules. This solvent shell is the result of intermolecular forces between solute and solvent.[12] These are called van-der-Waals forces and are classified in two distinct categories:

- The first category gathers the so-called directional, induction, and dispersion forces, which are non-specific.
- The second category comprises hydrogen bonding forces and the forces of transfer or electron-pair donor/ acceptor forces. To this group belong specific, directional forces.

The creation of a cavity is an endoergic effect, the magnitude of which depends on the forces holding the solvent molecules together and the size of the solute. The Gibbs free energy of reorganisation can be taken as zero, and so need to be considered no further. With the introduction of a solute molecule into the solvent cavity, a number of solute / solvent interactions will occur. These interactions are exoergic and aid the processes of solution. Both the cavity term and the solute / solvent interaction term will depend on the properties of the solute and the solvent under consideration. Hence to describe these effects for the general case of a number of solutes in a number of solvents, it is necessary to construct an equation that includes the relevant properties of both the solutes and the solvents.

Abraham, Kamlet and Taft[8,9,12-15] extended the LFERs of earlier workers[16] to involve solute / solvent interactions. Furthermore, they pointed out the necessity to consider both non-specific and specific solute / solvent interactions separately. This linear solvation energy relationship model has the general form shown in equation (3.1)

$$\text{Solute Property} = \text{constant} + \text{Cavity term(s)} + \text{Dipolarity/polarisability term(s)} + \text{Hydrogen bonding term(s)} \qquad (3.1)$$

3.1.2. The Solvatochromic Comparison Approach

Solvatochromism is an important approach to the problem of solute / solvent interactions. The term solvatochromism is used to describe the pronounced change in position of an UV/Vis absorption band, accompanying a change in the polarity medium. A hypsochromic (or blue) shift, with increasing solvent polarity, is usually called negative solvatochromism. The corresponding bathochromic (or red) shift is termed positive solvatochromism.

Kamlet, Taft and co-workers employed and further developed a solvatochromic comparison method to evaluate a β-scale of solvent hydrogen-bond acceptor (HBA) basicity, an α-scale of solvent hydrogen-bond donor (HBD) acidity, and a π*-scale of solvent dipolarity / polarisability using UV/Vis spectral data of solvatochromic compounds. [17-21] Magnitudes of enhanced solvatochromic shifts, $\Delta\Delta\tilde{\upsilon}$, in HBA solvents have been determined for 4-nitroaniline relative to homomorphic N,N,-diethyl-4-nitroaniline. Both standard compounds are capable of acting as HBA substrates in HBD solvents, but only 4-nitroaniline can act as a HBD substrate in HBA solvents. Taking the $\Delta\Delta\tilde{\upsilon}$-value of 2800 cm^{-1} for hexamethylphosphoric triamide (a strong HBA) as a single reference point ($\beta_1 = 1.00$), a β-scale of solvent HBA basicity for HBA solvents was developed. Using the same solvatochromic comparison method, i.e. the enhanced solvatochromic shift $\Delta\Delta\tilde{\upsilon}$, in HBD solvents for 4-nitroanisole and the pyridinium-N-phenoxide betaine dye, an α-scale of HBD acidity was evaluated. The same authors also introduced a π-scale of solvent dipolarity / polarisability. The π*-scale is so named because it is derived from solvent effects on the $\pi \rightarrow \pi^*$ electronic transitions of a variety of nitroaromatics. Solvent effects on the $\tilde{\upsilon}_{max}$ values of such solvatochromic indicators have been employed in the initial construction of the π*-scale, which was then expanded and refined by multiple least-squares correlations with additional solvatochromic indicators. In this way an averaged π*-scale of solvent dipolarity / polarizability has been established which measures the ability of the solvent to stabilise a charge or a dipole by virtue of its dielectric effect. A normalised range of 0.00 (cyclohexane) to 1.00 (dimethyl sulfoxide) for the π*-values of common solvents has been chosen so that, taken with the α-scale of solvent HBD acidity and the β-scale of solvent HBA basicity, these parameters can be used together in a multiparameter equation.

Using the solvatochromic solvent parameters α, β, π^*, the multiparameter equation (3.2) has been proposed for use in so-called linear solvation energy relationships (LSER). [18-20] Kamlet, Taft, Abraham et al. developed equation (3.2) to correlate the solubility and distribution behaviour of non-electrolyte solutes with solvent properties.[21,22]

$$SP = SP_0 + A\,\pi_1^*\,\pi_2^* + B\,\alpha_1\beta_2 + C\,\beta_1\alpha_2 + D\,(\delta^2_H)_1\,(V_2/100) \qquad (3.2)$$

The solute property, SP, can represent, for example, the logarithm of a rate or equilibrium constant, as well as a position of maximal absorption in an UV/Vis, IR spectrum; SP_0 is the regression value of this solute property in cyclohexane solvent. Here and elsewhere, the subscript 1 refers to the solvent and subscript 2 to the solute. *A, B, C* and *D* are the regression coefficients for the exoergic dipolarity / polarizability term, the exoergic hydrogen bonding terms of adduct formation between HBD solvents and HBA solutes (measured by α_1 and β_2) as well as between HBA solvents and HBD solutes (measure by β_1 and α_2) and the endoergic cavity term, respectively. V_2 is the molar volume of the solute, taken as its molecular mass divided by its liquid density at 293K. The δ^2_H term represents Hildebrand's solubility parameter[23] squared and corresponds to the cohesive pressure c, which characterised the energy associated with the intermolecular solvent / solvent interactions. Thus, δ^2_H is considered as a measure of the energy required to separate solvent molecules to provide a suitable sized cavity for the solute. Whereas δ^2_H measures the solvent's contribution to the cavity term, $V_2/100$ represents the solute's contribution to the cavity term.[24]

When dealing with the effects of different solvents on properties of a single solute, the factors relating to the solute can be subsumed into the regression coefficients of equation (3.3), and the following equation results:

$$SP = SP_0 + s.\pi_1^* + d.\delta_1 + a.\alpha_1 + b.\beta_1 + h.(\delta_H^2)_1 \qquad (3.3)$$

where δ is a discontinuous polarizability correction term[22] equal to 0.00 for non-chlorine substitued aliphatics, and 0.5 for poly-chlorinesubstituted aliphatics, and 1.0 for aromatic solvents.

Conversely, when dealing with solubilities or other properties of a set of different solutes in a single solvent, or with distributions of different solutes between a pair of solvents, the resulting equation (3.4) relates property SP specifically to the solute parameters V_2, π_2^*, α_2, and β_2.

$$SP = SP_0 + s.\pi_2^* + d.\delta_2 + a.\alpha_2 + b.\beta_2 + m.(V_2/100) \qquad (3.4)$$

One of the earliest tests of the above equation was its use in correlating retention in reversed phase liquid chromatography.[26] Soon thereafter it was used to study octanol / water partition coefficients and solubility in water.[27] Equation (3.4) works well when applied to processes in condensed phases, but for the processes of the type gas to condensed phase, Abraham devised a solute parameter, denoted log L^{16}, where L^{16} is the Ostwald solubility coefficient on n-hexadecane at 298 K. This L^{16} term, as $\log L^{16}$, is related to the endoergic work of creating a cavity in the solvent and the exoergic solute / solvent dispersion interactions.[28] Once added to equation (3.4), the term in solute volume was redundant.

$$SP = SP_0 + s.\pi_2^* + a.\alpha_2 + b.\beta_2 + l.\log L^{16} \qquad (3.5)$$

Equations (3.4) and (3.5) were revised further by replacing the term δ_2 with R_2, where R_2 is the solute polarisability.[29] In an attempt to overcome some of the downfalls encountered with the earlier Kamlet and Taft solute parameter[10,11], Abraham and co-workers developed a new dipolarity / polarisability scale, π_2^H and new hydrogen bond acidity and basicity parameters, α_2^H and β_2^H respectively.[30-32] Vx, the McGowan characteristic volume[33,34], was also preferred to V_2. The newly devised descriptors were combined linearly to give the following equations.[35]

$$SP = c + r.R_2 + s.\pi_2^H + a.\alpha_2^H + b.\beta_2^H + v.\ Vx \qquad (3.6)$$

$$SP = c + r.R_2 + s.\pi_2^H + a.\ \alpha_2^H + b.\beta_2^H + l.\ \log L^{16} \qquad (3.7)$$

3.2. The Abraham Solvation Equation

3.2.1. The Abraham Solute Parameters

As just mentioned, Abraham and co-workers drew up a list of solute descriptors to characterise solute / solvent interactions occurring during the solvation process. R_2 is an excess molar refraction in units of $(cm^{-3}.mol^{-1}/10)$[29], π_2^H is the solute dipolarity / polarizability[30], α_2^H and β_2^H are the solute hydrogen bond acidity[31] and basicity[32], respectively. Vx is the M◡Gowan characteristic volume in units of $(cm^{-3}.mol^{-1})/100$,[33,34] and $\log L^{16}$ is a descriptor where L^{16} is the solute gas-hexadecane partition coefficient at 298K.[28] The first four descriptors, R_2, π_2^H, α_2^H, β_2^H, can be regarded as measures of the tendency of a solute to undergo various solute / solvent interactions, all of which are energetically favourable, i.e. exoergic. On the other hand, $\log L^{16}$ and Vx are both measure of the size of a solute, so will be measure of the cavity term, see Figure 3.1. Further, since the size of the solute is related to general dispersion interactions, both $\log L^{16}$ and Vx also describe solute / solvent dispersion interactions.[35]

The development of the Abraham solvation equations (3.6) and (3.7) to analyse, correlate and predict particular property, SP, requires a knowledge of the relevant Abraham parameters. Thus, the solute parameters need to be identified. They were determine by Abraham et al. by means of a variety of methods mainly based on experimental measurements, and details of how they were initially obtained are covered next. Further in order to overcome the reliance on experimental data, alternative methods to obtain the Abraham solute descriptors have also been proposed and will be discussed later in this chapter. The best advance in this field has been the use of a group contribution approach by Platts and co-workers.[36] This technique allows rapid and efficient calculation of descriptors.

3.2.1.1. Log L^{16} Parameter

The solute descriptor, log L^{16}, initially formulated by Abraham et al.[28], characterises solute size and solute tendency to participate in solute-solvent interactions of the general London dispersion type. Log L^{16} is now a well-established descriptor in linear free energy relationship. Log L^{16} is defined through L^{16} the solute Ostwald solubility of a solute in n-hexadecane at 298K,

$$L^{16} = L = \frac{\text{concentration of solute in n - hexadecane}}{\text{concentration of solute in the gas phase}} \qquad (3.8)$$

Abraham chose n-hexadecane as a reference solvent for solute descriptor because n-hexadecane is a readily available non-polar liquid of well-defined structure. $LogL^{16}$ values were originally determined by direct GLC measurements on packed columns coated with n-hexadecane at 298K.[28] This approach is mainly limited to non-polar and polar volatile solutes at 298K, but more importantly interfacial absorption phenomenon contributes non-negligibly to the retention mechanism. So to overcome this drawback, Dallas and Carr[37] have made the use of open tubular fused silica capillary column for which interfacial absorption impact on retention mechanism is small. These direct approaches, however, are restricted to volatile and semi-volatile solutes that have a suitable retention time at 298K. The use of predictive model has allowed the determination of $logL^{16}$ values for less volatile solutes. For solutes too involatile at 298K, values of $logL^{16}$ can be obtained via GLC measurements on non-polar phases such as squalane or apiezon at elevated temperature.[38] GLC data for a series of solutes can be fitted to an equation of the form,

$$SP = SP_0 + r.R_2 + l.\log L^{16} \qquad (3.9)$$

Here, SP can be either the logarithm of the retention volume, or relative retention time or can be just the retention index, I. Thus for the large series of solutes studied by Dutoit[32] on a hydrocarbon phase at 383K, Abraham and co-workers constructed equation in the line of equation (3.9). So that further values of $logL^{16}$ can be calculated for any solute for which I, or I/10, is available. In this way, some 1500 $logL^{16}$ values have been obtained.[38]

$I/10 = 6.669 + 8.918\ R_2 + 20.002 \log L^{16}$ (3.10)

$n = 138, r^2 = 0.9995, sd = 0.449, F = 67950$

Here and elsewhere, n is the number of data point, r is the correlation coefficient, sd is the standard deviation and F the Fisher statistic.

3.2.1.2. The McGowan Characteristic Volume, Vx

The cavity term Vx (in $cm^3\ mol^{-1}/100$)[33,34] was chosen by Abraham because it is so straightforward to calculate by simple summation of bonds and atoms in the molecule. All bonds are considered equal, so that a double bond such as C=C or C=O or a triple bond such as C≡C are regarded as 'one bond'. An algorithm proposed by Abraham, allows the number of bonds in a molecule to be obtained simply:[33,34]

$B = N - 1 + R$ (3.11)

Where B is number of bonds, N is the total number of atoms and R is the total number of ring structures. Therefore:

$Vx = (\sum \text{atom contributions} - (6.56 \times B))/100$ (3.12)

Some typical values for atom contributions required for the calculation of Vx are given in Table 3.1.

Table 3.1. Atom contributions for calculation of Vx (in $cm^3\ mol^{-1}/100$)[a]

C = 16.35	P = 24.78	Se = 27.81
N = 14.39	S = 22.91	Br = 26.21
O = 12.43	Cl = 20.95	Sn = 39.35
F = 10.48	B = 18.32	Sb = 37.74
H = 8.71	Ge = 31.02	Te = 36.14
Si = 26.83	As = 29.42	I = 34.35

[a] For each bond between atoms, 6.56 $cm^3.mol^{-1}$ is to be subtracted.

3.2.1.3. The Solute Hydrogen-Bond Acidity Scale, $\Sigma\alpha_2^H$

Abraham[24] derived a new hydrogen bonding acidity scale, α_2^H, from thermodynamic measurements on 1:1 hydrogen bond complexation, and are related to the Gibbs free energy. The α_2^H parameter is a measure of hydrogen bond acid strength. This scale was drawn from values of equilibrium constants for the 1:1 complexation, $\log K_{1:1}$ of acids by reference bases (such as pyridine) in an inert solvent (tetrachloromethane) at 298K.

$$A\text{–}H + B \rightleftharpoons A\text{–}H\text{-----}B \qquad (3.13)$$

The hydrogen bond acids and reference bases were present in low concentration so that in solution they were both monomeric, nonassociated solutes. Abraham and co-workers study forty five reference bases which yielded forty five equations in which $\log K_{1:1}$ is the dependent variable, see equation (3.14).

$$\log K_{1:1}\text{ (series of acids against reference base B)} = L_B \log K^H_A + D_B \qquad (3.14)$$

where L_B and D_B characterise the base and values of $\log K^H_A$ describe the acids.

The general scale of hydrogen-bond acidity was set up by plotting a series of acid log K (against reference base x) versus a series of log K (against any other reference base), yielding a series of straight lines. The generated lines intersected at a point where $\log K^H_A = \log K_{1:1} = -1.1$, where $K_{1:1}$ is expressed in molar concentration units). This point of intersection correspond to the origin of the scale of hydrogen bond acidity and so solutes with no hydrogen bond acidity have $\log K^H_A$ equals to -1.1 units. Further, the various $\log K_{1:1}$ plots show family-independent behaviour, so that it was possible to obtain an 'average' hydrogen-bond acidity for solutes in tetrachloromethane, given as $\log K^H_A$. The origin of the scale was shifted to zero and compressed by transforming $\log K^H_A$ into a hydrogen bond acidity scale, α_2^H, simply via equation (3.15).[31]

$$\alpha_2^H = (\log K^H_A + 1.1)/4.636 \qquad (3.15)$$

The factor 4.636 is an empirical value, used simply to compress the acidity scale into a convenient working range. In bulk solvent the solute can form multiple hydrogen bonds with the surrounding molecules, therefore the 1:1 complexation no longer applies and the α_2^H values may not be relevant. In the event, Abraham found that the α_2^H values could actually be used as the solute overall or effective hydrogen-bond acidity descriptor, $\Sigma\alpha_2^H$, for most mono-acids. To obtain $\Sigma\alpha_2^H$ values a preliminary version of equations (3.7) was set up using α_2^H as the hydrogen-bond acid descriptor and was applied to various water-solvent partitions (SP is equal to the logarithmic value of water-solvent partition coefficients).[35,40] The α_2^H descriptor was then modified where necessary, in order to obtain the effective value, $\Sigma\alpha_2^H$. A new set of equations was then constructed, and the same process repeated until a self-consistent set of equations and $\Sigma\alpha_2^H$ values was given. Since the solutes in the water / solvent partitions are surrounded by solvent molecules, the overall hydrogen bond acidity scale is the actual scale required. It was observed that values of $\Sigma\alpha_2^H$ were constant along any homologous series, except perhaps for the first one or two members, so once a few values are established, values for the rest of the homologous series can be deduced. Multiple hydrogen bonding of a solute with several solvent molecules gives a higher $\Sigma\alpha_2^H$ value than α_2^H, obtained from a simple 1:1 complexation constant, see Table 3.3.

It should be noted that the work carried out to obtain an overall $\Sigma\alpha_2^H$ scale from the 1:1 α_2^H scale proceeded side-by-side with the calculation of $\Sigma\alpha_2^H$ that took place during the π_2^H calculation, see section 3.2.1.6. Therefore a constant check had to be made on the self-consistency of the derived $\Sigma\alpha_2^H$ values.

Finally, it is important to clarify that hydrogen bond acidity, which indicates the ability of a compound to donate a hydrogen bond is not related to Brønsted acidity of a compound, which, in turns, refers to loss of a proton. This fact is well illustrated in Table 3.2. It can be seen from their $\Sigma\alpha_2^H$ values that acetic acid and phenol do have similar ability to participate to hydrogen bond interaction. However, acetic acid is a typical Brønsted acid as defined by its pKa value[41] of 4.75 at 298 K. This is not the case for phenol whose pKa value[41] is given as 9.89 at 298 K.

Table 3.2. A comparison between α_2^H and $\Sigma\alpha_2^H$ values

Solute	α_2^H	$\Sigma\alpha_2^H$
n-Heptane	0.00	0.00
Ethanol	0.33	0.37
Pyrrole	0.41	0.41
Water	0.35	0.82
Acetic acid	0.55	0.61
Phenol	0.60	0.60

3.2.1.4. The Solute Hydrogen-Bond Basicity Scale, $\Sigma\beta_2^H$

Abraham[32,42] established a new hydrogen-bond basicity scale, β_2^H, for solutes using 1:1 hydrogen bond complexation equilibrium constants in tetrachloromethane. logK1:1 values for thirty four bases against a given reference acid were used to construct a scale of solute hydrogen bond basicity similar to the solute hydrogen bond acidity scale. The thirty-four equations in terms of logK were of the form,

$$\log K_{1:1} \text{ (series of bases B against reference acid A)} = L_A \log K^H_B + D_A \qquad (3.16)$$

L_A and D_A are now characteristics of the acid and $\log K^H_B$ describes the hydrogen bond basicity of a series of bases. These equations also gave straight lines passing through $\log K^H_B = \log K_{1:1} = -1.1$. The hydrogen bond basicity was defined by equation (3.17) where the factor 4.636 was chosen so that $\beta_2^H = 1$ for the strong hydrogen bond base hexamethylphosphotriamide.

$$\beta_2^H = (\log K_B^H + 1.1)/4.636 \qquad (3.17)$$

The β_2^H descriptor can be used as the overall or effective hydrogen-bond basicity descriptor, $\Sigma\beta_2^H$, for mono-bases with a few exceptions. A large number of $\Sigma\beta_2^H$ values were determined in a similar way as $\Sigma\alpha_2^H$.[35,40,42,43] A preliminary version of equation (3.7) was set up using β_2^H as the hydrogen-bond base descriptor and was applied to various water-solvent partitions. The β_2^H descriptor was then modified where necessary,

in order to obtain $\Sigma\beta_2^H$ values. A new set of equations was then constructed, and the same process repeated until a self-consistent set of equations and $\Sigma\beta_2^H$ values was given. Table 3.3. gives a comparison between β_2^H and $\Sigma\beta_2^H$ values.

Abraham introduced an additional hydrogen bond basicity term, $\Sigma\beta_2^0$, for solutes such as sulfoxides, anilines, pyridines, and some heterocyclic compounds in water / solvent partitioning systems where the solvent phase is quite aqueous. The latter include n-octanol, ethyl acetate, n-butyl acetate, diethyl ether and di-n-butyl ether. For non-aqueous phases such as chloroform, alkanes, benzene and the gas phase, the original $\Sigma\beta_2^H$ term can be used for all solutes.

Table 3.3. Comparison between β_2^H and $\Sigma\beta_2^H$ values

Solute	β_2^H	$\Sigma\beta_2^H$
n-Heptane	0.00	0.00
Diethyl ether	0.45	0.45
Butanone	0.18	0.51
Acetonitrile	0.44	0.32
Ethanol	0.44	0.48
Benzene	0.15	0.14
Pyridine	0.62	0.52
Dimethyl sulfoxide	0.78	0.88
Phenol	0.22	0.30
4-Methoxyaniline	0.45	0.65
1,4-Dioxane	0.47	0.64

3.2.1.5. The Excess Molar Refraction, R_2

The polarisability-correction descriptor, δ_2, used in equation 3.2, is only an empirical factor limited to one of three values, 0.5 for halogenated aliphatics, 1.0 for aromatics or 0.0 for all other compounds. A number of possible alternatives for δ_2 were considered by Abraham,[29] molar refraction (MR_X) being one of them. Molar refraction is often used as a measure of polarisability and can be defined as:

$$MR_X = 10[(n^2-1)/(n^2 + 2)]Vx \qquad (3.18)$$

Here n is the refractive index of a solute that is liquid at 298 K (for solids, the refractive index of the hypothetical liquid at 298 K can be calculated) and Vx is the McGowan characteristic volume in $(cm^3mol^{-1})/100$.[33,34] Because of the volume term in molar refraction, the latter always increase with increasing size. The refractive index function itself is rather better indication of the presence of polarisable electrons in a molecule; thus values of the refractive index are always larger for aromatic or halogenated aliphatic compounds than for other aliphatics.

The molar refraction of a solute in 'excess' of the molar refraction of an alkane of the same characteristic volume can be defined as R_2 ($10^{-1}cm^3mol^{-1}$), where:

$$R_2 = MR_X(\text{observed}) - MR_X(\text{for alkane of the same Vx}) \qquad (3.19)$$

By substracting the molar refraction for an alkane of the same characteristic volume, the dispersive component of molar refraction (already accounted in Vx and $logL^{16}$ in LFER correlations) is removed. R_2 provides a quantitative indication of polarisable n and π electrons. R_2 is an almost additive quantity that can easily be estimated for solids, and for structures in general, from fragment or substructure values.[33,35]

3.2.1.6. The Solute Dipolarity / Polarisability Scale, π_2^H

Originally π_2^* was taken as the solvent parameter π_1^* for non-associated liquids set out by Kamlet and Taft.[8,11,20,44,45] As π_1^* is experimentally accessible only for compounds that are liquid at 298 K, values of π_2^* had to be estimated for associated compounds such as acids, phenols, alcohols and amides as well as gaseous and solid solutes. In addition, there is present the inherent assumption that π_1^* is identical to π_2^* for non-associated liquids, but this may not always be the case. Furthermore, because of its spectroscopic origin, this parameter fails to be Gibbs energy related. It therefore seemed necessary to develop a method that would allow the determination of a dipolarity / polarisability scale, π_2^H, that would be free energy related and include all types of solute molecule.

Abraham and co-workers constructed the new dipolar / polarisability parameter, π_2^H, from the extensive sets of retention gas liquid chromatographic (GLC) data. The use of the McReynolds[46] and Patte *et al.*[47] chromatographic data provided π_2^H values for hundreds of solutes. π_2^H values for substituted aromatics, polyhalogenated aliphatics, nitroalkanes and nitriles were using the general solvation equation on Fellous *et al.* retention data for 17 stationary phases.[48]The π_2^H values for halogenated or polyhalogenated solutes were again obtained by the same method using retention data for various other stationary phases.[49,50]

3.2.1.7. Advances in the Abraham Solute Descriptor Determination

In this section is first presented a general procedure for solute descriptor determination based on experimental data. Then, attention is focused on the use of empirical methods, especially the group contribution method developed by Platts and co-workers.[36]

a) Estimation of Solute Descriptors from Experimental Data

The Vx descriptor can always be determined from the solute structure. Most of the time, R_2 is easily calculated. In such a case, π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$ and $\log L^{16}$ remain to be determined. Abraham developed a general procedure to simultaneously determine descriptor values. The method is based on the use of equation (3.8) and equation (3.9) applied to as many physicochemical properties as possible. In Table 3.4, are listed some typical physicochemical properties in use. The descriptor values for π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$ and $\log L^{16}$ are taken as the most statistically sound descriptor values that satisfy the various physicochemical properties, already calibrated through known solvation equations. In this way, a database of descriptors for some 3500 compounds has been established, see Table 3.5. The general method has been recently detailed for the determination of descriptors for terpenes[52] and buckminsterfullerene[53].

This general method for determination of descriptors from experimental data has been widely used in this thesis. This approach has allowed the estimation of Abraham solute descriptors for over 300 volatile organic compounds, more details will be found in chapters 7, 8 and 9.

Table 3.4. Water / solvent and gas / solvent processes used in the determination of solute descriptors.

Systems	dependent variable	SP
water / solvent partition	P: water/solvent partition coefficient	log P
HPLC	k': capacity factor	log k'
gas/solvent partition	L: gas/solvent partition coefficient	log L
GLC	I: retention index	I
	t_{rel}: relative retention time	log t_{rel}
	Vg: specific retention volume	log Vg

Table 3.5. Availability of solute descriptors

Descriptors	Number	Maximum value	Minimum value
R_2	3850	4.62	-1.37
π_2^H	3200	5.60	-0.54
$\Sigma\alpha_2^H$	4040	2.10	0.00
$\Sigma\beta_2^H$	2820	4.52	0.00
Vx	2200	8.56	-1.74
log L^{16}	4330	29.97	0.07

b) Estimation of Solute Descriptors from Empirical Methods

While methods based on experimental data deliver descriptors for most molecules, a number of drawbacks exists. First, one must physically obtain a sample of the solute of interest. Second, certain measurements may not be suited to certain types of solute. Third, the process of measurement is time-consuming, and laborious. Finally, this approach, based on experimental measurements, limits the possibility of using it in so-called high-throughput screening, the rapid evaluation of molecular properties for large libraries of compounds.[36] Consequently, a number of methods for the estimation of solute descriptors that do not require experimental data but that are often based on computed quantities, have been proposed. The various approaches are now presented,

attention is principally drawn to the group contribution approach used by Platts and co-workers[36]

- From Structure

In any homologous series, the value of π_2^H, $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ remain almost constant, apart form the first two members. Thus, homologs are dealt with ease. For branched compounds, the values of π_2^H often decreases by 0.03 unit for each branch, compared to the unbranched compounds, this is particularly applicable to aliphatic compounds. $LogL^{16}$ values along homologous series are well correlated with carbon number, N_C, and for any such series a plot of $logL^{16}$ versus N_C will yield to a good regression equation, from which further $logL^{16}$ values can be extrapolated.[38]

- from Solute Physicochemical Properties

Abraham and co-workers[32] put forward a reasonable correlation between the solute hydrogen bond acidity, α_2^H, and the Hammet inductive parameter, σ_I, for a few halogenated solutes,

$$\alpha_2^H = -0.114 + 0.992\,\sigma_I \qquad (3.20)$$

$n = 18, r = 0.91, sd = 0.02$

This equation is good enough to calculate additional α_2^H values for halogenated solutes.

Abraham[50] proposed the estimation of π_2^H values for chlorinated benzenes through the use of solute dipole moment, μ, and the number of chlorine atoms, N_{Cl}. The choice of these two solute properties was driven by the interest in dissecting π_2^H values into contributions from dipolarity (μ) and polarisability (N_{Cl}),

$$\pi_2^H = 0.538 + 0.0743\,N_{Cl} + 0.0353\mu \qquad (3.21)$$

$n = 13, r = 0.98, sd = 0.03$

- From Quantum Properties

Murray and Politzer[54] have developed a general approach that permits the analysis, correlation and prediction of Abraham hydrogen bonding parameters from a series of computed quantities evaluated on solute molecular surface. The authors have shown that there is a reasonable relationship between the calculated surface maxima, $V_{S,max}$, that describes the electrostatic potential associated with hydrogen atoms in the solute, and the descriptors α_2^H and $\Sigma\alpha_2^H$, see equations (3.22) and (3.23)

$$\alpha_2^H = -0.371 + 0.0257\ V_{S,max} \qquad (3.22)$$

$n = 15, r = 0.9685, sd = 0.05, F = 199.6$

there is just as good a correlation if $\Sigma\alpha_2^H$ is used instead,

$$\Sigma\alpha_2^H = -0.316 + 0.0246\ V_{S,max} \qquad (3.23)$$

$n = 15, r = 0.9731, sd = 0.04, F = 222.4$

Since the usual error in hydrogen bond parameters is around 0.03 units, it is possible to estimate further $\Sigma\alpha_2^H$ values, at least for monofunctional oxygen acids. Similarly, Murray and Politzer[54] have shown a sufficient correlation between the electrostatic potential minimum, V_{min}, and β_2^H for a series of oxygen bases and a series of nitrogen bases. The oxygen and nitrogen bases have to be taken separately. Here is an example for the oxygen bases,

$$\beta_2^H = -0.228 - 0.0134\ V_{min} \qquad (3.24)$$

$n = 16, r = 0.9554, sd = 0.065, F = 146.4$

The above equation can be used to calculate β_2^H values for oxygen bases. The authors have also tried to establish a similar relationship using now the overall basicity descriptors, $\Sigma\beta_2^H$ and $\Sigma\beta_2^0$, the results showed that these descriptors do not correlate well to V_{min}.[54]

Sevcik and co-workers[55] who made the use of a neural network approach to estimate π_2^H values. The authors took a number of structural and quantum mechanical

properties as input, combining them either linearly via multivariate linear regression analysis (MLRA) or nonlinearly via a feed-forward neural network.

c) Estimation from Group Contribution Methods

In group contribution approach, molecules are broken down to predefined fragments and their corresponding contributions are summed up to obtain the final descriptor values. Molecules however are never mere collection of fragments. Group contribution methods attempt to account for this by introducing different correction factors that are also considered additive. Consequently, all such methods rely on a basic equation of the type

$$\text{descriptor value} = \sum_{i=1}^{n} a_i f_i + \sum_{j=1}^{m} b_j F_j \qquad (3.25)$$

where f_i represents the fragmental contribution and a_i represents the number of occurrences of fragment type i, F_j represents the contribution and b_j represents the number of occurrences of correction factor j.[56]

The group contribution method was first applied to estimation of the Abraham descriptors by Sevcik and co-workers.[55] Their approach simply consists on adding contributions to logL[16] from a given set of fragments, the contribution being derived by MLRA. Recently, Platts and co-workers[36] have developed additive models for the estimation of Abraham's molecular descriptors R_2, π_2^H, $\Sigma\alpha_2^H$, $\Sigma\beta_2^H$, $\Sigma\beta_2^0$, and log L^{16}. From a database of between 2500 and 3500 values for each descriptor, Platts and co-workers were able to identify common substructures and, through a process of MLRA they evaluated contributions of each substructures to each descriptors. Their final model used 81 atom and functional group fragments for R_2, π_2^H, $\Sigma\beta_2^H$, $\Sigma\beta_2^0$, and log L^{16} and was able to reproduce experimentally derived results with correlation ranging from 0.95 to 0.99. However, $\Sigma\alpha_2^H$ required an entirely separate set of 51 fragments to be developed, resulting in a correlation coefficient of 0.97. Typically, errors of around 0.05-0.15 log unit (for values covering a range of 2-6 log units) were found.[36]

Of particular importance is the speed of calculation allowing rapid evaluation of molecular properties for large libraries of compounds. Once a list of SMILES strings is entered in the program, the descriptors can be calculated for up to 50 molecules per minute in a PC and for up to 700 molecules per minute using a UNIX version. When the

descriptors are known, computer calculation of properties from the regression equations of type of equation (3.6) and equation (3.7), is trivial, and so a number of properties of molecules can be estimated very rapidly from structure. This model has been trained and applied to several systems.[36,57,58]

A software, ABSOLV, based on the group contribution approach developed by Platts and co-workers, is now commercially available. ABSOLV uses a simplified Abraham solute descriptor notation, see Table 3.6. This notation will be used from now on.[59]

Table 3.6. Old and new notation of the Abraham solute descriptors

Old Notation	New Notation	Solute Descriptor
R_2	**E**	Excess molar refraction
π_2^H	**S**	Dipolarity / polarizability
$\Sigma\alpha_2^H$	**A**	Overall solute hydrogen bond acidity
$\Sigma\beta_2^H$	**B**	Overall solute hydrogen bond basicity
$\Sigma\beta_2^0$	$\mathbf{B^0}$	Amended basicity parameter for specific solute/systems
Vx	**V**	McGowan volume
$\log L^{16}$	**L**	Solute gas / hexadecane partition coefficient

3.2.2. Applications of the Abraham Solvation Equation

Using the simplified notation, the general solvation equations (3.6) and (3.7) become:

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + v.\mathbf{V} \qquad (3.26)$$

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (3.27)$$

where SP is a property of a series of solutes in a given system, and the independent variables are solute descriptors. The Abraham general solvation equations are examples of Linear Free Energy Relationships (LFERs) that correlate a physical, chemical

property (log SP) for a set of solutes with a corresponding set of solute physicochemical property descriptors (**E**, **S**, **A**, **B** and **V** / **L**). When applied to biological properties (logSP), the general equations will then refered to Quantitative Structure Activity Relationships (QSARs). The equation coefficients (*c*, *e*, *s*, *a*, *b* and *v* / *l*) obtained are dependent on the system under investigation and can be used to predict or estimate further values of the independent variable for a completely new solute providing the descriptors are known. In addition, the equation coefficients provide information on the phase system. For a partition between two condensed phases, equation (3.26) is used, equation coefficients will then refer to differences in physicochemical properties of the two phases. The *e*-coefficient is a measure of difference in phase polarisability and the *s*-coefficient is a measure of phase dipolarity / polarisability difference. The *a*-coefficient measures the difference in the two phases hydrogen-bond basicity (because an acidic solute will interact with a basic phase) and the *b*-coefficient is a measure of how the phases differ in hydrogen-bond acidity. The *v*-coefficient is a combination of exoergic dispersion forces that make a positive contribution to the *v*-coefficient and an endoergic cavity term that makes a negative contribution. The dispersion interaction nearly always dominates so that the *v*-coefficient is usually positive except for solution of gases and vapours in water. The *v*-coefficient is a useful measure of how the hydrophobicity of the two phases differs. Equation (3.27) is simpler and is applied to processes involving gas to condensed phase transfer. The *l*-coefficient is also resultant of dispersion and cavity effects and is usually positive. Since the *l*-coefficient varies between -0.21 for water at 298K and + 1.00 for n-hexadecane at 298K, it seems to be a suitable measure of condensed phase lipophilicity.[38]

It is important to note that for gas to condensed phase processes the *s*-, *a*- and *b*-coefficients must always be positive, or null, because interactions occurring between a condensed phase and a solute increase the solubility of the solute. The *e*-coefficient is an exception because it is tied to hydrocarbons as zero; hence phases containing fluorinated or chlorinated compounds give rise to negative *e*-coefficient. Therefore, the coefficients in the solvation parameter equation are not just fitting constants but must obey general chemical principles.[38]

The generality of the solvation equations is highlighted by the fact that they have been applied to a hugely diverse range of processes. Equation (3.26) has been employed for processes that take place in condensed phases, such as water-solvent partitions,[59,60] water-micelle partitions,[62] high performance liquid chromatography (HPLC),[63] normal

phase liquid chromatography,[64] microemulsion[65] and micellar[66] electro-kinetic chromatography, thin-layer chromatography,[67] solid phase extraction,[68] blood-brain distribution,[69,70] brain perfusion,[71] water-skin permeation,[70,64] tadpole narcosis[73]. Abraham and Le[74] have recently shown that a modified form of the Abraham Solvation Equation can be used to calculate and predict the solubility of solids and liquids in water. Equation (3.26) has been applied to the prediction gastrointestinal absorption values[75]. Equation (3.27) has been applied to a numerous gas / solvent partitions[61], to gas / biological phase partitions[76], and to a very large number of gas chromatographic systems[59]. Similarly, this equation has been used to correlate nasal pungency threshold[77], eye irritation threshold[78], and to predict respiratory irritation in mice[79]. Both equations (3.26) and (3.27) are well tried and tested equations.

3.3. Multiple Linear Regression Analysis

Abraham uses the classical multiple linear regression (MLR) analysis to generate the coefficients in the solvation equation. This is a common technique in statistics where the dependent variable, y, is linearly correlated to two or more independent variables to produce equation coefficients specific to the data set under analysis.[80] Once the coefficients are known, it is possible to predict values of y based on new values of the independent variables. The accuracy of the predicted y variable depends on the degree of scatter in the data. A method of least squares is usually used for determining the best fit for the linear line through the data.

Once a MLR analysis output is available it is essential to measure how reliable the relationship is, i.e. it is necessary to validate the model so any predicted values can be obtained with accuracy and confidence. Statistical methods used to do this include the standard deviation, sd, in the dependent variable, the correlation coefficient, r, the student's t-test and the Fisher F-statistic, F.

The standard deviation, sd, is the square root of the sample variance and is given by:

$$sd = \sqrt{\frac{\sum_{i=1}^{n} (y - \bar{y})^2}{n - k - 1}} \qquad (3.28)$$

$y-\bar{y}$ represent the deviations of individual results from the mean. Sd is dependant upon the number of samples (n) in the data set and also upon the number of descriptors (k) used in the model. The standard deviation has the same units as the property being measured. Standard deviation measures the spread of a distribution around the mean. A low sd value indicates a low spread, i.e. a good relationship, and a high sd value (close to 1) indicates that the data set contains a high distribution of points from the mean, which is unfavourable in MLR analysis.

The correlation coefficient, r, gives a measure of the success of the correlation of the dependent variable y against the independent variables x. The correlation coefficient is given by equation (3.29):

$$r = \sqrt{\frac{1 - sd^2 (n - 2)}{\sigma_y^2 \, n}} \qquad (3.29)$$

where sd is the standard deviation, n is the number of data points in the set and σ_y^2 is the variance in the y values. r is a measure of how closely the data set fits the relationship given by the MLR analysis and can range from –1 through to 1. A value of –1 or 1 indicates that the data set is explained by the correlation equation perfectly, while a value of zero means there is no relationship between the data set and the MLR analysis. A negative value of r may be interpreted as a poor correlation by an inexperienced eye so more often than not it is r^2 that is quoted in relation to multiple linear regression. r^2 has values of zero through to one and is basically an indicator of how well the regression analysis explains the relationship among the variables.

The correlation coefficient and standard deviation do not provide any statistical evidence that the relationship observed between the dependent variable and the independent variables did not occur by chance alone. Tests that investigate the significance of the regression coefficients are the Student's t-test and the Fisher F-statistic, F.

Students t-test assumes a normal distribution of errors and has set confidence levels, usually 95 or 99%. This gives a limit to the range of values acceptable at the specified confidence level. In MLR analysis, the t-test is performed on each individual variable to test their significance. Sometimes not all the variables are necessary and this would be indicated by the level of significance, and so may be removed. A value of 1

means the variable is highly significant in the regression equation, whereas a value of zero means that variable has no significant effect on the regression and can be removed.

The F-statistic is used to determine whether the observed relationship between the dependent and independent variables occurs by chance and is given by:

$$F = r^2 (n - \nu - 1) / (1 - r^2)\nu \qquad (3.30)$$

Here r is the correlation coefficient, n is the number of data points and ν is the degree of freedom ($\nu = p - 1$, where p is the number of variables). From equation (3.30) it is clear that as the number of data points and the correlation coefficient increase, so does F, and the larger the value of F the better the regression.

3.3.1. Limitations of Multiple Linear Regression Analysis

The main problem with MLR is its sensitivity to collinearities among the independent variables. Collinearities occur when there is a high degree of linear correlation between two or more of the independent variables. If MLR is applied to a data set with correlated variables, the calculated regression coefficients become unstable and uninterpretable. Some regression coefficients may be much larger than expected, or they may even have the wrong sign. It is therefore very important to make sure the variables used in MLR, i.e. the solute descriptors in the case of the Abraham Solvation Equation, are well defined and independent.

The spread of the explanatory variables needs to be as wide as possible for two reasons. One, to produce a general regression equation that 'explains' a varied set of data and two, to provide a large 'descriptor space'. Predictions should only be made within the descriptor space of the compounds used to set up the regression equation, so that the wider the spread of variables the greater the descriptor space. This will result in greater success when applying the equation to predicting further dependent variables. As shown earlier, the greater the number of data points, the greater the reliability of the correlation; a minimum of five data points per variable is suggested to achieve a statistically significant and reliable regression equation.

3.3.2. *Multiple Linear Regression Analysis and the Abraham Solvation Equation*

An in-house software program has been developed that utilises the SmartWare II (Informix Software, Inc, Kansas) database facilities and statistical tools to generate the equation coefficients in the solvation equation along with a set of statistics. The output allows the user to assess if the coefficients are significant, if the over all correlation is within acceptable limits and so on. Presented next is an example of a typical MLR analysis output generated during the development of a solvation equation for water / squalane partition at 298K.[81] In this study, 177 logP values were regressed against the 5 Abraham descriptors, viz. **E, S, A, B, V**. The four sections of the output are summarised in Table 3.7. and explained in the following text.

The first information given is the correlation coefficients between the variables, part 1) Table 3.7. It is important to have as lower correlation as possible, so as to remain within the limits set out by the use of multiple linear regression. High correlation between two descriptors means they are exhibiting collinearity, which is unfavourable. Usually a value of 0.7 or lower is acceptable. The next section, shown in section 2) in Table 3.7, provides information on the set of solute descriptors themselves. The mean is the mean value for each descriptor and sd is the standard deviation for that descriptor value. In this case, the higher the sd the better as this indicates a wide spread of descriptors values and hence a more general solvation. The section 3) gives the numerical values for the system coefficients, i.e. the constants *c*, *r*, *s*, *a*, *b*, and *v* in equations (3.6). The sd is the standard deviation in each coefficient and is important if the final solvation equation is used for back-calculation of descriptors so the error in the descriptor value can be determined. Whether or not the coefficients are statistically significant is indicated by the student's t-test, values of 0.95 or above are generally acceptable. If the t-test value is lower than 0.95 it may be desirable to remove that coefficient, as it is not as statistically significant as others may with higher t-test values and may improve the regression. The final section 4) in the output gives the overall statistical view of the MLR analysis. r indicates the quality of the regression. The overall standard deviation ($Y_{obs} - Y_{calc}$) of the whole regression is given, the smaller the standard deviation, the better the regression. The F-statistic shows the quality of the correlation and thus shows a goodness of fit. The F-statistic improves (increases) as the number of points in the data set increases and the sd decreases.

Table 3.7. Output of the regression analysis of log P^{SQ} and solute descriptors

1)	**Variables**	**E**	**S**	**A**	**B**	**V**	
	S	0.762					
	A	0.067	0.224				
	B	0.124	0.596	0.516			
	V	0.394	0.090	-0.169	-0.184		
	logP^{SQ} obs	0.103	-0.402	-0.603	-0.792	0.708	
2)	**Variables**	**E**	**S**	**A**	**B**	**V**	**logP^{SQ} obs**
	MEAN	0.333	0.362	0.061	0.195	0.898	2.410
	SD	0.391	0.318	0.141	0.194	0.279	2.145
3)	**Variables**	***e***	***s***	***a***	***b***	***v***	***c***
	COEFFS	0.810	-1.702	-3.626	-4.810	4.239	-0.119
	ST.DEV	0.064	0.091	0.094	0.106	0.046	0.043
	T	12.650	18.619	38.425	45.418	91.865	2.708
	TTEST	1.000	1.000	1.000	1.000	1.000	1.000
4)	**r^2**	0.996					
	sd	0.123					
	DOF	136					
	F	7692.2					

3.4. References

1. N. S. Isaacs, Physical Organic Chemistry – 2nd ed. Longman Scientific and Technical, Harlow, (1995).
2. L. P. Hammett, J. Chem. Soc., 59 (1937) 96.
3. W.P. Purcell, G.E. Bass, J.M. Clayton, in Strategy to Drug Design: a guide to Biological Activity, Wiley Interscience: New York, (1973).
4. A. Crum Brown Fraser, T.R. Fraser, Proc. Roy. Soc. Edinburgh, 6 (1867-1868) 228.
5. A. Crum Brown Fraser, T.R. Fraser, Proc. Roy. Soc. Edinburgh, 6 (1867-1868) 461.
6. K. H. Meyer, H. Gottlieb-Billroth, Z. Hoppe-Seyler's Physiol. Chem., 112 (1920) 55, K.
7. H. Meyer, H. Hemmi, Biochem. Z., 277 (1935) 39.
8. M. J. Kamlet, R. M. Doherty, J-L. M. Abboud, M. H. Abraham and R. W. Taft, Chem. Tech, (1986) 566.
9. M. H. Abraham, R. M. Doherty, M. J. Kamlet and R. W. Taft, Chem. Br., 22 (1986) 551.
10. M. J. Kamlet, R. M. Doherty, J-L. M. Abboud, M. H. Abraham and R. W. Taft, J. Pharm. Sci., 75 (1986) 338.
11. M. J. Kamlet, R. M. Doherty, M. H. Abraham, P. W. Carr, R. F. Doherty and R. W. Taft, J. Phys. Chem. 91 (1987) 1996.
12. C.Reichardt, Solvents and Solvent effects in organic Chemistry, second, revised and enlarged edition, VCH-Weinheim, (1990).
13. M.J. Kamlet, R.W. Taft and J.-L. M. Abboud, J. Am. Chem. Soc., 91 (1977) 8325.
14. M. H. Abraham, R. M. Doherty, M. J. Kamlet aqnd R. W. Taft, Chem. Br., 22 (1986) 551.
15. M. J. Kamlet, R. M. Doherty, M. H. Abraham, P. W. Carr, R. F. Doherty and R. W. Taft, J. Phys. Chem. 91 (1987) 1996.
16. I.A. Koppel and V. A. Palm, Advances in Linear Free Energy Relationships, Eds. N.B. Chapman and J. Shorter, Plenum Press, London, (1972).
17. M.J. Kamlet and R.W. Taft, J.Am.Chem.Soc., 98 (1976) 2886.
18. M.J. Kamlet, J.L. Abboud, R.W. Taft, J.Am.Chem. Soc., 99 (1977) 6027; ibid. 99 (1977) 8325.
19. J.-L. M. Abboud, M.J. Kamlet, R.W. Taft, Progr.Phys.Org.Chem., 13 (1981) 485.
20. R.W. Taft, J.-L. M. Abboud, M.J. Kamlet, M.H.Abraham, J.Solution Chem, 14 (1985) 153.
21. M.J. Kamlet, R.W. Taft, Acta Chem. Scand., Part B 40 (1986) 619.
22. M.J. Kamlet, M.H. Abraham, R.M. Doherty, R.W. Taft, J.Chem. Soc. 106 (1984) 464.
23. M.J. Kamlet, R.M. Doherty, M.H. Abraham, R.W. Taft, Quant. Struct. Act. Relat., (1988) 71.
24. M.J.Kamlet, J.-L.M. Abboud, M.H.Abraham, R.W. Taft, J.Org.Chem. 48 (1983) 2877.

25. J.H. Hildebrand, R.L. Scott, The Solubility of Non-Electrolytes, Dover Pub., N.Y, (1964).
26. R.W. Taft, J. S. Murray, Quantitative Treatments of Solute/Solvents Interactions, Theoritial and computational chemistry, Vol.1, Eds. P. Politzer, J.S. Murray, Elsevier Science B.V., (1994) 55.
27. M.J. Kamlet, R.M Doherty, M.H. Abraham, Y. Marcus and R.W. Taft, J.Phys. Chem., 92 (1988) 5244.
28. M.H. Abraham, P.L. Grellier and R.A. McGill, J.Chem.Soc., Perkin Trans. 2, (1987) 797.
29. M.H. Abraham, G.S. Whiting, R.M. Doherty and W.J. Shuely, J. Chem. Soc. Perkin Trans 2, (1990) 1451.
30. M.H. Abraham, G.S. Whiting, R.M. Doherty and W.J. Shuely, J. Chrom., 587 (1991) 213.
31. M.H. Abraham, P.P. Duce, P.L. Grellier, D.V. Prior, J.J. Morris, P.J. Taylor, J. Chem. Soc. Perkin Trans 2, (1989) 699.
32. M.H. Abraham, P.L. Grellier, D.V. Prior, J.J. Morris, P/J. Taylor, C. Laurence, M. Berthelot, Tetrahedron Lett., 30 (1989) 2571.
33. J. C. McGowan, J. Appl. Chem. Biotechnol., 34A (1984) 38.
34. M. H. Abraham and J. C. McGowan, Chromatographia, 1987, Vol. 23, 243.
35. M. H. Abraham, Chem. Soc. Revs., 22 (1993) 73.
36. J.A. Platts, D. Butina, M.H. Abraham, A. Hersey, J. Chem. Inf. Comput. Sci., 39 (1999) 835.
37. A.J. Dallas, P.W. Carr, J. Phys. Chem., 98 (1994) 4927.
38. M.H. Abraham, Quantitative Treatments of Solute / Solvent Interactions Theoritical and Computational Chemistry Vol.1, Eds. P. Politzer, J.S. Murray, Elsevier Sci., B.V., (1994) 83.
39. J.-C. Dutoit, J. Chromatogr., 555 (1991) 191.
40. M. H. Abraham and H. S. Chadha, Applications of a solvation equation to drug transport properties, in Lipophilicity in Drug Action and Toxicology, Eds. V. Pliska, B. Testa and H. van de Waterbeemd, VCH, Weinheim, Germany, (1996).
41. Handbook of Chemistry and Physics, Ed. R.C. Weast, 55th Edition, CRC Press Inc., Cleveland, (1974).
42. M. H. Abraham, P. L. Grellier, D. V. Prior, J. Morris and P. J. Taylor, J. Chem. Soc. Perkin Trans. 2, (1990) 521.
43. M. H. Abraham, J. Phys. Org. Chem., 6 (1993) 660.
44. M. H. Abraham, M. J. Kamlet, R. W. Taft and R. M. Doherty, Nature (London), 106 (1984) 464.
45. R. W. Taft, M. H. Abraham, G. R. Famini, R. M. Doherty, J-L. M. Abboud and M. J. Kamlet, J. Pharm. Sci., 74 (1985) 807.

46. W. O. McReynolds, Gas Chromatographic Retention Data, Preston Technical Abstracts, Evanston, IL, (1966).
47. F. Patte, M. Etcheto and P. Laffort, Anal. Chem., 54 (1982) 2239.
48. R. Fellous, L. Lizzani-Cuvelier and R. Luft, Anal. Chem. Acta, 174 (1985) 53.
49. M. H. Abraham and G. S. Whiting, J. Chromatogr., 594 (1992) 229.
50. M. H. Abraham, J. Chromatogr., 644 (1993) 95.
51. A.Leo, the Medicinal Chemistry Project Pomona College, Claremont, CA, 91711
52. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, M. Roses, E. Bosch, M.L. Diaz, Arch. Toxicol., 27 (1998) 227.
53. M.H. Abraham, C.E. Green, W.E. Acree, Jr., J. Chem. Soc., Perkin Trans. 2, (2000) 281.
54. J.S. Murray, P. Politzer, J. Chem. Res. (S), (1992) 110.
55. P.Havelec, J.G.K. Sevcik, J.Phys.Chem.Ref.Data, 25 (1996) 1483.
56. P. Buchwald, N. Bodor, Curr. Med. Chem., 5 (1998) 353.
57. J.A. Platts, M.H. Abraham, D. Butina, A. Hersey, J.Chem. Inf. Comput. Sci., 40 (2000) 71.
58. J.A. Platts, M.H. Abraham, Environ. Sci. Technol., 34 (2000) 318.
59. Absolv, syrius
60. M.H.Abraham, H.S,Chadha, G.S.Whiting and R.C.Mitchell, J.Pharm.Sci., 83 (1994) 1085.
61. M.H. Abraham, J.Le, W.E. Acree, Jr., C.E. Hernandez, L.E. Roy, J.Chem. Soc., Perkin Trans. 2, (1998) 2677.
62. M.H.Abraham, H.S.Chadha, J.P.Dixon, C.Rafols and C.Treiner, J.Chem.Soc. Perkin Trans.2, (1995) 887.
63. M.H.Abraham and M.Roses, J.Phys.Org.Chem., 7 (1994) 672.
64. F. Z. Oumada, M. Rosés, E. Bosch and M. H. Abraham, Anal. Chim. Acta, 382 (1999) 301.
65. M.H.Abraham, C.Treiner, M.Roses, C.Rafols and Y.Ishihama, J.Chromatogr.A, 752 (1996) 243.
66. S.K.Poole and C.F.Poole, Analyst, 122 (1997) 267.
67. M.H.Abraham, C.F.Poole and S.K.Poole, J.Chromatogr.A, 749 (1996) 201.
68. D.S.Siebert, C.F.Poole and M.H.Abraham, Analyst, 121 (1996) 511.
69. M.H.Abraham, H.S,Chadha and R.C.Mitchell, J.Pharm.Sci., 83 (1994) 1257.
70. M. H. Abraham, H. S. Chadha, F. Martins, R. C. Mitchell, M. W. Bradbury and J. A. Gratton, Pestic. Sci., 55 (1999) 78.
71. J.A.Gratton, M.H.Abraham, M.W.Bradbury and H.S.Chadha, J.Pharm.Pharmacol., 49 (1997) 1211.
72. M.H.Abraham, H.S,Chadha and R.C.Mitchell, J.Pharm.Pharmacol., 47 (1995) 8.
73. M.H.Abraham and C.Rafols, J.Chem.Soc. Perkin Trans.2, (1995) 1843.
74. M. H. Abraham and J. Le, J. Pharm. Sci., 88 (1999) 868.

75. Y.H.Zhao; J.Le; M.H.Abraham; A.Hersey; P.J.Eddershaw; C.N.Luscombe; D.Butina; G.Beck; B.Sherbourne; I.Cooper; J.A.Platts , J. Pharm. Sci., 90 (2001) 3484.
76. M.H. Abraham, P.K. Weathersby, J. Pharm Sci., 83 (1994) 1450.
77. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, Arch. Toxicol., 72 (1998) 227.
78. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, R. Kumarsingh, Toxicol. in Vitro, 12 (1998) 403.
79. Y. Alarie, G.D. Shaper, G.D. Nielsen, M.H. Abraham, Arch. Toxicol., 72 (1998) 125.
80. L. Eriksson and E. Johansson, Chemom. Intell. Lab. Syst., 34 (1996) 1.
81. M.H. Abraham, J.M.R. Gola, unpublished results.

Chapter 4 An Introduction to Gas Liquid Chromatography

4.0. Introduction

Thanks to its simplicity and rapidity, chromatography is nowadays a current technique to separate and analyse complex mixtures of compounds that is widely used in a large number of scientific domains such as biology, chemistry, the pharmaceutical industry, and environmental sciences. In addition to its ability to separate complex mixtures of compounds, chromatography gives insight for compound thermodynamic properties as well as stationary phase sorption capacity.

Chromatography is a method of separation in which the compounds to be separated are distributed between two phases one of which is stationary (the stationary phase) while the other (the mobile phase) percolates through it in a definite direction[1]. The stationary phase may be a solid for gas solid chromatography (GSE) or a liquid in gas liquid chromatography (GLC). In GLC, an involatile liquid is coated onto an inert support (packed column) or is bonded to the wall of an open tubular column. Chromatography consists in injecting a small sample in a chromatograph. The sample is rapidly vaporised and carried by the mobile phase, or carrier gas, to the head of the column. Separation is achieved by the fact that individual components in a mixture have different solubilities in the stationary phase. Those having high solubility in the stationary phase spend less time in the moving gas phase than those of low solubility. When the components elute they pass into a detector, where their presence is detected and the electrosignal produced is amplified to a convenient level. The detector produce a response, which is proportional to the amount of compound passing through it, in a form of a signal that can be measured.

The principal function of the gas chromatograph is to provide those conditions required for achieving a separation without disturbing the performance of the column in any way. The conditions involve a regulated flow of carrier gas to the column, an inlet system to vaporise and mix the sample with the carrier gas, a thermostated oven to optimise the temperature for the separation, a detector to monitor the separation and associated electronic components to control and monitor instrument conditions, and to record, manipulate and format the chromatogram[1,2]. Once instruments and conditions

required for achieving a good separation are set important data retention data can be retrieved from the gas chromatogram. Figure 4.1 shows a schematic diagram of the GLC process.

It is not the purpose of this chapter to present in detail the GLC instrumentation. A number of documentation on this subject can be found in the literature. For instance, Poole and Poole[1] have recently published a book entitled Chromatography Today. A large amount of information can also be retrieved from the book of Conner and Young[2] published in 1979. This chapter will actually focused on retention data, as these data will be used throughout this thesis.

In the first section of this chapter, the retention data are presented, then attention will be drawn to their relationships with some thermodynamic properties. A section of this chapter is devoted to the static headspace gas chromatographic method. Finally, it is pointed out that the solvation parameter developed by Abraham[1] is one of the models that provides information about the sorption properties of the stationary phase.

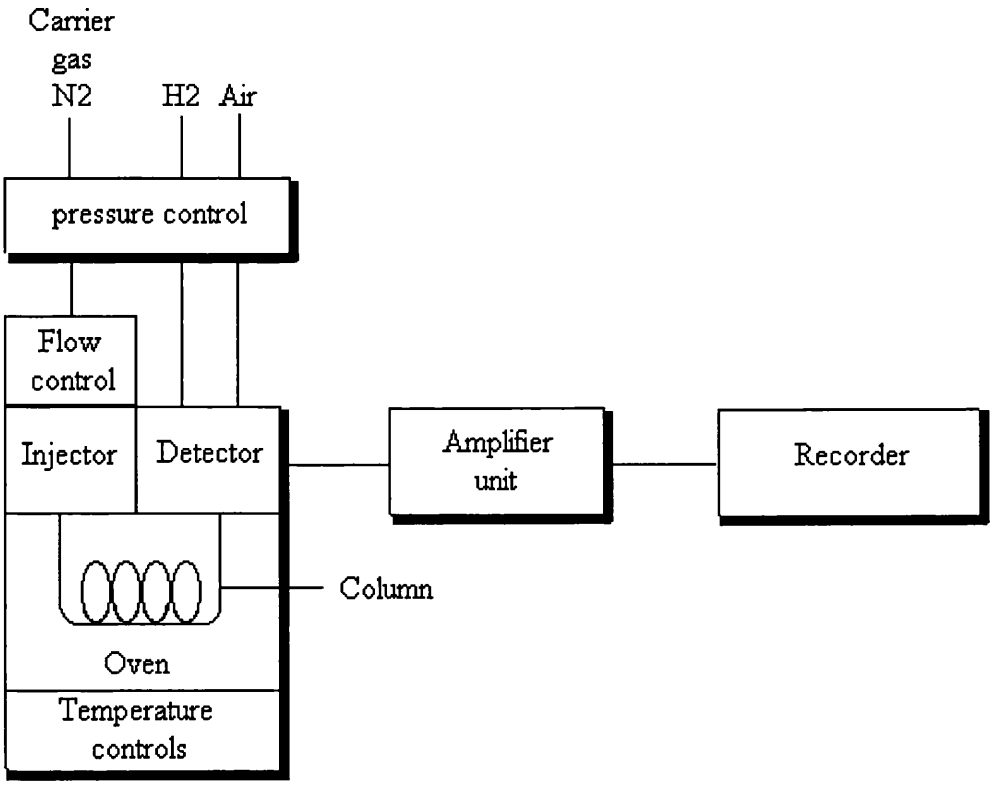


Figure 4.1. Process of Chromatography

4.1. Gas Chromatographic Retention Data

The chromatography process takes place as a result of repeated sorption/desorption acts during the movement of the sample components along the stationary bed. Distribution of the components of a mixture between the mobile and the stationary phases depends on the molecular interaction between the solutes and both the phases. Intermolecular forces control chromatography; polar and dispersion interactions forces make major contributions to the overall interactions. Polar forces include dipole-dipole interactions and hydrogen bonding. Dispersion forces are produced as a result of dipoles formed between electrons and nuclei interacting on the polarisable electronic system of other atoms. Dispersion forces are relatively weak but are still the main interactive forces in non-polar solvent such as hydrocarbons[1,2].

4.1.1. Thermodynamic

The GLC process involves establishment of an equilibrium in which the components distribute between a mobile phase and a stationary phase:

$$C_m \underset{}{\overset{K}{\rightleftharpoons}} C_s \qquad (4.1)$$

$$\text{With } K = \frac{C_s}{C_m} \qquad (4.2)$$

Here C_s and C_m are the concentrations of the solute in the stationary phase and mobile phase respectively and K is the partition coefficient or distribution coefficient equivalent to the Ostwald solubility coefficient, L. The larger the K value the greater the affinity of the solute for the stationary phase. The value of K varies for different solutes and as a result, each of the components of a mixture of compounds in a homologous series will proceed through the chromatographic column at different speed. Ideally K is constant over a wide range of solute concentrations. Thus, C_s is directly proportional to C_m. Chromatography carried out under conditions in which K is more or less constant is termed linear chromatography. The partial molar Gibbs free energy of solution, ΔG^0, for a solute at infinite dilution in the stationary phase can be obtained directly from K:

$$\Delta G^0 = -R\,T \ln K \qquad (4.3)$$

Here and elsewhere, T is the column temperature (K) and R is the universal gas constant ($dm^3.atm.K^{-1}.mol^{-1}$).

Although gas-liquid partitioning is the dominant retention mechanism in gas-liquid chromatography it is not the only possible retention mechanism. It may be accompanied by interfacial adsorption at the gas-liquid interface and by adsorption at the support or capillary column wall.

4.1.2. Retention Data

The information obtained from a chromatographic experiment is contained in the chromatogram, a record of the concentration or mass profile of the sample components as a function of the movement of the mobile phase, Figure 4.2. The horizontal axis represents time and it also represents value of gas if the flow rate is constant. The vertical axis represents the quantity of substance, which is proportional to the response.

Compounds that do not interact with the stationary phase, e.g. methane, will be eluted at the column dead time, holdup time, t_M, equivalent to the time required for an unretained solute to reach the detector from the point of injection[1]. The retention time of a solute, t_R, is the time the average molecule of solute takes to travel the whole length of a chromatographic column and is measured to the midpoint of the elution curve. A part, t_M, of this time is required by all solutes simply to pass through the mobile phase from inlet to outlet.

Gas liquid chromatography mainly consists of measurements of the retention times t_M and t_R from which a number of retention data can be derived.

$\mathbf{t_{R'}}$, the adjusted retention time defines the amount of time the solute spends in the stationary phase

$$t_{R'} = t_R - t_M \qquad (4.4)$$

$\mathbf{V_{R'}}$, the adjusted retention volume corresponds to the volume of carrier gas passing through the stationary phase during $t_{R'}$

$$V_{R'} = F.t_{R'} \qquad (4.5)$$

where F refers to the corrected value of the carrier gas, see equation (4.8).

V_N, the net retention volume, is the adjusted retention volume corrected for mobile phase compressibility

$$V_N = J.V_{R'} \qquad (4.6)$$

The term J represents the gas compressibility factor that is calculated according to equation (4.9).

V_G, the specific retention volume, is the net retention volume at the column temperature for unit weight of stationary phase

$$V_G = V_N \cdot \frac{1}{W_L} \qquad (4.7)$$

In the above equation, W_L is weight of stationary phase in gram.

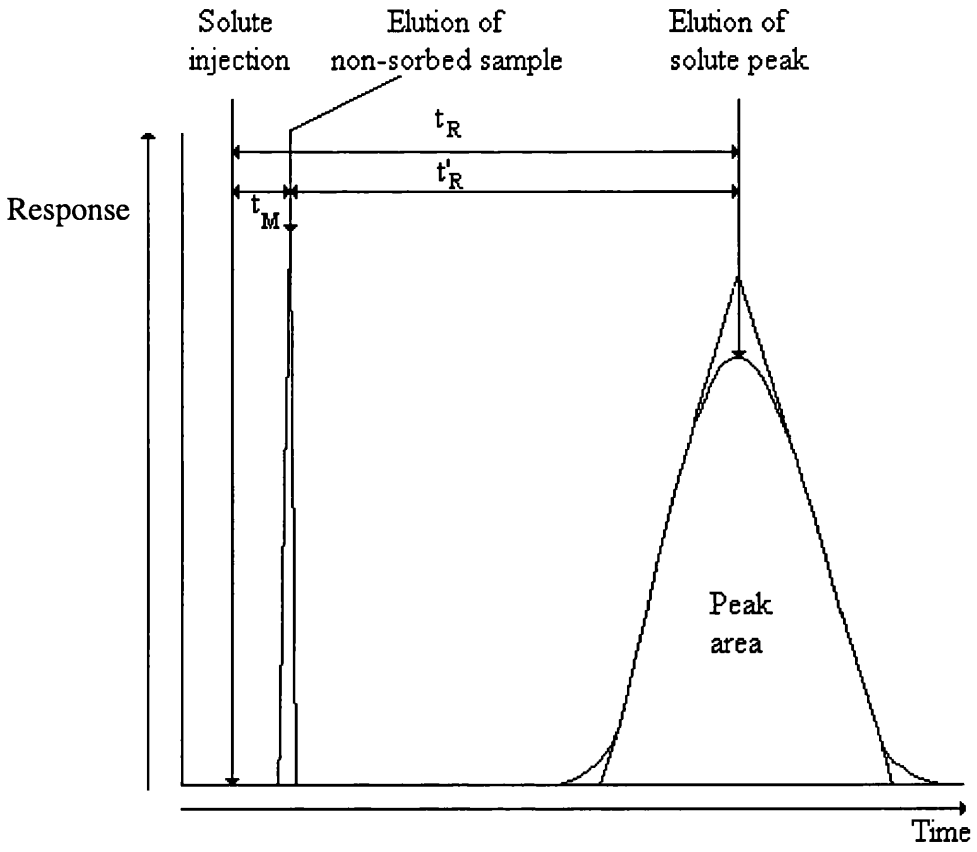


Figure 4.2. Example of Chromatogram

4.1.3. Gas Flow Rate and Correction Factors

4.1.3.1 Carrier Gas Flow Rate

A soap-film meter is undoubtedly the best instrument for measuring gas flows in the region of 1-1000 $cm^{-3}.min^{-1}$. The observed flow rate is corrected for saturation vapour pressure of water, which is little affected by the presence of the detergent and also for the difference in temperature between the column and flow meter, as indicated in the following equation:

$$F = F_a \cdot \left[\frac{T}{T_a}\right]\left[1-\frac{P_w}{P_a}\right] \qquad (4.8)$$

where F is the corrected value of the carrier gas flow rate, F_a the flow rate at the column outlet, T_a the ambient temperature (K), P_a the ambient temperature (Torr), and P_w the vapour pressure of water (Torr) at T_a.

4.1.3.2. Compressibility Correction Factor

Under average chromatographic conditions liquids can be considered incompressible, but not so for gases, and in gas chromatography elution volumes are corrected to a mean column pressure by multiplying them by the gas compressibility factor:

$$J = \frac{3}{2}\left[\frac{P^2-1}{P^3-1}\right] \qquad (4.9)$$

where P is the relative pressure (P_i/P_o), P_i the column inlet pressure and P_o the column outlet pressure, The column inlet pressure is usually measured with a pressure gauge at the head of the column. The gauge actually reads the pressure drop across the column; thus, the inlet pressure used for calculating P in equation (4.9) is the value read from the gauge plus the value for P_o. The compressibility correction factor, J, is always equal to, or less than unity, the reason being that the solute moves faster as it migrates along the column due to increasing velocity of gas as it expands.

4.1.3.3. Gas Imperfection Correction Factor

For the most exact work it may be necessary to allow for non ideal behaviour of the gas phase by applying a virial correction. At moderate column pressure drops and for carrier gases that are insoluble in the stationary phase equation (4.10) is a reasonable approximation:

$$\ln V_N = \ln V_N(0) + \beta.P_o.J_3^4 \qquad (4.10)$$

$$\text{with } \beta = (2B_{12} - V_1) / RT_c \text{ and } J_3^4 = \frac{3}{4}.\left[\frac{P^4 - 1}{P^3 - 1}\right] \qquad (4.11), (4.12)$$

where $V_N(0)$ is the net retention volume at zero column pressure drop, B_{12} the second interaction virial coefficient of the solute with the carrier gas, V_1 the solute molar volume at infinite dilution in the stationary phase and R the universal gas constant. Virial corrections are usually made only when it is desired to calculate exact thermodynamic constants from retention volume measurements. When converting the measured retention volume into activity coefficient, use of the approximate equation (4.7) typically yields a value of activity coefficient 1-5 per cent below the more accurate value given by equation (4.12).

4.1.4. Determination of Thermodynamic Constants

Gas chromatography is now a widely used technique for determining solution thermodynamic properties. This method has several advantages such as small sample size requirement, the ability to measure properties of impure samples and provides easy variation of temperature. Measurements are usually made at infinite dilution, Henry's law region. Thereby, the solute molecules are not sufficiently close to interact between each other, and the environment of each may be only regarded as solvent molecules. Hence, not only the partition coefficient, K, but also the activity coefficient, γ_1, can be assumed to have a constant value. K and γ_1 values can be obtained as follows.

The activity coefficient and the specific retention volume are related by equation (4.13).

$$V_G = \frac{R}{M_2 \gamma_1 P_1^o} \qquad (4.13)$$

where M_2 is the molecular weight of the solvent, P_1^0 the saturation vapour pressure of the pure solute at the given temperature.

The gas liquid partition coefficient is evaluated from the specific retention volume using equation (4.14).

$$K = V_G \rho_L \qquad (4.14)$$

where ρ_L is the liquid phase density at the column temperature.

4.1.5. Kovats Retention Index

Another important value derived from gas chromatographic measurement is the Kovats retention index, I. The retention index elaborated by Kovats uses the homologous series of n-alkanes as a standard; the retention index expresses the retention of a compound relative to n-alkanes analysed under similar, isothermal conditions. The retention index can be defined as the carbon number of a hypothetical n-alkane multiplied by 100 which would have exactly the same retention characteristics (adjusted retention time, adjusted, specific or net retention volume) as the compound of interest[3]. The fundamental equation for the retention index is

$$I_s^{st.ph.} = 100.z + 100.\frac{\left[\log X_{(s)} - \log X_{(z)}\right]}{\left[\log X_{(z+1)} - \log X_{(z)}\right]} \qquad (4.15)$$

where z is the carbon number of n-alkane eluting immediately before the substance of interest denoted by s, and z+1 the carbon number of the n-alkane eluting immediately after the substance x. X is the retention value used in the calculation, e.g. $t_{R'}$, V_G or V_N. Thus the retention index of a substance is expressed on a uniform scale with increased precision in the determination due to the use of two closely eluting standards for the experimental measurements.

4.1.5.1 Reference Compounds

The retention indices of n-alkanes used as reference compounds are defined as follows for any stationary phase and at any column temperature:

$$I(n\text{-}Pz) = 100\ z \qquad (4.16)$$

where n-Pz represents an n-alkane with z carbon atoms. Thus, for example, the retention index of n-pentane is 100*5 = 500 and that of n-decane is 100*10 = 1000 on any stationary phase at any column temperature.

Alkyl ethers and alkanoic acid methyl esters are often preferred to n-alkanes These standards of intermediate polarity are less likely to be retained by absorption on polar phase and give more reliable results.

4.1.5.2. Kovats Index and Temperature

The retention index is temperature dependent and when an index value is required at another temperature it can be obtained by interpolation using an Antoine-type function:

$$I\ (T) = A + \frac{B}{T + C} \qquad (4.17)$$

where T is the column temperature (K) and A,B and C are experimentally derived constants[4]. The curve according to such Antoine-like equation can have a significant linear portion, the length of which depends mainly on the polarity of the substance examined, on the stationary phase applied and on their interactions[5].

For mixture of wide boiling point range the determination of retention indices under isothermal conditions would be time consuming and unnecessarily restrictive. A temperature program method is then required. Under temperature program conditions an approximately linear relationship exist between the elution temperature of n-alkanes and their carbon number. An expression equivalent to equation (4.15) can be given for linear temperature program conditions by replacing the retention data X by the elution temperature, equation (4.18)

$$I_s^{st.ph.} = 100.z + 100.\frac{\left[\log T_{R(s)} - \log T_{R(z)}\right]}{\left[\log T_{R(z+1)} - \log T_{R(z)}\right]} \qquad (4.18)$$

where T_R is the elution temperature (K) and the subscripts, s, z, z+1, are identified in equation (4.15). Temperature programmed retention indices are more sensitive to the chromatographic conditions than isothermal indices and are generally of lower accuracy. The temperature program indices are influenced by the time, the temperature of any isothermal period prior to programming, the temperature program rate, and the method of carrier gas flow control, since the viscosity of the mobile phase increases with temperature[4].

4.1.5.3. The Unified Retention Time

A large number of retention index values are available in the literature. However, their comparison can show some differences. Under favourable circumstances the reproducibility of retention index between different laboratories is within one unit for nonpolar phases and within a few index units for polar phases. Therefore, in order to get rid of the unsatisfactory inter-laboratory reproducibility, Dimov[6] proposed the so-called unified retention index at any temperature, UI_T. Its value could be calculated by the following equation:

$$UI_T = UI_0 + (dUI/dT)\,T \qquad (4.19)$$

Where UIo is the value of UI_T at 273K. The unified retention index concept is a linear regression of the retention data published by various authors at different temperatures. The values of UI_T obtained and its temperature increment were considered as reliable if the data included in the regression matrix were from two authors and at three temperatures at least. UI_T and dUI/dT are statistical values and hence are more reliable than any individual experimental retention values. Dimov, Papozova and more recently Skribc have calculated the unified retention indices and temperature increment of the unified retention indices for a large number of hydrocarbons on squalane[6-9].

4.1.5.4. Connection between the Physicochemical Quantities and Kovats Retention Index

In practice, the following relationship exits between the specific retention volume, V_G and the retention index, I:

$$\log V_G(s)_T = \log V_G(z) + \left[\frac{I_s^{st.ph.}(T) - 100\,z}{100}\right].b_T^{st.ph.} \qquad (4.20)$$

Once V_G values are obtained, they can easily be transformed into partition coefficient value, K, equation (4.14). The $b_T^{st.ph.}$ value is calculated from the plot of the $\log t_{R'}(z)$ versus carbon number (z) relationship for n-alkane:

$$\log t_{R'}(z) = b_T^{st.ph.}\, z + a \qquad (4.21)$$

This correlation is generally linear except for the first few members of the homologous series. Further the temperature dependence of $b_T^{st.ph.}$ can be approximately described with the following equation:

$$b_T^{st.ph.} = (D/T) + E \qquad (4.22)$$

where D and E are experimentally derived constant.

4.1.5.5 Retention Index Differences

In additional to the individual retention index values, the differently formed retention index differences are of great significance because of their physicochemical meaning. Some of these are given below.

$$\Delta I(T) = I_s^{p}(T) - I_s^{Np}(T) \qquad (4.23)$$

Where ΔI is the difference between the retention indexes of a single substance (s) measured on two different stationary phases (p and Np) at an identical, isothermal column temperature (T); p refers to a polar stationary phase whereas Np is a non-polar stationary phase such as squalane. The ΔI values can be approximately used for the calculation of individual relationships, to follow their variations, for the characterisation of the polarity of stationary phases, or to predict the values of retention data, etc[4].

The dI values are applied in molecular structure investigations.

$$dI(T) = I^{st.ph.}_{s(2)}(T) - I^{st.ph.}_{s(1)}(T) \qquad (4.24)$$

Where dI is the difference between the retention indices of two substance, s(1) and s(2), on the same stationary phase (st. ph.) at an identical, isothermal column temperature (T).

4.2. Qualitative and Quantitative Analyses

4.2.1. Qualitative Analysis

A chromatogram provides information regarding the complexity; number of components, peak height or areas, and identity of the components in a mixture. Identification of components or qualitative analysis of a complex mixture depends on the fact that the retention time or retention volume of any components remains constant under fixed chromatographic conditions (column dimensions, column packing material, and column temperature and gas carrier flow rate). However, the certainty of identification based solely on retention is not always satisfactory. More reliable information leading to identification of components can be provided by means of hyphenated systems in which chromatographic and spectroscopic techniques are combined. The principal hyphenated techniques are gas chromatography interfaced to mass spectrometry (GC/MS) and Fourier infrared spectrometry (GC/FTIR). Gas chromatography combined with spectroscopic detectors was recently reviewed by Ragunathan et al[10].

4.2.2. Quantitative Analysis

Quantitative analysis depends on the fact that the area under the peak is proportional to the quantity of the component present, see equation (4.25) Firstly, several methods are available for calculating peak areas. Manual methods include the product of peak height and width at half height, triangulation, cut and weight[1]. These techniques are labour intensive and often lack of precision Nowadays, they are often neglected in favour of electronic integrators and microcomputers that are rapid and can report peak heights and peak areas for even complex chromatograms[1].

$$a_G = A_G f_A v_G \quad (4.25)$$

where a_G is the concentration of a solute A in the gas phase, A_G is the chromatographic peak area of the substance A. Further, f_A is a coefficient expressing the sensitivity of the chromatographic detector for A and v_G is the volume of sample of the gas phase injected into the column.

Four techniques are commonly used to convert peak height or area information into relative composition data for the sample. These are normalisation method, the external standard method, the internal standard method and the method of standard additions[1,11-13]. The internal standard is presented here.

An internal standard is a substance that is added to the sample at the earliest possible point in an analytical scheme to compensate for sample losses occurring during sample preparation and final chromatographic measurements. The substance most commonly used as internal standards include analogues, homologues, isomers, enantioners. Inert n-alkanes can be used as well. The internal standard should be incorporated to the matrix in exactly the same way as the analyte. The internal standard and analyte should elute close together, respond to the detection system in a similar way, and be present in nearly equal concentrations.

4.2.3 Headspace Gas Liquid Chromatographic Method

The static headspace gas liquid chromatographic method is an important technique to determine both the nature and concentration of the volatile analytes present in a sample. It is actually a convenient and indirect way to study volatile organic compounds present in non-volatile matrices without the necessity of carrying out liquid or solid extraction. Before going any further, let us define some important terms. Headspace refers to the gas phase or vapour phase in contact and in equilibrium with a liquid or solid sample and its investigation is the headspace analysis. According to Kolb[14], the matrix is the bulk of the sample that contains the volatile compounds.

The static headspace gas chromatography (HS-GC) technique is a two-step method. First, the sample is filled into a close container. The vial is then thermostated at constant temperature. Part of the volatile compounds will distribute between the sample and the gas phase until the equilibrium is reached. Then, an aliquot of the headspace is

introduced into the gas chromatograph. The carrier gas stream transports the aliquot into the column where it is analysed in the usual way. The information obtained from this one aliquot is used to determine the nature and concentration of the volatile present in the original sample. A schematic of the static HS-GC method is given in Figure 4.3.

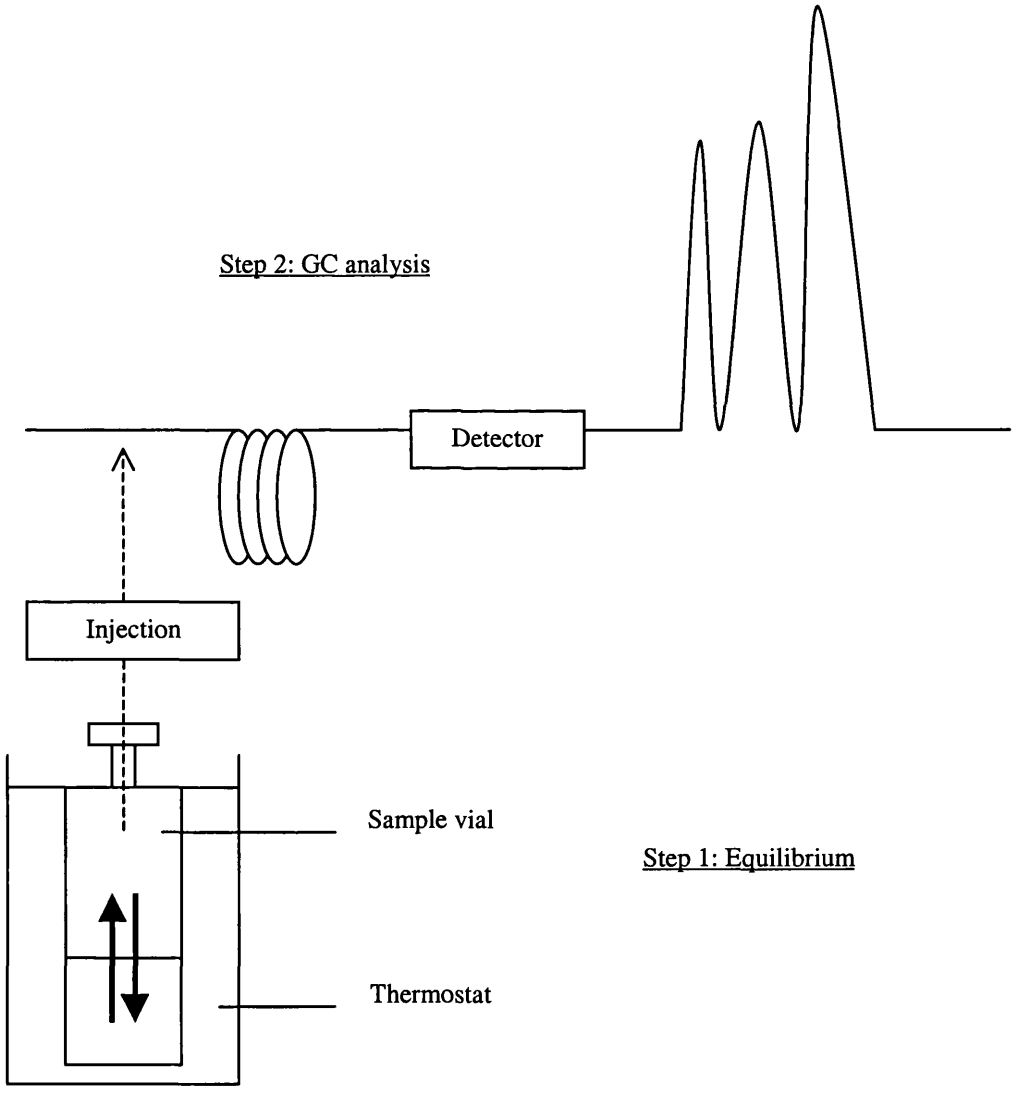


Figure 4.3. Principle of the static headspace gas chromatographic method

4.2.3.1. Theory

Consider a closed vessel containing a solution of solute A in a solvent. This analyte will distribute between the gas and liquid phase as shown in Figure 4.4:

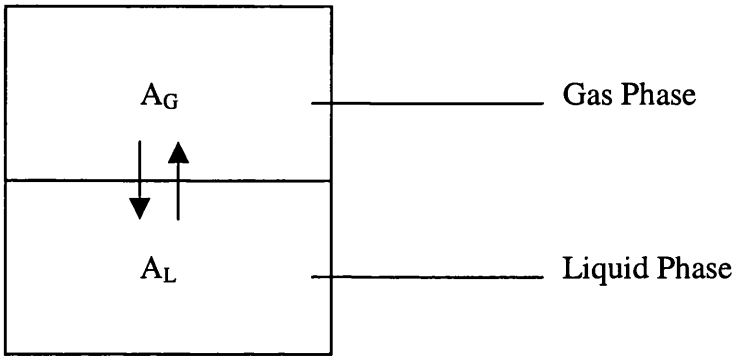


Figure 4.4. Equilibrium between the gas (G) phase and the liquid (L) phase

The distribution of the component A between the gas (G) and the liquid (L) phase is the accordance with the gas / liquid partition coefficient, or Ostwald solubility coefficient, L, which is defined as follows

L = [molar conc. in the liquid phase] / [molar conc. in the gas phase] (4.26)

Therefore, for A:

$$L = a_L / a_G \quad (4.27)$$

The concentration in the liquid phase is then:

$$a_L = L_A \cdot a_G \quad (4.28)$$

Furthermore, the chromatographic peak area of a substance in the gas phase, A_G, is proportional to the concentration of its substance in the gas phase according to equation (4.25). Replacing a_G by $A_G^A \cdot f^A \cdot v_G$ in equation (4.28), one obtains:

$$a_L = L_A \cdot A_G^A \cdot f^A \cdot v_G \quad (4.29)$$

Therefore, gas chromatographic analysis of the headspace above a solution enables the concentration of solute A in solvents to be obtained.

Headspace analysis method is a simple and straightforward method. However, it is very likely to be adversely influenced by several experimental variables such as the sample preparation, the headspace sampling and the chromatographic analysis. These drawbacks can be overcome by setting up an internal headspace analysis method. A standard substance is now added to the system to compensate for sample losses occurring during sample preparation and final chromatographic analysis[14].

Consider a new system composed of a standard D and an analyte A; both components will distribute between the gas and liquid phase as shown in Figure 4.5:

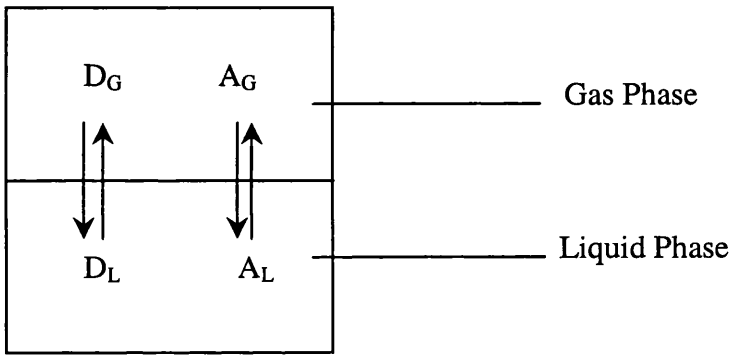


Figure 4.5. D and A in equilibrium between the gas and the liquid phase.

According to equation (4.25), the concentration of the standard in the liquid phase can be estimated from gas chromatographic analysis. Then, from equation (4.25),

$$d_L = L_D \cdot A_G^D \cdot f^D \cdot v_G \qquad (4.30)$$

where L_D, A_G^D and f^D are the distribution constant, the chromatographic peak area of the substance D, and the sensitivity of the detector for D respectively.
Combining equations (4.25) and (4.30), one obtains:

$$a_L / d_L = ((L_A \cdot f^A)/(L_D \cdot f^D)).(A_G^A / A_G^D) \qquad (4.31)$$

The v_G term is cancelled out. Hence, this method does not require exactly reproducible sample volumes, and allows the use of common as well as special syringes to remove the headspace.
Finally,

$$a_L / d_L = k.(A_G^A / A_G^D) \qquad (4.32)$$

where k is the constant of proportionality including the following variables: the sample preparation, the headspace sampling, the chromatographic conditions, the detector sensitivity.

4.2.3.2. Applications

Thanks to its simplicity, the headspace method has found a large number of applications. The technique is widely used as a fingerprint or to facilitate the investigation of volatile components present in a complex sample. For instance, the method is applied to investigate essential oils and perfume composition in fragrance chemistry, to characterise the volatile components of wine, cheese, plants and so forth, and to determine toxic impurities in the environment and ethanol in blood. It is also used to determine trace composition of volatile organic compounds. In addition to the direct analysis of the sample for its quantitative and/or qualitative composition, the headspace technique can be used for physicochemical measurements. They permit the determination of several parameters, such as vapour pressure of pure compounds, activity coefficients, and partition coefficients and reactions rates[14].

4.3. Characterisation of Stationary Phases

Retention data are commonly used to characterise stationary phase properties, viz. solvent strength (polarity) and selectivity, see section 4.1.3.1. Various approaches to characterise solvent strength and selectivity of a stationary phase have been proposed. Early studies in this field were carried out by Rorhschneider[15], followed by McReynolds[16].These techniques together with more recent methods are reviewed in detail elsewhere and only the solvation parameter model is presented here.

The Abraham general solvation equation for gas to solvent partition process provides information about the contribution of defined intermolecular interactions to the sorption properties of the stationary phase. The equation, presented in more detail in chapter 2, takes the form of:

$$SP = c + e\mathbf{E} + s\mathbf{S} + a\mathbf{A} + b\mathbf{B} + l\mathbf{L} \qquad (4.33)$$

Here SP refers to gas chromatographic data for a series of solutes on a given stationary phase, for example log V_G or log $t_{R'}$ or I. The solute parameters, **E**, **S**, **A**, **B** and **L**, characterise the solute effects on various solute-stationary phase interactions occurring throughout the chromatographic process. Hence, the coefficients in equation (4.33) correspond to the complimentary effect of the stationary phase on these interactions. The regression equation coefficients encode the stationary phase properties. The coefficients can be regarded as system constants that define the stationary phase. The reference for such characterisation will be the gas phase, because all GC data refers to transfer from the gas phase to the stationary phase.

As pointed out in chapter 2, the system constant can be interpreted as follows:

- The *e*-coefficient shows the tendency of the phase to interact with solutes through π- and n-electron pairs.
- The *s*-coefficient gives the tendency of the phase to interact with dipolar/polarizable solutes
- The *a*- and *b*-coefficients denote the hydrogen bond basicity and hydrogen bond acidity of the phase, respectively.
- The *l*-coefficient is a measure of the hydrophobicity of the stationary phase.

The coefficients in equation (4.33) are not only fitted constants but must obey general chemical principles. It should be pointed out that the regression equations remain the same, except for the c-constant, no matter whether the dependent variable is log K, or log V_G or even log $t_{R'}$. However, the system constants change distinctly with temperature since the intermolecular interactions in general decrease with an increase of the temperature. The effect of temperature on regression coefficient values is reported elsewhere[17].

Abraham and co-workers have applied equation (4.33) to the classification of stationary phases[17]. Examples of system constants for non-ionic stationary phases at

395-397 K are listed in Table 4.2. The constants give information about the nature of the stationary phase of interest. From Table 4.1, it can be seen that SQ, SE-30 and OV-11 are rather non-polar (*s*) with non or weak hydrogen bond basicity (*a*). On the other hand, OV-330, OV-275 and CW20M are both dipolar/polarisable and of considerable hydrogen bond basicity. It is rare to encounter hydrogen bond acid stationary phases. The phenolic stationary phase of Abraham[18] is both strong hydrogen bond acid (*b*) and hydrogen bond base (*a*). This result agrees well with the structure of the stationary phase H10 that contains strongly basic sulfone functionality as shown in Figure 4.6. Similarly, the alcohol phase, PSF6, is a strong hydrogen bond acid but also has no hydrogen bond basicity. This is in line with the presence of the 1,1,1,3,3,3-hexafluoropropanoyl moiety.

The solvation model can be used to deduce important chemical properties of stationary phases from the regression coefficients. This is an important tool to quantify selectivity differences between stationary phases. As pointed out by Abraham et al.[17], this method can be used to identify phases with redundant properties and then to replace them with phases with properties that would enhance the separations of complex mixtures, for example. Besides, chemometric classification procedures have been used to sort stationary phases properties that have been put forward by the solvation model. This technique has been reported in detail in a review[17].

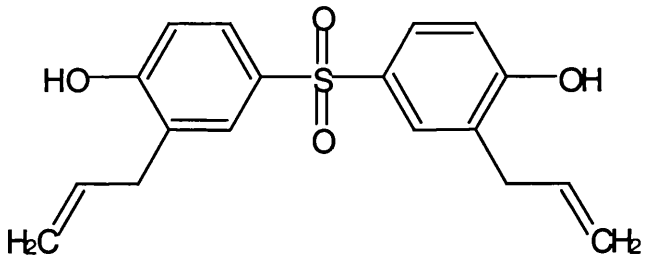


Figure 4.6. Structure of the phase H10.

Table 4.1. System constants for non-ionic stationary phases (395-397K)[a].

Stationary Phase	*c*	*e*	*s*	*a*	*b*	*l*
SQ	-0.222	0.129	0.011	0.000	0.000	0.583
SE-30	-0.194	0.024	0.190	0.125	0.000	0.498
OV-11	-0.303	0.097	0.544	0.174	0.000	0.516
OV-330	-0.430	0.104	1.056	1.419	0.000	0.481
OV-275	-0.909	0.206	2.080	1.986	0.000	0.294
PFS6	-0.510	-0.360	0.820	0.000	1.110	0.540
DOP	-0.275	0.000	0.797	1.004	0.000	0.571
CW20M	-0.560	0.317	1.256	1.883	0.000	0.447
DEGS	-0.669	0.197	1.668	2.246	0.000	0.411
HP10	-0.568	-0.051	1.323	1.266	1.457	0.418

[a] Adapted from Ref. 17

4.4. References

1. C.F. Poole, S.K. Poole, in Chromatography Today, Elsevier Sci. Pub. B.V, Amsterdam, (1991).
2. J.R. Conder, C.L. Young, in Physical Measurement by Gas Chromatography, Wiley, New York, NY, (1979).
3. E.Sz. Kovats, Helv. Chim. Acta, 41 (1958) 1915.
4. M.V. Budahegyi, E.R.Lombosi, T.S.Lombosi, S.Y. Meszaros, Sz. Nyiredy, G. Tarjan, I. Timar and J.M. Takacs, J.Chromatogr., 271 (1983) 213.
5. J. Takacs, M.Rockenbauer and I.Olacsi, J. Chromatogr., 42 (1969) 19.
6. N. Dimov, J.Chromatog., 366 (1985) 3447.
7. D.Papazova, R.Milina, Cromatographia, 25 (1988) 177.
8. B.D. Skribc, J.D. Cvejanov, L.S.Pavic-Suzuki, Chromatographia, 42 (1996) 660.
9. D. Papzova, N. Dimov, J.Chromatogr., 356 (1986) 320.
10. N.Ragunathan, K.A. Krock, C.Klauvun, T.A. Sasaki and C.L. Wilkins, J.Chromatogr. A, 856 (1999) 349-397.
11. J.A. Perry, in Introduction to Analytical Gas Chromatography. History, principles and Practice, Dekker, New York, NY, (1981).

12. M.L. Lee, F.J. Yang, K.D. Bartle, in Open Tubular Column Gas Chromatography: Theory and Practice, Wiley, New York, NY, (1984).
13. E. Katz, Quantitative Analysis using Chromatographic Techniques, Wiley, New York, NY, (1987).
14. B.Kolb and L.S.Ettre, Static Headspace Gas Chromatography: Theory and Practice, Wiley-CH, New York, NY, (1997).
15. L. Rohrschneider, J.Chromatogr.Sci, 11 (1973) 160, L.Rohrschneider, Chromatographia, 38 (1994) 679.
16. W.O. McReynolds, J.Chromatogr.Sci., 8 (1970) 685.
17. M.H.Abraham, C.F.Poole, S.K. Poole, J.Chromatogr. A, 842 (1999) 79.
18. M.H. Abraham, J.Andobnian-Haftvan, I. Hamerton, C.F.Poole, T.O.Koolie, J.Chromatogr. 646 (1993) 351.

Chapter 5 Aims of the Present Work

The perceived effect of VOCs can be divided into odor and sensory irritation. Bio-Assays for odor detection thresholds, ODT, nasal pungency thresholds, NPT, and eye irritation thresholds, EIT, are thus very relevant to the assessment of indoor air quality. However, it is quite impractical for bio-assays on odor and sensory irritation to be conducted on the large number of VOCs that could be encountered indoor, using panels of human subjects. One method that will aid to overcome this problem is the application of quantitative structure activity relationship, QSARs. The Abraham general solvation equation will be used to achieve this aim. The solvation equation coefficients, *c, e, a, b,* and *l,* are generated by multiple linear regression of the threshold value sets against corresponding sets of the five Abraham solute descriptors.

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L}$$

Where SP is one VOC property, **E** is the VOC excess molar refraction, **S** is the dipolarity / polarizability, **A** and **B** are the solute overall hydrogen-bond acidity and basicity, and **L** is defined through log L^{16} where L^{16} is the Ostwald solubility in n-hexadecane at 298 K.

Algorithms that will be developed in this work could only be used to predict NPT, ODT values for VOCs if solute descriptors are available. The first aim of this work (Chapter 6, Chapter 7 and Chapter 8) will be to calculate the solute descriptors for a number of VOCs such as refrigerants and terpenes using two possible methods:

(i) Descriptor assignment using gas / solvent and water / solvent partition measurements.

(ii) Descriptor assignment using reversed-phase HPLC and GLC data.

The second objective of this work (Chapter 9 and Chapter 10) will be to develop mathematical models to predict the chemosensory effects of VOCs in humans. Nasal pungency threshold (NPT) and odour detection threshold (ODT) values are available for a series of VOCs that cover a large range of solute properties such dipolarity, hydrogen bonding capacity and lipophilicity. Each of these sets of biological data is regressed

against a number of solute descriptors to obtain QSARs for ODT and NPT. The solvation equations obtained will provide information on the physico-chemical interactions involved in either the odor or nasal pungency process as well as characterise the system under investigation. Once VOC descriptors are available, the Abraham general solvation equation may be applied to predict the solute property of interest (Chapter 11).

The final goal of this work (Chapter 12) will be to investigate the interaction between two VOCs at proximity of the chemosensory receptor areas. Then, a headspace gas chromatographic method is devised to determine the extent of complexation between hydrogen bond donors and hydrogen bond acceptors in octan-1-ol, model of the receptor area. These measurements will allow the prediction of the percentage of association between two VOCs at the proximity of the chemosensory receptor areas.

Chapter 6 Characterisation of Squalane and Apolane; Calculation of Further **L**-Values

6.0 Introduction

The Abraham general solvation equation (6.1) is of great interest in understanding physicochemical and biochemical phenomena in which solutes distribute between the gas phase and a condensed phase.[1] As explained in chapter 3, equation (6.1) consists on a linear combination of five solvation descriptors which represent the solute physicochemical properties. **E** is an excess molar refraction, **S** is the dipolarity / polarizability, **A** and **B** are the overall hydrogen-bond acidity and basicity. Finally, **L** is defined through $\log L^{16}$, where L^{16} is the solute Ostwald solubility coefficient on n-hexadecane at 298K.[1]

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (6.1)$$

The dependent variable, SP, is the logarithmic value of 'some property' of a series of solute in a given phase system. The regression coefficients, *c*, *e*, *s*, *a*, *b*, and *l* are found by multiple linear regression analysis, MLRA, and reflect the complementary properties of the solvent phase or biophase. Equation (6.1) has been used to analyse and predict numerous gas / solvent partitions, gas / biophase systems, and to a large number of gas chromatographic systems. It is then a well-trained and tested equation. However, the use of equation (6.1) depends on determination of the solute descriptors that are mostly derived from experimental gas liquid chromatographic data, see chapter 3. They have been obtained for more than 3000 solutes and compiled in an in-house database. Out of this database, **L** is the descriptor for which least values are available. Therefore, in order to make a wider use of equation (6.1), **L** values need to be assigned for greater number of solutes.[1]

The solute descriptor, **L**, initially formulated by Abraham et al., characterises the solute size and its tendency to participate in solute / solvent interactions of the general London dispersion type.[2] **L** is now a well-established descriptor in linear free energy relationship. Hence, it is not surprising that several studies on **L** determination

have been carried out. **L** values were originally measured on n-hexadecane stationary phase at 298K.[2] However, this technique is limited to volatile and semi-volatile solutes and is often replaced by a number of alternative methods.[3-6] These approaches based either on experimental or empirical data were presented in detail in chapter 3.

Other hydrocarbon-like stationary phases, such as squalane, apiezon and apolane, have been proposed as substitutes for n-hexadecane.[5] The main advantage of these hydrocarbons is that a larger spread of solutes can be analysed; experiments with such solvent phases are not limited to volatile or semi-volatile compounds but also include non-volatile ones. Two works in this field are presented here. First, Weckwerth and co-workers[3] have recently shown that gas / apolane partition coefficient values, L^{87}, correlate with the descriptor **L**. Apolane, a highly branched nonpolar hydrocarbon synthesised by Kovats,[7] is a stable non-volatile stationary phase that can be used over the temperature range of 300-553K. The authors measured L^{87} and L^{16} values for 139 solutes by open-tubular capillary GLC at 313K and 298K respectively, and put forward a strong relationship between $\log L^{87}$ and $\log L^{16}$, or **L**, see equation (6.2). Then, knowing the L^{87} values, it is easy to calculate the corresponding **L** values. However, it is important to note that only a few **L** values have been derived from gas / apolane partition coefficient values, so far.

$$\text{Log}L^{16} = \mathbf{L} = 0.175\ (0.024) + 1.1004\ (0.0082)\ \log L^{87} \qquad (6.2)$$

$n = 139$, $r^2 = 0.992$, $sd = 0.093$

Here and elsewhere, n is the number of data points, r is the overall correlation coefficient, and sd is the standard deviation in the dependent variable. The sd values for the coefficients are given in parentheses.

This recent approach is closely related to the Abraham alternative method for estimation of **L**.[5] Equation (6.1) provides actually a number of options for the direct determination of **L**. Values of **L** can be obtained through GLC measurements on non-polar solvent phases in which the *l*-coefficient is large. GLC retention data can then be fitted to an equation of the form,

$$SP = c + e.\mathbf{E} + l.\mathbf{L} \qquad (6.3)$$

Here, SP can be the logarithmic value of retention volume, or relative retention index, or gas / solvent partition coefficients, or SP can be the retention index, I. Since **E** can be calculated with sufficient accuracy from structure or calculated directly from the refractive index of liquids, this method is an easy way to obtain **L** value for any given solutes. This method based on the correlation of retention properties on low polarity phases other than n-hexadecane, is a well-established method that has led to a considerable amount of **L** values.[5]

In this work, the Abraham method for estimation of **L** is favoured. Equations similar to equation (6.3) have been developed for gas / squalane partitions, L^{SQ}, at 298K and gas / apolane partitions, L^{87}, at 298K and 313K. From these equations, some 146 new L values were determined. An equation for water / squalane partitions at 298K was also developed according to the Abraham solvation equation for processes within condensed phases[1], see equation (6.4).

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + v.\mathbf{V} \qquad (6.4)$$

This equation for water / squalane partitions together with the equation for gas / squalane systems were compared with similar equations for water / alkanes and gas / alkanes previously developed by Abraham and co-workers.[2]

6.1. Construction of Solvation Equations for Gas / Alkane and Water / Alkane Partition Process

6.1.1. Construction of an Equation for log L^{SQ}

Squalane (2,6,10,15,19,23-hexamethyltetracosane) is a stable non-polar and non-volatile solvent that can be used over the temperature range of 293-423 K. Thanks to these properties, squalane has been extensively used, and considerable quantities of thermodynamic properties for a wide range of solutes have been accumulated. A survey of the literature showed that there were enough data on squalane at 298K to set up a statistically significant regression equation similar to equation (6.3) where the dependent variable, SP, is the gas / squalane partition coefficient, L^{SQ}. Next, are presented the various methods in use for the calculation of L^{SQ} values from literature data.

6.1.1.1 Calculation of Gas / Squalane Partition Coefficient, L^{SQ}

Gas / squalane partition coefficient, L^{SQ}, at 298K for 396 varied solutes have been obtained from the literature. Some of these L^{SQ} values were available as such,[8-14] others were calculated from the reported activity coefficient[15,16], γ, and Henry's law coefficient[17,18], H^{SQ}, at 298K. A large number of L^{SQ} values were derived from gas chromatographic data, such as the solute specific retention volume[9,10,19,20], V_G, and the unified retention index at 298K[21-24], UI_{298}. L^{SQ} values for five solids were calculated from their solubility in squalane. The transformation of these several constants into gas / squalane partition coefficients is now covered.

First, activity coefficient and Henry's law coefficient values were transformed into L^{SQ} values by means of the following equations.

$$L^{SQ} = \frac{R\,T\,\rho_{SQ}}{\gamma\,P^0\,M_{SQ}} \qquad (6.5)$$

$$L^{SQ} = \frac{R\,T\,\rho_{SQ}\,1000}{H^{SQ}\,M_{SQ}} \qquad (6.6)$$

Here and elsewhere, ρ_{SQ} and M_{SQ} are the density ($g.cm^{-3}$) and the molecular weight ($g.mol^{-1}$) of the pure squalane . R is the gas constant ($dm^3.atm.K^{-1}.mol^{-1}$), T is the system temperature (K), and P^0 is the vapour pressure of the pure solute (atm). Additional data were required to make the use of these equations. An average value of squalane density, ρ_{SQ}, at 298K was obtained from literature.[10,11] (ρ_{SQ} = 0.80745 $g.cm^{-3}$). P^0 values at 298K were evaluated from Antoine equation constants found in references 25 to 30.

Secondly, a large number of L^{SQ} values was obtained from gas chromatographic data. As previously pointed out in Chapter 4, L^{SQ} is connected to the solute specific retention volume, V_G, at the column temperature.

$$L^{SQ} = V_G \rho_{SQ} \qquad (6.7)$$

Note that if V_G is the specific retention volume corrected to 273K then

$$L^{SQ} = \frac{V_G\,T.\rho_{SQ}}{273} \qquad (6.8)$$

Further L^{SQ} values were then calculated from reported V_G values at 298K using the above equations. Some V_G values at 298K were estimated from V_G values measured at either 295K and 303K.[19,20] These estimated values are notified in Table 6.1.

L^{SQ} values were also calculated from the unified retention index at 298K, UI_{298}. UI_{298} values were calculated from equation (6.9):

$$UI_{298} = UI_0 + 298\,(dUI / dT) \qquad (6.9)$$

where UI_0 is the value of UI_T at 273K , dUI /dT is the index increment with the analysis temperature (usually given as dUI/283K). UI_{298} values are based on statistical analyses of all the existing experimental retention index values on squalane. Dimov, Papazova

and more recently Skrbic published UI_0 and temperature increment values for 333 hydrocarbons on squalane. Pompe and Novic have reported an extensive data set of 381 organic compounds with known retention indices at 343 K, taken from literature.[32] Although this data set is believed to be the largest data set available in the literature, it was decided not to use it in this work because of various inconsistencies regarding the stationary phase and column temperature. Squalane and squalene were both cited as the stationary phase. Further looking back at the original references, it appeared that the retention indices values reported in the article were measured at either 323 or 343K.

In practice, the following relationship exists between the specific retention volume, V_G and the retention index, I.[33]

$$\log V_G = \log V_G(z) + \left[\frac{I - 100\,z}{100}\right].b_T^{SQ} \qquad (6.10)$$

where z is the solute carbon number, b_T^{SQ} is the slope of the plot of log $t_{R'}(z)$ ($t_{R'}$ is the adjusted retention time) versus carbon number for n-alkanes at 298K. In cases where the unified retention index at 298K, UI_{298}, is used instead of I, equation (6.11) becomes

$$\log V_G = \log V_G(z) + \left[\frac{UI_{298} - 100\,z}{100}\right].b_{298}^{SQ} \qquad (6.11)$$

Furthermore,

$$\log V_G = \log V_G(z) + UI_{298}\left(\frac{b_T^{SQ}}{100}\right) - z\,b_{298}^{SQ} \qquad (6.12)$$

Finally, replacing log V_G by (log L^{SQ} - log ρ_S) in equation (6.12), one obtains

$$\log L^{SQ} = A\,UI_{298} + C \qquad (6.13)$$

where $A = b_{298}^{SQ} / 100$ and $c = -z\,b_{298}^{SQ} + \log V_G(z) + \log \rho_s$

UI_{298} and L^{SQ} values were available for 32 solutes as listed in Table 6.2. The following correlation was obtained and was used to transform the remaining 301 UI_{298} values into their corresponding L^{SQ} values with a reasonable accuracy.

$$L^{SQ} = 0.005\ (\pm 3.10^{-5})\ UI_{298} - 0.452\ (\pm 0.021) \qquad (6.14)$$

$n = 32,\ r^2 = 0.999,\ sd = 0.02,\ F = 25566.06$

A value for the $b_{298}{}^{SQ}$ constant can be calculated from the slope of equation (6.14). It is interesting to note that this value, 0.500, is quite similar to the one given in reference 32 in which a value of 0.468 was actually proposed.

Finally, the partition coefficient of a solid between the gas phase and the squalane phase, L^{SQ}, can be obtained from the solubility of the solid in $mol.dm^{-3}$ in squalane, C_{SQ}, and the gaseous concentration, C_G, as shown in equation (6.15), provided that certain conditions are fullfilled.[34] The major condition being that the same solid phase must be in equilibrium with the saturated solution in squalane.

$$L^S = \frac{C_S}{C_G} \quad \text{or} \quad \log L^S = \log C_S - \log C_G \qquad (6.15)$$

with

$$\text{Log}\ C_G = \log P^0 - \log RT \qquad (6.16)$$

Acree and co-workers measured the solubility of five solids in squalane[25], C_{SQ} (mol/l), at 298K. L^{SQ} was then obtained from C_{SQ} and the solid concentration in the gas phase, C_G, as explained in equations (6.15) and (6.16). L^{SQ}, C_{SQ} and C_G values are given in Table 6.3.

Gas / squalane partition coefficients, L^{SQ}, of 394 nonpolar and polar organic compounds, spanning a wide range of functional groups, dipolarities and hydrogen bonding capabilities and of two inorganic gases, were calculated using one of the above methods or a combination them. Hence, average L^{SQ} values were considered for some compounds of the set.

Table 6.1. Estimation of $\log V_G$ at 298K.

Compounds	Obs V_G (293K)[a]	Obs V_G (303K)[b]	Calc V_G (298K)	error[e]
3-Methylhexane	1099.0	-	882.24[c]	2.56
2,3-Dimethylpentane	1018.0	-	817.98[c]	2.53
3,3-Dimethylpentane	859.4	-	692.15[c]	2.58
2,2,3-Trimethylbutane	686.2	-	554.74[c]	2.78
2,2,4-Trimethylpentane	1252.0	-	1003.63[c]	2.70
trans-2-Hexene	632.1	-	511.81[c]	2.87
cis-3-Hexene	354.4	-	291.49[c]	3.47
2-Methyl-2-pentene	440.6	-	359.88[c]	3.26
Mesitylene	-	15400.0	20839.6[d]	182.93
n-Propylbenzene	-	10600.0	14325.2[d]	123.68
Isopropylbenzene	-	7830.0	10565.8[d]	89.53
Chlorobenzene	-	3100.0	4146.4[d]	31.49

[a] Ref. 19
[b] Ref. 20
[c] Calculated from V_G (298K) = 0.793 V_G(293K) + 10.323, n=6, r^2 = 0.999, sd = 6.20
[d] Calculated from V_G (298K) = 1.357 V_G(293K) + 60.76, n=7, r^2 = 0.999, sd = 16.09
[e] standard error of the predicted value

Table 6.2. Calculation of Log L^{SQ} from UI_{298} values from equation (6.14)

Solute	obs. $\log L^{SQ}$	Ref.	UI_{298}	Ref.	calc. $\log L^{SQ}$	obs.$\log L^{SQ}$ - calc.$\log L^{SQ}$
2-Methylbutane	1.91	16	474.24	24	1.92	-6.54E-03
2-Methylpentane	2.40	16	569.45	21	2.40	1.84E-03
3-Methylpentane	2.47	16	582.98	21	2.46	1.53E-03
2,2-Dimethylbutane	2.21	16	534.74	21	2.22	-1.55E-02
2,3-Dimethylbutane	2.37	16	564.97	21	2.37	-3.61E-04
2-Methylhexane	2.89	16	666.17	21	2.88	5.95E-03
3-Methylhexane	2.95	19	675.24	21	2.93	1.97E-02
3-Ethylpentane	2.97	16	684.65	21	2.97	4.72E-04
2,2-Dimethylpentane	2.68	16	624.35	21	2.67	5.48E-03
2,3-Dimethylpentane	2.91	19	669.90	21	2.90	1.36E-02
2,4-Dimethylpentane	2.70	16	629.02	21	2.69	5.35E-03
3,3-Dimethylpentane	2.84	19	655.61	21	2.83	1.26E-02

Solute	obs. $\log L^{SQ}$	Ref.	UI_{298}	Ref.	calc. $\log L^{SQ}$	obs.$\log L^{SQ}$ - calc.$\log L^{SQ}$
Cyclopentane	2.37	9,16	562.39	21	2.36	8.66E-03
Cyclohexane	2.85	9	658.27	21	2.84	4.18E-03
1-Pentene	1.95	16	480.24	24	1.95	2.99E-03
1-Hexene	2.47	16	581.65	21	2.46	1.71E-02
trans-2-Hexene	2.47	19	597.31	21	2.54	-6.23E-02
cis-3-Hexene	2.46	19	592.20	21	2.51	-4.58E-02
2-Methyl-1-pentene	2.47	16	579.59	21	2.45	1.98E-02
2-Methyl-2-pentene	2.56	19	597.96	21	2.54	1.69E-02
1-Heptene	2.97	16	681.32	21	2.96	1.72E-02
1-Octene	3.46	16	780.17	21	3.45	8.16E-03
Cyclohexene	2.89	16	666.07	23	2.88	5.56E-03
3-Hexyne	2.69	10	630.08	23	2.70	-9.73E-03
Benzene	2.72	9,11,16	629.87	21	2.70	2.17E-02
Toluene	3.24	9,16	739.14	21	3.25	-8.23E-03
Ethylbenzene	3.68	16	828.20	21	3.69	-1.09E-02
o-Xylene	3.85	16	862.70	21	3.86	-9.00E-03
m-Xylene	3.75	16	845.66	21	3.78	-2.79E-02
p-Xylene	3.73	16	842.48	21	3.76	-3.53E-02
Propylbenzene	4.16	24	917.79	21	4.14	1.69E-02
Isopropylbenzene	4.02	23	889.54	20	4.00	2.60E-02

Table 6.3. Determination of $\log L^{SQ}$ values from solubility data [a]

Solute	$\log C_{SQ}$	$\log C_G$	$\log L^{SQ}$
trans-Stilbene	-1.316	-8.580	7.260
Acenaphthene	-0.517	-6.900	6.380
Pyrene	-1.132	-9.650	8.520
Biphenyl	-0.380	-6.280	5.900
Diphenyl sulfone	-2.431	-11.030	8.600

[a] Data taken from Ref. 25.

6.1.1.2. Regression Analyses

Here and elsewhere, multiple linear regression analysis and statistical tests were performed on an in-house software using the SmartWare II (informix Software, Inc, Kansas) database and statistical tools. The solute descriptors used in the solvation parameter model were from the same in-house software.

The Abraham solvation parameters were available for a set of 255 compounds among the large set of 396 solutes. The total of 255 compounds is listed in Table 6.9, see section 6.4. Some 15 compounds were excluded from the 255 data set because preliminary analysis showed that they were out of the range, mostly because of experimental problems. For instance, Nitta et al.[9] pointed out that the log L^{SQ} values for acetonitrile was not reliable enough because of the poor solubility of this solute in squalane and, on the other hand, its high absorption at the interface. The 15 outliers are listed in Table 6.9 as Nos 241-255. A total of 241 compounds were left for the final analysis. The descriptor space used to determine the individual models is defined by **L** = -1.200 to 6.469, E = 0.000 to 1.604. The space covered by each descriptors as well as by the log L^{SQ} values are given in Figures 6.1-6.3 It can be seen that both the **L** and log L^{SQ} variables have a normal distribution over a range of 9 log units. This is condition required for establishing good linear correlation. This is not the case for the **E** descriptor. Most of the **E** values are between 0.000 and 1.

24 compounds were selected in a random order to form the test set, to give 217 compounds as the training set. The test set that is given in Table 6.4. The 217 log L^{SQ} values were regressed against the corresponding Abraham parameters, **E** and **L**.

$$\log L^{SQ} = -0.111\ (\pm 0.009) + 0.076\ (\pm 0.010)\ \mathbf{E} + 0.994\ (\pm 0.003)\ \mathbf{L} \qquad (6.17)$$

$n = 217,\ r^2 = 0.999,\ sd = 0.038,\ F = 89104.16$

The predictive ability of the above equation (6.1) can be probed through the test set of 24 compounds given in Table 6.5, where the observed and calculated log L^{SQ} values for equation (6.17) are listed. The standard error of the predicted value for the 24 compound test set is 0.047 log units over a range of 7 log units. Finally, the training set and test set were combined to obtain equation (6.18) for the total of 241 compounds.

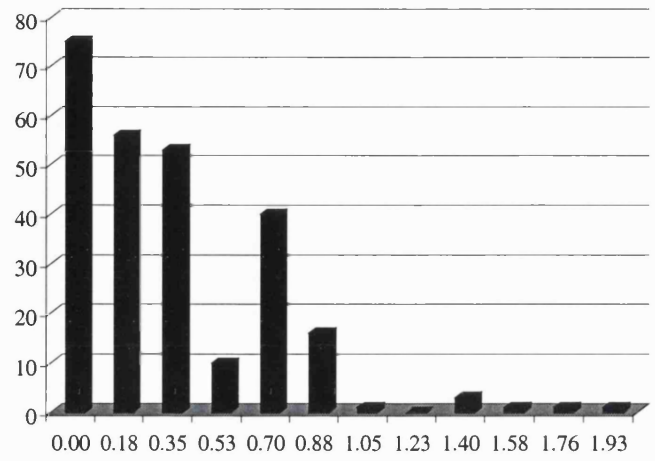


Figure 6.1. Distribution of the descriptor **E**

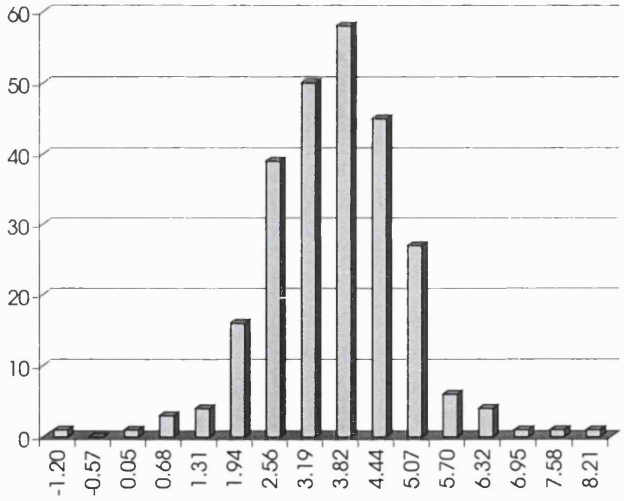


Figure 6.2. Distribution of the descriptor **L**

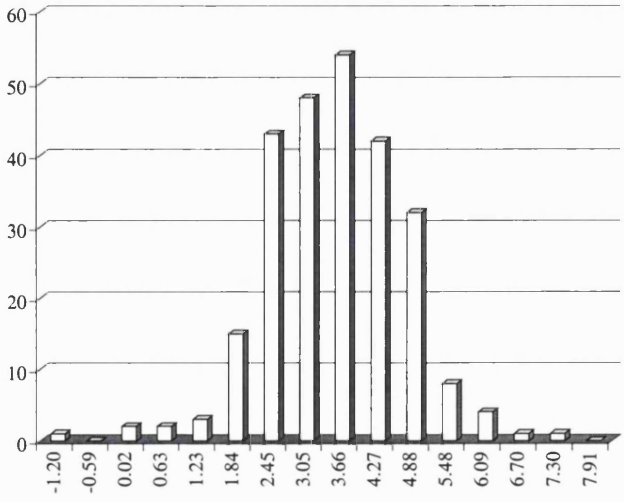


Figure 6.3. Distribution of the dependent variable log L^{SQ}

Table 6.4. Test set of compounds [a,b]

Compounds	E	L	obs $logL^{SQ}$	calc $logL^{SQ}$	obs. $logL^{SQ}$ - calc. $logL^{SQ}$
Hydrogen	0.000	-1.200	-1.195	-1.304	0.109
2-Methylpropane	0.000	1.409	1.320	1.290	0.030
Hexane	0.000	2.668	2.560	2.541	0.019
2,3-Dimethylpentane	0.000	3.016	2.890	2.887	0.003
2,4-Dimethylhexane	0.000	3.319	3.180	3.188	-0.008
2,3,4-Trimethylpentane	0.000	3.481	3.270	3.349	-0.079
2,4-Dimethylheptane	0.000	3.758	3.640	3.624	0.016
2,2,4-Trimethylhexane	0.000	3.605	3.460	3.472	-0.012
2,2,3,3-Tetramethylpentane	0.000	3.880	3.770	3.746	0.024
1,3-trans-Dimethylcyclopentane	0.156	3.075	2.930	2.957	-0.027
1,3-trans-Dimethylcyclohexane	0.190	3.655	3.530	3.537	-0.007
Pent-1-ene	0.093	2.047	1.930	1.931	-0.001
trans-Hex-3-ene	0.126	2.659	2.480	2.542	-0.062
3,3-Dimethylbut-1-ene	0.037	2.201	2.030	2.080	-0.050
2-Methylbuta-1,3-diene	0.313	2.101	1.990	2.001	-0.011
Penta-1,4-diene	0.185	1.998	1.800	1.889	-0.089
3-Methylcyclohexene	0.360	3.319	3.150	3.215	-0.065
Dichloromethane	0.387	2.019	1.860	1.925	-0.065
Pentan-3-one	0.154	2.811	2.740	2.695	0.045
Propan-1-ol	0.236	2.031	1.910	1.926	-0.016
Hexan-1-ol	0.210	3.610	3.490	3.493	-0.003
m-Xylene	0.623	3.839	3.760	3.752	0.008
Butylbenzene	0.600	4.730	4.650	4.636	0.014
2-Propyltoluene	0.664	4.766	4.670	4.677	-0.007
1,4-Dimethyl-2-ethylbenzene	0.693	4.824	4.750	4.737	0.013
α-Methylstyrene	0.851	4.292	4.270	4.220	0.050
Chlorobenzene	0.718	3.657	3.620	3.579	0.041

[a] Observed log L^{SQ} values taken from Table 6.9 in section 6.4.
[b] Log L^{SQ} values calculated on equation (6.17)

The output from the MLR analysis of log L^{SQ} with the solute descriptors, **E** and **L**, for the gas / squalane system is summarised in Table 6.5, detail on the output table can be found in Chapter 3.

$$\log L^{SQ} = -0.108\ (\pm 0.008) + 0.080\ (\pm 0.010)\ \mathbf{E} + 0.992\ (\pm 0.003)\ \mathbf{L} \qquad (6.18)$$

$n = 241$, $r^2 = 0.999$, $sd = 0.038$, $F = 98441.48$

The coefficients in equation (6.18) together with the t-statistic values confirm the fact that **L** mainly influences the gas / squalane partition processes. A plot of observed L^{SQ} against calculated L^{SQ} from equation (6.18) is shown in Figure 6.4.

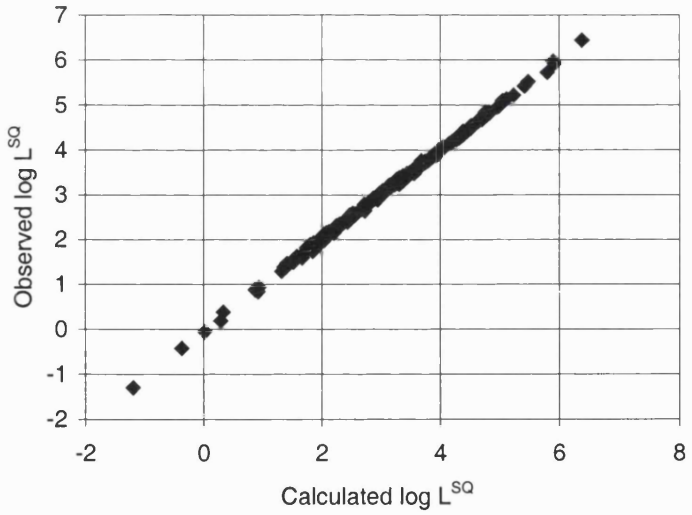


Figure 6.4. Plot of observed against calculated L^{SQ} on equation (6.18).

Table 6.5. Output of the regression analysis of log L^{SQ} and solute descriptors

1)	**Variables**	**E**	**L**	
	L	0.552		
	logL^{SQ} obs.	0.567	0.999	
2)	**Variables**	**E**	**L**	**logL^{SQ} obs**
	MEAN	0.266	3.318	3.206
	SD	0.302	1.120	1.126
3)	**Variables**	***e***	***l***	***c***
	COEFFS	0.080	0.992	-0.108
	ST.DEV	0.010	0.003	0.008
	T	8.000	365.47	13.137
	TTEST	1.000	1.000	1.000
4)	**r^2**	0.999		
	SD	0.039		
	DOF	239		
	F	98441.46		

6.1.2. Construction of an Equation for log P^{SQ}

In the previous section, attention was drawn to the determination of log L^{SQ} values from literature data and on the development of a solvation equation. Next, these log L^{SQ} values were combined with corresponding gas / water partition coefficient values, log L^{W}, through equation (6.19) to obtain log P^{SQ} values for the transfer of solutes from water to squalane.

$$\log P^{S} = \log L^{S} - \log L^{w} \qquad (6.19)$$

Log L^W values were available for 153 solutes. Most of these values have been published elsewhere.[35] Some log L^W values were calculated from the solubility of compound in water[36] and in the gas phase as explained in equation (6.20).

$$L^W = \frac{C_W}{C_G} \quad \text{or} \quad \log L^W = \log C_W - \log C_G \qquad (6.20)$$

Values of vapour pressure were taken from references 25-30. Log L^W and log P^{SQ} values are listed in Table 6.10, see section 6.3. After a preliminary analysis, six compounds were pronounced outliers (2,2,4-trimethylpentane, n-nonane, 4-methyloctane, 3,3-diethylpentane, sec-butylbenzene, and sec-pentylbenzene). P^{SQ} values for the remaining 137 compounds were regressed against the Abraham solvation descriptors for water to solvent transfer process. The output from the MLR analysis of log P^{SQ} with solute descriptors for the water / squalane system is summarised in Table 3.7 in chapter 3.

$$\log P^{SQ} = 0.119\ (\pm 0.043) + 0.810\ (\pm 0.064)\ \mathbf{E} - 1.702\ (\pm 0.091)\ \mathbf{S} - 3.626\ (\pm 0.094)\ \mathbf{A} - 4.810\ (\pm 0.106)\ \mathbf{B} + 4.239\ (\pm 0.046)\mathbf{V} \qquad (6.21)$$

$n = 137,\ r^2 = 0.996,\ sd = 0.123,\ F = 7692.2$

The statistical analysis of equation (6.21) is fairly good, in view of the fact that the error in the experimental P^{SQ} values will include errors in both L^{SQ} and L^W values. As a result, there are fewer outliers in the above equation than in equation (6.20) since the larger sd value in equation (6.21) 'masks' the majority of the outliers in this analysis. A plot of observed versus calculated log P^{SQ} values is shown in Figure 6.5.

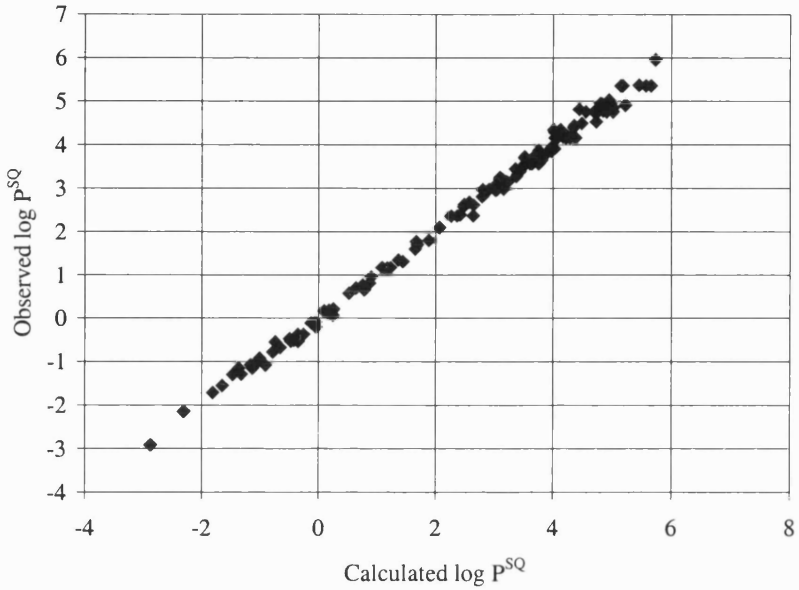


Figure 6.5. Plot of observed against calculated log P^{SQ} on equation (6.21)

6.1.3. Solubility of Gases and Vapours in Apolane at 313K

Apolane, or Apolane-87, is a highly branched nonpolar hydrocarbon developed by Kovats et al.[7] Apolane is stable, nonvolatile liquid that can be used over the temperature range of 303-553 K. Recently, Weckwerth and co-workers measured gas / apolane coefficients, L^{87}, of 162 nonpolar and polar organic solutes.[3] L^{87} values were measured by open tubular capillary gas chromatography at 313K. The Abraham solvation descriptors were available for 157 compounds, see Table 6.10 in section 6.4. A first regression analysis, highlighted nine outliers listed as Nos 149-157 in Table 6.10. The log L^{87} values for the remaining 148 compounds were regressed against the corresponding **E** and **L** descriptor values. The output for equation (6.22) is given in Table 6.6.

$$\log L^{87} = -0.152\ (\pm 0.011) + 0.192\ (\pm 0.013)\ \mathbf{E} + 0.879\ (\pm 0.003)\ \mathbf{L} \qquad (6.22)$$

n= 148, r^2 = 0.999, sd = 0.044, F = 46138

The above equation is statistically better than the one proposed by Weckwerth an co-workers[3]. These authors developed a straight correlation between **L** and log L^{87}, see equation (6.2). Here, it can be seen that the insertion of the **E** descriptor improves the correlation.

Table 6.6. Output of the regression analysis of log L^{87} and solute descriptors

1)	**Variables**	**E**	**L**	
	L	0.405		
	logL^{87} obs.	0.449	0.998	
2)	**Variables**	**E**	**L**	**logL^{87} obs**
	MEAN	0.329	3.470	2.965
	SD	0.311	1.224	1.103
3)	**Variables**	***e***	***l***	***c***
	COEFFS	0.192	0.879	-0.152
	ST.DEV	0.013	0.003	0.011
	T	15.020	271.300	13.721
	TTEST	1.000	1.000	1.000
4)	**r^2**	0.999		
	SD	0.044		
	DOF	145		
	F	46138		

6.1.4. Results and Discussion

A large data set of 396 logarithmic values of gas / squalane partition coefficient at 298K, $\log L^{SQ}$, has been obtained from the literature. 240 $\log L^{SQ}$ values have been regressed against the Abraham descriptors for gas to solvent partition processes. The *a*-, *b*- and *s*-coefficients are zero, as desired in a non-polar and non-hydrogen bond donor or acceptor stationary phase.

$$\log L^{SQ} = -0.108\ (0.008) + 0.080\ (0.010)\ \mathbf{E} + 0.992\ (0.003)\ \mathbf{L} \qquad (6.18)$$

$n = 240$, $r^2 = 0.999$, $sd = 0.038$, $F = 98441.48$

The model is statistically sound and chemically sensible. Similar equations were set up by Abraham and co-workers for a variety of alkanes, see Table 6.7, for the regression coefficients.[35,37,38] Comparison of the coefficients in Equation (6.18) and those listed in Table 6.7 shows that there is little difference in solubility properties amongst alkane solvents. The coefficients are in general similar to those for n-hexadecane solvent, for which the *e*-, *s*-, *a*-, and *b*-coefficients were fixed to zero and the l-coefficient to one by definition.[1]

More importantly, solvation equations for gas / squalane partitions at various temperature are available in the literature, see Table 6.8.[39-41] Abraham et al. developed an equation for gas / squalane partition at 393 K.[39] More recently, Poole and co-workers analysed the effect of temperature on the gas / squalane partition process.[39] The authors measured L^{SQ} values on a squalane packed column over the temperature range 334-394K and made the use of these experimental data to develop four solvation equations. It is important to note that the authors observed that the interfacial adsorption at low temperature contributes to the retention of hydrocarbons. They found actually that the value of the *b*-coefficient at 334K accounts for specific solute adsorption at the liquid / solid interface. This specific interaction weakens and becomes almost non-existent at temperature higher than 354K.[40] Similarly, these effects are slightly present in the solvation equation proposed by Abraham, Ballantine and Callihan[41] at 393K. In this example, the authors made the use of MacReynold's specific retention volume values, $\log V_G$, on squalane at 393K, see equation (6.23)

Table 6.7. Water / alkane and gas / alkane partition processes at 298 K

solvent	Ref	*c*	*e*	*s*	*a*	*b*	*v / l*
Water / alkane partition processes							
n-Pentane	a	0.369	0.386	-1.568	-3.535	-5.215	4.514
n-Hexane	a	0.361	0.579	-1.723	-3.599	-4.764	4.344
Cyclohexane	b	0.159	0.784	-1.678	-3.740	-4.929	4.577
n-Heptane	a	0.325	0.670	-2.061	-3.317	-4.733	4.543
n-Octane	a	0.223	0.642	-1.647	-3.480	-5.067	4.526
iso-Octane	c	0.288	0.382	-1.668	-3.639	-5.000	4.561
n-Nonane	a	0.240	0.619	-1.713	-3.532	-4.921	4.482
n-Decane	a	0.160	0.585	-1.734	-3.435	-5.078	4.582
n-Hexadecane	b	0.087	0.667	-1.617	-3.587	-4.869	4.433
Squalane	d	0.119	0.810	-1.702	-3.626	-4.810	4.239
Gas / alkane partition processes							
n-Pentane	a	0.335	-0.276	0.000	0.000	0.000	0.968
n-Hexane	a	0.292	-0.169	0.000	0.000	0.000	0.979
Cyclohexane	b,f	0.163	-0.110	0.000	0.000	0.000	1.013
n-Heptane	a	0.275	-0.162	0.000	0.000	0.000	0.983
n-Octane	a	0.215	-0.049	0.000	0.000	0.000	0.967
iso-Octane	b,f	0.275	-0.244	0.000	0.000	0.000	0.972
n-Nonane	a	0.200	-0.145	0.000	0.000	0.000	0.980
n-Decane	a	0.156	-0.146	0.000	0.000	0.000	0.989
n-Hexadecane	b	0.000	0.000	0.000	0.000	0.000	1.000
Squalane	d	-0.108	0.080	0.000	0.000	0.000	0.992
Apolane	d,e	-0.152	0.191	0.000	0.000	0.000	0.879

[a] Ref. 37.
[b] Ref. 35.
[c] Ref. 38.
[d] The present work
[e] water / apolane partition at 313 K.

$$\log V_G = -0.29 + 0.080\,\mathbf{E} + 0.080\,\mathbf{S} + 0.110\,\mathbf{A} + 0.674\,\mathbf{L} \qquad (6.23)$$

Further, from equation (6.23) one can write,

$$\log L^{SQ} = (\log \rho_{SQ} - 0.29) + 0.080\,\mathbf{E} + 0.080\,\mathbf{S} + 0.110\,\mathbf{A} + 0.674\,\mathbf{L} \qquad (6.24)$$

Therefore *e*-, *s*- and *a*-coefficients in equation (6.23) can be directly compared to any equation similar to equation (6.24). On the other hand, the *c*-coefficient cannot be compared because it includes the constant log ρ_{SQ}. In figures 6.6 and 6.7 are given a plot of *e*- and *l*-coefficient values versus 1/T in K-1, respectively. First, a scattered plot is obtained for the *e*-coefficient. This result does not agree with the one given by Poole and co-workers, who observed a linear dependence between the *e*-coefficient and the column temperature. However, it has to be kept in mind that Poole and co-workers used their own measured log L^{SQ} values to set up their equations[40] and not literature data as was done in this work or in Abraham and co-workers' works.[39,41] Furthermore, the E descriptor is not the main factor that influences the gas / squalane partition process as can be seen from the small *e*-coefficient. On the other hand, a linear trend exists between the *l*-coefficient values and the reciprocal values of the column temperature, 1/T. The *l*- coefficient at 298K agrees very well with the *l*-coefficient obtained at other temperatures. It was also interesting to include the *e*- and *l*-coefficients for gas / apolane at 313K into figures (6.6) and (6.7). The *e*-coefficient is much higher than any of the *e*-coefficient values for gas / squalane systems. However, the *l*-coefficient fits quite well.

In Table 6.8, are listed the regression coefficients for water / alkane partitions. These coefficients refer to differences between the aqueous phase and the alkane solvents.[1] Regression coefficients for water / squalane partitions agree well with the analyses of regression coefficients for the various water / alkanes previously reported. It was found that water has a hydrogen-bonding acidity, described by the *b*-coefficient, of around 4.95 units and an hydrogen–bonding basicity, characterized by the a-coefficient, of 3.54 units. Further, the dipolarity/polarisability of water can be taken as 1.71units. These three values are, of course, based on the premise that **S**, **A**, and **B** are all zero for the alkane solvents. The *e*-coefficients are positive as the alkanes are more prone to interact through dispersion interactions than is water. Similarly, the *v*-coefficient

represents an 'alkane-like scale' and it would not be expected to have *v*-coefficients larger than 4.7 units.[1] The regression coefficients describing the water / squalane system are in line of those for a variety of water / alkane systems. The only difference lies with the e-coefficient, that is larger in the case where solute distribute between water and squalane. This result is correct since squalane is a highly branched hydrocarbon and it is expected to give rise to larger dispersion interactions.

Table 6.8. System constants for squalane and Apolane

Solvent	T (K)	c	e	s	a	b	l	Ref
Squalane	298	-0.108	0.080	0.000	0.000	0.000	0.992	a
Squalane	334	-0.213	0.041	0.000	0.000	0.061	0.859	b
Squalane	354	-0.171	0.076	0.000	0.000	0.000	0.715	b
Squalane	374	-0.194	0.105	0.000	0.000	0.000	0.635	b
Squalane	394	-0.221	0.138	0.000	0.000	0.000	0.584	b
Squalane	393	-0.290	0.080	0.080	0.110	0.000	0.674	c
Squalane	393	-0.222	0.129	0.000	0.000	0.000	0.583	d
Apolane	313	-0.152	0.192	0.000	0.000	0.000	0.879	e

[a] This work, b Ref. 39, c Ref. 41, d Ref. 39
[e] This work, system constant for Apolane

In the previous section, was presented the elaboration of solvation equation for gas / squalane and gas / apolane systems. Now it is shown how these equations can be used to determine **L** value for a large number of compounds.

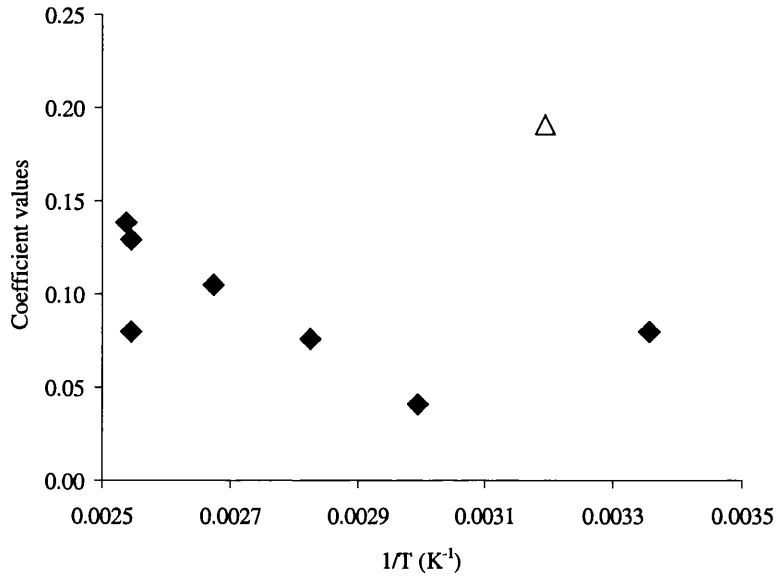


Figure 6.6. Plot of *e*-coefficient values against the temperature, T (K). Closed symbol is for coefficient values for gas / squalane process, open symbol is for coefficient value for gas / apolane process.

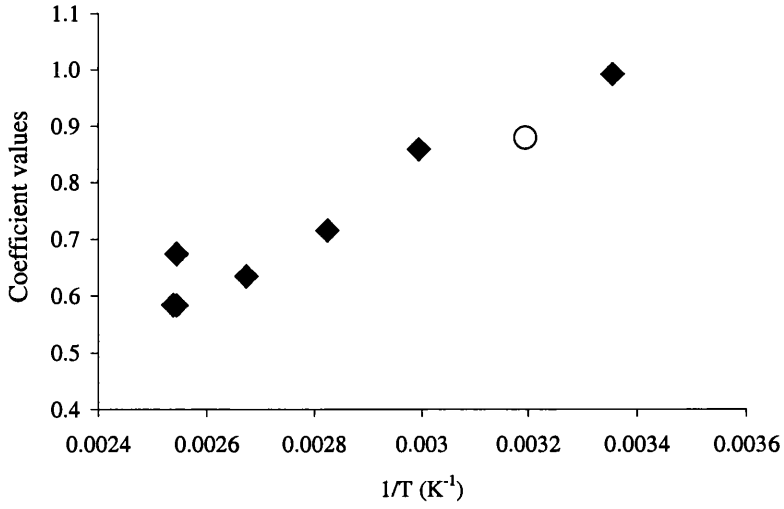


Figure 6.7. Plot of *l*-coefficient values against the temperature, T (K). Closed symbol is for coefficient values for gas / squalane process, open symbol is for coefficient value for gas / apolane process.

6.2. Determination of New Values of **L** from Gas / Alkane Partition Coefficients

The following models are suitable for the determination of L values.

$$\log L^{SQ} = -0.108\ (0.008) + 0.080\ (0.010)\ \mathbf{E} + 0.992\ (0.003)\ \mathbf{L} \qquad (6.18)$$

n= 240, r^2 = 0.999, sd = 0.038, F = 98441.48

and

$$\log L^{87} = -0.152\ (\pm 0.011) + 0.192\ (\pm 0.013)\ \mathbf{E} + 0.879\ (\pm 0.003)\ \mathbf{L} \qquad (6.22)$$

n= 148, r^2 = 0.999, sd = 0.044, F = 46138

The **E** descriptor value was calculated directly from the refractive index of liquids at 293K.[1] Refractive index values were mostly retrieved from chemical supplier catalogues. If the refractive index value was not available, it was estimated from a chemistry drawing package, (ACD/Chemsketch, version 4.55, 05.2000, Ontario). This package contains a built function that can calculate the refractive index value, from which a value of **E** can be estimated. The uncertainty in the determination of **L** values from equation (6.18) and (6.22) is given by (sd/l) and is 0.038 log units for equations (6.18) and 0.050 log units for equations (6.22). This is of the same order as the general error in the expected **L** descriptor [42], viz. 0.02 log units. Calculated **L** values are given in Table 6.11 in section 6.4.

6.3. Conclusion

Literature data was used to obtain gas / squalane and water / squalane partition coefficient values at 298K for a large number of volatile organic compounds. These data were then used to establish solvation equations on the line of equation (6.1) and (6.2). It was shown that the two equations obtained agree with equations previously obtained for other gas / alkane and water / alkane systems. Similarly, an equation for gas / apolane partition process at 313K was developed.

Further values of **L** were calculated from the equations for gas / squalane and gas / apolane systems. Some 141 new values of **L** were obtained from the equation for gas / squalane partition process. Only 5 new values of **L** were calculated from the equation for gas / apolane partition process. A total of 146 new descriptor values for a variety of volatile organic compounds were determined in this work.

6.4. Appendix

Table 6.9. Observed retention index and partition coefficient values together with calculated partition coefficient values for gas / squalane and water / squalane systems at 298K.

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc log$L^{SQ\,a}$	residual [b]	logL^{w}	Ref.	obs log$P^{SQ\,c}$	calc log $P^{SQ\,d}$	residual [d]
1	Hydrogen			-1.195	18	-1.191	-0.004	-1.72	35	0.525	0.579	-0.054
2	Carbon dioxide			0.016	17	0.058	-0.042	-0.08	35	0.096	0.171	-0.075
3	Methane			-0.372	18	-0.321	-0.051	-1.46	35	1.088	1.176	-0.088
4	Ethane			0.332	18	0.489	-0.157	-1.34	35	1.672	1.774	-0.102
5	Propane			0.934	18	1.043	-0.109	-1.44	35	2.374	2.371	0.003
6	Butane			1.519	16	1.604	-0.084	-1.52	35	3.039	2.968	0.071
7	2-Methylpropane	36.53	23	1.317	23	1.399	-0.082	-1.7	35	3.017	2.968	0.049
8	Pentane			2.053	10,11,16	2.147	-0.093	-1.7	35	3.753	3.565	0.188
9	2-Methylbutane	47.42	24	1.893	10,16	1.999	-0.105	-1.75	35	3.643	3.565	0.078
10	Hexane			2.563	9,10,11,16	2.649	-0.086	-1.82	35	4.383	4.163	0.22
11	2-Methylpentane	56.94	21	2.378	16,21	2.485	-0.107	-1.84	35	4.218	4.163	0.056
12	3-Methylpentane	58.3	21	2.447	16,21	2.563	-0.116	-1.84	35	4.287	4.163	0.124
13	2,2-Dimethylbutane	53.47	21	2.194	16,21	2.335	-0.141	-1.84	35	4.034	4.163	-0.128
14	2,3-Dimethylbutane	56.5	21	2.355	16,21	2.477	-0.123	-1.72	35	4.075	4.163	-0.088
15	Heptane			3.056	9,11,16	3.15	-0.095	-1.96	35	5.016	4.76	0.256
16	2-Methylhexane	66.62	21	2.869	16,21	2.98	-0.111	-2.15	35	5.019	4.76	0.259
17	3-Methylhexane	67.52	21	2.922	18,21	3.022	-0.101	-1.99	35	4.912	4.76	0.152
18	3-Ethylpentane	68.47	21	2.96	16,21	3.069	-0.109					
19	2,2-Dimethylpentane	62.44	21	2.658	16,21	2.776	-0.119	-1.904	35	4.562	4.76	-0.198
20	2,3-Dimethylpentane	66.99	21	2.892	18,21	2.995	-0.103	-1.85	35	4.742	4.76	-0.018
21	2,4-Dimethylpentane	62.9	21	2.681	16,21	2.789	-0.108	-2.08	35	4.761	4.76	0.001
22	3,3-Dimethylpentane	65.56	21	2.819	18,21	2.925	-0.106	-1.88	35	4.699	4.76	-0.061
23	2,2,3-Trimethylbutane	63.62	21	2.721	18,21	2.897	-0.176					
24	Octane			3.548	16	3.651	-0.102	-2.11	35	5.658	5.357	0.301

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{W}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
25	2-Methylheptane	76.46	21	3.354	21	3.455	-0.102					
26	3-Methylheptane	77.15	21	3.389	21	3.485	-0.096	-2.18	35	5.569	5.357	0.212
27	4-Methylheptane	76.65	21	3.363	21	3.458	-0.095					
28	2,2-Dimethylhexane	71.83	21	3.118	21	3.238	-0.120					
29	2,3-Dimethylhexane	75.85	21	3.323	21	3.426	-0.104					
30	2,4-Dimethylhexane	73.08	21	3.181	21	3.295	-0.114					
31	2,5-Dimethylhexane	72.76	21	3.165	21	3.285	-0.120	-2.02	35	5.185	5.357	-0.173
32	3,3-Dimethylhexane	74.00	21	3.228	21	3.335	-0.107					
33	3,4-Dimethylhexane	76.80	21	3.371	21	3.534	-0.163					
34	3-Ethylhexane	77.13	21	3.388	21	3.494	-0.106					
35	2-Methyl-3-ethylpentane	75.79	21	3.320	21	3.434	-0.115					
36	3-Methyl-3-ethylpentane	76.92	21	3.377	21	3.477	-0.100					
37	2,2,3-Trimethylpentane	73.32	21	3.193	21	3.301	-0.108					
38	2,2,4-Trimethylpentane	68.73	21	2.980	18,21	3.084	-0.103	-1.554	35	4.534	5.357	-0.823
39	2,3,3-Trimethylpentane	75.39	21	3.299	21	3.404	-0.105					
40	2,3,4-Trimethylpentane	74.86	21	3.272	21	3.456	-0.184	-1.88	35	5.152	5.357	-0.205
41	2,2,3,3-Tetramethylbutane	71.97	21	3.125	21	3.242	-0.117					
42	Nonane			4.042	16	4.152	-0.111	-2.31	35	6.352	5.955	0.397
43	2-Methyloctane	86.43	21	3.862	21	3.938	-0.076					
44	3-Methyloctane	86.97	21	3.890	21	3.970	-0.080					
45	4-Methyloctane	86.23	21	3.852	21	3.933	-0.081	-2.611	35	6.463	5.955	0.509
46	3-Ethylheptane	86.60	21	3.871	21	3.964	-0.093					
47	4-Ethylheptane	85.64	21	3.822	21	3.916	-0.094					
48	2,2-Dimethylheptane	81.47	21	3.609	21	3.712	-0.103					
49	2,3-Dimethylheptane	85.36	21	3.808	21	3.897	-0.090					
50	2,4-Dimethylheptane	82.10	21	3.641	21	3.731	-0.090					
51	2,6-Dimethylheptane	82.66	21	3.670	21	3.753	-0.083					

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc $logL^{SQ}$ [a]	residual [b]	$logL^{w}$	Ref.	obs $logP^{SQ}$ [c]	calc log P^{SQ} [d]	residual [d]
52	3,4-Dimethylheptane	85.57	21	3.818	21	3.907	-0.089					
53	3,5-Dimethylheptane	83.25	21	3.700	21	3.799	-0.099					
54	4,4-Dimethylheptane	82.48	21	3.661	21	3.743	-0.083					
55	2-Methyl-3-ethylhexane	84.11	21	3.744	21	3.823	-0.079					
56	2-Methyl-4-ethylhexane	82.34	21	3.654	21	3.733	-0.08					
57	3-Methyl-3-ethylhexane	84.94	21	3.786	21	3.862	-0.076					
58	3-Methyl-4-ethylhexane	85.11	21	3.795	21	3.872	-0.077					
59	2,2,3-Trimethylhexane	81.78	21	3.625	21	3.735	-0.111					
60	2,2,4-Trimethylhexane	80.41	21	3.555	21	3.579	-0.024					
61	2,2,5-Trimethylhexane	77.48	21	3.406	21	3.542	-0.136	-2.328	35	5.734	5.955	-0.221
62	2,3,3-Trimethylhexane	83.52	21	3.714	21	3.805	-0.091					
63	2,3,4-Trimethylhexane	84.26	21	3.751	21	3.854	-0.103					
64	2,3,5-Trimethylhexane	81.02	21	3.586	21	3.698	-0.111					
65	3,3,4-Trimethylhexane	84.78	21	3.778	21	3.863	-0.086					
66	3-Ethyl-2,2-dimethylpentane	81.71	21	3.621	21	3.713	-0.092					
67	3-Ethyl-2,3-dimethylpentane	86.57	21	3.869	21	3.942	-0.073					
68	3-Ethyl-2,4-dimethylpentane	83.21	21	3.698	21	3.801	-0.103					
69	2,2,3,3-Tetramethylpentane	84.60	21	3.769	21	3.852	-0.084					
70	2,2,3,4-Tetramethylpentane	81.42	21	3.606	21	3.711	-0.105					
71	2,2,4,4-Tetramethylpentane	76.97	21	3.38	21	3.487	-0.107					
72	2,3,3,4-Tetramethylpentane	85.21	21	3.8	21	3.882	-0.082					
73	3,3-Diethylpentane	87.09	21	3.896	21	3.985	-0.089	-1.63	35	5.526	5.955	-0.429
74	Cyclopentane	56.24	21	2.346	9,16,21	2.457	-0.111	-0.88	35	3.226	3.148	0.078
75	1,1-Dimethylcyclopentane	66.93	21	2.867	21	3.006	-0.138					
76	1,2-cis-Dimethylcyclopentane	71.66	21	3.109	21	3.248	-0.139					
77	1,2-trans-Dimethylcyclopentane	68.54	21	2.95	21	3.075	-0.126					
78	1,3-cis-Dimethylcyclopentane	67.92	21	2.918	21	3.042	-0.124					

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{w}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
79	1,3-trans-Dimethylcyclopentane	68.22	21	2.933	21	3.052	-0.118					
80	Ethylcyclopentane	72.93	21	3.174	21	3.298	-0.125					
81	1-Methyl-1-ethylcyclopentane	78.84	21	3.475	21	3.585	-0.109					
82	Propylcyclopentane	82.56	21	3.665	21	3.774	-0.109	-1.56	35	5.225	4.909	0.316
83	Cyclohexane	65.83	21	2.828	9,21	2.940	-0.112	-0.90	35	3.728	3.779	-0.051
84	Methylcyclohexane	71.99	21	3.140	9,21	3.293	-0.154	-1.21	35	4.350	4.395	-0.046
85	1,1-Dimethylcyclohexane	77.98	21	3.431	21	3.554	-0.123					
86	1,2-cis-Dimethylcyclohexane	82.19	21	3.646	21	3.766	-0.120	-1.16	35	4.806	4.954	-0.148
87	1,2-trans-Dimethylcyclohexane	79.59	21	3.513	21	3.606	-0.093					
88	1,3-cis-Dimethylcyclohexane	77.95	21	3.430	21	3.506	-0.076					
89	1,3-trans-Dimethylcyclohexane	79.98	21	3.533	21	3.627	-0.094					
90	1,4-cis-Dimethylcyclohexane	79.85	21	3.527	21	3.633	-0.106					
91	1,4-trans-Dimethylcyclohexane	77.96	21	3.430	21	3.511	-0.081	-1.55	35	4.980	4.933	0.048
92	Ethylcyclohexane	82.75	21	3.685	16,21	3.847	-0.162	-1.148	35	4.833	4.940	-0.107
93	Ethene	17.72	23	0.292	12,23	0.286	0.006	-0.94	35	1.232	1.171	0.061
94	Propene	28.66	23	0.916	23	0.938	-0.022	-0.97	35	1.886	1.799	0.087
95	But-1-ene	38.48	23	1.417	23	1.517	-0.100	-1.01	35	2.427	2.394	0.033
96	cis-But-2-ene	41.65	24	1.578	25	1.724	-0.145					
97	trans-But-2-ene	40.65	24	1.527	24	1.651	-0.124					
98	2-Methylprop-1-ene	38.33	23	1.409	23	1.566	-0.157	-0.86	35	2.269	2.362	-0.093
99	Pent-1-ene	48.02	24	1.928	16,24	2.032	-0.103	-1.23	35	3.158	2.986	0.173
100	cis-Pent-2-ene	50.47	21	2.028	21	2.200	-0.172	-0.96	35	2.988	3.025	-0.037
101	trans-Pent-2-ene	50.07	21	2.008	21	2.170	-0.163	-0.99	35	2.998	3.012	-0.015
102	3-Methylbut-1-ene	44.86	24	1.742	24	1.918	-0.176	-1.34	35	3.082	3.092	-0.010
103	2-Methylbut-1-ene	48.74	24	1.940	24	2.081	-0.14					
104	2-Methylbut-2-ene	51.40	21	2.075	21	2.212	-0.136	-0.96	35	3.035	2.974	0.061
105	Hex-1-ene	58.16	21	2.448	16,21	2.553	-0.105	-1.16	35	3.608	3.571	0.037

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{w}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
106	cis-Hex-2-ene	60.31	21	2.530	21	2.664	-0.134					
107	trans-Hex-2-ene	59.73	21	2.487	18,21	2.635	-0.148					
108	cis-Hex-3-ene	59.22	21	2.470	18,21	2.644	-0.174					
109	trans-Hex-3-ene	59.29	21	2.478	21	2.639	-0.162					
110	2-Methylpent-1-ene	57.96	21	2.440	16,21	2.572	-0.131	-1.08	35	3.520	3.532	-0.012
111	3-Methylpent-1-ene	54.94	21	2.256	21	2.407	-0.151					
112	4-Methylpent-1-ene	54.80	21	2.249	21	2.401	-0.152	-1.41	36	3.659	3.709	-0.050
113	2-Methylpent-2-ene	59.80	21	2.530	18,21	2.667	-0.137					
114	cis-3-Methylpent-2-ene	60.21	21	2.525	21	2.685	-0.161					
115	cis-4-Methylpent-2-ene	55.53	21	2.286	21	2.426	-0.139					
116	trans-4-Methylpent-2-ene	56.20	21	2.320	21	2.447	-0.126					
117	2-Ethylbut-1-ene	59.20	21	2.474	21	2.604	-0.13					
118	2,3-Dimethylbut-1-ene	55.74	21	2.297	21	2.437	-0.139					
119	3,3-Dimethylbut-1-ene	50.52	21	2.030	21	2.185	-0.154					
120	Hept-1-ene	68.13	21	2.951	21,16	3.040	-0.089	-1.22	35	4.171	4.179	-0.008
121	cis-Hept-2-ene	70.16	21	3.032	21	3.186	-0.154					
122	trans-Hept-2-ene	69.86	21	3.017	21	3.156	-0.139	-1.23	35	4.247	4.201	0.046
123	cis-Hept-3-ene	69.01	21	2.974	21	3.119	-0.145					
124	trans-Hept-3-ene	68.80	21	2.963	21	3.102	-0.139					
125	Oct-1-ene	78.02	21	3.446	16,21	3.542	-0.096	-1.41	35	4.856	4.778	0.078
126	cis-Oct-2-ene	80.05	21	3.537	21	3.656	-0.119					
127	2,4,4-Trimethylpent-2-ene	71.39	21	3.095	21	3.225	-0.13	-1.35	35	4.445	4.816	-0.371
128	Non-1-ene	88.12	21	3.948	21	4.043	-0.095	-1.51	35	5.458	5.372	0.086
129	2-Methylbuta-1,3-diene	49.63	23,24	1.986	23,24	2.083	-0.097	-0.5	35	2.486	2.582	-0.096
130	cis-Penta-1,3-diene	52.31	24	2.122	24	2.261	-0.139					
131	trans-Penta-1,3-diene	51.44	24	2.078	24	2.231	-0.153					
132	Penta-1,4-diene	46.04	24	1.802	24	1.982	-0.18	-0.69	35	2.492	2.632	-0.139

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{w}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
133	Hexa-1,5-diene	56.22	24	2.321	24	2.431	-0.110	-0.77	35	3.091	3.217	-0.125
134	Cyclopentene	54.58	23,24	2.239	23,24	2.382	-0.143	-0.41	35	2.649	2.369	0.280
135	1-Methylcyclopentene	64.20	23	2.728	23	2.841	-0.112					
136	Cyclohexene	66.61	23	2.868	16,23	2.996	-0.128	-0.27	35	3.138	3.019	0.120
137	1-Methylcyclohexene	75.67	23	3.313	23	3.455	-0.141	-0.49	35	3.803	3.647	0.157
138	3-Methylcyclohexene	72.50	23	3.152	23	3.292	-0.141					
139	4-Methylcyclohexene	72.75	23	3.164	23	3.300	-0.136					
140	Cyclopenta-1,3-diene	51.79	24	2.096	24	2.203	-0.107					
141	Hex-1-yne	58.55	23	2.44	23	2.491	-0.050	-0.21	35	2.650	2.618	0.032
142	Hept-1-yne	68.67	23	2.952	10,23	2.977	-0.025	-0.44	35	3.392	3.326	0.066
143	Oct-1-yne	78.25	23	3.445	23	3.495	-0.050	-0.52	35	3.965	3.937	0.028
144	Oct-2-yne	84.48	23	3.763	23	3.821	-0.058					
145	Non-1-yne	88.23	23	3.954	23	3.989	-0.035	-0.78	36	4.734	4.530	0.204
146	Dec-1-yne	98.22	23	4.463	23	4.504	-0.040					
147	Dodec-1-yne	118.2	23	5.483	23	5.616	-0.133					
148	Dichloromethane			1.863	16	2.001	-0.139	0.96	35	0.903	0.954	-0.052
149	Trichloromethane			2.444	16	2.459	-0.015	0.79	35	1.654	1.603	0.050
150	Tetrachloromethane			2.75	9,16	2.799	-0.049	-0.06	35	2.810	2.976	-0.167
151	3-Chloroprop-1-yne			1.991	10	2.062	-0.071					
152	Diethylether			1.958	11	2.000	-0.043	1.17	35	0.788	0.660	0.127
153	Tetrahydrofuran			2.505	11	2.615	-0.109	2.55	35	-0.045	-0.203	0.158
154	1,4-Dioxane			2.800	11	2.869	-0.068	3.71	35	-0.910	-1.083	0.173
155	Propanone			1.666	9,11	1.628	-0.700	2.79	35	-1.124	-1.113	-0.011
156	Butanone			2.221	11	2.269	-0.049	2.72	35	-0.499	-0.475	-0.024
157	Pentan-2-one			2.734	15	2.734	0.000	2.58	35	0.154	0.137	0.016
158	Pentan-3-one			2.743	15	2.790	-0.047	2.50	35	0.243	0.180	0.063
159	3-Methylbutan-2-one			2.563	11	2.672	-0.109	2.38	35	0.183	0.182	0.001

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{w}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
160	Heptan-2-one			3.674	15	3.732	-0.058	2.23	35	1.444	1.316	0.129
161	Butan-2,3-dione			2.206	11	2.232	-0.026					
162	Methyl acetate			1.886	9,11	1.896	-0.010	2.30	35	-0.414	-0.452	0.038
163	Butyl acetate			3.307	15	3.244	-1.100	1.94	35	1.367	1.349	0.018
164	Proprionitrile			2.036	11	2.066	-0.030	2.82	35	-0.784	-0.776	-0.008
165	Nitromethane			1.820	8	1.876	-0.056	2.95	35	-1.130	-1.157	0.027
166	Methanol			0.866	9	0.961	-0.095	3.74	35	-2.874	-2.918	0.044
167	Ethanol			1.365	9	1.472	-0.107	3.67	35	-2.305	-2.143	-0.162
168	Propan-1-ol			1.911	9	2.015	-0.103	3.56	35	-1.649	-1.554	-0.094
169	Propan-2-ol			1.670	9	1.750	-0.080	3.48	35	-1.810	-1.711	-0.099
170	Butan-1-ol			2.445	9	2.581	-0.136	3.46	35	-1.015	-0.967	-0.048
171	2-Methylpropan-1-ol			2.291	9	2.394	-0.103	3.3	35	-1.009	-0.921	-0.088
172	Butan-2-ol			2.256	9	2.319	-0.063	3.39	35	-1.134	-1.110	-0.024
173	2-Methylpropan-2-ol			1.915	9	1.947	-0.032	3.28	35	-1.365	-1.158	-0.207
174	Pentan-1-ol			3.000	9	3.082	-0.082	3.35	35	-0.350	-0.374	0.024
175	Pentan-2-ol			2.744	9	2.818	-0.074	3.22	35	-0.476	-0.531	0.055
176	Pentan-3-ol			2.768	9	2.838	-0.07	3.19	35	-0.422	-0.512	0.090
177	2-Methylbutan-2-ol			2.508	9	2.610	-0.102	3.25	35	-0.742	-0.549	-0.193
178	Hexan-1-ol			3.487	13	3.583	-0.096	3.23	35	0.257	0.216	0.041
179	Hexan-3-ol			3.23	13	3.318	-0.088	2.98	35	0.250	0.071	0.179
180	Heptan-1-ol			3.965	13	4.084	-0.119	3.09	35	0.875	0.815	0.060
181	Heptan-2-ol			3.71	14	3.803	-0.093					
182	Benzene	62.99	21	2.694	9,11,16,21	2.699	-0.005	0.63	35	2.064	2.091	-0.027
183	Toluene	73.91	21	3.231	9,16,21	3.296	-0.065	0.65	35	2.581	2.681	-0.100
184	Ethylbenzene	82.82	21	3.679	16,21	3.746	-0.067	0.58	35	3.099	3.257	-0.158
185	o-Xylene	86.27	21	3.854	16,21	3.906	-0.051	0.66	35	3.194	3.165	0.029
186	m-Xylene	84.57	21, 24	3.759	16,21,24	3.806	-0.047	0.61	35	3.149	3.200	-0.051

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc log$L^{SQ\,a}$	residual [b]	logL^{w}	Ref.	obs log$P^{SQ\,c}$	calc log $P^{SQ\,d}$	residual [d]
187	p-Xylene	84.25	21	3.739	21	3.806	-0.067	0.59	35	3.149	3.864	-0.715
188	Propylbenzene	91.78	24	4.146	20,24	4.195	-0.049	0.39	35	3.756	3.864	-0.108
189	Isopropylbenzene	88.95	21	4.007	20,24	4.050	-0.042	0.22	35	3.787	3.832	-0.045
190	1,2,3-Trimethylbenzene	98.67	24	4.487	24	4.527	-0.041	0.89	35	3.597	3.585	0.012
191	1,2,4-Trimethylbenzene	96.59	24	4.380	24	4.403	-0.023	0.63	35	3.750	3.629	0.121
192	1,3,5-Trimethylbenzene	95.31	24	4.317	20,24	4.307	0.01	0.66	35	3.657	3.674	-0.017
193	2-Ethyltoluene	94.62	24	4.280	24	4.309	-0.03	0.76	35	3.52	3.696	-0.176
194	3-Ethyltoluene	93.02	24	4.198	24	4.239	-0.041					
195	4-Ethyltoluene	93.44	24	4.219	24	4.253	-0.034	0.70	35	3.519	3.724	-0.205
196	Butylbenzene	101.85	24	4.649	24	4.691	-0.042	0.29	35	4.359	4.441	-0.082
197	Isobutylbenzene	96.37	24	4.369	24	4.463	-0.094	-0.12	35	4.489	4.493	-0.004
198	sec-Butylbenzene	96.74	24	4.388	24	4.469	-0.081	0.33	35	4.058	4.447	-0.389
199	tert-Butylbenzene	95.65	24	4.333	24	4.376	-0.044	0.32	35	4.013	4.346	-0.333
200	1,2-Diethylbenzene	101.82	24	4.647	24	4.693	-0.046					
201	1,3-Diethylbenzene	100.94	23,24	4.602	23,24	4.647	-0.045					
202	1,4-Diethylbenzene	101.48	24	4.645	24	4.693	-0.048	0.52	35	4.126	4.351	-0.225
203	1,2,4,5-Tetramethylbenzene	108.02	24	4.963	24	4.988	-0.025					
204	1,2,3,5-Tetramethylbenzene	108.64	24	4.995	24	5.010	-0.015					
205	1,2,3,4-Tetramethylbenzene	109.7	24	5.049	24	5.133	-0.084					
206	2-Propyltoluene	102.33	23,24	4.671	23,24	4.725	-0.055					
207	3-Propyltoluene	101.51	23,24	4.629	23,24	4.671	-0.042					
208	4-Propyltoluene	101.57	24	4.633	24	4.695	-0.062	0.50	35	4.133	4.333	-0.200
209	2-Isopropyltoluene	99.46	24	4.527	24	4.695	-0.168					
210	3-Isopropyltoluene	99.14	24	4.510	24	4.518	-0.008					
211	4-Isopropyltoluene	99.5	23,24	4.529	23,24	4.552	-0.023	0.50	35	4.029	4.289	-0.260
212	1,2-Dimethyl-4-ethylbenzene	104.49	24	4.783	24	4.833	-0.05					
213	1,3-Dimethyl-4-ethylbenzene	104.41	24	4.779	24	4.806	-0.027					

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc log$L^{SQ\,a}$	residual [b]	logL^{w}	Ref.	obs log$P^{SQ\,c}$	calc log $P^{SQ\,d}$	residual [d]
214	1,3-Dimethyl-5-ethylbenzene	102.97	24	4.706	24	4.72	-0.015					
215	1,4-Dimethyl-2-ethylbenzene	103.93	24	4.755	24	4.784	-0.029					
216	Pentylbenzene	111.06	24	5.118	24	5.187	-0.069	0.17	35	4.948	5.034	-0.086
217	sec-Pentylbenzene	105.36	24	4.762	24	4.869	-0.107	0.13	36	4.632	5.042	-0.410
218	tert-Pentylbenzene	104.08	24	4.827	24	4.727	0.100					
219	4-tert-Butyltoluene	105.31	24	4.825	24	4.897	-0.072					
220	1,4-Diisopropylbenzene	113.5	23	5.243	23	5.272	-0.029					
221	1,3,5-Triethylbenzene	116.99	24	5.421	24	5.465	-0.045					
222	1,3,5-Triisopropylbenzene	126.78	23	5.92	23	5.976	-0.056					
223	Styrene	85.67	23	3.823	23	3.823	0.000	0.91	35	2.913	2.98	-0.067
224	α-Methylstyrene	94.44	23	4.271	23	4.254	0.017	0.91	35	3.361	3.452	-0.091
225	3-Methylstyrene	96.18	23	4.359	23	4.336	0.023					
226	4-Methylstyrene	96.83	23	4.392	23	4.360	0.032					
227	Allylbenzene	90.45	23	4.067	23	4.099	-0.032					
228	Biphenyl			5.900	25	5.965	-0.065	1.95	35	3.950	3.899	0.051
229	Naphthalene	110.86	23	5.108	23	5.112	-0.004	1.73	35	3.378	3.278	0.100
230	2-Methylnaphthalene	124.42	23	5.800	23	5.718	0.081	1.83	35	3.970	3.846	0.124
231	Indene	98.82	23	4.494	23	4.515	-0.021					0.000
232	Indane	99.03	23	4.505	23	4.549	-0.044	1.07	35	3.435	3.382	0.053
233	Acenaphthene			6.380	25	6.416	-0.036	2.36	35	4.020	3.908	0.112
234	Chlorobenzene			3.618	20	3.617	0.001	0.82	35	2.798	2.813	-0.015
235	Phenol			3.691	15	3.733	-0.042	4.85	35	-1.159	-1.077	-0.082
236	2-Methylphenol			4.185	15	4.181	0.004	4.31	35	-0.125	-0.110	-0.015
237	3-Methylphenol			4.254	15	4.272	-0.018	4.60	35	-0.346	-0.532	0.186
238	4-Methylphenol			4.246	15	4.274	-0.028	4.50	35	-0.254	-0.373	0.119
239	Furan			1.846	11	1.81	0.036	0.67	35	1.174	1.164	0.010
240	Thiophene			2.734	11	2.796	-0.062	1.04	35	1.694	1.701	-0.007

Nos	Solute	$UI_{298}/10$	Ref.	obs log L^{SQ}	Ref.	calc logL^{SQ} [a]	residual [b]	logL^{w}	Ref.	obs logP^{SQ} [c]	calc log P^{SQ} [d]	residual [d]
241	2,2-Dimethylpropane	41.08	23	1.550	23	1.801	-0.251	-1.97	35	3.520	3.707	-0.187
242	Methylcyclopentane	62.46	21	2.656	16,21	2.794	-0.138	-1.17	35	3.826	3.714	0.112
243	Isopropylcyclopentane	80.61	21	3.565	21	3.750	-0.185	-1.87	35			
244	Acetaldehyde			1.249	11	1.129	0.120	2.57	35	-1.321	-1.440	0.080
245	Propionaldehyde			1.862	15	1.708	0.154	2.52	35	-0.658	-0.675	0.017
246	Butyraldehyde			2.279	15	2.159	0.12	2.33	35	-0.051	-0.085	0.034
247	Ethyl acetate			2.340	11	2.196	0.144	2.16	35	0.180	0.149	0.031
248	Propyl acetate			2.811	15	2.696	0.115	2.05	35	0.761	0.769	-0.008
249	Isopropyl acetate			2.58	15	2.422	0.158	1.94	35	0.640	0.694	-0.054
250	Ethyl propanoate			2.839	11	2.684	0.155	1.97	35	0.869	0.799	0.070
251	Acetonitrile			1.384	9	1.636	-0.252	2.85	35	-1.466	-1.301	-0.165
252	Diethylsulfide			2.877	11	3.001	-0.124					
253	Pentamethylbenzene	125.31	24	5.845	24	5.712	0.133					
254	trans-Stilbene			7.264	25	7.468	-0.204	2.78	35	4.484	4.496	-0.012
255	Pyrene			8.518	25	8.879	-0.361	3.50	35	5.018	4.853	0.165

[a] Calculated from equation (6.20).
[b] Observed log L^{SQ} - calculated log L^{SQ}.
[c] Calculated from equation (6.21).
[d] Observed log P^{SQ} - calculated log P^{SQ}.

Table 6.10. Data for gas / apolane partition process at 313K.[a,b]

Nos	Solute	E	L	obs log L^{87}	calc log L^{87}	obs-calc
1	Pentane	0.000	2.162	1.714	1.752	-0.038
2	Hexane	0.000	2.668	2.176	2.197	-0.021
3	2-Methylpentane	0.000	2.503	2.055	2.052	0.003
4	Heptane	0.000	3.173	2.628	2.641	-0.013
5	2,4-Dimethylpentane	0.000	2.809	2.322	2.321	0.001
6	Octane	0.000	3.677	3.071	3.084	-0.013
7	2,5-Dimethylhexane	0.000	3.308	2.754	2.759	-0.005
8	2,3,4-Trimethylpentane	0.000	3.481	2.876	2.912	-0.036
9	Nonane	0.000	4.182	3.529	3.528	0.001
10	Decane	0.000	4.686	3.977	3.971	0.006
11	Undecane	0.000	5.191	4.422	4.415	0.007
12	Dodecane	0.000	5.696	4.863	4.860	0.003
13	Tridecane	0.000	6.200	5.302	5.303	-0.001
14	Tetradecane	0.000	6.705	5.759	5.747	0.012
15	Cyclopentane	0.263	2.477	2.092	2.079	0.013
16	Cyclohexane	0.305	2.964	2.477	2.515	-0.038
17	Cycloheptane	0.350	3.704	3.092	3.175	-0.083
18	Cyclooctane	0.413	4.329	3.619	3.736	-0.117
19	Ethylcyclohexane	0.263	3.877	3.262	3.310	-0.048
20	Pent-1-ene	0.093	2.047	1.646	1.668	-0.022
21	trans-Hex-2-ene	0.122	2.655	2.101	2.208	-0.107
22	Dichloromethane	0.387	2.019	1.760	1.700	0.060
23	Trichloromethane	0.425	2.480	2.132	2.113	0.019
24	Tetrachloromethane	0.458	2.823	2.412	2.421	-0.009
25	1,2-Dichloroethane	0.416	2.573	2.229	2.193	0.036
26	1,1,1-Trichloroethane	0.369	2.733	2.302	2.324	-0.022
27	1-Chloropropane	0.216	2.202	1.846	1.828	0.018
28	1-Chlorobutane	0.210	2.722	2.291	2.284	0.007
29	1-Chloropentane	0.208	3.223	2.769	2.725	0.044
30	Dibromomethane	0.714	2.886	2.500	2.525	-0.025
31	Tetrabromomethane	1.190	4.557	4.047	4.086	-0.039
32	Diethylether	0.041	2.015	1.616	1.630	-0.014
33	Methyl-tert-butylether	0.024	2.372	1.908	1.941	-0.033
34	Dipropylether	0.008	2.954	2.449	2.450	-0.001
35	Diisopropylether	-0.060	2.530	2.056	2.064	-0.008
36	Dibutylether	0.000	3.924	3.325	3.301	0.024
37	Diisobutylether	0.000	3.485	2.941	2.915	0.026
38	Dipentylether	0.000	4.875	4.199	4.138	0.061
39	Dihexylether	0.000	5.938	5.077	5.072	0.005
40	Tetrahydrofuran	0.289	2.636	2.171	2.224	-0.053
41	Tetrahydropyran	0.275	2.906	2.520	2.459	0.061
42	1,4-Dioxane	0.329	2.892	2.390	2.457	-0.067
43	Ethanol	0.246	1.485	1.231	1.203	0.028

Nos	Solute	E	L	obs log L^{87}	calc log L^{87}	obs-calc
44	Propan-1-ol	0.236	2.031	1.660	1.682	-0.022
45	Propan-2-ol	0.212	1.764	1.485	1.442	0.043
46	Butan-1-ol	0.224	2.601	2.137	2.181	-0.044
47	Butan-2-ol	0.217	2.338	1.936	1.948	-0.012
48	2-Methylpropan-1-ol	0.217	2.413	2.018	2.014	0.004
49	2-Methylpropan-2-ol	0.180	1.963	1.595	1.611	-0.016
50	Pentan-1-ol	0.219	3.106	2.605	2.624	-0.019
51	3-Methylbutan-1-ol	0.192	3.011	2.479	2.535	-0.056
52	Hexan-1-ol	0.210	3.610	3.018	3.065	-0.047
53	Hexan-2-ol	0.187	3.340	2.815	2.823	-0.008
54	Heptan-1-ol	0.211	4.115	3.491	3.510	-0.019
55	Octan-1-ol	0.199	4.619	3.925	3.951	-0.026
56	Cyclopentanol	0.427	3.241	2.723	2.782	-0.059
57	Cyclohexanol	0.460	3.758	3.157	3.243	-0.086
58	2,2,2-Trifluoroethanol	0.015	1.224	0.942	0.930	0.012
59	1,1,1,3,3,3-Hexafluoropropan-2-ol	-0.240	1.392	1.018	1.029	-0.011
60	Acetaldehyde	0.208	1.230	0.936	0.972	-0.036
61	Propionaldehyde	0.196	1.815	1.441	1.484	-0.043
62	Butyraldehyde	0.187	2.270	1.904	1.882	0.022
63	2-Methylproprionaldehyde	0.144	2.120	1.801	1.742	0.059
64	Pentanal	0.163	2.851	2.363	2.389	-0.026
65	Hexanal	0.146	3.357	2.821	2.831	-0.010
66	Heptanal	0.140	3.865	3.264	3.276	-0.012
67	Octanal	0.160	4.361	3.710	3.716	-0.006
68	Nonanal	0.150	4.826	4.156	4.123	0.033
69	Propanone	0.179	1.696	1.357	1.376	-0.019
70	Butanone	0.166	2.287	1.982	1.893	0.089
71	Pentan-2-one	0.143	2.755	2.356	2.301	0.055
72	Pentan-3-one	0.154	2.811	2.357	2.352	0.005
73	3-Methylbutan-2-one	0.134	2.692	2.176	2.243	-0.067
74	Hexan-2-one	0.136	3.262	2.772	2.745	0.027
75	Heptan-2-one	0.123	3.760	3.203	3.181	0.022
76	Octan-2-one	0.108	4.257	3.623	3.615	0.008
77	Nonan-2-one	0.119	4.735	4.081	4.037	0.044
78	Cyclopentanone	0.373	3.221	2.706	2.754	-0.048
79	Cyclohexanone	0.403	3.792	3.161	3.262	-0.101
80	Proprionitrile	0.162	2.082	1.662	1.712	-0.050
81	1-Cyanopropane	0.188	2.548	2.139	2.127	0.012
82	2-Cyanopropane	0.142	2.452	1.956	2.034	-0.078
83	1-Cyanobutane	0.177	3.108	2.573	2.618	-0.045
84	1-Cyanopentane	0.166	3.608	3.017	3.055	-0.038
85	Ethylamine	0.236	1.677	1.452	1.370	0.082
86	Propylamine	0.225	2.141	1.815	1.776	0.039
87	Butylamine	0.224	2.618	2.299	2.196	0.103
88	Diethylamine	0.154	2.395	2.030	1.986	0.044

Nos	Solute	E	L	obs log L^{87}	calc log L^{87}	obs-calc
89	Hexylamine	0.197	3.655	3.111	3.102	0.009
90	Triethylamine	0.101	3.040	2.523	2.543	-0.020
91	Ethyl formate	0.146	1.845	1.425	1.501	-0.076
92	Propyl formate	0.132	2.433	1.975	2.015	-0.040
93	Methyl acetate	0.142	1.911	1.530	1.558	-0.028
94	Ethyl acetate	0.106	2.314	1.990	1.906	0.084
95	Propyl acetate	0.092	2.819	2.402	2.347	0.055
96	Isopropyl acetate	0.055	2.546	2.205	2.100	0.105
97	Butyl acetate	0.071	3.353	2.853	2.813	0.040
98	Isobutyl acetate	0.052	3.161	2.673	2.640	0.033
99	Pentyl acetate	0.067	3.844	3.287	3.244	0.043
100	Isopentyl acetate	0.051	3.740	3.122	3.149	-0.027
101	Hexyl acetate	0.056	4.351	3.722	3.687	0.035
102	Ethyl propanoate	0.087	2.807	2.398	2.336	0.062
103	Ethyl butanoate	0.068	3.271	2.792	2.740	0.052
104	Isobutyl isobutanoate	0.000	3.885	3.284	3.267	0.017
105	N,N-Dimethylformamide	0.367	3.173	2.609	2.711	-0.102
106	Nitromethane	0.313	1.892	1.663	1.574	0.089
107	Nitroethane	0.270	2.414	2.023	2.025	-0.002
108	1-Nitropropane	0.242	2.894	2.416	2.442	-0.026
109	Benzene	0.610	2.786	2.480	2.417	0.063
110	Toluene	0.601	3.325	2.926	2.890	0.036
111	Ethylbenzene	0.613	3.778	3.297	3.290	0.007
112	o-Xylene	0.663	3.939	3.463	3.441	0.022
113	m-Xylene	0.623	3.839	3.378	3.346	0.032
114	p-Xylene	0.613	3.839	3.372	3.344	0.028
115	Propylbenzene	0.604	4.230	3.678	3.686	-0.008
116	Butylbenzene	0.600	4.730	4.099	4.125	-0.026
117	1,2-Diethylbenzene	0.688	4.732	4.154	4.144	0.010
118	1,3-Diethylbenzene	0.637	4.686	4.154	4.093	0.061
119	1,2-Dipropylbenzene	0.665	5.492	4.794	4.807	-0.013
120	1,4-Dipropylbenzene	0.620	5.603	4.894	4.896	-0.002
121	1,2-Dibutylbenzene	0.645	6.377	5.586	5.582	0.004
122	1,4-Dibutylbenzene	0.600	6.588	5.764	5.759	0.005
123	1,2-Dipentylbenzene	0.630	7.572	6.638	6.630	0.008
124	1,4-Dipentylbenzene	0.585	7.490	6.557	6.549	0.008
125	Naphthalene	1.340	5.161	4.631	4.646	-0.015
126	Fluorobenzene	0.477	2.788	2.422	2.394	0.028
127	Chlorobenzene	0.718	3.657	3.218	3.204	0.014
128	1,2-Dichlorobenzene	0.872	4.518	3.924	3.991	-0.067
129	1,4-Dichlorobenzene	0.825	4.435	3.980	3.909	0.071
130	Bromobenzene	0.882	4.041	3.597	3.573	0.024
131	Iodobenzene	1.188	4.502	4.070	4.037	0.033
132	Methylphenylether	0.708	3.890	3.450	3.407	0.043
133	Benzyl alcohol	0.803	4.221	3.753	3.716	0.037

Nos	Solute	E	L	obs log L^{87}	calc log L^{87}	obs-calc
134	2-Phenylethanol	0.811	4.628	4.043	4.076	-0.033
135	4-Phenylbutan-1-ol	0.811	5.049	4.487	4.446	0.041
136	Phenol	0.805	3.766	3.407	3.316	0.091
137	2-Methylphenol	0.840	4.218	3.739	3.721	0.018
138	3-Methylphenol	0.822	4.310	3.764	3.798	-0.034
139	4-Methylphenol	0.820	4.312	3.761	3.799	-0.038
140	Benzaldehyde	0.820	4.008	3.504	3.532	-0.028
141	Acetophenone	0.818	4.501	3.956	3.965	-0.009
142	Benzonitrile	0.742	4.039	3.479	3.544	-0.065
143	Phenylacetonitrile	0.751	4.570	3.989	4.013	-0.024
144	Aniline	0.955	3.934	3.506	3.493	0.013
145	N-Methylaniline	0.948	4.478	3.998	3.970	0.028
146	N,N-Dimethylaniline	0.957	4.701	4.223	4.168	0.055
147	Pyridine	0.631	3.022	2.620	2.629	-0.009
148	Nitrobenzene	0.871	4.557	3.992	4.025	-0.033
149	Cyclodecane	0.474	5.34	4.43	4.637	-0.207
150	1,1,2,2-Tetrachloroethane	0.595	3.803	3.141	3.309	-0.168
151	Tetrachloroethene	0.639	3.584	3.359	3.125	0.234
152	Methanol	0.278	0.97	0.936	0.757	0.179
153	Acetonitrile	0.237	1.739	1.185	1.425	-0.240
154	Methyl formate	0.192	1.285	0.861	1.017	-0.156
155	N,N-Dimethylacetamide	0.363	3.717	2.977	3.189	-0.212
156	Dimethylsulfoxide	0.522	3.459	2.69	2.992	-0.302
157	3-Phenylpropan-1-ol	0.821	4.663	4.263	4.108	0.155

a Observed log L^{87} values taken from Ref. 3.
b Log L^{87} values calculated on equation (6.22).

Table 6.11. Estimation of **L** values from log L^{SQ} and **E** values

Nos	Solute	**E**	Ref.	log L^{SQ}	Ref.	Calc L^c
	Aliphatic alkanes					
1	1-trans-2-Dimethylpropane	0.29	a	2.04	23	2.15
2	3,3,5-Trimethylheptane	0.00	a	4.05	21	4.19
3	2,2,4-Trimethylheptane	0.00	a	3.90	21	4.04
	Cyclic alkanes					
4	Ethylcyclopropane	0.32	a	2.17	23	2.27
5	1-cis-2-Dimethylcyclopropane	0.29	a	2.20	23	2.31
6	1,2,2-Trimethylcyclopropane	0.24	a	2.37	23	2.48
7	1,1,2,2-Tetramethylcyclopropane	0.19	a	2.70	23	2.82
8	Ethylcyclobutane	0.19	b	2.70	23	2.82
9	1,1,3-Trimethylcyclopentane	0.17	a	3.19	23	3.31
10	1-trans-2-cis-4-Trimethylcyclopentane	0.15	a	3.27	21	3.39
11	1-trans-2-cis-3-Trimethylcyclopentane	0.15	a	3.31	21	3.43
12	1,trans-2,cis- 3-Trimethylcyclopentane	0.15	a	3.31	23	3.43
13	1,1,2-Trimethylcyclopentane	0.17	a	3.37	21	3.49
14	1-cis-2-trans-4-Trimethylcyclopentane	0.15	a	3.42	21	3.54
15	1-cis-2-trans-3-Trimethylcyclopentane	0.15	a	3.45	23	3.57
16	1-Methyl-cis-3-ethylcyclopentane	0.19	a	3.47	23	3.59
17	1-Ethyl-trans-2-methylcyclopentane	0.19	a	3.51	21	3.63
18	1-Methyl-trans-3-ethyl-cyclopentane	0.19	a	3.51	23	3.63
19	1-cis-2-cis-3-Trimethylcyclopentane	0.15	a	3.55	21	3.67
20	1-Methyl-cis-2-ethylcyclopentane	0.19	a	3.65	21	3.77
21	1,1,3-Trimethylcyclohexane	0.14	a	3.74	21	3.87
22	1-Methyl-cis-3-n-propylcyclopentane	0.18	a	3.93	23	4.05
23	1-Methyl-1-n-propylcyclopentane	0.20	a	3.94	23	4.06
24	1-Methyl-trans-2-n-propylcyclopentane	0.18	a	3.94	23	4.06
25	1-Methyl-trans-3-n-propylcyclopentane	0.18	a	3.95	23	4.07
26	1-Methyl-cis-2-n-propylcyclopentane	0.18	a	4.08	23	4.20
27	n-Butylcyclopentane	0.24	a	4.11	23	4.23
28	s-Butylcyclopentane	0.30	a	4.16	23	4.28
29	n-Pentylcyclopentane	0.24	a	4.64	23	4.76
	Aliphatic alkenes					
30	4,4-Dimethyl-1-pentene	0.05	b	2.63	21	2.76
31	3-Methyl-trans-2-pentene	0.16	b	2.68	21	2.80
32	2,3-Dimethyl-2-butene	0.21	b	2.73	21	2.84
33	3,3-Dimethylpent-1-ene	0.09	b	2.73	21	2.85
34	2,3,3-Trimethylbut-1-ene	0.11	b	2.74	21	2.86
35	4,4-Dimethyl-cis-2-pentene	0.18	a	2.77	21	2.89
36	3,4-Dimethyl-1-pentene	0.14	b	2.78	21	2.90
37	2,4-Dimethylpent-1-ene	0.09	b	2.79	21	2.91
38	2,4-Dimethyl-2-pentene	0.12	b	2.81	21	2.93
39	3-Methyl-1-hexene	0.12	a	2.82	21	2.94
40	3-Ethyl-1-pentene	0.12	a	2.83	21	2.95

Nos	Solute	E	Ref.	log L^{SQ}	Ref.	Calc L^{c}
41	2-Methyl-trans-3-hexene	0.17	a	2.84	21	2.96
42	2,3-Dimethylpent-1-ene	0.12	b	2.85	21	2.97
43	5-Methyl-1-hexene	0.12	a	2.85	21	2.97
44	4-Methyl-cis-2-hexene	0.17	a	2.88	21	3.00
45	4-Methyl-trans-2-hexene	0.17	a	2.88	21	3.00
46	4-Methyl-1-hexene	0.12	a	2.89	21	3.01
47	3,4-Dimethyl-cis-2-pentene	0.13	a	2.95	21	3.07
48	2-Methyl-1-hexene	0.11	b	2.99	21	3.11
49	2-Ethylpent-1-ene	0.13	b	3.01	21	3.13
50	2-Methyl-2-hexene	0.18	a	3.06	21	3.18
51	3-Methyl-cis-2-hexene	0.18	a	3.06	21	3.18
52	3-Ethyl-2-pentene	0.18	b	3.08	21	3.20
53	2,3-Dimethyl-2-pentene	0.21	b	3.11	21	3.23
54	2,2-Dimethyl-cis-3-hexene	0.17	a	3.16	21	3.28
55	trans-Oct-4-ene	0.11	b	3.49	21	3.62
56	cis-Oct-4-ene	0.13	b	3.51	21	3.63
57	cis-Oct-3-ene	0.13	b	3.51	21	3.64
58	trans-Oct-3-ene	0.12	b	3.52	21	3.65
59	trans-Oct-2-ene	0.12	b	3.57	21	3.70
60	cis-4-Nonene	0.17	a	3.97	21	4.09
61	trans-4-Nonene	0.17	a	3.98	21	4.10
62	cis-3-Nonene	0.17	a	3.98	21	4.10
63	trans-3-Nonene	0.17	a	3.99	21	4.11
64	trans-2-Nonene	0.17	a	4.04	21	4.17
65	cis-2-Nonene	0.17	a	4.05	21	4.18
	Cyclic alkenes					
66	3-Methyl-1-cyclopentene	0.41	a	2.62	23	2.72
67	3-Allyl-1-cyclopentene	0.90	a	3.09	23	3.15
68	3-Ethyl-1-cyclopentene	0.38	a	3.14	23	3.24
69	1-Ethyl-1-cyclopentene	0.42	a	3.31	23	3.41
70	3-Isopropyl-1-cyclopentene	0.40	a	3.52	23	3.62
71	1-Isopropyl-1-cyclopentene	0.45	a	3.59	23	3.69
72	1-Allyl-1-cyclopentene	0.96	a	3.64	23	3.70
73	3-Propyl-1-cyclopentene	0.41	a	3.61	23	3.71
74	1-Propyl-1-cyclopentene	0.41	a	3.73	23	3.83
75	3-Ethyl-1-cyclohexene	0.34	a	3.74	23	3.85
76	4-Ethyl-1-cyclohexene	0.34	a	3.75	23	3.86
77	1-Ethyl-1-cyclohexene	0.39	a	3.82	23	3.93
78	3-Isobutyl-1-cyclopentene	0.36	a	3.91	23	4.02
79	1-Isobutyl-1-cyclopentene	0.40	a	3.96	23	4.07
80	3-Butyl-1-cyclopentene	0.37	a	4.08	23	4.19
81	1-Isopropyl-1-cyclohexene	0.42	a	4.10	23	4.21
82	3-Isopropyl-1-cyclohexene	0.34	a	4.12	23	4.23
83	1-Butyl-1-cyclopentene	0.41	a	4.20	23	4.31
84	3-Propyl-1-cyclohexene	0.34	a	4.20	23	4.31

Nos	Solute	E	Ref.	log L^{SQ}	Ref.	Calc L^{c}
85	4-Propyl-1-cyclohexene	0.34	a	4.21	23	4.32
86	1-Propyl-1-cyclohexene	0.38	a	4.23	23	4.34
87	3-Isopentyl-1-cyclopentene	0.35	a	4.37	23	4.48
88	1-Isopentyl-1-cyclopentene	0.39	a	4.39	23	4.50
89	1-Isobutyl-1-cyclohexene	0.39	a	4.46	23	4.57
90	3-Pentyl-1-cyclopentene	0.36	a	4.55	23	4.66
91	1-Pentyl-1-cyclopentene	0.41	a	4.67	23	4.78
92	4-Butyl-1-cyclohexene	0.34	a	4.67	23	4.79
93	3-Butyl-1-cyclohexene	0.34	a	4.67	23	4.79
94	1-Butyl-1-cyclohexene	0.38	a	4.70	23	4.81
	Dienes					
95	3-Methyl-1,2-butadiene	0.30	b	2.19	23	2.29
96	1,2-Pentadiene	0.31	b	2.26	23	2.36
97	2,3-Pentadiene	0.34	b	2.28	23	2.38
98	1-trans-4-Hexadiene	0.24	b	2.53	23	2.64
99	1-cis-4-Hexadiene	0.24	b	2.55	23	2.66
100	1-trans-3-Hexadiene	0.35	b	2.67	23	2.77
101	2-Methyl-1,3-trans-pentadiene	0.27	a	2.74	23	2.85
102	4-Methyl-1,3-pentadiene	0.27	a	2.74	23	2.85
103	3-Methyl-1-cis-3-pentadiene	0.27	a	2.79	23	2.90
104	3-Methyl-1-trans-3-pentadiene	0.27	a	2.80	23	2.91
105	trans-2,trans-4-Hexadiene	0.40	b	2.83	23	2.93
106	cis-2,trans-4-Hexadiene	0.40	b	2.87	23	2.97
	Aliphatic alkynes					
107	3,3-Dimethylbut-1-yne	0.05	b	1.89	10	2.01
108	Hex-3-yne	0.22	b	2.69	10,23	2.80
109	Hex-3-yne	0.22	b	2.76	10,23	2.87
110	Hex-2-yne	0.24	b	2.83	23	2.94
111	Hept-3-yne	0.23	b	3.21	23	3.33
112	Hept-2-yne	0.24	b	3.32	23	3.44
113	Oct-4-yne	0.21	b	3.64	23	3.76
114	Oct-3-yne	0.21	b	3.68	23	3.80
115	Non-4-yne	0.22	b	4.12	23	4.24
116	Non-3-yne	0.20	b	4.15	23	4.27
117	Non-2-yne	0.23	b	4.26	23	4.38
118	Dec-4-yne	0.22	a	4.58	23	4.70
119	Dec-5-yne	0.19	b	4.58	23	4.71
120	Dec-3-yne	0.22	a	4.62	23	4.74
121	Dec-2-yne	0.22	a	4.74	23	4.86
122	Undec-1-yne	0.14	b	4.93	23	5.06
123	Undec-4-yne	0.22	b	5.04	23	5.17
124	Undec-5-yne	0.19	b	5.04	23	5.17
125	Undec-3-yne	0.22	b	5.09	23	5.22
126	Undec-2-yne	0.22	b	5.22	23	5.35
127	Dodec-6-yne	0.23	b	5.49	23	5.62

Nos	Solute	E	Ref.	log L^{SQ}	Ref.	Calc L[c]
128	Dodec-5-yne	0.23	b	5.50	23	5.63
129	Dodec-4-yne	0.23	b	5.52	23	5.65
130	Dodec-3-yne	0.23	b	5.57	23	5.70
131	Dodec-2-yne	0.57	b	5.70	23	5.80
	Halogenated alkynes					
132	3-Bromoprop-1-yne	0.57	b	2.38	10	2.47
133	3-Chloro-3-methylbut-1-yne	0.26	b	2.37	10	2.48
	Aliphatic Alcohols					
134	Heptan-4-ol	0.18	b	3.69	14	3.78
135	Octan-3-ol	0.18	b	4.19	14	4.29
136	Octan-4-ol	0.16	b	4.16	14	4.25
137	2,3-Dimethylhexan-2-ol	0.17	a	3.85	14	3.95
138	3,4-Dimethylhexan-2-ol	0.15	a	4.04	14	4.13
139	3,5-Dimethylhexan-3-ol	0.17	a	3.77	14	3.86
140	3-Methylheptan-3-ol	0.19	b	3.96	14	4.05
	Miscellaneous					
134	1,1,1-Trifluoropropanone	-0.05	b	1.00	11	1.13
135	Trimethylsilacetylene	-0.20	b	2.02	10	2.16
136	Chloroacetone	0.38	b	2.59	11	2.69
137	Diethylsulfite	0.19	b	2.88	11	3.00
138	Ethyl thioacetate	0.44	b	3.16	11	3.26
139	Diethylselenide	0.52	b	3.19	11	3.28
140	1-Ethyl-4-n-propylbenzene	0.63	b	5.03	24	5.13
141	1-Ethyl-4-tert-butylbenzene	0.58	a	5.16	23	5.26
142	1,2-Dibutylbenzene	0.64	a	5.59	3	6.38
143	1,4-Dibutylbenzene	0.63	a	5.76	3	6.59
144	1,2-Dipentylbenzene	0.60	a	6.64	3	7.57
145	1,4-Dipentylbenzene	0.58	a	6.56	3	7.49
146	Cyanopropane	0.14	a	1.96	3	2.37

[a] **E** values calculated from experimental refractive index at 293K.
[b] **E** values calculated from estimated refractive index at 293K.
[c] For compounds 1-141, **L** values back-calculated from equation (6.18), for compounds 142-146, **L** values back-calculated from equation (6.22).

6.5. References

1. M.H. Abraham, Chem. Soc. rev., 22 (1993) 73.
2. M.H. Abraham, P.L. Grellier, R.A. McGill, J.Chem.Soc., Perkin Trans. 2, (1987) 797.
3. J.D. Weckwerth, P.W. Carr, M.F. Vitha, Asad Nasehzadhec, Anal. Chem., 70 (1998) 3712.
4. A.J. Dallas, P.W. Carr, J. Phys. Chem., 98 (1994) 4927.
5. M.H. Abraham in 'Quantitative Treatments of Solute / Solvent Interactions, Theoretical and Computational Chemistry' Vol. 1, Eds. Politzer, J.S. Murray, Elsevier Sci., B.V., (1994) 83.
6. P. Haveleck, J.G.K. Sevsick, J. Phys. Chem. Ref. Data, 25 (1996) 1483.
7. G. Defayes, D.F. Fritz, T. Gorner, G. Huber, C. De Reyff, E.J. Kovats, Chromatogr. 500 (1990) 139.
8. J.H. Park, A. Hussan, P. Couasnon, D. Fritz, P.W. Carr, Anal. Chem., 59 (1987) 1970.
9. T. Nitta, K. Morinaga, T. Katayana, Ind. Eng. chem. Fundam., 21 (1982) 396.
10. R. Queignec, B. Wojtkowiak, Bull. Soc. Chim. Fr., (1970) 860.
11. R. Queignec, M. Cabanestos-Queignec, J. Chromatogr., 209 (1981) 345.
12. IUPAC Solubility Data Series.
13. L.M. Romero, M.R. Filgueira, L.G. Gagliardi, A.M. Nardillo, R.C. Castells, Phys. Chem. Chem. Phys., 1 (1999) 3351.
14. R.C. Castells, L.M. Romero, A.M. Nardillo, J. Chrom A, 898 (2000) 103.
15. G.M. Janini, L.A. Qaddora, J. Liq. Chromatogr., 9 (1986) 39.
16. D.C. Locke, J.Chromatogr., 35 (1968) 24.
17. C.-P. Chai, M.E. Paulaitis, J. Chem. Eng. Data, 26 (1981) 277.
18. C.C. Chappellow, J.M. Prausnitz, AIChe J., 20 (1974) 1097.
19. D.H. Desty, W.T. Swanton, J. Phys. chem., 65 (1961) 766.
20. L.A. Quaddora, G.M. Janini, J. Chem. Eng. Data, 31 (1986) 392.
21. N. Dimov, J. Chromatogr., 347 (1985) 366.
22. D. Papazova, R.Milina, Chromatographia, 25 (1988) 177.
23. B.D. Skribc, J.D. Cvejanov, Lj.S. Pavic-Suzuki, Chromatographia, 42 (1996) 660.
24. D. Papazova, N. Dimov, J. Chromatogr., 356 (1986) 320.
25. K.M. De Fina, T. Van, W.E. Acree, Jr., Unpublished results.
26. TRC Thermodynamic Tables, Texas A and M University System, College Station, Texas.
27. T. Boublik, The Vapour Pressure of Pure Substances. Selected values of Temperature Dependence of some Pure Substances in the Normal and Low Pressure Region., Eds. T. Boublik, V. Fried, E. Ha'la, Elsevier, Amsterdam, 1973.
28. R.R. Dreishbach, Physical Properties of Chemical Compounds, ACS, Washington D.C., 1955 (Vol.1), 1959 (Vol.2), 1961 (Vol.3).

29. J.A. Riddick, W.B. Bunger, 'Organic Solvents, Vol. 2' in Techniques of Chemistry, Ed. E. Weissberger, 3rd Ed., Wiley-Interscience, New-York, 1970.
30. R.M. Stephenson, S. Malanowski, 'Handbook of the Thermodynamics of Organic Compounds', Elsevier, Amsterdam, 1987.
31. R.C. Wilhoit, B.J. Zwolinski, 'Handbook of Vapor Pressures and Heats of Vaporization of Hydrocarbons', Thermodynamic Research Center, College Station, Texas, USA, 1971.
32. M. Pompe, M. Novic, J. Chem. Inform. Comp. Sci., 39 (1999) 59.
33. M.V. Budahegyi, E.R. Lombosi, T.S. lombosi, S.Y. Mesraros, Sz. Nyiredy, G. Tarjan, I. Timar, J.M. Takacs, J. Chromatogr., 271 (1983) 213.
34. M.H. Abraham, C.E. Green, W.E. Acree, Jr., C.E. Hernandez, L.E. Roy, J.Chem. Soc., Perkin Trans. 2, (1998) 2677.
35. M.H. Abraham, H.S. Chadha, G.S. Whithing, R.C. Mitchell, J.Pharm. Sci., 83 (1994) 1085.
36. V.N. Viswanadhan, A.K. Ghose, U.C. Singh, J.J. Wendoloski, J.Chem. Inf. Comp. Sci., 39 (1999) 405.
37. M.H. Abraham, C.E. Green, J. Platts, Unpublished work.
38. M.H. Abraham, H.S. Chadha, in 'Lipophilicity in Drug Action and Toxicology', Eds. V. Pliska, B. Testa and H. Van de Waterbeemd, VCh, Weinheim, (1996).
39. M.H. Abraham, C.F. Poole, S.K. Poole, J.Chromagr.A, 842 (1999) 79.
40. Q. Li, C.F. Poole, W. Kiridena, W. W. Koziol, Analyst, 125 (2000) 2180.
41. M.H. Abraham, D.S. Ballantine, B.K. Callihan, J. Chromatogr., 878 (2000) 115.
42. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W. S. Cain, M. Roses, E. Bosch, M.L. Diaz, J.Chem. Soc., Perkin Trans. 2, (1998) 2405.

7.0. Introduction

Organofluorocarbon fluids are stable, non-flammable, non-corrosive and non-explosive. Thanks to these physical properties, they have been used in numerous applications[1]. The chlorofluorocarbons (CFCs) are used as refrigerants, aerosol propellants, foaming agents, solvents and cleaning agents[1]. However, due to their effect on ozone depletion, a complete ban on their production by the year 2000 has been scheduled by the 1987 Montreal protocol and its latest amendments[2]. As a result, attention has turned on the development and assessment of hydrochloro-fluorocarbon (HCFCs), hydrofluorocarbons (HFCs) and perfluorocarbons (PFCs). HCFCs have much reduced ozone depletion potentials compared to CFCs, nevertheless their production in developed countries is due to be phased out by the year 2030. HFCs and PFCs have zero ozone depletion potentials. Although, CFCs, HCFCs, HFCs and PFCs are still widely used in everyday life, many physicochemical and biochemical properties are not known. Hence it would be of great interest if it were possible to predict such properties.

The solvation descriptor method of Abraham[3] based on equation (7.1) and (7.2) is one of the most useful approaches for the analysis and prediction of solute effects in chemical and biochemical systems.

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + v.\mathbf{V} \qquad (7.1)$$

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (7.2)$$

where SP is the logarithmic value of some property for a series of solutes. The use of the above equations depends on the availability of the solute descriptors, **E**, **S**, **A**, **B**, **V**, and **L**. Theory of the Abraham solvation equations together with the general methods of obtaining the solute descriptors in equation (7.1) and equation (7.2) were detailed in chapter 3. In brief, **V** can easily be calculated.[3] For gaseous compounds at 293K, **E** is better added to the list of descriptors to be determined, viz. **S**, **A**, **B**, and **L**. Equations on

the line of equation (7.1) and equation (7.2) are set out for processes in which values for a given solute are available. The only unknowns are the solute descriptors themselves, and are obtained by a least squares method that assigns values of descriptors that minimise the differences between the observed and calculated SP values.

In the present work, this approach was used to calculate the descriptors for a series of 18 organofluorocarbon fluids, classed as refrigerants, including CFCs, HCFCs, HFCs and PFCs; names and abbreviations are in Table 7.1.

Table 7.1. List of refrigerants.

Solute	Refrigerant code
Difluoromethane	R32
Trifluoromethane	R23
Tetrafluoromethane	R14
Trichlorofluoromethane	R11
Dichlorodifluoromethane	R12
Chlorotrifluoromethane	R13
1,1-Difluoroethane	R152a
1,1,1,2-Tetrafluoroethane	R134a
Pentafluoroethane	R125
1,1-Dichloro-1-fluoroethane	R141b
1-Chloro-1,2-difluoroethane	R142b
1-Chloro-1,1,2-trifluoroethane	R133a
1-Chloro-1,2,2,2-tetrafluoroethane	R124
1,2-Dichloro-1,1,2,2-tetrafluoroethane	R114
1-Chloro-1,1,2,2,2-pentafluoroethane	R115
1,1,1,2,3,3,3-Heptafluoropropane	R227
1-Chloro-2,2-difluoroethene	R1122
Perfluoropropene	R1216

7.1. General Method for Descriptor Determination

7.1.1. Determination of Partition Coefficient Values from Literature Data

Arlt and co-workers have recently measured Henry's law coefficient, H^S, for the 18 refrigerants listed in Table 7.1.[4,5] Values of H^S were determined in the solvents (S) octan-1-ol (dry octanol), n-nonane, N-methylpyrrolidone (dry NMP) and dimethylformamide (dry DMF) and in water and water/octan-1-ol mixture (wet octanol). The Henry's law coefficient, H^S can easily be transformed into the gas-solvent partition coefficient, L^S:

$$L^S = R.\ T.\ \rho_s .1000 / H^S.M_s \qquad (7.3)$$

where R is the gas constant ($dm^3.bar.K^{-1}.mol^{-1}$), T is the temperature (K), and ρ_s and M_s are the density ($g.cm^{-3}$) and the molecular weight ($g.mol^{-1}$) of the pure solvent. Values of ρ_S at T were taken from the literature.[6] The factor 1000 is needed to keep L^S as a dimensionless data. Equation (7.3) can be simplified as follows:

$$L^S = A / H^S \qquad (7.4)$$

A is the (R. T. ρ_S. 1000 / M_S) term, see Table 7.2 for values of A used in this work.

Table 7.2. Values of A used in equation (7.4)

Solvent	A
Wet Octanol	136.16
DMF	315.82
Dry Octanol	154.34
NMP	253.93
Nonane	198.47
Water (298K)	1354.00
Water (310K)	1403.23

Now, if the gas-water (W) and gas-solvent (S) partitions, $logL^W$ and $logL^S$, are known, then the water-solvent partition coefficient, $logP_S$, can be deduced from:

$$logP_S = logL_S - logL_w \qquad (7.5)$$

Values of log L^S, log P^S and log L^W obtained from the data of Arlt and co-workers [4,5] are given in Table 7.3.

In addition, Eger and co-workers[7] determined gas-olive oil partition coefficients, log L^{oil}, for a number of HFCs and PFCs at 310K. Values of log L^w at 310K are available, see Table 7.3, so that $logP^{oil}$ at 310K can be obtained from equation (7.5).

Finally, a number of directly determined water-solvent partition coefficients, P^S, for R14, R134a, R125 and R114 were obtained either from the Medchem97 database [8] or literature surveys [9-12], and the corresponding log L^S values were deduced from equation (7.4). Note that 'dry' solvents are solvents that are miscible with water and for which log P^S values are for the hypothetical partition between water and the pure dry solvent, obtained from equation (7.5). The 'wet' solvents are those for which partitions have been obtained by direct experiments in which the solvent is saturated with water.

7.1.2. Solvation Equations

Equations in log P^S and log L^S have been constructed for equations (7.1) and (7.2); the coefficients for equation (7.1) are given in Table 7.4, and the coefficients for equations (7.2) are listed in Table 7.5. Then, for a given solute, if any values of log P^S and log L^S are known in several systems, the entire set of equations can be solved to yield to set of descriptors.

Table 7.3. Gas/solvent and water/solvent partition coefficients calculated from Henry's law coefficients

Code	$LogL^{w298}$	$LogL^{w310}$	$logL^{woct}$	$logP^{woct}$	$logL^{oct}$	$LogP^{oct}$	$LogL^{NMP}$	$logP^{NMP}$	$logL^{DMF}$	$logP^{DMF}$	$logL^{non}$	$logP^{non}$
R32	0.235	0.124	0.427	0.192	0.413	0.178	1.242	1.007	1.241	1.007	0.267	0.033
R23	-0.510	-0.622	0.130	0.640	0.148	0.658	0.993	1.502	0.632	1.141	-0.046	0.463
R14	-2.306	-2.386	-0.950	1.356	-0.913	1.385	-1.324	0.982	-1.142	1.164	-0.590	1.716
R11	-0.451	-0.641	1.809	2.260	1.885	2.335	1.875	2.326	1.851	2.302	2.083	2.534
R12	-1.129	-1.275	0.920	2.049	0.982	2.111	0.978	2.107	0.994	2.123	1.233	2.362
R13	-1.670	-1.777	0.005	1.679	0.043	1.711	-0.174	1.496	-0.059	1.610	0.338	2.008
R152a	0.090	-0.041	2.369	0.690	0.782	0.692	1.479	1.390	1.553	1.463	0.791	0.701
R134a	-0.408	-0.547	0.648	1.056	0.647	1.055	1.442	1.849	1.616	2.024	0.558	0.996
R125	-1.059	-1.203	0.391	1.450	0.419	1.477	1.159	2.218	1.126	2.185	0.290	1.348
R141b	-0.148	-0.336	1.804	1.952	1.949	2.097	-	-	-	-	-	-
R142b	-0.449	-0.605	1.168	1.617	1.189	1.638	1.562	2.011	1.528	1.977	1.274	1.724
R133a	1.011	-0.165	1.451	1.446	1.476	1.471	-	-	-		1.348	1.344
R124	-0.569	-0.737	1.180	1.749	1.220	1.789	1.749	2.318	1.786	2.355	1.104	1.674
R114	-1.652	-1.788	1.126	2.778	1.181	2.833	1.093	2.745	1.165	2.817	1.474	3.126
R115	-2.116	-2.212	0.227	2.344	1.903	2.395	-0.045	2.071	1.127	2.212	0.643	2.579
R227	-1.468	-1.636	0.754	2.222	0.821	2.289	1.606	3.074	1.537	3.005	-	-
R1122	-0.380	-0.544	1.136	1.516	1.236	1.616	1.656	2.035	1.463	1.843	-	-
R1216	-1.772	-1.895	0.209	1.981	0.160	1.932	-	-	-	-	0.582	2.354

Table 7.4. Regression coefficients in equation (7.1) for partition from water at 298 K

Process	*c*	*r*	*s*	*a*	*b*	*v*	n	r	sd	F	Ref.
Water to solvent											
Gas phase	-0.994	0.577	2.549	3.813	4.841	-0.869	408	0.998	0.151	16 810	14
Gas phase, 310 K	-0.966	0.698	2.412	3.393	4.577	-1.072	82	0.995	0.156	1271	15
Olive oil, 310 K	-0.011	0.577	-0.800	-1.470	-4.921	4.173	174	0.997	0.145	5841	16
Dry DMF	0.105	0.317	0.462	1.154	-4.843	3.757	68	0.996	0.140	1525	17
Dry NMP	-0.071	0.686	0.455	1.547	-5.068	3.899	65	0.995	0.155	1221	17
Benzene	0.142	0.464	-0.588	-0.309	-0.625	4.491	213	0.996	0.143	5317	17
Tetrachloromethane	0.212	0.602	-1.234	-3.515	-4.528	4.552	173	0.998	0.119	15658	17
Dry methanol	0.329	0.299	-0.671	0.080	-3.389	3.512	93	0.994	0.156	1440	18
Dry ethanol	0.208	0.409	-0.959	0.176	-3.645	3.928	64	0.995	0.173	1205	19
Dry propanol	0.148	0.436	-1.098	0.389	-3.893	4.036	76	0.998	0.130	2892	20
Dry butanol	0.152	0.438	-1.177	0.096	-3.916	4.122	88	0.997	0.125	2719	21
Dry pentanol	0.080	0.521	-1.294	0.208	-3.908	4.208	59	0.998	0.112	2597	21
Dry hexanol	0.044	0.470	-1.153	0.083	-4.057	4.249	46	0.999	0.114	3775	21
Dry heptanol	-0.226	0.491	-1.258	0.035	-4.155	4.415	38	0.999	0.081	2333	21
Dry decanol	0.008	0.485	-0.974	0.015	-3.798	3.945	45	0.999	0.123	3843	21
Dry octanol	0.013	0.550	-1.205	-0.020	-4.262	4.253	99	0.999	0.103	9536	17
Wet octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	613	0.997	0.116	23162	22
Hexane	0.361	0.579	-1.723	-3.599	-4.764	4.344	173	0.994	0.207	2721	23
Heptane	0.325	0.670	-2.061	-3.317	-4.733	4.543	173	0.992	0.254	2281	23
Octane	0.223	0.642	-1.647	-3.480	-5.067	4.526	149	0.992	0.205	1702	23
Nonane	0.240	0.619	-1.713	-3.532	-4.921	4.482	64	0.991	0.123	6417	23
Decane	0.228	0.621	-1.550	-3.535	-5.359	4.533	62	0.998	0.144	2790	23
Hexadecane	0.087	0.667	-1.617	-3.587	-4.869	4.433	370	0.998	0.124	20236	22
Cyclohexane	0.127	0.816	-1.731	-3.778	-4.905	4.646	170	0.997	0.131	5512	22

Table 7.5. Regression coefficients in equation (7.2) for partition from the gas phase at 298 K

Process	*c*	*r*	*s*	*a*	*b*	*l*	n	r	sd	F	Ref.
Gas to Solvent											
Water	-1.272	0.822	2.743	3.904	4.814	-0.231	392	0.996	0.175	10229	14
Water, 310 K	-1.328	1.058	2.568	3.658	4.533	-0.248	84	0.992	0.178	863	15
Olive oil, 310 K	-0.230	0.009	0.795	1.353	0.000	0.888	141	0.998	0.087	9508	17
Dry DMF	-0.161	-0.179	2.327	4.756	0.000	0.808	72	0.998	0.115	3581	17
Dry NMP	-0.293	0.253	2.210	5.094	0.000	0.817	72	0.996	0.128	1921	17
Benzene	0.107	-0.313	1.053	0.457	0.169	1.020	175	0.999	0.119	12570	17
Tetrachloromethane	0.282	-0.303	0.460	0.000	0.000	1.047	173	0.998	0.119	15658	17
Dry methanol	-0.001	-0.196	1.117	3.671	1.501	0.771	93	0.998	0.134	3680	17,18
Dry ethanol	0.012	-0.221	0.819	3.636	1.249	0.854	74	0.998	0.145	3534	19,21
Dry propanol	-0.028	-0.175	0.648	4.022	1.043	0.869	77	0.999	0.120	6073	20,21
Dry butanol	-0.039	-0.276	0.539	3.781	0.995	0.934	92	0.998	0.158	5099	21
Dry pentanol	-0.042	-0.277	0.526	3.779	0.983	0.932	61	1.000	0.076	19143	21
Dry hexanol	-0.035	-0.298	0.626	3.726	0.729	0.936	46	1.000	0.089	17171	21
Dry heptanol	-0.062	-0.168	0.429	3.541	1.171	0.927	38	1.000	0.067	23045	21
Dry octanol	-0.071	-0.119	0.443	3.689	0.589	0.933	99	0.999	0.030	9535	21,24
Wet octanol	-0.222	0.088	0.701	3.478	1.477	0.851	395	0.994	0.210	6363	17
Hexane	0.292	-0.169	0.000	0.000	0.000	0.979	119	0.998	0.102	15683	23
Heptane	0.275	-0.162	0.000	0.000	0.000	0.983	109	0.999	0.088	19486	23
Octane	0.215	-0.049	0.000	0.000	0.000	0.967	105	0.999	0.098	17429	23
Nonane	0.200	-0.145	0.000	0.000	0.000	0.980	55	0.992	0.174	6310	23
Decane	0.156	-0.143	0.000	0.000	0.000	0.989	60	1.000	0.065	26396	23
Hexadecane	0.000	0.000	0.000	0.000	0.000	1.000	-	-	-	-	22
Cyclohexane	0.216	0.000	-0.179	0.000	0.000	1.019	114	0.998	0.115	12839	22,23

[a] Coefficients are defined as such

Now it is important that the equations in Table 7.4 (and equations in Table 7.5) contain different information that can be used to obtain descriptors. An approach of comparison of solvation equations is that of Ishihama and Asakawa.[13] The authors regard the five-descriptor equations as a line in five dimensions. Then for two equations, the angle, θ, between the lines is a measure of how close the equations are. In order to calculate values of θ for the 24 water / solvent equations, one equation has to be taken as the standard and the first equation in Table 7.4 and Table 7.5 was chosen as the arbitrary standard. The values of θ are given in Table 7.6. There are huge variations in θ, and so the equations in Table 7.4 as well as those in Table 7.5 are sufficiently different to be able to obtain descriptors for the refrigerants in Table 7.1.

Table 7.6. List of θ values

Water / Solvent equations	θ	Gas/ Solvent equations	θ
Gas phase	0.00	Water	0.00
Gas phase, 310 K	3.21	Water, 310 K	2.54
Olive oil, 310 K	140.33	Olive oil, 310 K	53.86
Dry DMF	119.71	Dry DMF	47.77
Dry NMP	117.22	Dry NMP	47.42
Benzene	107.49	Benzene	62.37
Tetrachloromethane	146.83	Tetrachloromethane	84.53
Dry methanol	128.39	Dry methanol	32.11
Dry ethanol	127.72	Dry ethanol	36.67
Dry propanol	127.16	Dry propanol	41.04
Dry butanol	129.35	Dry butanol	42.16
Dry pentanol	128.38	Dry pentanol	42.36
Dry hexanol	129.31	Dry hexanol	44.62
Dry heptanol	129.63	Dry heptanol	39.79
Dry decanol	129.48	Dry octanol	46.78
Dry octanol	130.74	Wet octanol	33.27
Wet octanol	128.67	Hexane	93.08
Hexane	150.38	Heptane	93.03
Heptane	149.07	Octane	92.28
Octane	149.48	Nonane	92.93

Water / Solvent equations	θ	Gas/ Solvent equations	θ
Nonane	149.73	Decane	92.91
Decane	149.98	Hexadecane	91.94
Hexadecane	149.62	Cyclohexane	95.90
Cyclohexane	148.99		

7.1.3. Methodology

A MS Excel' 97 with the function Solver was used to calculate the set of descriptors for the refrigerants listed in Table 7.1. This computer program contains the coefficients of equations (7.1) and (7.2) for several water / solvent and gas / solvent systems. The dependent variable were put into Solver and were referenced as log $P_{(obs)}$ and log $L_{(obs)}$. Calculated **V** values are also required by the program. The function Solver examines a large number of possible descriptor values and finally selects those that give the smallest standard deviation in observed and calculated dependent variable values. The standard of the estimate, sd (P+L) is defined as follows:

$$sd\,(P+L) = \sqrt{\frac{\sum(\log P_{(obs)} - \log P_{(calc)})^2 + \sum(\log L_{(obs)} - \log L_{(calc)})^2}{n-1}} \qquad (7.6)$$

where log $P_{(calc)}$ and log $L_{(calc)}$ are the calculated partition measurements and n is the number of partition measurements used in the Solver analysis.

The function Solver is a straightforward and rapid computer program, which avoids a large number of laborious calculations. However, this is only a statistical function that does not consider the physicochemical properties of the solute. Quite often the Solver results have to be adjusted to agree with the chemical features of the compound. For instance, a '**A**' value is sometimes attributed after a 'Solver analysis' to a compound that does not contain any hydrogen atom. In such a case, unfitted descriptors values, e.g. $\mathbf{A} \neq 0$, are replaced by expected descriptor values, e.g. $\mathbf{A} = 0$, which are fixed in Solver, and another statistical analysis is run. A new set of descriptor is obtained with a new standard deviation that is very often close to the one originally obtained.

In the present work of descriptor determination for a range of refrigerants, such

replacements have occurred. In one hand, **A** values were fixed through equation (7.11) and in the other hand **A** and **B** were fixed to zero. These cases are further explained in the following part, R23-trifluoromethane as an example.

Sixteen partition measurements were used to determine descriptor values for R23. **V** was calculated as described in chapter 3 section 3.2.1.2. These data were entered into Solver. A first run of this program identified outlying log P^S values: the given value of 0.993 for the gas / dry NMP system and the value of 1.503 for the water / dry NMP system were in some way out of line. As a result, these values were not included for the calculation of descriptors. A second run of the function Solver with now fourteen partition measurements led to the following results

E	-0.427	n	14
S	0.208	sd (P+L)	0.047
A	0.092	n_{logP}	7
B	0.035	sd (P)	0.027
L	-0.246	n_{logL}	7
V	0.303	sd (L)	0.064

where n_{logP} is the number of water / solvent partition measurements. Similarly, $n_{log\ L}$ is the number of gas / solvent partition measurements. sd (P) and sd (L) are the standard deviation of the estimated log P and log L, respectively. The above calculated **A** value was found smaller than the one estimated through the relationship between **A** and σ_I, see section 7.2.3. Therefore, the estimated **A** value (0.11) was fixed into Solver to finally obtain the set of descriptors:

E	-0.427	n	14
S	0.173	sd (P+L)	0.051
A	0.110	n_{logP}	7
B	0.036	sd (P)	0.051
L	-0.274	n_{logL}	7
V	0.302	sd (L)	0.066

The 'fixed' **A** value in Solver slightly increased the sd (P+L) value and did not greatly affect the determination of **E**, **S**, **B** and **L** values. All the descriptor values were in the same range before and after the substitution of the **A** value.

As previously said, the function Solver examines all the possible descriptor values that can describe the observed partition measurements and chooses those that are statistically the best. To illustrate that, several 'Solver analyses' were run with different

fixed **A** values. The sd (P+L) values obtained through these studies were then plotted against the corresponding **A** values to get a curve, as shown in Figure 7.1.

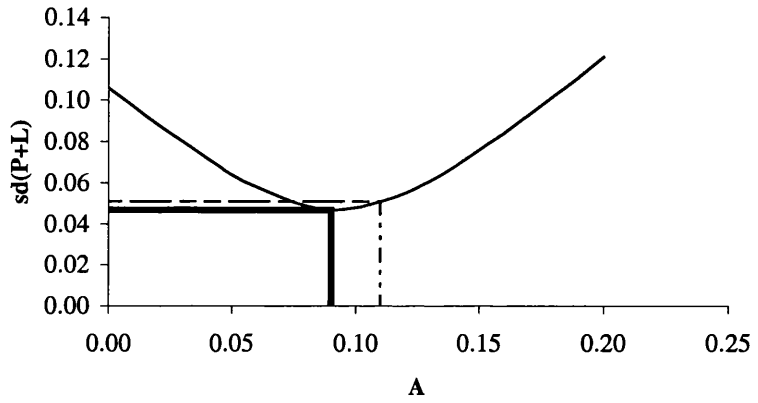


Figure 7.1. Plot of sd (P+L) against **A**

Several **A** values could lead to small standard deviation in Solver. The statistical **A** value is the one that led to the smallest standard deviation of the estimate (line in bold). The difference between the sd (P+L) values derived from the statistical and the 'chemical' (dashed line) **A** values is not significant. Further, such difference can be attributed to experimental error.

With the above descriptor values assigned to R23, the fourteen dependent variables could be reproduced with a standard deviation of 0.051 log units. The calculated values of log P^S (from equation 7.1) and log L^S (from equation 7.2) are given in Table 7.7, together with the residuals. The standard deviation is 0.051 for the seven calculated and observed log P^S values, 0.066 log units for the seven observed and calculated log L^S values. It appears clearly from Table 7.7 that the observed values for log $P^{\text{dry NMP}}$ and log $L^{\text{dry NMP}}$ are out of line.

Since enough dependent variables were obtained, descriptors for all the 18 refrigerants listed in Table 7.1 were calculated using the same approach as the one used for R23.

Table 7.7. Comparison between observed and calculated partition measurements after 'Solver analysis' for R23.

Solvent	Obs. log L^S	Calc.[a] Log L^S	Residual [b]	Obs. log P^S	Calc.[c] Log P^S	Residual [b]
Octanol wet	0.131	0.068	0.063	0.640	0.695	-0.055
Dry DMF	0.632	0.647	-0.015	1.141	1.153	-0.012
Dry octanol	0.149	0.231	-0.082	0.659	0.697	-0.038
Olive oil, 310K [7]	-0.092	-0.173	0.091	0.531	0.530	0.001
Dry NMP	0.993[d]	0.340	0.653	1.503[d]	0.897	0.606
Nonane	0.046	-0.007	0.053	0.463	0.463	0.000
Water, 298K	-0.510	-0.469	-0.041	-0.510	-0.453	-0.057
Water, 310K	-0.622	-0.685	0.063	-0.622	-0.617	-0.004

[a] Calculated from equation (7.2)
[b] Obs. log L^S - calc. log L^S, or obs. log P^S - calc. log P^S
[c] Calculated from equation (7.1)
[d] Omitted

7.2. Results and Discussion

The final list of solute descriptors for 18 organofluorocarbons, classed as refrigerants, is given in Table 7.8. It is not easy to calculate the error in the descriptors obtained by our method, because all the descriptors are calculated simultaneously. However, we can take the errors as those suggested before[25], viz.; 0.03 unit for **S** , **A** and **B** , and 0.02 unit for **L**.

The physicochemical properties of organic molecules can be dramatically affected by fluorination. The fluorine atom is the smallest halogen atom of the periodical table. This atom has a high ionisation potential, and a relatively low polarizability. These characteristics lead to weak intermolecular interactions; low surface energies and low refractive indexes for perfluorocarbons. The extreme electronegativity of fluorine atom insures that it will always be electron withdrawing inductively when bonded to carbon atom. C-F is a very polarized bond, which has a high ionic character. Further, the strength of each C-F bond increases, and it shortens, as the number of fluorine on the carbon increases. According to the fluorine atom properties, fluorinated compounds have unusual behaviour.

Table 7.8. Descriptors of the refrigerants

Code	E	S	A	B	L	V
R32	-0.316	0.487	0.065	0.052	0.040	0.2849
R23	-0.427	0.173	0.110	0.034	-0.274	0.3026
R14	-0.550	-0.250	0.000	0.000	-0.819	0.3203
R11	0.207	0.240	0.000	0.070	1.950	0.6344
R12	0.027	0.125	0.000	0.000	1.124	0.5297
R13	-0.247	-0.046	0.000	0.000	0.209	0.4250
R152a	-0.250	0.498	0.040	0.050	0.517	0.4258
R134a	-0.410	0.342	0.060	0.040	0.317	0.4612
R125	-0.510	-0.019	0.105	0.064	0.100	0.4789
R141b	0.084	0.430	0.005	0.054	1.920	0.6530
R142b	-0.080	0.240	0.060	0.056	1.081	0.5482
R133a	-0.160	0.350	0.060	0.080	1.176	0.5659
R124	-0.309	0.170	0.097	0.071	0.904	0.5860
R114	-0.190	0.050	0.000	0.000	1.427	0.7060
R115	-0.360	-0.120	0.000	0.000	0.543	0.6013
R227	-0.557	0.012	0.070	0.030	0.688	0.6552
R1122	-0.340	0.285	0.150	0.000	0.723	0.5052
R1216	-0.500	-0.166	0.000	0.100	0.337	0.5945

Inspection of the descriptors shows that the refrigerants interact with the neighbouring solvent molecules through very weak interaction forces. These forces become smaller as the number of fluorine atoms in the molecule increases. The **E** term describes the polarizability of a solute; **E** values for most of the refrigerants are small and even negative. The negative value means that such compounds have less polarizability ability than the corresponding hydrocarbons, for which **E** is zero.[26] This is in agreement with characteristics of the fluorine atom. Even though hydrogen and fluorine atoms have almost the same atomic size, the latter shields the carbon-backbone due to its electron withdrawing capability. The **S** descriptor measures the ability of a molecule to stabilise a neighbouring charge or dipole. The PFCs of the set, R14 and R1216, have negative **S** values -0.250 and -0.166, respectively. Hence R14 is less dipolar than the homologous hydrocarbon (CH_4, **S** = 0). R1216 is more polar than octafluoropropane (**S** = -0.45) due to presence of a double bond in R1216. In partially fluorinated alkanes, HFCs, the dipolar effect is more important than in PFCs because of the net C-F dipole that is absent in PFCs. It was also found that organofluorocarbons are not important hydrogen bond acids or bases. Finally, the **L** descriptor is a measure of the

dissolution of the gaseous solute into n-hexadecane. A large positive value means that the dissolution of the solute in n-hexadecane is easy, and therefore that the solute can be classed as lipophilic. A negative value for **L** suggests that the dissolution in n-hexadecane is unfavoured and that the solute has little lipophilic character. This is the case for the perfluorocarbon R14 and the hydrofluorocarbon R23. The remaining compounds have small **L** values, and hence are slightly lipophilic.

For some classes of solute, it is possible to estimate various descriptors, as now shown.

*7.2.1. Prediction of **E** -Values for n-Halogenated Alkanes*

E, the excess molar refractivity, describes the polarizability of a solute. This property correlates well with the size of atoms contained in a molecule, and also with the electron distribution. There are over 200 **E** values of n-halogenated alkanes available in our database. To these, can be added the **E** values obtained in this work. A data set of 220 **E** values was then obtained, see Table 7.210. 22 n-halogenated alkanes were selected in a random order to form the test set, to give 178 compounds as the training set. 22 compounds of the test set are listed as Nos 199-220 in Table 7.10 in section 7.4. A regression of **E** against the number of iodine, bromine, chlorine and fluorine atoms, N_I, N_{Br}, N_{Cl} and N_F respectively, in halogenated n-alkanes, yields equation 7.7.

$$\mathbf{E} = 0.633\ (\pm 0.014)\ N_I + 0.326\ (\pm 0.005)\ N_{Br} + 0.139\ (\pm 0.005)\ N_{Cl} - 0.098\ (\pm 0.002)\ N_F \qquad (7.7)$$

n = 178, s.d = 0.082

Here and elsewhere, n is the number of data points, and sd is the standard deviation in the independent variable. The sd values for the coefficient are given in the parentheses. Because the intercept was constraint to be zero in equation (7.7) no values of the correlation coefficient; r, or of the Fisher F-statistic; F, are given. These statistical constants have no meaning under this circumstance. As expected, the number of carbon atoms was found to be not significant. The predictive ability of the above equation (7.7) can be probe through the test set of 22 compounds. A plot of the observed **E** values

against calculated **E** values are given in Figure 7.2. The standard error of the predicted value for the 22 compounds is 0.061 over a range of 1.3 units. Finally, equation (7.8) was developed for the total of 222 compounds.

$$\mathbf{E} = 0.643\ (\pm 0.013)\ N_I + 0.328\ (\pm 0.005)\ N_{Br} + 0.140\ (\pm 0.003)\ N_{Cl} - 0.098\ (\pm 0.002)\ N_F \quad (7.8)$$

n = 220, s.d = 0.082

Regression coefficients in equation (7.8) are chemically correct. As expected, iodine, bromine and chlorine atoms influence positively the ability of any given compound to interact with neighbouring molecules through n and π electrons. These atoms are large atoms and therefore the nuclear control is less, the electron distribution is flabbier and then the polarizability is greater. On the other hand, fluorine atoms negatively influence the **E** value. Finally, Equation (7.8) could certainly be used to estimate further values of **E**.

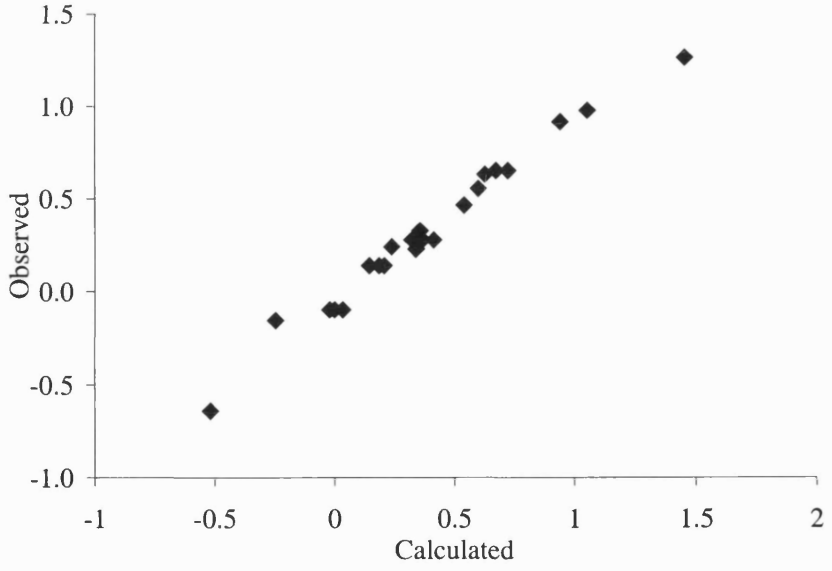


Figure 7.2. Plot of Observed versus calculated **E** values for the test set compounds

7.2.2. Analyses of **S***-Values*

The **S** descriptor is a blend of polarisability and dipolarity, therefore in order to estimate **S** values for chlorinated and fluorinated n-alkanes (with n<3), it is useful to dissect **S** values into contributions from these two factors. As previously pointed out in chapter 3, Abraham[27] proposed the estimation of new **S** values for chlorinated benzenes through the use of the dipole moment, μ, and the number of chlorine atoms, N_{Cl}. In this work, the squareddipole moment, μ_o^2, was preferred because this constant yields better statistical results than straight dipole moment value. 39 **S** values for chlorinated and fluorinated n-alkanes were available in the in-house database, see Table 7.11 in section 7.4. Dipole moment values were taken from literature.[28] The following equation was obtained

$$\mathbf{S} = 0.108\ (0.007)\mu_o^2 + 0.108\ (0.006)\ N_{Cl} - 0.053\ (0.007)N_F \qquad (7.9)$$

n = 39, sd. = 0.091

Again, no r or F values are given because the equation was constrained to have a zero intercept. Equation (7.9) is just about good enough to use to estimate further values of **S**.

7.2.3. Estimation of **A** *-Values*

In solvation situations **A** the "effective" or "summation" solute hydrogen bonding acidity must be considered.[29] Here the solute is surrounded by solvent molecules and undergoes multiple hydrogen bonding. In a previous study, Abraham and co-workers[31] highlighted a relationship between the 1:1 hydrogen bond acidity; α_2^H, and the Hammet inductive parameter; σ_I for a few halogenated compounds, see chapter 3 section 3.2.1.8.

The theory behind the α_2^H was well presented in chapter 3. σ_I is the Hammet substituent constant characteristic of a substituent inductive effect. This parameter was derived from the well known Hammett equation. Charton proposed σ_I values for 294

substituents.[30] The Abraham approach was used to establish a correlation between **A** and σ_I for few CFCs and HFCs. The same type of relationship between **A** and σ_I was expected. **A** values calculated in this work were plotted against σ_I values recently determined by Taylor[31], see Table 7.9. Hence, it appeared that a correlation between these parameters could be drawn:

$$\Sigma\alpha_2^H = -0.0535 + 0.389\ \sigma_I \qquad (7.10)$$

$r = 0.977$, s.d = 0.03, n = 10

Such equation (7.10) can be used to estimate further values of **A** for aliphatic halogenated compounds providing that σ_I value are available.

Table 7.9. Calculated **A** and σ_I values

Code	**A**	σ_I [a]
R32	0.065	0.32
R23	0.110	0.43
R152a	0.040	0.25
R134a	0.060	0.30
R125	0.105	0.39
R141b	0.005	0.14
R142b	0.060	0.30
R133a	0.060	0.29
R124	0.097	0.37
R227	0.070	0.31

[a] From Ref. 33.

7.3. Conclusion

In this work, partition measurements taken from literature were used to determine the Abraham descriptors for a series of 18 organofluorocarbons, classed as refrigerants. Predictive equations were also developed to estimate further values of **E**, **S**

and **A** for halogenated n-alkanes. Finally, descriptor values can be used to predict partition measurement values for any water / solvent and gas / solvent systems, provided that coefficient values in equation (7.1) and (7.2) are known. More interestingly, the chemosensory properties of the organofluorocarbons can be estimated, as covered in chapter 11.

7.4. Appendix

Table 7.10. Observed and calculated **E** values

Nos	Compounds	Obs **E**	Calc **E** [a]	Obs.**E** - Calc. **E**
1	Decafluorobutane	-1.030	-0.978	-0.052
2	1-Bromoperfluoroheptane	-0.943	-1.138	0.195
3	Octafluoropropane	-0.900	-0.782	-0.117
4	1-Iodoperfluorooctane	-0.890	-1.019	0.129
5	1-Bromoperfluorohexane	-0.782	-0.943	0.161
6	1-Iodoperfluorohexane	-0.637	-0.628	-0.009
7	Hexafluoroethane	-0.610	-0.587	-0.023
8	1,1,1,2,3,3,3-Heptafluoroethane	-0.557	-0.684	0.127
9	Tetrafluoromethane	-0.550	-0.391	-0.159
10	Pentafluoroethane	-0.510	-0.489	-0.021
11	1-Iodoperfluorobutane	-0.488	-0.432	-0.056
12	Trifluoromethane	-0.427	-0.293	-0.134
13	1,1,1,2-Tetrafluoroethane	-0.410	-0.391	-0.019
14	Chloropentafluoroethane	-0.360	-0.348	-0.012
15	Difluoromethane	-0.316	-0.196	-0.120
16	1-Chloro-1,1,1,2-tetrafluoroethane	-0.286	-0.251	-0.035
17	1,1,2,2-Tetrafluoroethane	-0.280	-0.391	0.111
17	1,1-Difluoroethane	-0.250	-0.196	-0.054
19	1-Iodoperfluoropropane	-0.200	-0.041	-0.159
20	1,2-Dichlorotetrafluoroethane	-0.190	-0.110	-0.080
21	1-Chloro-1,1,2-trifluoroethane	-0.160	-0.153	-0.007
22	1,1,1-Trifluoro-2,2-dichloroethane	-0.160	-0.012	-0.148
23	1-Chloro-1,2-difluoroethane	-0.080	-0.055	-0.025
24	Fluorochloromethane	-0.080	0.043	-0.123
25	1-Bromo-1,2,2,2-tetrafluoroethane	-0.070	-0.063	-0.007
26	1-Bromo-2,2,3,3-Tetrafluoropropane	-0.070	-0.063	-0.007
27	1-Iodoperfluoroethane	-0.021	0.155	-0.176
28	1-Fluoroheptane	-0.010	-0.098	0.088
29	Difluorochloromethane	0.000	-0.055	0.055
30	Ethane	0.000	0.000	0.000
31	Methane	0.000	0.000	0.000
32	Propane	0.000	0.000	0.000
33	Butane	0.000	0.000	0.000

Nos	Compounds	Obs E	Calc E [a]	Obs.E - Calc. E
34	Pentane	0.000	0.000	0.000
35	Hexane	0.000	0.000	0.000
36	Heptane	0.000	0.000	0.000
37	Octane	0.000	0.000	0.000
38	Nonane	0.000	0.000	0.000
39	Decane	0.000	0.000	0.000
40	Undecane	0.000	0.000	0.000
41	Dodecane	0.000	0.000	0.000
42	1-Fluoropentane	0.002	-0.098	0.100
43	2-Fluoropropane	0.004	-0.098	0.102
44	1,1,1-Trifluoro-2-chloroethane	0.010	-0.153	0.163
45	1,1,2-Trifluorotrichloroethane	0.010	0.128	-0.117
46	1,1,2-Trichlorotrifluoroethane	0.010	0.128	-0.117
47	1-Fluorobutane	0.017	-0.098	0.115
48	1,1,1-Trichlorotrifluoroethane	0.017	0.128	-0.110
49	Difluorodichloromethane	0.027	0.085	-0.058
50	1-Bromo-2,2,2-trifluoroethane	0.041	0.035	0.006
51	Fluoroethane	0.052	-0.098	0.150
52	1,2-Dibromo-1,1,2,2-tetrafluoroethane	0.062	0.266	-0.204
53	1-Bromo-2,2,2-trifluoropropane	0.063	0.035	0.028
54	Fluoromethane	0.066	-0.098	0.164
55	1-Fluoro-1,1-dichloroethane	0.084	0.173	-0.099
56	1-Bromo-1-chloro-2,2,2-trifluoroethane	0.102	0.176	-0.074
57	2-Chloro-2-methylpropane	0.142	0.140	0.002
58	2-Chloro2,3-dimethylpentane	0.146	0.140	0.006
59	2-Chloro-3-methylbutane	0.154	0.140	0.014
60	1-Chloro-2,2-dimethylpropane	0.166	0.140	0.026
61	1-Chloro-2,2-dimethylbutane	0.166	0.140	0.026
62	2-Chloro-2-methylbutane	0.172	0.140	0.032
63	1,4-Dibromo-1,1,2,2-tetrafluoropropane	0.173	0.266	-0.093
64	2-Chloropropane	0.177	0.140	0.037
65	1-Chloro-3,3-dimethylbutane	0.179	0.140	0.039
66	2-Chloropentane	0.179	0.140	0.039
67	3-Chloropentane	0.175	0.140	0.045
68	2-Chlorobutane	0.179	0.140	0.049
69	1-Chloro-2-methylpropane	0.191	0.140	0.051
70	1-Chlorooctane	0.191	0.140	0.051
71	3-Chloro-2,2-dimethyl butane	0.191	0.140	0.051
72	1-Chloroheptane	0.194	0.140	0.054
73	2-Chloro-2,3-dimethylbutane	0.196	0.140	0.056
74	2-Chloro-2,3-dimethylpentane	0.200	0.140	0.060
75	1-Chlorohexane	0.201	0.140	0.061
76	Fluorotrichloromethane	0.207	0.324	-0.117
77	Fluorodichloromethane	0.210	0.173	0.027
78	1-Chlorobutane	0.210	0.140	0.070

Nos	Compounds	Obs E	Calc E [a]	Obs.E - Calc. E
79	1-Chloropropane	0.216	0.140	0.076
80	Difluorochlorobromomethane	0.220	0.273	-0.053
81	Chloroethane	0.227	0.140	0.087
82	1,2-Difluorotetrachloroethane	0.227	0.366	-0.139
83	1,1-Difluorotetrachloroethane	0.230	0.366	-0.136
84	3-Chloro-2,2-dimethylpentane	0.234	0.140	0.094
85	1-Chloro-4-fluorobutane	0.237	0.043	0.194
86	1-Iodo-2,2,2-trifluoroethane	0.243	0.350	-0.107
87	1-Chloro-2-fluoroethane	0.245	0.043	0.202
88	Chloromethane	0.249	0.140	0.109
89	Dibromodifluoromethane	0.272	0.461	-0.179
90	1-Iodo-3,3,3-trifluoropropane	0.272	0.350	-0.078
91	1,1-Dichloro-3-methylbutane	0.287	0.281	0.006
92	2-Bromododecane	0.296	0.328	-0.032
93	1,1-Dichloroethane	0.322	0.281	0.041
94	2-Bromooctane	0.322	0.328	-0.006
95	1-Bromododecane	0.325	0.328	-0.003
96	1-Bromo-3,3-dimethylbutane	0.327	0.328	-0.001
97	1-Bromodecane	0.331	0.328	0.003
98	1-Bromo-5-fluoropentane	0.331	0.231	0.100
99	2-Bromopropane	0.332	0.328	0.004
100	2-Bromopentane	0.333	0.328	0.005
101	1-Bromononane	0.336	0.328	0.008
102	2-Bromopentane	0.338	0.328	0.010
103	1-Bromooctane	0.339	0.328	0.011
104	1-Bromo-3-methylbutane	0.342	0.328	0.014
105	1-Bromoheptane	0.343	0.328	0.015
106	2-Bromobutane	0.344	0.328	0.016
107	Fluorochlorobromomethane	0.345	0.371	-0.026
108	1,3-Dichloro-3-methylbutane	0.348	0.281	0.067
109	1-Bromohexane	0.349	0.328	0.021
110	1-Bromo-4-fluorobutane	0.351	0.231	0.120
111	3-Bromopentane	0.352	0.328	0.024
112	2,3-Dichlorobutane	0.355	0.281	0.074
113	1-Bromo-4-methylpentane	0.356	0.328	0.028
114	1-Bromo-2-methylbutane	0.358	0.328	0.030
115	1-Bromo-2-methylpentane	0.359	0.328	0.031
116	1-Bromobutane	0.360	0.328	0.032
117	1-Bromo-3-methylpentane	0.360	0.328	0.032
117	1-Bromo-1-fluoroethane	0.362	0.231	0.131
119	1,10-Dichlorodecane	0.366	0.281	0.085
120	Bromoethane	0.366	0.328	0.038
121	1-Bromopropane	0.366	0.328	0.038
122	1,1,1-Trichloroethane	0.369	0.421	-0.052
123	2,2-Dichloropropane	0.369	0.281	0.088
124	1,2-Dichlorobutane	0.369	0.281	0.088

Nos	Compounds	Obs E	Calc E [a]	Obs.E - Calc. E
125	1,2-Dichloropropane	0.371	0.281	0.090
126	3-Bromo-3-methylpentane	0.377	0.328	0.049
127	Dichloromethane	0.387	0.281	0.106
128	Bromomethane	0.399	0.328	0.071
129	1,3-Dichloropropane	0.408	0.281	0.127
130	1,4-Dichlorobutane	0.413	0.281	0.132
131	1,2-Dichloroethane	0.416	0.281	0.135
132	1,5-Dichloropentane	0.421	0.281	0.140
133	Trichloromethane	0.425	0.421	0.004
134	2,2,3-Trichlorobutane	0.453	0.421	0.032
135	Tetrachloromethane	0.458	0.562	-0.104
136	Fluorodibromomethane	0.495	0.559	-0.064
137	1,1,2-Trichloroethane	0.499	0.421	0.078
138	Dibromochlorofluoromethane	0.522	0.700	-0.178
139	1,2,3-Trichlorobutane	0.529	0.421	0.108
140	1,1,1,2-Tetrachlorethane	0.542	0.562	-0.020
141	1,2,3-Trichloropropane	0.547	0.421	0.126
142	Fluoroiodomethane	0.560	0.546	0.014
143	1-Bromo-2-chloroethane	0.569	0.469	0.100
144	1-Chloro-2-bromoethane	0.572	0.797	-0.225
145	2-Bromo-1-chlorobutane	0.573	0.469	0.104
146	1-Bromo-4-chlorobutane	0.576	0.469	0.107
147	Dichlorobromomethane	0.593	0.609	-0.016
148	1,1,2,2-Tetrachloroethane	0.595	0.562	0.033
149	1-Iodooctane	0.606	0.643	-0.037
150	1-Iodoheptane	0.608	0.643	-0.035
151	2-Iodobutane	0.610	0.643	-0.033
152	1-Iodo-3-methylbutane	0.610	0.643	-0.033
153	1-Iodohexane	0.615	0.643	-0.028
154	2-Iodo-2-methylbutane	0.615	0.643	-0.028
155	1-Iodopentane	0.621	0.643	-0.022
156	2-Iodopropane	0.622	0.643	-0.021
157	2-Iodopentane	0.622	0.643	-0.021
158	1,2,3,3-Tetrachlorobutane	0.628	0.562	0.066
159	3-Iodopentane	0.629	0.643	-0.014
160	1-Iodo-1-methylbutane	0.630	0.643	-0.013
161	1-Iodopropane	0.634	0.643	-0.009
162	Bromotrichloromethane	0.637	0.750	-0.113
163	Iodoethane	0.640	0.643	-0.003
164	Pentachloroethane	0.648	0.702	-0.054
165	1,1-Dibromoethane	0.653	0.657	-0.004
166	1,1-Dibromopropane	0.660	0.657	0.003
167	Iodomethane	0.676	0.643	0.033
168	Hexachloroethane	0.680	0.843	-0.163
169	1,1-Dibromo-2,2-dimethylpropane	0.680	0.657	0.023
170	1,3-Dibromo-2,2-dimethylpropane	0.681	0.657	0.024

Nos	Compounds	Obs E	Calc E [a]	Obs.E - Calc. E
171	1,2-Dibromobutane	0.698	0.657	0.041
172	1,2-Dibromopropane	0.710	0.657	0.053
173	Fluorotribromomethane	0.712	0.888	-0.176
174	Dibromomethane	0.714	0.657	0.057
175	1,4-Dibromopentane	0.720	0.657	0.063
176	1,5-Dibromopentane	0.725	0.657	0.068
177	1,4-Dibromobutane	0.733	0.657	0.076
178	1,2-Dibromoethane	0.747	0.657	0.090
179	1,2,2,3,4-Pentachlorobutane	0.766	0.702	0.064
170	Chlorodibromomethane	0.775	0.797	-0.022
171	Dichlorodibromomethane	0.832	0.938	-0.106
172	Chloroiodomethane	0.845	0.784	0.061
173	1,1,2,3,4,4-Hexachlorobutane	0.858	0.843	0.015
174	1,2-Dibromo-1,1-dichloroethane	0.906	0.938	-0.032
175	1,2-Dibromo-1,2-dichloroethane	0.950	0.938	0.012
176	1,1,1,2,2,3,3-Heptachloropropane	0.970	0.983	-0.013
177	Tribromomethane	0.974	0.985	-0.011
178	Iodotrichloromethane	0.982	1.065	-0.083
179	1,3,3-Tribromobutane	0.986	0.985	0.001
190	1,1,2,2,3,4,4-Heptachlorobutane	0.998	0.983	0.015
191	2,2,3-Tribromobutane	1.007	0.985	0.022
192	1,2,4-Tribromobutane	1.011	0.985	0.026
193	Tetrabromomethane	1.190	1.314	-0.124
194	1,5-Diiodopentane	1.325	1.287	0.038
195	1,4-Diiodobutane	1.340	1.287	0.053
196	1,1,2,2-Tetrabromoethane	1.343	1.314	0.029
197	1,1-Diiodoethane	1.357	1.287	0.070
198	1,2,2,3-Tetrabromobutane	1.378	1.314	0.064
199	1-Iodo-1H,1H,2H,2H-perfluorooctane	-0.519	-0.628	0.109
200	Trifluorochloromethane	-0.247	-0.153	-0.094
201	1-Fluorooctane	-0.020	-0.098	0.078
202	1-Fluorohexane	0.000	-0.098	0.098
203	1-Fluoropropane	0.034	-0.098	0.132
204	4-Chloro-2,2-dimethylpentane	0.146	0.140	0.006
205	1-Chloro-3-methylbutane	0.176	0.140	0.046
206	1-Chloropentane	0.208	0.140	0.068
207	1-Iodo-2,2,3,3-tetrafluoropropane	0.239	0.252	-0.013
208	1,1-Dichlorobutane	0.324	0.281	0.043
209	1-Bromo-1-fluoropropane	0.341	0.231	0.110
210	1-Bromopentane	0.356	0.328	0.028
211	1,3-Dichlorobutane	0.367	0.281	0.086
212	1,4-Dichlorobutane	0.413	0.281	0.132
213	Chlorobromomethane	0.541	0.469	0.072
214	1,2,2,3-Tetrachlorobutane	0.600	0.562	0.038
215	1-Iodobutane	0.628	0.643	-0.015
216	1,3-Dibromobutane	0.673	0.657	0.016

Nos	Compounds	Obs E	Calc E [a]	Obs.E - Calc. E
217	1,3-Dibromopropane	0.723	0.657	0.066
217	1,4-Difluorooctachlorobutane	0.941	0.928	0.013
219	1,2,3-Tribromobutane	1.052	0.985	0.067
220	Diiodomethane	1.453	1.287	0.166

[a] Calculated from equation (7.8).

Table 7.11. Squared dipole moment and **S** values for halogenated n-alkanes used in equation (7.9)

Nos	Solute	μ^2_o	Obs **S**	Calc **S** [a]	Obs **S** - Calc **S**
1	Hexafluoroethane	0.00	-0.430	-0.309	-0.121
2	Chloropentafluoroethane	0.16	-0.120	-0.133	0.013
3	Trifluorochloromethane	0.23	-0.046	-0.023	-0.023
4	Pentafluoroethane	2.43	-0.019	0.005	-0.024
5	Ethane	0.00	0.000	0.000	0.000
6	Methane	0.00	0.000	0.000	0.000
7	Propane	0.00	0.000	0.000	0.000
8	1,2-Dichlorotetrafluoroethane	0.31	0.050	0.041	0.009
9	Difluorodichloromethane	0.30	0.125	0.143	-0.017
10	1,1,2-Trifluorotrichloroethane	0.28	0.130	0.196	-0.066
11	1-Chloro-1,2,2,2-tetrafluoroethane	2.16	0.170	0.134	0.036
12	Trifluoromethane	2.66	0.173	0.132	0.051
13	Fluorotrichloromethane	0.22	0.240	0.292	-0.052
14	Difluorochloromethane	2.13	0.250	0.234	0.016
15	1,1,1,2-Tetrafluoroethane	4.24	0.342	0.252	0.090
16	Fluoromethane	3.42	0.350	0.317	0.032
17	Fluoroethane	3.76	0.350	0.354	-0.004
17	1-Chloro-1,1,2-trifluoroethane	3.61	0.350	0.341	0.009
19	1-Chloro-1,2-difluoroethane	4.41	0.380	0.479	-0.099
20	Fluorodichloromethane	1.66	0.400	0.341	0.059
21	1,1-Dichloro-2,2,2-trifluoroethane	1.85	0.400	0.258	0.142
22	Chloroethane	4.00	0.400	0.538	-0.138
23	1,1,1-Trichloroethane	2.89	0.410	0.632	-0.222
24	1,1-Dichloro-1-fluoroethane	4.04	0.430	0.597	-0.167
25	Chloromethane	3.50	0.430	0.484	-0.054
26	Difluoromethane	3.92	0.487	0.320	0.167
27	Trichloromethane	1.02	0.490	0.430	0.060
28	1,1-Difluoroethane	5.11	0.498	0.448	0.050
29	Dichloromethane	3.24	0.570	0.563	0.007
30	1,1,1,2-Tetrachlorethane	1.66	0.630	0.606	0.024
31	1,2-Dichloroethane	3.24	0.640	0.563	0.077
32	Pentachloroethane	0.85	0.660	0.625	0.035
33	Hexachloroethane	0.00	0.680	0.640	0.040
34	1,3-Dichloropropane	4.33	0.800	0.680	0.120

Nos	Solute	μ^2_o	Obs **S**	Calc **S** [a]	Obs **S** - Calc **S**
35	Tetrafluoromethane	0.00	-0.250	-0.206	-0.044
36	1,2-Dichloro-1,1,2,2-tetrafluoroethane	0.45	0.050	0.056	-0.006
37	Tetrachloromethane	0.00	0.380	0.427	-0.047
38	1,1-Dichloroethane	4.00	0.490	0.645	-0.155
39	1,1,2,2-Tetrachloroethane	1.66	0.760	0.606	0.154

[a] Calculated from equation (7.9).

7.5. References

1. R. E. Banks, B. E. Smart and J. C. Tatlow, Organofluorine chemistry, principles and commercial applications, Plenum press, New-York, (1994)
2. S. D'Souza, J. Aerosol. Med., 8 (1995) 513.
3. M. H. Abraham, Chem. Soc. Revs, 22 (1993) 72.
4. S. Maaβen, H. Knapp, W. Arlt, Fluid Phase Equilib., 116 (1996) 354.S. Maaβen, Dissertation (1995), TU Berlin
5. A. Reichl, Dissertation (1995), TU Berlin
6. J.A. Riddick and W. B. Bunger, Techniques of chemistry volume II Organic solvents physical properties and methods of purification, Third Edition, Wiley-Interscience, New-York, (1970)
7. E. I. Eger, J. Liu, D. D. Koblin, M. J. Laster, S. Taheri, M. J. Halsey, P. Ionescu, B. S. Chortkoff, T. Hudlicky, Anaesth. Analg., 74 (1994) 245.
8. A. Leo, The Medicinal Chemistry Project Pomona College, Claremont, CA 91711
9. E. Wilhehm, R. Battino, Chem.Rev., 74 (1973) 1.
10. Bo, R. Battino, E.Wilhehm, J.Chem.Eng.Data ,38 (1993) 611.
11. P.J. Hesse, R.Battino, P.Scharlin, E.Whilem, J.Chem.Eng.Data, 41 (1996) 195.
12. D.P. Walsh, Unpublished results
13. Y. Ishihama, N. asakawa, J. Pharm. Sci., 88 (1999) 1305.
14. M.H. Abraham, J. Andonian-Haftan, G.S. Whiting, A.Leo, R.W. Taft, J. Chem. Soc. Perkin Trans. 2, 2 (1994) 1777.
15. M.H. Abraham, J.M.R. Gola, Unpublished Results.
16. M.H. Abraham, J. Phys. Chem., 6 (1994) 1994.
17. M.H.Abraham, Unpublished Results.
18. M.H. Abraham, G.S. Whiting, P.W. Carr, H. Ouyang, J.Chem. Soc., Perkin Trans. 2, (1998) 1385.
19. M.H. Abraham, G.S. Whiting, W.J. Shuely, R.M. Doherty, Can. J. Chem., 76 (1998) 703.

20. M.H. Abraham, J.Le, W.E. Acree, Jr., P.W. Carr, J. Phys. Org. Chem., 12 (1999) 675.
21. M.H. Abraham, J.Le, W.E. Acree, Jr., Coll. Czech. Chem. Comm., 64 (1999) 1748.
22. M.H. Abraham, H.S. Chadha, G.S. Whiting, R.C. Mitchell, J. Pharm. Sci., 83 (1994) 1085.
23. M.H. Abraham, C.E. Green, J. A. Platts, unpublished work.
24. M.H. Abraham, J.Le, W.E. Acree, Jr., A.J. Dallas, in preparation.
25. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, M. Roses, E. Bosh, M.L. Diaz, J. Chem. Soc. ,Perkin Trans. 2, (1998) 2405.
26. M.H. Abraham, G. S. Whiting, W.J. Shuely, J. Chem. Soc., Perkin Trans. 2, (1990) 1451.
27. M.H. Abraham, J. Chrom., 644 (1993) 95.
28. A.L. McClellan, Tables of Experimental Dipole Moments, W.H. Freeman & Co., London, (1963).
29. M.H. Abraham, P.P. Duce, P.L. Grellier, D.V. Prior, J.J. Morris and P. J. Taylor, J. Chem. Soc., Perkin Trans. 2, (1989) 699.
30. M. Charton, Prog. Phys. Org. chem., 13 (1981) 119.
31. P.J. Taylor, Personal communication

Chapter 8 Solvation Properties of Terpenes

'It is probable that from the beginning of his history, man has been interested in the diverse and fragrant odours associated with certain plants. From such a time, then, he may be said to have taken an interest in terpenoids chemistry.' Paul de Mayo, 1959.

8.0. Introduction

8.0.1. Generality

Essential oils generally contain volatile organic compounds, VOCs, which constitute over 90% of the whole oil. This fraction includes the monoterpene, sesquiterpene and diterpene hydrocarbons and their oxygenated derivatives. Terpenes are a class of natural products or hydrocarbons having a structural relationship to isoprene, that is 2-methylbuta-1,3-diene.[1,5] The most common monoterpenes are derivatives of geraniol, the main constituent of geranium oil. Limonene composes over 90% of lemon oil; α-pinene is found in the oil of rosemary. Menthol is the principal component of peppermint oil. The sesquiterpenes include cedranes from various species of woods, caryophyllanes and humulenes. Quaianes and quaianolides are isolated from fungi, marine organisms, or plants.[4,5]

Terpenoids are oxygenated derivatives of hydrocarbons or new compounds structurally related to isoprene. They are produced by a wide variety of plants, animals and microorganisms and are often included in the class of terpenes.[6] The uses to which nature puts terpenoids can be grouped into three classes; functional, defence and communication. The commercial uses of terpenoids reflect their natural uses. Terpenoids will be found in drugs such as Taxol, structural materials such as rubber, pest control and of course, in cosmetics and fragrances.[5] Over 5000 structurally determined terpenes are known.[6]

Biological formation of terpenes occurs when two molecules of acetic acid give rise to mevalonic acid. The latter is converted into isopentenyl pyrophosphate, which can be isomerised enzymically into pentenyl pyrophosphate. Coupling of these two isomeric materials gives geranyl pyrophosphate as shown in Figure 8.1. Addition of

further molecules leads to higher terpenes. Terpenes are grouped according to the specific number of isoprene units (C_5H_8) in a molecule, e.g., monoterpenes ($C_{10}H_{16}$) contain two units; sesquiterpenes ($C_{15}H_{24}$) contain three; there are also diterpenes ($C_{20}H_{32}$), triterpenes ($C_{30}H_{48}$); and tetraterpenes ($C_{40}H_{64}$).[1-6] Figure 8.2 shows how the isoprene units and the original backbone can be traced out in a number of terpenes that are important in perfumery. Sometimes skeletal rearrangements occur which make this process more difficult and fragmentation or degradation reactions can reduce the number of carbon atoms so that empirical formula does not contain a simple multiple of five unit carbons.[5]

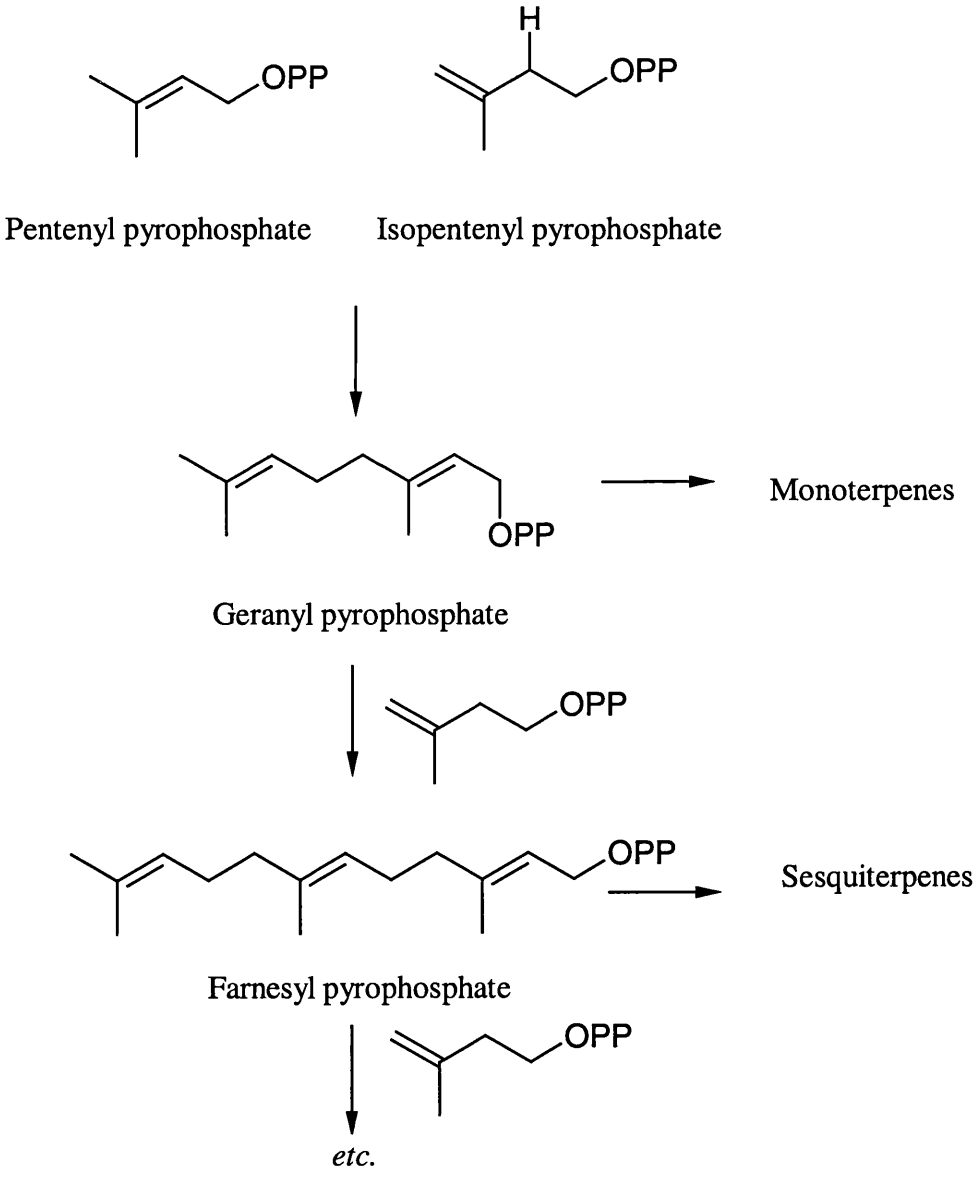


Figure 8.1. Biological formation of terpenes

Terpenes constitute a fairly large family of alcohols, hydrocarbons, ketones ... These substances may be divided into classes having zero, one, two rings ... Furthermore, the most distinctive feature of terpene might be the facility with which they undergo cyclisation and rearrangement. They also represent a framework upon which very many general chemical reactions may be studied. These include examples of the relationship of stereochemistry, stability and other physicochemical properties. For instance, menthol (1-methyl-4-isopropylcyclohexan-3-ol) is a mono-cyclic terpene which possesses three asymmetric carbon atoms and therefore exists in eight stereoisomeric forms, see Figure 8.3. Since they are pairs of diastereomers, their physical properties differ.[5] Thus, at atmospheric pressure, D,L-menthol boils at 489 K, D,L-neomenthol at 485 K, D,L-isomenthol at 491 K and D,L-neoisomenthol at 487 K.

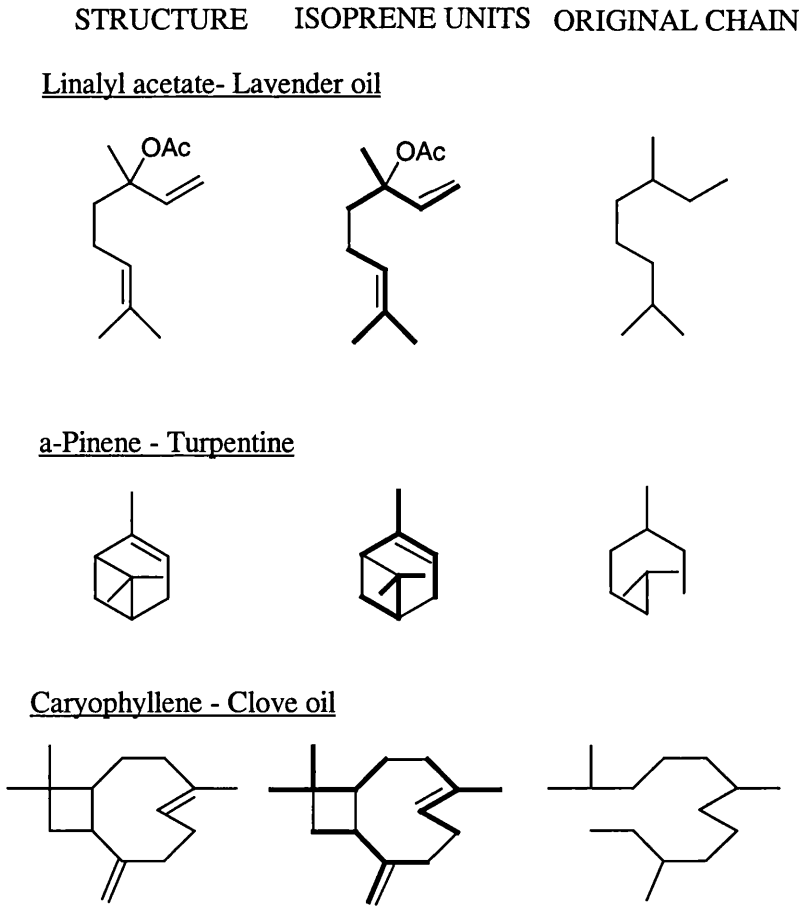


Figure 8.2. Structure, isoprene units (in bold) and original chain of three important terpenes.

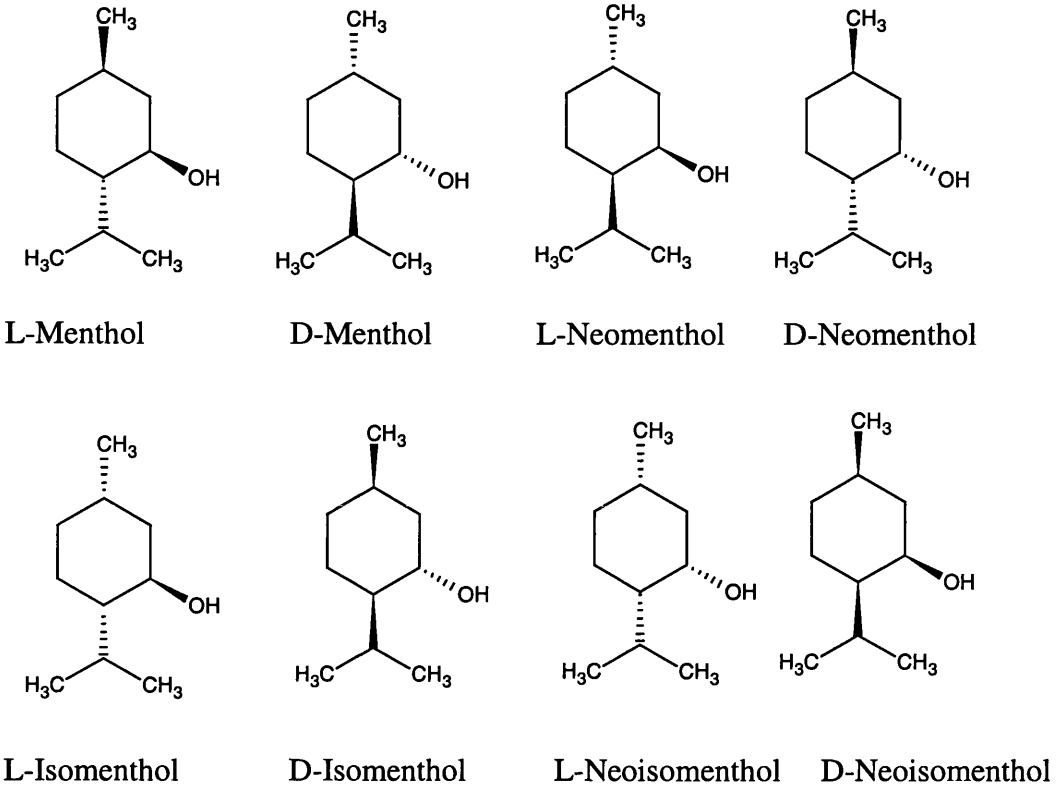


Figure 8.3. The Eight Diastereoisomers of Menthol

8.0.2. Exposure to Terpenes

Terpenes are important indoor air pollutants. Andersen estimated that 30% of the air contaminants in new homes are terpenes.[7] Exposure to terpenes may occur in various ways, including living in a home of pine wood, living in a area surrounded by pines trees, using scented cleaners, using paints and paint cleaners containing terpentines. The use of pine resins for building materials, creosote bushes and so on may also contribute to human exposure to terpenes.[4,5] Numerous studies have been carried out to assess the health impacts due to exposure to terpenes. Those spread from odor annoyance, headache and irritation to asthma and eczema.[4]

Presence of terpene compounds indoors could lead to unwanted sensory reactions.[8] Thus mouse bioassays[9] have been carried out to evaluate the sensory irritation potency and human bioassays[8] have been developed to measure olfactory and trigeminal sensitivity. Kasanen et al.[10] have recently stressed that all pinenes, i.e. L- and

D-α-pinene and L- and D-β-pinene, induced with some variance sensory irritation in mice. D-α-pinene and D-β-pinene had almost equal potency as sensory irritants and were four to five times more potent sensory irritants than L-β-pinene. L-α-pinene was found to be almost inactive. In humans exposed to D-α-pinene, a statistical significant exposure relationship was found for sensory irritation of the eyes, nose and throat at concentration of 2, 40 and 80 parts per million (ppm, by volume).[11] During exposure to 80 ppm of turpentine, consisting mainly of pinenes, the volunteers experienced discomfort in the throat and airways.[12] For olfaction, enantioselectivity of some monoterpenes is also established. Some of them such as limonene, citronellal and carvone can be discriminated by the human nose as having different odors[13] and odor thresholds[14].

Cometto-Muniz and Cain have recently measured olfactory and trigeminal sensitivity on humans to seven representative terpenes; Δ-3-carene, cumene, p-cymene, linalool, 1,8-cineole, menthol and geraniol.[8] Odor thresholds ranged from 0.1 ppm for geraniol to 1.7 ppm for Δ-3-carene, whereas nasal pungency thresholds lay about three orders of magnitude above odor thresholds, ranging from 235 ppm for cineole to 1 636 ppm for Δ-3-carene. Eye irritation fell into register with nasal pungency thresholds.[8]

8.0.3. Physicochemical Properties of Terpenes

Now the number of terpenes that could be present in indoor air is so large that it is simply not possible to measure olfactory and trigeminal sensitivity on humans for more than a small fraction. The Abraham solvation equation applied to nasal pungency thresholds (NPT) is of great help but requires the knowledge of the Abraham solvation descriptors.[15] Descriptor values for a few terpenes are available in the Abraham database.[16] Still a large number of terpenes need to be analysed.

Abraham and co-workers have recently reported the solvation properties of thirty-two terpenes.[15] To gain the values for the descriptors **S**, **A**, **B** and **L**, the authors combined gas liquid chromatographic (GLC) data, high performance liquid chromatographic (HPLC) data in the reverse phase mode and other solute physicochemical properties, such as water / octanol partition coefficient, log P^{oct}. Values for the descriptors **E** and **V** could be obtained from simple programs for any

given compounds. Data for gas to condensed phases processes, e.g. GLC data, can be used to set up solvation equations similar to equation (8.1).

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (8.1)$$

Abraham and other researchers[17-20] have shown that equation (8.2) satisfactorily correlates HPLC capacity factors, as log k', and so it is possible to use such correlative equations to determine descriptors.

$$SP = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + v.\mathbf{V} \qquad (8.2)$$

Descriptor values obtained via the above method were found to be in good agreement with those for similar compounds. Furthermore, they were obtained with a small general error in **S**, **A**, and **B** of 0.030 log units and 0.025 log units for **L**.

The above method is highly dependent upon the availability of physicochemical data for any given terpene. GLC data are very abundant, thus number of equations in the line of equation (8.1) can be developed. Conversely, data for condensed processes, such as HPLC or log P^{oct} , are scarcely reported. For instance, Fichan et al.[21] and Li et al.[22] measured log P^{oct} and water solubility for only a few monoterpenes. Griffin et al. have measured HPLC retention factors and water solubility for a large set of 57 various terpenes.[23,24] It is believed to be the largest data set ever reported. The general lack of data limits the use of equations similar to equation (8.2) in the task of descriptor determination. Therefore, difficulties in obtaining an experimental value for the **B** descriptor arise. Alternative methods are then envisaged.

In this chapter, are presented two approaches for descriptor determination. In the first section, it is explained how descriptor values are obtained from HPLC and GLC data. Then, it is shown how they are retrieved from GLC data only.

8.1. Determination of Solvation Properties from HPLC and GLC data

Griffin, Wyllie and Markham have recently measured the HPLC capacity factor, k', for 64 VOCs, 57 of which are terpenes.[23] HPLC analysis of samples was carried out using a C-18 stationary phase with at least four methanol-water ratio mixtures as eluents. The percentages of modifier, viz. methanol, used in their study were 50, 55, 60, 65, 70 and 75%, see Table 8.11 in section 8.4.

Griffin et al. studied five types of terpene structure: acyclics, 3.1.1. bicyclics, p-menthane, carenes and 2.1.1. bicyclics, see Figure 8.4. Various functionality groups were also investigated; alcohols, acetates, epoxides, ketones, and ethers. Solvation descriptors were initially available for a maximum of 37 VOCs among the 64 depending on the HPLC system investigated. Structure of most of the substances investigated are given in section 8.4.

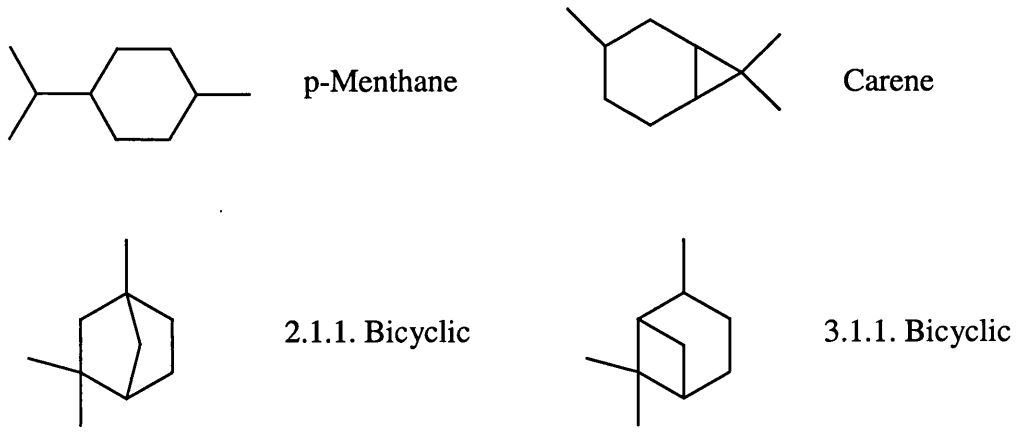


Figure 8.4. Backbone structure

8.1.1. Methodology

8.1.1.1. Development of Solvation Equations

The method adopted here was to set up solvation equations for the six HPLC systems by means of MS Excel' 97 software with the 'Data Analysis' tool. Equations for the various HPLC systems were combined with those for other processes, such as GLC processes, in order to determine descriptor values for the remaining VOCs. The set of descriptors that best describes the various processes could then be used to adjust the solvation equations for HPLC process. As explained in detail in chapter 7, regression coefficients were input data into an Excel spreadsheet together with the corresponding dependent variable. Then, descriptor values that best described the set of dependent variables input were obtained by triggering the function 'Solver'. The process was repeated in a round-robin procedure, until the descriptor values were determined for the entire set of VOCs, and until the equation coefficients for the HPLC processes had also converged.

8.1.1.2. Determination of Descriptor Values from Structure

Note that no GLC data were available for myrtanol, perilla aldehyde, perilla alcohol, dihydrocarvone, trans-p-menth-6-ene-2,8-diol and car-3-en-2-one. In such a case, values of the descriptors **S**, **A**, **B** and **L** were estimated from structure. The structures of these terpenes were compared with structure of compounds for which descriptor values are already known because the closer the structure, the closer the descriptor values. Then the consistency of the estimated descriptors was tested through the use of the HPLC equations. When necessary, descriptor values were adjusted in order to obtain better agreement between calculated and observed HPLC data. Although this approach gave good results for most of the terpenoids listed above, problems were encountered with myrtanylamine, another terpene for which no GLC data were available. Therefore, myrtanylamine was not considered any further in this work. Additional data are needed in order to determine the Abraham descriptors for this particular compound.

8.1.2. Results and Discussion

8.1.2.1. Analysis of Solvation Equations for HPLC

The system constants and statistics for the fit of the solvation equation to the experimental data are summarised in Table 8.1 for the C-18 column with 50-75% v/v methanol in water as the mobile phase. Myrtanylamine was not included in any data sets because of lack of reliable descriptor values. Furthermore, 1,8-cineole is an outlier in equations (8.3), (8.4) and (8.5). The standard deviation values for the coefficients are given in parentheses. The *e*-coefficient of the independent variable **E** was found to be not significant, and was not included in the final regression analyses. The results generally indicate a good model fit and a standard error in the estimate of the retention factor of about 0.060-0.085 log units, a figure in agreement with previous results for reversed phase liquid chromatography.[17,18] However, equation (8.8) is not as statistically sound as the other equations. This can be due to a lack of compounds in the data set, only nineteen compounds were used. It was then decided not to use this equation for the determination of descriptor values.

The coefficients summarised in Table 8.1 reflect the different properties of the mobile phase and the stationary phase. The negative *s*-coefficient indicates that the stationary phase is less dipolar than the mobile phases. The negative *a*-coefficient shows that the mobile phases are more hydrogen bond basic than the stationary phase. the large *b*-value indicates that the mobile phases are much more acidic than the stationary phase. Finally the positive *v*-coefficient shows that the stationary phase is much more hydrophobic than the mobile phase. All these results are in good agreement with the chemical structure of the mobile and stationary phases.

Variation of coefficient values with systems exits. The *s*-coefficient varies between -0.996 and -0.674, and the *v*-coefficient values oscillate between 1.841 and 1.046. The *a*- and *b*-coefficient fluctuate less. To facilitate the comparison between the equations, the regression coefficients were divided by the coefficient *v* of the volume term **V**. In Table 8.2 are collected the ratios for all the equations in Table 8.1, together with the absolute error in the ratio given in parentheses. The normalised variable were plotted against the corresponding percentage of methanol, see Figure 8.5. The ratio *a/v*, *b/v*, and *s/v* for log k' 55-70% methanol are remarkably constant, when the absolute

error for each ratio are considered. Similar results were reported by Abraham and co-workers in their previous studies on C-18 stationary phases with aqueous methanol eluents.[17,18]

Table 8.1. System constant for various mixtures of methanol in water on a C-18 stationary phase

		System Constant					Statistics				
Eq.	SP	*c*	*s*	*a*	*b*	*v*	n	r^2	sd	F	DOF
(8.3)	**log k' (50)**	**0.621**	**-0.996**	**-0.273**	**-1.411**	**1.841**	**33**	**0.977**	**0.060**	**296**	**28**
		(0.090)	(0.063)	(0.054)	(0.106)	(0.076)					
(8.4)	**log k' (55)**	**0.703**	**-0.981**	**-0.198**	**-1.399**	**1.623**	**43**	**0.950**	**0.075**	**181**	**38**
		(0.090)	(0.059)	(0.063)	(0.111)	(0.088)					
(8.5)	**log k' (60)**	**0.593**	**-0.961**	**-0.230**	**-1.333**	**1.542**	**48**	**0.951**	**0.085**	**208**	**43**
		(0.088)	(0.067)	(0.069)	(0.105)	(0.073)					
(8.6)	**log k' (65)**	**0.499**	**-0.848**	**-0.237**	**-1.555**	**1.501**	**63**	**0.974**	**0.081**	**543**	**58**
		(0.077)	(0.053)	(0.065)	(0.074)	(0.061)					
(8.7)	**log k' (70)**	**0.475**	**-0.832**	**-0.305**	**-1.358**	**1.310**	**61**	**0.975**	**0.073**	**544**	**56**
		(0.071)	(0.050)	(0.063)	(0.070)	(0.056)					
(8.8)	**log k' (75)**	**0.524**	**-0.674**	**0.000**	**-1.189**	**1.046**	**19**	**0.833**	**0.085**	**25**	**15**
		(0.210)	(0.307)	(0.000)	(0.366)	(0.190)					

Table 8.2. Ratio coefficient for Griffin et al.[23] data set.

Eq.	SP	*s/v*	*a/v*	*b/v*
(8.3)	**log k' (50)**	**-0.541**	**-0.148**	**-0.766**
		(0.041)	(-0.030)	(-0.066)
(8.4)	**log k' (55)**	**-0.604**	**-0.122**	**-0.862**
		(-0.049)	(-0.039)	(-0.083)
(8.5)	**log k' (60)**	**-0.623**	**-0.149**	**-0.864**
		(-0.053)	(-0.045)	(-0.079)
(8.6)	**log k' (65)**	**-0.565**	**-0.158**	**-1.036**
		(-0.042)	(-0.044)	(-0.065)
(8.7)	**log k' (70)**	**-0.635**	**-0.233**	**-1.037**
		(-0.047)	(-0.049)	(-0.069)

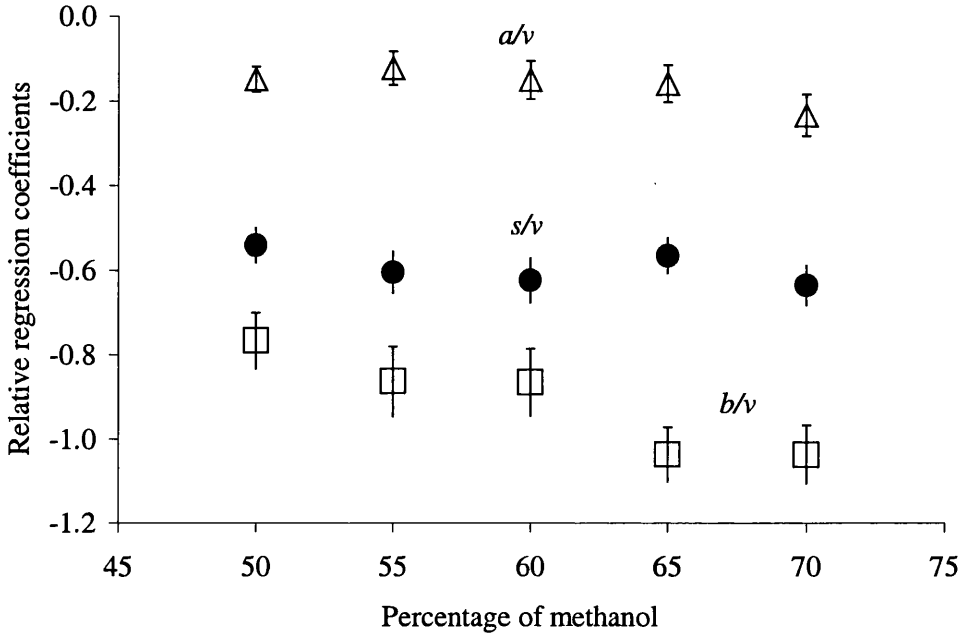


Figure 8.5. Relative regression coefficients (● *s/v*, Δ a/v, and □ b/v) against the percentage of methanol.

8.1.2.2 Solvation Properties for Terpenes

a- Methodology

The descriptor values for the thirty terpenes investigated in this section were determined according to the general procedure presented in chapter 7. The case of α-terpineol is emphasised in Table 8.3. This table lists the log k' values reported by Griffin et al.[23], together with a number of GLC data as well as other physicochemical properties that were used to determine descriptors for α-terpineol. For instance, Li and co-workers[22] measured the solubility of α-terpineol in water at 298K, log S_W = -2.337 with S_W in mol.dm^{-3}. The authors[22] also determined the vapour pressure at 298K, the gaseous concentration, C_G, was then calculated, log C_G = -5.643 with C_G in mol.dm^{-3}. Thus, combining log S_W and log C_G, the air / water partition coefficient, log L^W, was found to be 3.306. Water / octanol partition coefficient at 298K was reported in the literature as

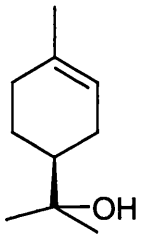
$\log P^{oct} = 2.980$.[22] Then a value of $\log L^{oct}$ was calculated, $\log L^{oct} = \log P^{oct} + \log L^{W} = 6.286$. Finally, an experimental value of the acetonitrile / hexane partition coefficient at 298K was also available, $\log P^{H/A} = -0.440$.[25] All these values are given in Table 8.3. There was a final set of 22 equations for which the system constants were available. The set of 22 equations was solved using the function 'Solver' as already presented in detail in Chapter 7. With fixed values of **E** and **V**, the values for **S**, **A**, **B** and **L** that best reproduced the 22 dependent variables were obtained. At the heading of Table 8.3, the chemically significant descriptor set that best describes the dependent variables is summarised. Calculated dependent variables are also given; they reproduce the experimental observation with an standard deviation, sd, value of 0.078.

Calculations on the above lines were carried out for the remaining compounds; results are given in Table 8.4. Noteworthy, HPLC and GLC data used in this work were generally self-consistent.

b- Analysis of Error in Descriptor Determination

Because the analysis yields the one set of descriptors that best reproduce the observations, there are no statistics as regards standard deviation in the values of the descriptors. Application of the method of 'leave-one-out' can rectify this[15]. Let's consider the case of α-terpineol as an example. Twenty-two equations were used in the calculation of the values for **S**, **A**, **B**, **L**. The first equation, equation (8.3), is left out and a set of descriptors is calculated from the remaining equations. Then the second equation, equation (8.4), is left out and a set of descriptors calculated. Each equation is left out in turn, and so 22 sets of descriptors were achieved. The statistics shown in Table (8.5) were then derived from these sets. The 'leave-one-out' values are almost exactly the same as those obtained from the entire 22 solvent set, but the internal self-consistency of the calculations can now be assess through, for example, the sd values for each individual descriptor. These sd values are very small indeed.

The 'leave-one-out' method was applied to randomly selected compounds viz. 1,8-cineole, verbenone and citronellol. For 1,8-cineole the standard deviation values are 0.003 (**S**), 0.016 (**B**) and 0.005 (**L**). In the case of verbenone, the standard deviation values for the descriptors **S**, **B** and **L** are 0.025, 0.012 and 0.013 respectively. Note that for 1,8-cineole and verbenone, the **A** value was fixed to zero as these substances do not



α–Terpineol

Table 8.3. The dependent variables for processes in Tables 8.1, 8.8 and 8.12 for α-terpineol; calculation of descriptors

E = 0.553, **S** = 0.640, **A** = 0.170, **B** = 0.600, **L** = 5.230, **V** = 1.4247, sd = 0.060

Eq.	Solvent	SP	Obs.	Calc.	Obs.	Calc.
Within condensed phase processes						
(8.3)	50%MeOH	log k'(50)	1.702	1.652	-	-
(8.4)	55%MeOH	log k'(55)	1.501	1.474	-	-
(8.5)	60%MeOH	log k'(60)	1.314	1.297	-	-
(8.6)	65%MeOH	log k'(65)	1.124	1.102	-	-
(8.7)	70%MeOH	log k'(70)	0.952	0.983	-	-
(8.10)	Water solubility	log S_W	-2.334	-2.619	-	-
(8.49)	Gas phase	log P^{air}	3.306	3.320	-	-
(8.50)	Wet Octanol	log P^{oct}	2.980	2.980	-	-
(8.52)	Hexane / Acetonitrile	log $P^{H/A}$	-0.440	-0.413	-	-
Gas to condensed phase processes						
(8.14)	OV1	I/100	-	-	11.72	11.72
(8.15)	OV101	I/100	-	-	11.85	11.83
(8.22)	SE 30-393K	I/100	-	-	11.79	11.74
(8.23)	SE 54	I/100	-	-	11.97	11.95
(8.26)	BP-5	I/100	-	-	11.99	11.93
(8.28)	DB-5	I/100	-	-	11.89	11.89
(8.30)	DDP	I/100	-	-	12.41	12.50
(8.34)	Carbowax 20M	I/100	-	-	16.84	16.84
(8.36)	Carbowax 20M	I/100	-	-	16.61	16.65
(8.38)	PEG 20M	I/100	-	-	10.82	10.81
(8.48)	HP-5MS	log $t_{R'}$	-	-	1.218	1.198
(8.53)	Water	log L^W	-	-	3.306	3.425
(8.54)	Wet Octanol	log L^{oct}	-	-	6.286	6.214

Table 8.4. Descriptors for terpenes from HPLC and GLC data;

Calculated log P^{air}, calculated log P^{16}. Observed and calculated water solubility.

	E	S	A	B	L	V	log P^{air}_{calc} [a]	log P^{16}_{calc} [b]	log L^{16}_{calc} [c]	Obs log S_W [d]	Calc log S_W [e]
3.1.1. Bicyclic											
Myrtenal	0.689	0.910	0.000	0.420	5.210	1.273	2.650	2.674	5.324	-2.581	-2.845
Myrtenol	0.643	0.700	0.300	0.450	5.224	1.316	3.340	1.951	5.291	-2.238	-2.672
Verbenone	0.640	1.000	0.000	0.620	5.197	1.273	3.819	1.522	5.341	-1.278	-2.119
Myrtanol	0.599	0.600	0.350	0.550	5.290	1.359	3.697	1.607	5.305	-2.469	-2.467
Verbenol	0.611	0.590	0.220	0.620	5.134	1.316	3.559	1.567	5.126	-2.314	-2.217
p-Menthane derivatives											
Carveol	0.655	0.680	0.240	0.590	5.360	1.382	3.688	1.816	5.504	-1.978	-2.527
Perilla aldehyde	0.725	0.850	0.000	0.460	5.540	1.339	2.654	2.892	5.546	-2.634	-2.988
Isomenthol	0.366	0.530	0.220	0.600	5.246	1.468	3.036	2.270	5.306	-2.774	-2.569
Perilla alcohol	0.674	0.710	0.210	0.550	5.550	1.382	3.467	2.084	5.551	-2.303	-2.681
1,4-Cineole	0.365	0.290	0.000	0.690	4.633	1.359	2.115	2.527	4.642	-2.411	-2.140
1,8-Cineole	0.378	0.340	0.000	0.750	4.674	1.359	2.541	2.162	4.703	-2.231	-1.941
α-Terpineol	0.553	0.640	0.170	0.600	5.230	1.425	3.320	2.157	5.477	-1.926	-2.619
α-Terpinolene	0.593	0.310	0.000	0.150	4.984	1.323	-0.285	5.116	4.831	-4.501	-3.989
Carvacrol	0.824	0.790	0.540	0.360	5.560	1.339	4.134	1.604	5.737	-2.258	-2.954
Dihydrocarveol	0.531	0.600	0.220	0.640	5.300	1.425	3.541	1.881	5.422	-2.164	-2.437

	E	S	A	B	L	V	log P^{air}_{calc} [a]	log P^{16}_{calc} [b]	log L^{16}_{calc} [c]	Obs log S_W [d]	Calc log S_W [e]
Limonene oxide	0.443	0.580	0.000	0.530	4.856	1.316	2.162	2.698	4.860	-2.640	-2.500
Terpinen-4-ol	0.531	0.455	0.185	0.678	5.285	1.425	3.221	2.058	5.279	-2.015	-2.389
Dihydrocarvone	0.475	0.750	0.000	0.640	5.292	1.382	3.089	2.201	5.290	-2.332	-2.338
Piperitone	0.559	0.890	0.000	0.640	5.485	1.382	3.495	2.030	5.524	-1.957	-2.373
Pulegone	0.587	0.785	0.000	0.560	5.495	1.382	2.856	2.608	5.464	-2.312	-2.693
t-p-Menth-6-ene-2,8-diol	0.776	1.000	0.650	0.900	6.500	1.483	7.549	-1.152	6.397	-1.021	-1.743
Carenes											
Car-2-ene	0.511	0.220	0.000	0.110	4.412	1.257	-0.699	5.111	4.412	-5.020	-3.851
Car-3-en-2-one	0.626	1.000	0.000	0.620	5.290	1.273	3.811	1.512	5.324	-1.602	-2.105
2.1.1. Bicyclic											
Fenchol	0.688	0.420	0.220	0.570	5.020	1.359	2.891	2.327	5.218	-2.264	-2.642
Isoborneol	0.650	0.520	0.220	0.680	5.150	1.359	3.656	1.605	5.261	-2.662	-2.235
Fenchone	0.427	0.600	0.000	0.570	4.820	1.316	2.397	2.461	4.858	-2.128	-2.349
Acyclic											
Citronellol	0.351	0.440	0.360	0.600	5.432	1.533	3.275	2.194	5.469	-2.844	-2.706
Others											
Isoeugenol	1.172	1.000	0.250	0.500	6.200	1.354	4.428	1.925	6.353	-2.180	-3.096
Methyleugenol	0.953	0.800	0.000	0.610	6.310	1.495	3.249	3.088	6.336	-2.526	-3.253
Linalool oxide	0.326	0.300	0.200	1.170	4.855	1.483	5.096	-0.019	5.077	-0.668	-0.904

a) Calculated on equation (8.49). b) Calculated on equation (8.51). c) Calculated on equation (8.11). d) From Ref. 6.
e) Calculated on equation (8.10).

have any hydrogen bond donor ability. The standard deviation values for citronellol are somewhat higher than those for 1,8-cineole and verbenone; 0.06 (**S**), 0.04 (**A**), 0.02 (**B**) and 0.02 (**L**). The standard deviation values obtained in this work were in line with those obtained by Abraham and co-workers[15]. The authors suggested an estimate of 0.03 as the general error in **S**, **A** and **B**, and 0.025 as the general error in **L**.

Table 8.5. Calculation of solvation descriptors by the method of 'leave-one-out' for α-terpineol

Descriptor	S	A	B	L
Max value [a]	0.618	0.140	0.596	5.221
Min value [a]	0.689	0.180	0.615	5.248
Mean value [a]	0.634	0.174	0.608	5.237
sd [a]	0.014	0.007	0.003	0.009
All equations [b]	0.640	0.170	0.600	5.230

[a] By the 'leave-one-out' method.
[b] All 22 equations used.

c- Descriptor Consistency

Descriptor values obtained for the thirty terpenes were input data in the predictive solvation equations for water / gas phase partition process, P^{gas}, and water / hexadecane partition process, P^{16}. The coefficients for these equations are given in Table 8.12 in section 8.4. Now, since log L^{16} values are related to log P^{gas} and log P^{16}, according to equation (8.9), it was then possible to calculate log L^{16} values for the set of terpenes studied.

$$\log L^{16} = \log P^{16} + \log P^{gas} \qquad (8.9)$$

These calculated log L^{16} are highly dependent upon the values allocated to the descriptors **E**, **S**, **A**, **B**, and **V**. An easy way to assess for the reliance on these descriptors is to compare log L^{16} values calculated on equation (8.9), $\mathbf{L}_{calc}$, with log L^{16} values determined through the 'Solver' analysis, **L**. A plot of **L** against $\mathbf{L}_{calc}$ values is shown in Figure 8.6. A reasonable agreement between the two sets of data was observed. This indicates that the values for the descriptors **E**, **S**, **A**, **B**, **V** and **L** are in good harmony.

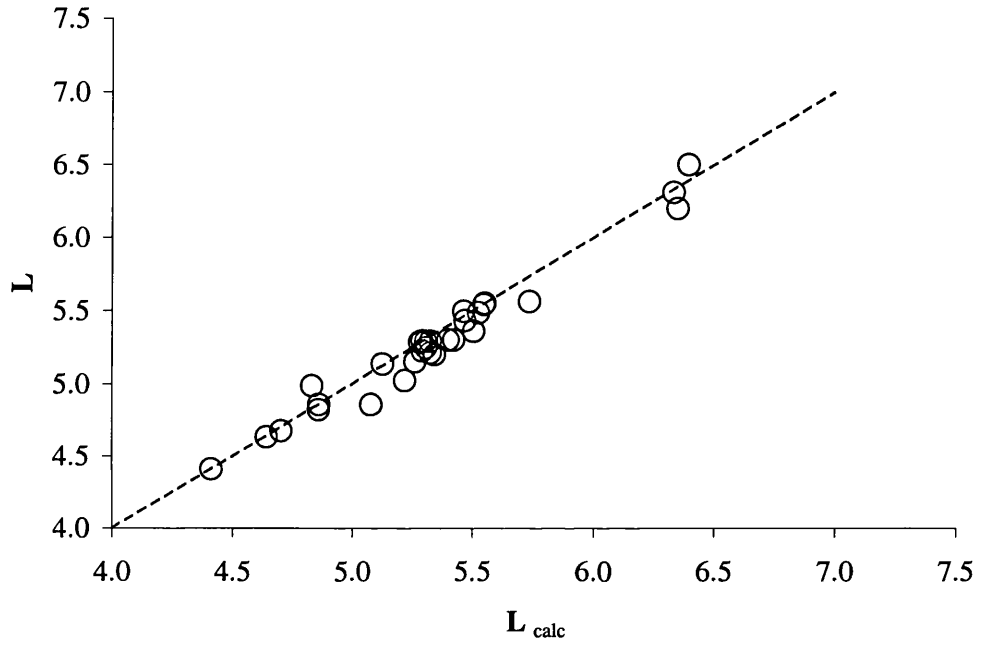


Figure 8.6. Plot of **L** values against $\mathbf{L}_{calc}$ values calculated on equation (8.9).
--- Identity line

d- Descriptor Values Analysis

The determined descriptor values were compared with those for compounds with similar functionality. In general a reasonable agreement was observed. For instance, solvation properties of α-terpinolene and limonene, another monocyclic di-alkene, are highly comparable. The unsaturated cyclic ketone, fenchone, has a similar **B** value to cyclohexanone (**B** = 0.56) but is less dipolar / polarisable. Isoborneol and isomenthol have the same properties as their corresponding isomers, borneol (**S** = 0.52, **A** = 0.28, **B** = 0.68) and menthol (**S** = 0.50, **A** = 0.23, **B** = 0.58).

Some revealing trends are present, such as the decreased acidity found on moving from primary (myrtanol **A** = 0.35, citronellol **A** = 0.36) to secondary

(dihydrocarveol **A** = 0.22) and tertiary (α-terpineol **A** = 0.17) alcohols. A departure from the usual descriptor values for cyclic ether also appears. The basicity of 1,4-cineole (**B** = 0.69) and 1,8-cineole (**B** = 0.76) is much larger than the one for tetrahydropyran (**B** = 0.40). Conversely, the latter is more dipolar than the two terpenes. It is quite surprising result, such large **B** values for cyclic ether have never been measured before.

e- Comparison between Calculated and Observed Descriptor Values.

The group contribution method for the estimation of the Abraham descriptors **E**, **S**, **A**, **B**, and **L** developed by Platts and co-workers[26] was applied to the thirty substances investigated in this work. Thus, SMILES string of each studied compound was entered in the program (UNIX version). Results are given in Table 8.13 in section 8.4.

The largest standard error between calculated and observed values (0.137) was found for **L**, purely a result of the much larger range of values covered here. The **S** and **B** descriptors have slightly smaller standard error values, 0.103 and 0.107 respectively. The **E** descriptor has a standard error equal to 0.08. The error in **A** is only 0.05, though only compounds with no zero values were considered, i.e. sixteen terpenes instead of thirty. It should be pointed out that individual molecules might have considerably larger errors than the standard error values. For instance, the **E** value for linalool oxide is in error by 0.300, the **S** value for isoeugenol is in error by 0.26 and an error of 0.25 in the **B** value is observed for 1,8-cineole. Note that some observed E values were obtained through the use of an estimated refractive index. The large standard error for **L** is due mainly to the large error in the t-p-menth-6-ene-2,8-diol value for which the individual error is 0.640. This can be due to the fact that there are not many compounds with the functionality 'diols' in the database used to develop the calculation method. Despite the very small standard error in **A**, calculated **A** values are almost always larger than the observed ones. No difference in hydrogen bond acidity for primary, secondary and tertiary alcohols are noticeable in calculated **A** values though there is a clear distinction in observed **A** values. In general the agreement between calculated and observed descriptor values is reasonable.

f- Estimation of Descriptor **B** from Hydrogen Bond group Constants

Abraham and Platt[27] have recently put forward an attractive method to predict values for **B** by means of hydrogen bond structural group constants, s**B**. These constants effect the propensity of functional groups to act as hydrogen bond bases and have the property of additivity providing that the functional groups are not interacting. Hydrogen bond structural group constants have been calculated on aliphatic series and aromatic series among others as follows. Because **B** for aliphatic compounds is zero, the basicity structural constant, s**B**, is taken as **B** (s**B** = **B**); but since **B** for benzene is 0.14 unit, the structural constant for aromatic compounds is taken as s**B** = **B** – 0.14. Analyses of the in-house database led the authors[27] to produce an extensive list of these functional group constants. Some of these values are reported in Table 8.13 in section 8.4.

This approach was utilised to predict values of **B** for most of the terpenes investigated in the present work, see Table 8.5 for the results. The omitted terpenes were myrtanol, perilla aldehyde, perilla alcohol, dihydrocarvone, trans-p-menth-6-ene-2,8-diol and car-3-en-2-one because their **B** values was obtained from stucture. Linalool oxide and isoeugenol were also omitted because their functional groups, see structures in section 8.4, can interact through hydrogen bonding interaction. Some functional group constants were derived from the in-house database and added to the already important list of Abraham and Platts[27], see Table 8.14 in section 8.4. The case of α-terpineol is explained. α-terpineol contains a tertiary alcohol and a closed double bond on a ring, so that **B** for α-terpineol can be taken as **B** = s**B** (tertiary alcohol) + s**B** (closed double bond on a ring) = 0.60 + 0.10 = 0.70.

With the exception of 1,4-cineole and 1,8-cineole for which observed values much larger than the calculated ones, calculated and observed values of **B** are generally in good agreement. Excluding 1,4-cineole and 1,8-cineole, the standard deviation between observed and calculated values of **B** was found to be 0.07 units. Even though the amount of error is larger than the general error in **B** determination, sd = 0.03 units see part b of this section, this is still a rather small error. Thus, it can be concluded that the method for determining value of B from functional group can be applied to terpene compounds.

Table 8.6. Observed and calculated values for B for terpenes using the structural constants in Table 8.13.

Terpenes	Obs. B	Calc. B	Terpenes	Obs. B	Calc. B
Myrtenal	0.42	0.50	Dihydrocarveol	0.64	0.67
Myrtenol	0.45	0.48	Limonene Oxide	0.53	0.45
Verbenone	0.62	0.50	Terpinen-4-ol	0.68	0.60
Verbenol	0.62	0.48	Piperitone	0.64	0.51
Carveol	0.59	0.64	Pulegone	0.56	0.51
Isomenthol	0.60	0.56	Car-2-ene	0.11	0.10
1,4-Cineole	0.69	0.40	Fenchol	0.57	0.60
1,8-Cineole	0.75	0.40	Isoborneol	0.60	0.60
α-Terpineol	0.60	0.70	Fenchone	0.57	0.52
α-Terpinolene	0.15	0.15	Citronellol	0.60	0.58
Carvacrol	0.36	0.30	Methyleugenol	0.61	0.48

g- Prediction of Physicochemical Properties for Terpenes using the Abraham Descriptors

Once descriptor values are known, they can be used to predict any physicochemical property for which a solvation equation has been set up. Interestingly, Griffin and al.[24] have measured the solubility in water, log S_W, of the thirty terpenes investigated. These experimental values can be compared with the ones calculated on the solvation equation for the prediction of solute water solubility,[28] equation (8.10). The Abraham descriptors in Table 8.4 are input data in equation (8.10), observed and calculated log S_W values are also given in Table 8.4.

$$\log S_W = 0.394 - 0.954\,\mathbf{E} + 0.318\,\mathbf{S} + 1.157\,\mathbf{A} + 3.255\,\mathbf{B} - 0.786\,\mathbf{A*B} - 3.329\,\mathbf{V} \qquad (8.10)$$

$n = 1071$, $r^2 = 0.880$, $sd = 0.671$, $F = 1401$

where **A*B** describes hydrogen bond acid and base interaction in the solid. The standard error between observed and calculated log S_W values, was found to be 0.453 log units.

8.2. Determination of Solvation Properties from GLC Data

Determination of solvation properties from GLC data is a common approach that has already led to the characterisation of a large number of solutes.[15] Here, this method is applied to the determination of solute descriptors for 84 terpenes. Retention properties of numerous terpenes can be found in the literature.[25,29-51] Retention data such as net retention time, $t_{R'}$, the retention index, I, and a relative gas / stationary phase partition coefficient , L', were obtained from a large number of literature sources. They were generally measured using open tubular columns under temperature-programmed conditions, except data in references 39 and 49 that are isothermal retention indices. These retention data can be used to develop solvation equations on the lines of equation (8.1) with no *b*-coefficient because no hydrogen bond acid stationary phase is available on the market. Once the solvation equations are obtained, they can be used to determine solvation properties of terpenes.

8.2.1. Development of Solvation Equations for GLC Data

8.2.1.1. Stationary Phases Properties

Stationary phases with a poly(siloxane) or poly(ethyleneglycols) backbone structure are stationary phases on which terpenes are commonly analysed. It was therefore not surprising to find a large number of retention data on such stationary phases in the literature. Composition and commercial name of all the stationary phases considered in this work are given on Table 8.7. The phases were of various physicochemical properties. First, the majority of them were inert, non-polar and non-hydrogen bond donor stationary phases, e.g. poly(methylsiloxane) and (5%-phenyl) methylpolysiloxane. The latter is similar to poly(methylsiloxane) except that 5% of the methyl group bonded to the siloxane backbone are substituted with phenyl groups. Though the phenyl contribution makes it slightly more polar than poly(methylsiloxane), this phase is still considered to be a relatively non-polar stationary phase. Secondly, some phases were of medium polarity, e.g. di-n-decylphthalate, and DB-1701. The DB-1701 polymer is a (14%cyanopropylphenyl)methyl-polysiloxane; in other words, 14% of the backbone silicon atoms have no dimethyl functionalities. Instead, the two methyl

groups are replaced by one cyanopropyl group and one phenyl group. Finally, a large number of data were measured on highly dipolar phases, such as polyethylene glycols.

Table 8.7. Name and composition of stationary phases

Commercial Name / Abbreviations	Cross-Linked Stationary Phases
DB-1, HP-1, SE-30, CP-Sil 5, BP-1	100% Poly(dimethylsiloxane)
OV1	100% Poly(dimethylsiloxane) gum
OV101	100% Poly(dimethylsiloxane) fluid
SF 96	100% (methylsilicone)
BP-5, HP-5, DB-5, HP-5MS, SE-54	5% phenyl goups poly(methylphenylsiloxane)
DB-1701	14% Cyanopropylphenyl groups poly(cyanopropylphenylmethylsiloxane)
Carbowax 20M, PEG 20M, DB-WAX, CP-WAX 52 CB, Carbowax 1540, BP-20	Polyethylene glycol
DDP	Di-*n*-decylphthalate
Zonyl E-7	Fluoroalkyl ester of pyromellitic acid
DEGS	Diethylene glycol succinate
TCEP	1,2,3-Tris(2-cyanoethoxy)propane
PPE	Polyphenyl ether 6 rings
Octakis	6-O-methyl-2,3-di-O-pentyl-γ-Cyclodextrin

8.2.1.2. Retention Data

Temperature programmed indices, I, were reported in the majority of articles reviewed. Most of these I values[25,29-42,44-51] were measured on various stationary phases with n-alkanes standard in the usual way. For the PEG phase in reference 32, I values were calculated with methyl esters as the standards. Some authors such as Allahverdiev et al.[50] and Koukos et al.[51] reported relative retention time, $t_{R'}$.

Laffort and co-workers[49] published I values at 393K for 240 substances on five stationary phases, Zonyl E 7, carbowax 1540, TCEP, PPE, and DEGS. Abraham and co-workers derived a relative gas / stationary phase partition coefficient, L', from these I values according to equation (8.11).

$$\log L' = \log L^{dec} + \left(\frac{I-1000}{100}\right) b_T^{st.ph.} \qquad (8.11)$$

where L^{dec} is the gas / stationary phase partition coefficient for n-decane and $b_T^{st.ph.}$ is the factor of variation of log $t_{R'}$ with carbon number for n-alkanes. Note that for this particular example, the logarithm of gas / stationary phase partition coefficient, log L' was used throughout instead of the retention index I.

8.2.1.3. Solvation Equations for GLC Data

In the present work, thirty-one solvation equations were developed using an in-house software using Smart Ware II (Informix software, Inc., Kansas.). These equations are equations (8.12)-(8.29), (8.31)-(8.40) and (8.46)-(8.48) in Table 8.7, The solute descriptors used in the various solvation equations were from the same in-house software. Regression coefficients in equations (8.30), (8.41)-(8.45) in Table 8.8 were previously obtained by Abraham and co-workers.[52] The regression coefficients obtained are listed in Table 8.8 together with their statistics. In general, the regression equations are all statistically sound, with r values between 0.999 and 0.970.

Regression coefficient values depends upon the physicochemical properties of the phases. For instance, large value for the *s*-coefficient will identify phase of high polarity / polarisability ability and small value for the *a*-coefficient will characterise a phase of weak hydrogen bond basicity.

An analysis of the regression coefficients in equations (8.12)-(8.40) in Table 8.8 was carried out to check whether they agree with the chemical structure of the phases. Stationary phases of low, medium and high polarity were used in this work. The method of Ishihama and Asakawa,[53] presented in Chapter 7 was utilised to classify the equations. The approach is simply based on the production of an angle θ whose value shows how close the equations are. Equation (8.12) for which the I values were measured on a non-polar column was taken has the standard. The histogram in Figure 8.7. provides a convenient way to visualise the similarity or difference between the regression coefficients in equations (8.12)-(8.40). Thus, equations (8.12)-(8.24) describe phases of low polarity and equations (8.32)-(8.40) characterise phases of high polarity. The difference between phases of low and medium polarity is less noticeable. It can be concluded that regression coefficients in equations (8.12)-(8.40) in Table 8.8 are in good agreement with the chemical structure of the phase.

Table 8.8. Solvation equations for GLC processes

Ref.	Eq.	SYSTEM	SP	c	e	s	a	l	n	r	SD	F
29	(8.12)	HP-1	I/100	**0.442**	**0.258**	**0.937**	**0.621**	**2.015**	**50**	**0.997**	**0.224**	**2697**
				0.126	*0.097*	*0.110*	*0.254*	*0.021*				
30-31	(8.13)	HP-1	I/100	**0.839**	**0.000**	**0.649**	**1.255**	**1.968**	**84**	**0.999**	**0.127**	**21475**
				0.058	*0.000*	*0.044*	*0.067*	*0.008*				
32	(8.14)	OV1	I/100	**0.632**	**-0.424**	**1.087**	**0.964**	**2.001**	**167**	**0.999**	**0.138**	**25194**
				0.064	*0.042*	*0.044*	*0.079*	*0.006*				
33	(8.15)	OV 101	I/100	**0.538**	**-0.224**	**1.011**	**1.009**	**2.026**	**354**	**0.998**	**0.017**	**26427**
				0.035	*0.037*	*0.041*	*0.060*	*0.006*				
34	(8.16)	OV101	I/100	**0.711**	**-0.244**	**0.991**	**0.581**	**1.981**	**39**	**0.999**	**0.080**	**15447**
				0.058	*0.063*	*0.064*	*0.139*	*0.008*				
35	(8.17)	CP-Sil 5	I/100	**0.658**	**-0.396**	**1.035**	**1.199**	**1.988**	**48**	**0.999**	**0.105**	**5668**
				0.073	*0.068*	*0.072*	*0.141*	*0.014*				
36	(8.18)	BP-1	I/100	**0.784**	**-0.203**	**0.968**	**0.744**	**1.971**	**38**	**0.999**	**0.060**	**18448**
				0.041	*0.047*	*0.059*	*0.133*	*0.007*				
37	(8.19)	DB-1	I/100	**0.168**	**0.000**	**1.294**	**1.468**	**2.052**	**54**	**0.990**	**0.347**	**820.2**
				0.277	*0.000*	*0.204*	*0.221*	*0.044*				
38	(8.20)	DB-1	I/100	**0.642**	**-0.212**	**0.916**	**1.124**	**1.996**	**80**	**0.999**	**0.084**	**34516**
				0.037	*0.052*	*0.038*	*0.047*	*0.006*				
39	(8.21)	SE 30-373K	I/100	**0.849**	**-0.059**	**0.805**	**0.601**	**1.960**	**38**	**0.997**	**0.111**	**1472**
				0.115	*0.077*	*0.064*	*0.132*	*0.027*				
39	(8.22)	SE 30-393K	I/100	**1.128**	**-0.389**	**1.098**	**0.955**	**1.910**	**44**	**0.995**	**0.108**	**979**
				0.174	0.091	0.097	0.114	0.032				
32	(8.23)	SE-54	I/100	**0.621**	**-0.377**	**1.423**	**0.898**	**2.003**	**173**	**0.999**	**0.134**	**28039**
				0.044	*0.040*	*0.043*	*0.076*	*0.006*				
40	(8.24)	SE-54	I/100	**0.636**	**0.000**	**1.251**	**0.849**	**1.987**	**95**	**0.999**	**0.017**	**25100**
				0.067	*0.000*	*0.057*	*0.091*	*0.008*				
41	(8.25)	SF-96	I/100	**0.673**	**-0.227**	**1.269**	**0.000**	**1.990**	**19**	**0.999**	**0.096**	**2463**
				0.137	*0.129*	*0.169*	*0.000*	*0.024*				
42	(8.26)	BP-5	I/100	**0.703**	**-0.724**	**1.658**	**1.101**	**1.983**	**21**	**0.999**	**0.049**	**29796**
				0.049	*0.061*	*0.056*	*0.078*	*0.065*				
43	(8.27)	BPX5	I/100	**0.617**	**0.230**	**1.157**	**1.933**	**2.005**	**28**	**0.999**	**0.109**	**2918**
				0.120	*0.102*	*0.078*	*0.189*	*0.019*				
26	(8.28)	DB-5	I/100	**0.794**	**-0.326**	**1.257**	**0.768**	**1.974**	**36**	**0.999**	**0.070**	**9008**
				0.065	*0.071*	*0.057*	*0.149*	*0.011*				
44	(8.29)	HP-5	I/100	**0.417**	**-0.652**	**1.739**	**0.705**	**2.020**	**37**	**0.999**	**0.104**	**12595**
				0.088	*0.065*	*0.079*	*0.084*	*0.011*				
45	(8.30)	DDP	I/100	**1.050**	**0.143**	**1.691**	**2.547**	**1.886**	**113**	**0.996**	**0.182**	**3306.7**
				0.073	*0.082*	*0.075*	*0.110*	*0.018*				
46	(8.31)	DB-1701	I/100	**0.773**	**-0.729**	**2.653**	**3.064**	**1.958**	**59**	**0.996**	**0.190**	**1489**
				0.125	*0.130*	*0.106*	*0.137*	*0.026*				
30	(8.32)	HP-Innowax	I/100	**0.700**	**1.321**	**5.515**	**10.636**	**1.984**	**54**	**0.999**	**0.208**	**4836**
				0.113	*0.112*	*0.118*	*0.148*	*0.015*				
36	(8.33)	BP-20	I/100	**0.468**	**0.703**	**5.777**	**8.710**	**2.021**	**38**	**0.998**	**0.200**	**2917**
				0.138	*0.157*	*0.199*	*0.444*	*0.025*				
43	(8.34)	Carb 20M	I/100	**0.334**	**0.844**	**5.989**	**9.798**	**2.023**	**33**	**0.998**	**0.285**	**1487**
				0.263	*0.294*	*0.294*	*0.329*	*0.035*				

Ref.	Eq.	SYSTEM	SP	c	e	s	a	l	n	r	SD	F
41	**(8.35)**	**Carb 20M**	**I/100**	**0.704**	**0.403**	**6.371**	**0.000**	**1.986**	**19**	**0.998**	**0.207**	**1839**
				0.211	*0.286*	*0.411*	*0.000*	*0.028*				
33	**(8.36)**	**Carb 20M**	**I/100**	**0.259**	**1.469**	**5.468**	**7.877**	**2.059**	**358**	**0.996**	**0.300**	**12006**
				0.065	0.068	0.075	0.101	0.011				
34	**(8.37)**	**PEG 20M**	**I/100**	**0.525**	**0.822**	**5.771**	**9.300**	**2.000**	**36**	**0.999**	**0.130**	**6008**
				0.101	*0.118*	*0.106*	*0.234*	*0.011*				
32	**(8.38)**	**PEG 20M**	**I/100**	**-5.585**	**1.021**	**5.653**	**8.848**	**2.047**	**165**	**0.999**	**0.206**	**12959**
				0.073	0.064	0.066	0.122	0.010				
47	**(8.39)**	**DB-WAX**	**I/100**	**0.704**	**0.032**	**5.840**	**8.992**	**1.983**	**40**	**0.999**	**0.178**	**4504**
				0.139	*0.110*	*0.170*	*0.248*	*0.017*				
48	**(8.40)**	**DB-WAX**	**I/100**	**0.610**	**1.034**	**5.618**	**8.802**	**1.995**	**40**	**0.998**	**0.236**	**2436**
				0.160	0.181	0.204	0.226	0.020				
49	**(8.41)**	**Carb 1540**	**log K**	**-1.983**	**0.259**	**1.254**	**2.022**	**0.423**	**204**	**0.997**	**0.066**	**7218**
				0.015	*0.023*	*0.020*	*0.026*	*0.004*				
49	**(8.42)**	**DEGS**	**log K**	**-1.797**	**0.348**	**1.597**	**1.874**	**0.388**	**204**	**0.997**	**0.061**	**9174**
				0.014	*0.021*	*0.019*	*0.023*	*0.003*				
49	**(8.43)**	**PPE**	**log K**	**-2.472**	**0.133**	**0.884**	**0.597**	**0.539**	**204**	**0.998**	**0.052**	**11408**
				0.012	*0.018*	*0.016*	*0.021*	*0.003*				
49	**(8.44)**	**TCEP**	**log K**	**-1.664**	**0.242**	**1.939**	**1.894**	**0.366**	**204**	**0.998**	**0.060**	**10885**
				0.014	*0.021*	*0.018*	*0.023*	*0.003*				
49	**(8.45)**	**Zonyl E 7**	**log K**	**-1.979**	**-0.441**	**1.471**	**0.771**	**0.431**	**204**	**0.996**	**0.071**	**5853**
				0.015	*0.024*	*0.022*	*0.027*	*0.004*				
50	**(8.46)**	**CP-Wax 52 CB**	**log $t_{R'}$**	**-1.129**	**-0.592**	**0.906**	**0.000**	**0.516**	**7**	**0.987**	**0.037**	**36**
				0.299	*0.461*	*0.242*	*0.000*	*0.084*				
50	**(8.47)**	**Cyclodextrin**	**log $t_{R'}$**	**-0.230**	**0.000**	**0.359**	**0.000**	**0.272**	**7**	**0.970**	**0.029**	**31**
				0.225	*0.000*	*0.119*	*0.000*	*0.051*				
51	**(8.48)**	**HP-5MS**	**log $t_{R'}$**	**0.118**	**0.000**	**0.233**	**0.000**	**0.178**	**13**	**0.995**	**0.021**	**465**
				0.040	*0.033*	*0.000*	*0.000*	*0.009*				

*Note that in order to effect a suitable weighting of equations with I as the dependent variable in equation (8.1), with equations in which log L', or log $t_{R'}$ is the dependent variable, all retention indices were divided by a factor of 100.

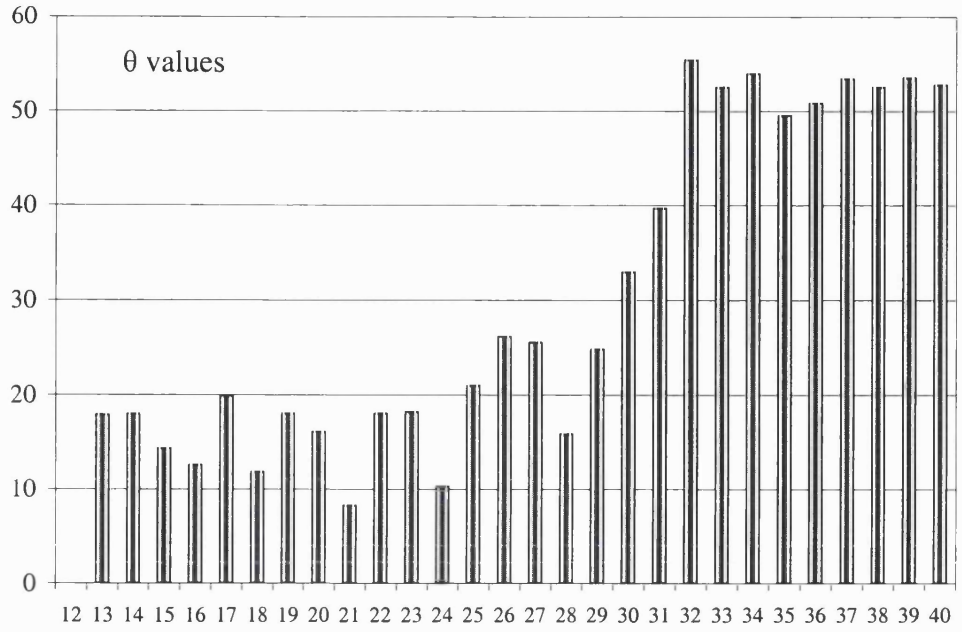


Figure 8.7. Distribution of θ values according to the phase properties. Numbers on the *x* axis refer to the equation number in Table 8.6.

8.2.2. Solvation Properties of Terpenes

The equations listed in Table 8.8 are suitable for the determination of the values for most descriptors. Not all the processes need to be used in any given cases. **E** and **V** are known for any new terpenes. The lack of a *b*-coefficient value in any of the equations means that no **B** value can be achieved by this method. Hence, only the three descriptors, **S**, **A**, and **L,** remain to be determined.

8.2.2.1. Descriptor Values for α–Copaene

The case of α-copaene is presented here. The **E** value was calculated from an experimental refractive index value at 293K, **E** = 0.663. In Table 8.9, are listed the GLC data used to obtain the descriptor values for **S**, **A** and **L**. These retention data were measured on both non-polar and polar stationary phases, this offers a good range of polarity. The set of ten equations was solved. At the heading of Table 8.9 the set of descriptors that best reproduces the ten dependent variables is supplied. Calculated

dependent variables are also given, they are reproduced with an standard deviation, sd, value of 0.073.

The value of **B** was calculated from hydrogen bond structural group constants, s**B**, as presented in section 8.1.2.2. The constants utilised in this study are given in section 8.4. Accordingly, an estimated value for **B**, **B** =s**B** = 0.100, was allocated to α-copaene.

Using a similar approach as the one for α-copaene, it was possible to determine descriptor values for 83 terpenes. Values for the **B** descriptor are estimated ones using data in Table 8.14. No analysis of error was carried out. However based on previous work by Abraham and co-workers[9], a general error in **S**, **A** of 0.03 log units and 0.025 log units in **L** was proposed. Results are given in Table 8.10. Structures for some of these terpenes are displayed in section 8.4.

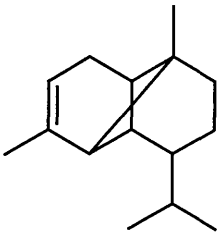
8.3. Conclusion

In this study, HPLC measurements and GLC data taken from the literature were used to determine the values of the descriptor **S**, **A**, **B** and **L** for 114 terpenes. In general The descriptor values assigned to the terpenes are not unusual, and fall within the range of values for compounds with same functionality.

Some interesting trends appear in this study. First, the hydrogen bond acidity ability decreases from primary alcohol (**A** = 0.35) to secondary (**A** = 0.22) and tertiary alcohol (**A** = 0.17). Secondly, the hydrogen bond basicity of 1,4-cineole and 1,8-cineole was found to be much larger than the usual one for cyclic ethers. For these two compounds, it would be worth to carry out some water / solvent partition measurements in order to check the reliance of their unusual ability.

It was shown that observed descriptor values are in rather good agreement with those estimated through the use of the contribution approach of Abraham and Platts[26]. Interestingly, the hydrogen bond structural group constant method of Abraham and Platt[28] can be used to estimate further **B** values.

Descriptor values can now be used to predict physicochemical properties; solubility in water was estimated for a set of thirty terpenes with an error of 0.46 log units. In chapter 11 chemosensory effects of terpenes will be calculated.



α-Copaene

Table 8.9. The dependent variables for processes in Table 8.6 for α-copaene; calculation of descriptors

E = 0.630, **S** = 0.100, **A** = 0.000, **B** = 0.100, **L** = 6.610, **V** = 1.8533, sd = 0.073

Eq.	Solvent	SP	Obs.	Calc.
(8.15)	OV101	I/100	13.98	13.89
(8.24)	SE-54	I/100	13.85	13.89
(8.25)	Sf-96	I/100	13.90	13.82
(8.27)	BPX5	I/100	14.12	14.60
(8.29)	HP-5	I/100	13.91	13.86
(8.32)	HP-Innowax	I/100	15.19	15.20
(8.36)	Carbowax 20 M	I/100	15.19	15.31
(8.37)	PEG 20	I/100	14.86	14.84
(8.40)	DB-Wax	I/100	14.51	14.60
(8.41)	HP-5MS	$\log t_{R'}$	1.354	1.318

Table 8.10. Solvation descriptors for terpenes.

Compounds	E	S	A	B	L	V
Alloaromadendrene	0.756	0.200	0.000	0.200	6.981	1.853
Anis alcohol	0.883	1.100	0.440	0.750	5.300	1.116
Anis aldehyde	1.022	1.335	0.000	0.750	5.246	1.073
α-Bisabolol	0.691	0.450	0.180	0.510	7.195	2.086
Bornyl acetate	0.397	0.550	0.000	0.600	5.844	1.657
Bornyl butanoate	0.533	0.420	0.000	0.550	6.848	1.938
Bornyl formate	0.500	0.600	0.000	0.500	5.552	1.516
Bornyl pentanoate	0.533	0.380	0.000	0.550	7.354	2.079
Bornyl propionate	0.541	0.420	0.000	0.550	6.383	1.798
β-Bourbonene	0.774	0.050	0.000	0.100	6.725	1.853
δ-Cadinene	0.650	0.350	0.000	0.200	7.139	1.919
γ-Cadinene	0.658	0.350	0.000	0.200	7.134	1.919
α-Cadinol	0.640	0.500	0.180	0.600	7.696	2.064
Calamenene	0.758	0.500	0.000	0.200	7.044	1.876
Camphene hydrate	0.597	0.450	0.190	0.550	5.232	1.457
α-Cedrene	0.709	0.150	0.000	0.100	6.786	1.853
α-Cedrene oxide	0.706	0.590	0.000	0.500	7.351	1.912
β-Cedrene	0.746	0.150	0.000	0.100	6.877	1.853
Cedrol	0.836	0.500	0.160	0.600	7.480	1.955
Citronellal	0.289	0.670	0.000	0.500	5.030	1.490
Citronellyl acetate	0.204	0.640	0.000	0.500	6.043	1.831
Citronellyl butanoate	0.160	0.580	0.000	0.500	6.925	2.113
Citronellyl formate	0.229	0.680	0.000	0.480	5.661	1.690
Citronellyl propionate	0.189	0.580	0.000	0.500	6.522	1.972
α-Copaene	0.635	0.100	0.000	0.100	6.610	1.853
α-Cubenene	0.516	0.160	0.000	0.100	6.473	1.853
Cuminaldehyde	0.859	0.900	0.000	0.450	5.449	1.296
p-Cymen-8-ol	0.837	0.760	0.270	0.590	5.101	1.340
β-Damascenone	0.776	0.700	0.000	0.600	6.300	1.578
Dihydrojasmone	0.462	0.960	0.000	0.400	6.208	1.523
Dihydrolinalool	0.334	0.382	0.250	0.420	5.023	1.533
Dihydromyrcenol	0.249	0.370	0.270	0.450	4.676	1.533
Dihydrosafrole	0.794	0.930	0.000	0.600	5.710	1.289
Ethyl tiglate	0.271	0.640	0.000	0.420	3.981	1.126
α-Fenchene	0.556	0.120	0.000	0.100	4.380	1.258
Fenchyl acetate	0.329	0.480	0.000	0.600	5.529	1.657
Geranyl butanaote	0.317	0.640	0.000	0.500	7.065	2.070

Compounds	E	S	A	B	L	V
Geranyl formate	0.419	0.760	0.000	0.480	5.755	1.647
Geranyl isobutanoate	0.322	0.560	0.000	0.500	6.898	2.070
Geranyl phenylacetate	0.876	1.040	0.000	0.550	9.210	2.396
Geranyl propionate	0.336	0.660	0.000	0.500	6.645	1.929
D-Germacrene	0.563	0.330	0.000	0.300	7.000	1.985
Globulol	0.661	0.500	0.200	0.600	7.346	1.955
α-guaiene	0.756	0.100	0.000	0.200	6.970	1.919
α-Gurgujene	0.737	0.100	0.000	0.100	6.762	1.853
Heliotropine	0.990	1.600	0.000	0.520	5.477	1.023
α-Humulene	0.772	0.170	0.000	0.350	7.022	1.985
α-Ionone	0.700	0.700	0.000	0.600	6.470	1.761
β-Ionone	0.892	0.650	0.000	0.600	6.841	1.761
cis-Jasmone	0.687	0.970	0.000	0.500	6.125	1.480
Ledol	0.661	0.540	0.180	0.600	7.384	1.955
Linalyl butanaote	0.269	0.550	0.000	0.550	6.442	2.070
Linalyl formate	0.296	0.680	0.000	0.550	5.420	1.647
Linalyl isobutanoate	0.258	0.480	0.000	0.580	6.246	2.070
Linalyl isopentanoate	0.263	0.480	0.000	0.580	6.726	2.211
Linalyl octanoate	0.253	0.470	0.000	0.550	8.392	2.683
Linalyl pentanoate	0.274	0.500	0.000	0.550	6.942	2.211
Linalyl propionate	0.277	0.550	0.000	0.550	6.033	1.929
Methyl anthranilate	1.124	1.100	0.220	0.700	5.823	1.172
Methyl tiglate	0.332	0.680	0.000	0.400	3.630	0.985
α-Muurolene	0.572	0.350	0.000	0.250	7.058	1.919
γ-Muurolene	0.658	0.350	0.000	0.200	6.982	1.919
Nerolidol	0.545	0.500	0.200	0.510	7.134	2.152
β-Ocimene	0.603	0.300	0.000	0.300	4.829	1.388
Patchouli alcohol	0.875	0.450	0.200	0.700	7.751	1.955
Phenyl ethyl tiglate	0.810	1.020	0.000	0.420	7.049	1.734
Propyl anthranilate	0.923	1.160	0.220	0.550	6.545	1.433
cis-Rose oxide	0.350	0.480	0.000	0.550	4.950	1.425
trans-Rose oxide	0.350	0.480	0.000	0.550	5.030	1.425
Sabinene	0.441	0.260	0.000	0.080	4.430	1.257
trans-Sabinene hydrate	0.822	0.400	0.160	0.550	4.830	1.359
Safrole	0.894	1.028	0.000	0.450	5.634	1.246
Santene	0.483	0.180	0.000	0.100	4.068	1.165
α-Selinene	0.823	0.280	0.000	0.200	7.109	1.533
b-Terpineol	0.563	0.600	0.200	0.700	5.066	1.4247

Compounds	E	S	A	B	L	V
Tetrahydrocitral	0.124	0.540	0.000	0.510	4.958	1.919
Tetrahydrogeraniol	0.180	0.440	0.320	0.560	5.228	1.576
Tetrahydrolinalool	0.167	0.340	0.170	0.510	4.928	1.576
Tetrahydromyrcenol	0.115	0.360	0.230	0.510	4.830	1.576
α-Thujene	0.362	0.160	0.000	0.100	4.280	1.316
Thujone	0.369	0.650	0.000	0.500	4.887	1.257
thujopsene	0.774	0.180	0.000	0.150	6.880	1.853
Tonalid	0.698	0.920	0.000	0.700	8.482	2.173
Tricyclene	0.303	0.150	0.000	0.000	4.211	1.192

8.4. Appendix

Table 8.11. Experimental HPLC capacity factors measured by Griffin et al.[23]

Solute	log k'(50)	log k'(55)	log k'(60)	log k'(65)	log k'(70)	log k'(75)
Acyclic						
Citronellal	1.994	1.739	1.524	1.308	1.080	
Citronellol	2.030	1.766	1.542	1.312	1.105	
Geraniol	1.826	1.588	1.378	1.163	0.976	
Geranyl acetate			1.915	1.653	1.424	1.178
Linalool	1.803	1.576	1.369	1.160	0.972	
Linalyl acetate			1.850	1.594	1.368	1.127
Nerol	1.783	1.567	1.350	1.131	0.954	
Neryl acetate			1.885	1.629	1.401	1.159
3.1.1. Bicyclic						
Myrtanylamine	0.868	0.752	0.581	0.352	0.187	
Myrtenal	1.503	1.303	1.127	0.954	0.800	
Myrtenol	1.638	1.420	1.235	1.044	0.873	
Verbenone	1.056	0.879	0.728	0.581	0.453	
Myrtanol		1.619	1.419	1.130	0.944	
β-Pinene				1.968	1.745	1.501
Verbenol	1.595	1.376	1.193	1.006	0.839	
(+)-α-Pinene				2.062	1.816	1.575
(-)-α-Pinene				2.062	1.810	1.575
p-Menthane						
p-Menthane						
Carveol	1.587	1.382	1.198	1.011	0.844	
(-)-Limonene				2.029	1.787	1.556
(-)-Menthol				1.265	1.040	
Menthone		1.628	1.509	1.221	1.033	
Perilla aldehyde	1.602	1.402	1.223	1.044	0.890	
Isomenthol		1.683	1.558	1.232	1.032	
(+)-Limonene				2.029	1.788	1.557
α-Terpinyl acetate			1.892	1.650	1.431	1.208
Carvone	1.404	1.240	1.068	0.897	0.747	
Perilla alcohol	1.609	1.392	1.214	1.027	0.859	
1,4-Cineole		1.561	1.422	1.220	1.043	
1,8-Cineole	1.417	1.249	1.058	0.890		
α-Terpinene				1.976	1.743	1.510
α-Terpineol	1.702	1.501	1.314	1.124	0.952	
α-Terpinolene				2.000	1.776	1.542
Carvacrol	1.642	1.437	1.239	1.051	0.885	
Dihydrocarveol	1.688	1.517	1.320	1.120	0.951	
Dihydrocarvone	1.576	1.384	1.203	1.021	0.869	
γ-Terpinene				2.012	1.770	1.538
Limonene oxide	1.662	1.465	1.290	1.112	0.937	
p-Cymene			1.952	1.783	1.514	1.296
Terpinen-4-ol	1.714	1.533	1.343	1.154	0.985	
Thymol	1.642	1.512	1.296	1.117	0.906	
Piperitone	1.418	1.211	1.048	0.876	0.729	
Pulegone	1.590	1.398	1.225	1.050	0.884	

Solute	log k'(50)	log k'(55)	log k'(60)	log k'(65)	log k'(70)	log k'(75)
t-p-menth-6-ene-2,8-diol	0.840	0.699	0.553	0.404		
Menthyl acetate				1.782	1.541	1.305
Carenes						
Car-2-ene				2.019	1.761	1.524
Car-3-en-2-one	1.170	0.981	0.833	0.677	0.538	
Car-3-ene				2.032	1.790	1.555
2.1.1. Bicyclics						
Borneol		1.508	1.364	1.080	0.896	
Camphene				1.953	1.721	1.492
Camphor	1.398	1.221	1.066	0.901	0.751	
Fenchol		1.605	1.461	1.167	0.978	
Isoborneol		1.593	1.437	1.110	0.932	
Fenchone		1.518	1.203	0.946	0.784	
Others						
Isoeugenol	1.355	1.149	0.975	0.798	0.635	
Eugenol	1.224	1.110	0.924	0.769	0.593	
Methyleugenol	1.596	1.377	1.188	1.002	0.842	
Benzyl alcohol	0.534	0.456	0.337	0.229	0.113	
Phenol	0.530	0.457	0.333	0.227	0.108	
benzaldehyde	0.729	0.648	0.519	0.409	0.295	
Benzene	1.181	1.100	0.961	0.838	0.718	
Toluene	1.498	1.402	1.232	1.091	0.950	
Cyclohexene		1.680	1.505	1.348	1.177	1.001
Linalool oxide		1.104	0.949	0.734	0.575	

Table 8.12. System constants for various processes at 298K

Eq.	Process	*c*	*e*	*s*	*a*	*b*	*v/l*	Ref
Processes within condensed phases								
(8.49)	Gas Phase	-0.994	0.577	2.549	3.813	4.841	-0.869	55
(8.50)	Wet Octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	16
(8.51)	Hexadecane	0.087	0.667	-1.617	-3.587	-4.869	4.433	55
(8.52)	Acetonitrile / hexane	-0.008	0.346	-1.600	-1.190	-0.880	0.877	56
Gas to condensed phase processes								
(8.53)	Water	-1.272	0.822	2.743	3.904	4.814	-0.231	55
(8.54)	Wet octanol	-0.222	0.088	0.701	3.478	1.477	0.851	16

Table 8.13. Calculated[a] and experimental descriptors for a series of terpenes.

Terpenes	E		S		A		B		L	
	exp.	calc.	exp.	calc.	exp.	calc.	exp.	calc.	exp.	calc.
Myrtenal	0.689	0.780	0.910	0.850	0.000	0.000	0.420	0.510	5.210	5.160
Myrtenol	0.643	0.820	0.700	0.670	0.320	0.350	0.450	0.470	5.224	5.360
Verbenone	0.640	0.790	0.950	0.820	0.000	0.000	0.600	0.500	5.264	5.130
Myrtanol	0.599	0.660	0.600	0.560	0.350	0.350	0.550	0.470	5.290	5.220
Verbenol	0.611	0.800	0.670	0.640	0.180	0.350	0.450	0.490	5.097	5.140
Carveol	0.655	0.770	0.680	0.620	0.240	0.350	0.590	0.450	5.360	5.290
Perilla aldehyde	0.725	0.750	0.850	0.830	0.000	0.000	0.460	0.470	5.540	5.310
Isomenthol	0.366	0.460	0.530	0.450	0.220	0.350	0.600	0.450	5.246	5.050
Perilla alcohol	0.674	0.780	0.710	0.660	0.210	0.350	0.550	0.430	5.550	5.520
1,4-Cineole	0.365	0.520	0.290	0.420	0.000	0.000	0.690	0.520	4.633	4.770
1,8-Cineole	0.378	0.510	0.340	0.420	0.000	0.000	0.750	0.500	4.674	4.730
α–Terpineol	0.553	0.620	0.550	0.560	0.200	0.350	0.650	0.470	5.300	5.240
α-Terpinolene	0.593	0.650	0.310	0.410	0.000	0.000	0.150	0.110	4.984	4.920
Carvacrol	0.824	0.930	0.790	0.800	0.540	0.550	0.360	0.450	5.560	5.480
Dihydrocarveol	0.531	0.610	0.600	0.500	0.220	0.350	0.640	0.450	5.300	5.140
Dihydrocarvone	0.475	0.600	0.750	0.690	0.000	0.000	0.640	0.470	5.292	5.140
Limonene oxide	0.443	0.660	0.580	0.480	0.000	0.000	0.530	0.500	4.856	4.830
Terpinen-4-ol	0.531	0.620	0.455	0.560	0.185	0.350	0.678	0.450	5.285	5.240
Piperitone	0.559	0.600	0.890	0.750	0.000	0.000	0.640	0.470	5.485	5.190
Pulegone	0.587	0.630	0.785	0.770	0.000	0.000	0.560	0.450	5.495	5.390
t-p-Menth-6-ene-2,8-diol	0.776	0.770	1.000	0.850	0.650	0.690	0.900	0.780	6.500	5.860
Car-2-ene	0.511	0.650	0.220	0.350	0.000	0.000	0.110	0.170	4.412	4.510
Car-3-en-2-one	0.626	0.790	1.000	0.820	0.000	0.000	0.620	0.500	5.290	5.130
Fenchol	0.688	0.650	0.420	0.520	0.220	0.350	0.570	0.500	5.020	5.030
Isoborneol	0.650	0.650	0.520	0.520	0.220	0.350	0.680	0.500	5.150	5.030
Fenchone	0.427	0.640	0.600	0.710	0.000	0.000	0.570	0.520	4.820	5.030
Isoeugenol	1.172	1.030	1.000	0.840	0.250	0.290	0.500	0.620	6.200	5.930
Methyleugenol	0.953	0.950	0.800	0.940	0.000	0.000	0.610	0.610	6.310	5.940
Citronellol	0.351	0.490	0.440	0.350	0.360	0.430	0.600	0.430	5.432	5.290
Linalool oxide	0.326	0.630	0.300	0.570	0.200	0.260	1.170	0.850	4.855	5.090

[a] By the method of Platts and Abraham, Ref. 26.

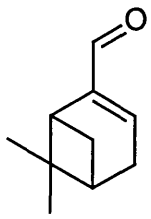
Table 8.14. Values of **sB** for functional groups, used to calculate descriptors for terpenes.

	sB
Alkanes	
Aliphatic alkane	0.00
Cyclic alkane	0.00
Alkenes	
Aliphatic alkenes	0.10
cis-Diene	0.24
trans-Diene	0.20
Monocyclic alkene-closed double bond	0.10
Monocyclic alkene-open double bond	0.10
Monocyclic Diene - two closed double bonds	0.15
Monocyclic Diene - one open and one closed double bonds	0.15
Phenyl group (C_6H_5-)	0.15
Alcohol group (-OH)	
* On aliphatic alkane	
OH (1°)	0.48
OH (2°)	0.56
OH (3°)	0.60
HO-$(CH_2)_4$-OH	0.90
* On monocyclic alkanes	
Monocyclic alkane	
OH (2°)	0.57
OH (3°)	0.50
*On dicyclic alkane	
OH (2°)	0.73
* On aliphatic alkenes	
-CH=CH-CH_2-OH	0.48
-CH=CH-CH(R)(OH)-	0.54
-CH=CH-C(R)(R)(OH)	0.51
-CH=CH-CH_2-CH(R)(OH)	0.55
-CH=CH-CH_2-C(R)(R)(OH)	0.41
*On aromatic ring	
C_6H_5-OH	0.16
C_6H_5-CH_2OH	0.42
Ether group (-O-)	
aliphatic ether	
-O-C(R)(R)(R)	0.63
-O-CH(R)(R)	0.45
-O-CH_2(R)	0.44
CH(R)(R)-O-C(R)(R)(R)	0.45
-O-C(R)=C(R)(R)	0.41
-O-CH_2-C(R)=C(R)(R)	0.45
-O-CH_3	0.49

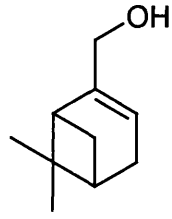
	sB
Cyclic ether	0.40
Epoxide	0.35
Ester group (-O-C(=O)-)	
-O-C(=O)H	0.38
-O-C(=O)CH_3	0.45
-O-C(=O)C_6H_5	0.46
Lactone- cyclic ester	0.51
Aldehyde group (-CH=O)	
Aliphatic	0.45
-CH=CH-C(=O)H	0.50
b position of an aromatic ring	0.64
Ketone group (-C(=O)-)	
Aliphatic	0.51
Cyclic	0.52
α position of a double bond	0.51
β position of an aromatic ring	0.52

R represents an alkyl group

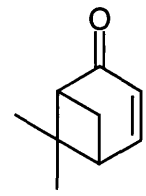
Terpenes investigated in section 8.1.



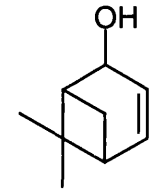
Myrtenal



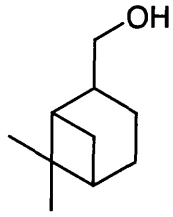
Myrtenol



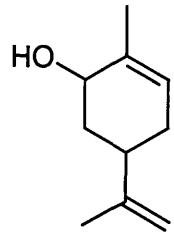
Verbenone



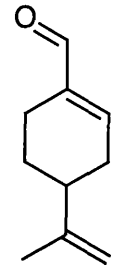
Verbenol



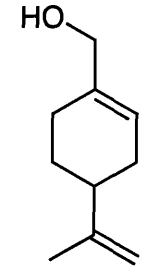
Myrtanol



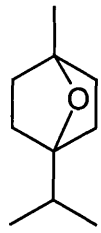
Carveol



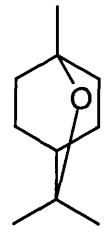
Perilla aldehyde



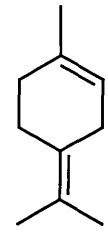
Perilla alcohol



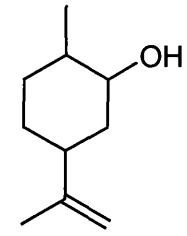
1,4-Cineole



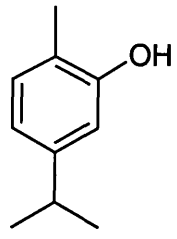
1,8-Cineole



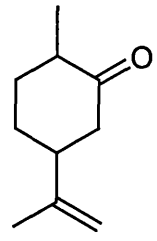
α-Terpinolene



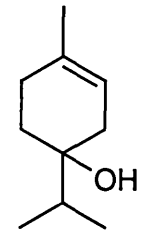
Dihydrocarveol



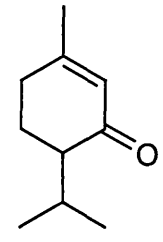
Carvacrol



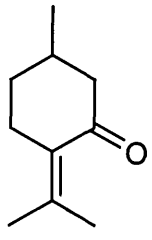
Dihydrocarvone



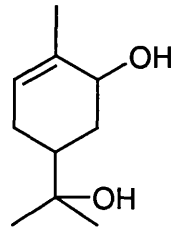
Terpinen-4-ol



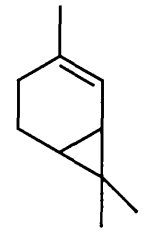
Piperitone



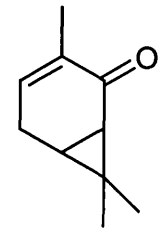
Pulegone



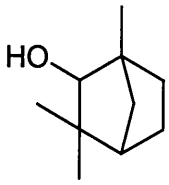
t-p-Menth-6-ene-2,8-diol



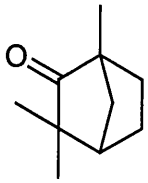
Car-2-ene



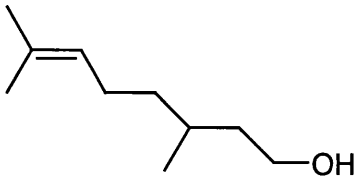
Car-3-en-2-one



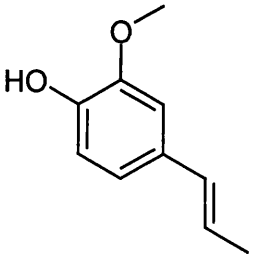
Fenchol



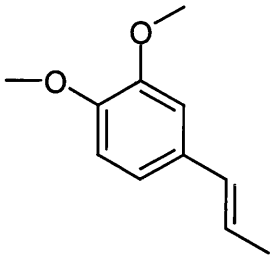
Fenchone



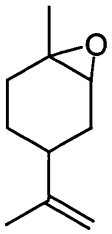
Citronellol



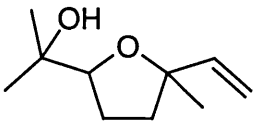
Isoeugenol



Methyleugenol

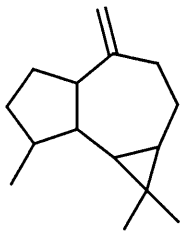


Limonene oxide

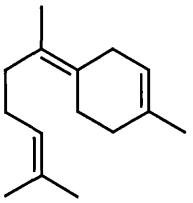


Linalool oxide

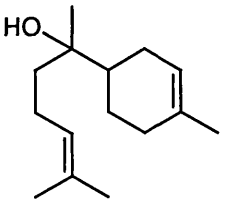
Terpenes investigated in section 8.2.



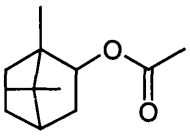
Alloaromadendrene



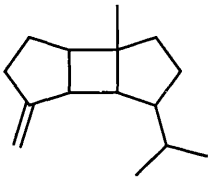
α-Bisabolene



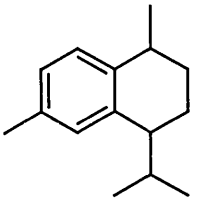
α-Bisabolol



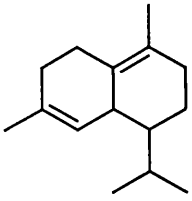
Bornyl acetate



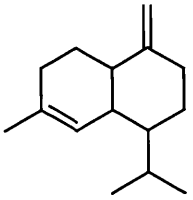
β-Bourbonene



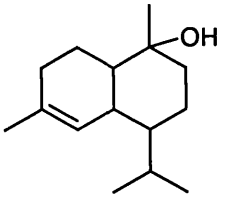
Calamenene



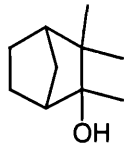
δ-Cadinene



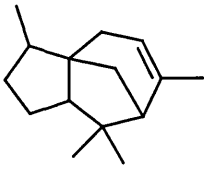
γ-Cadinene



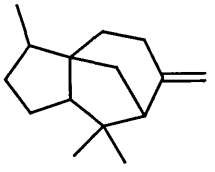
α-Cadinol



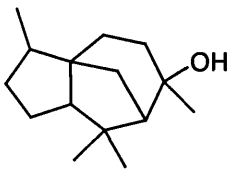
Camphene hydrate



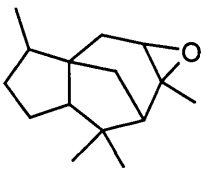
α-Cedrene



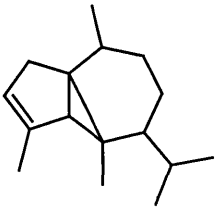
β-Cedrene



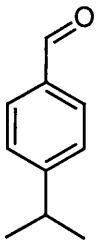
Cedrol



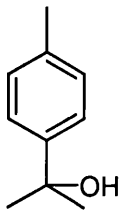
α-Cedrene oxide



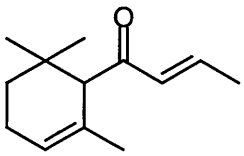
α-Cubebene



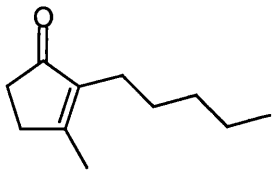
Cuminaldehyde



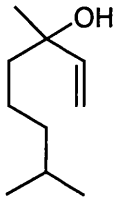
p-Cymen-8-ol



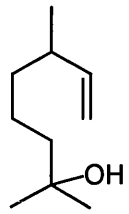
β-Damascenone



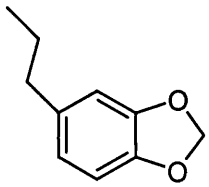
Dihydrojasmone



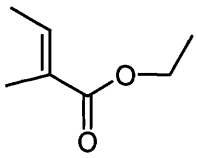
Dihydrolinalool



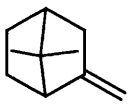
Dihydromyrcenol



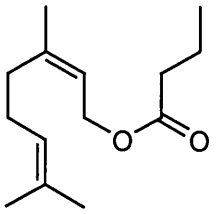
Dihydrosafrole



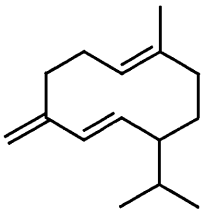
Ethyl tiglate



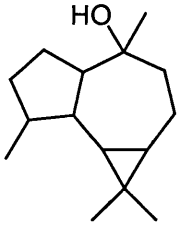
α-Fenchene



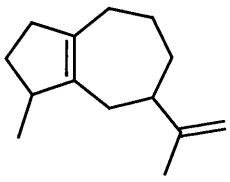
Geranyl butanoate



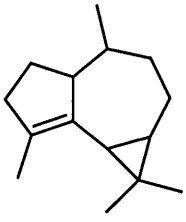
D-Germacrene



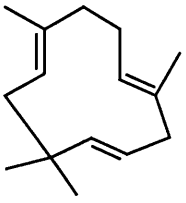
Globulol



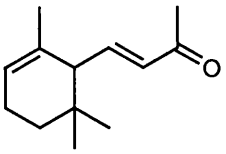
α-Guaiene



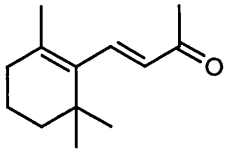
α-Gurgujene



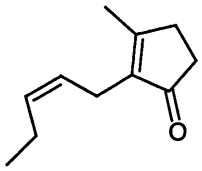
α-Humulene



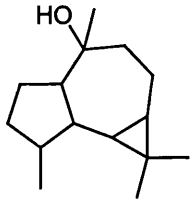
α–Ionone



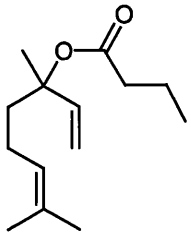
β–Ionone



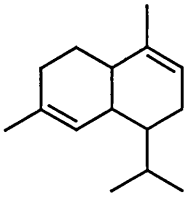
cis-Jasmone



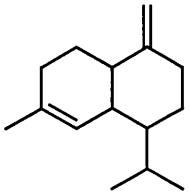
Ledol



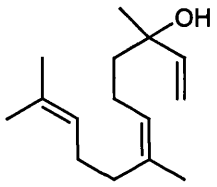
Linalyl butanoate



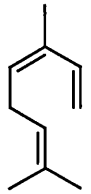
α-Muurolene



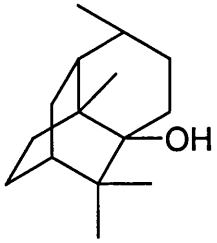
γ-Muurolene



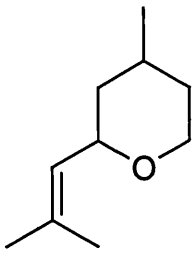
Nerolidol



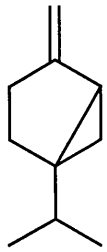
β-Ocimene



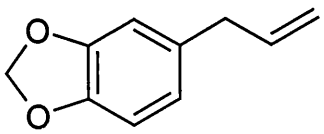
Patchouli alcohol



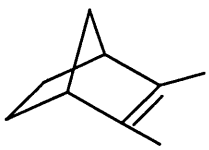
Cis-Rose oxide



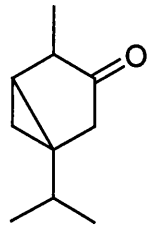
Sabinene



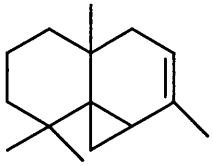
Safrole



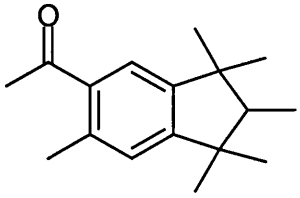
Santene



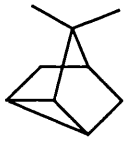
Thujone



Thujospene



Tonalid



Tricyclene

8.5. References

1. P. De Mayo, The Chemistry of Natural Products, Vol. II, Mono- and Sesquiterpenoids, Interscience Pub, Inc., New York, (1959).
2. J.L. Simonsen, The Terpenes, Vol. II, the Dicyclic Terpenes and their Derivatives, University press, Cambridge, (1949).
3. J.L Simonsen, D.H.R. Barton, The Terpenes, Vol. III, the Sesquiterpenes, Diterpenes and their derivatives, University Press, Cambridge, (1952).
4. W.J. Rea, Chemical Sensitivity Vol.2, Sources of Total Body Load, Lewis Publishers, (1992) 979.
5. C.S. Sell, in The Chemistry of Fragrance, Eds D.H. Pybus, C.S. Sell, The Royal Society of Chemistry, Cambridge, (1999) 24.
6. Anon, Mc Graw-Hill Encyclopeadia of Sciences and Technology, Vol. 13, Mc Graw-Hill, New York, (1982) 533.
7. I. Andersen, Atmos. Environ., 6 (1972) 275.
8. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, R. Kumarsingh, Pharmacol. Biochem. Behav., 60 (1998) 765.
9. Y. Alarie, M. Shaper, G.D. Nielsen, M.H. Abraham, SAR/QSAR Envir. Res., 6 (1996) 151.
10. J.-P. Kasanen, A.-L. Panasen, P. Panasen, J. Liesivuori, V. M. Kosma, Y. Alarie, Arch. Toxicol., 72 (1998) 514.

11. A.A. Falk, M.T. Hagberg, A.E. Lof, E.M. Wigaeus-Hejlm, Z. Wang, Scand. J. Work Environ. health, 16 (1990) 372.
12. A Falk Filipson, Occup. Environ. Med., 53 (1996) 100.
13. L. Friedman, J.G. Miller, Science, 172 (1972) 1044.
14. J.Ide, T. nakamoto, T. Moriizumi, Sens. Actuators A, 49 (1995) 73.
15. M.H. Abraham, R. Kumarsingh, J. E. Cometto-Muniz, W.S. Cain, M. Roses, E. Bosch, M.L. Diaz, J.Chem. Soc., Perkin Trans 2, (1998) 2405.
16. M.H. Abraham Unpublished results.
17. M.H. Abraham, M. Roses, J.Phys.Org.Chem., 7 (1994) 672.
18. M.H. Abraham, M.Roses, C.F. Poole and S.K. Poole, J. Phys. Org. Chem, 10 (1997) 358.
19. D.S. Seibert, C.F. Poole, J.High Resolut. Chromatogr., 18 (1995) 26.
20. P.T. Jackson, M.R. Schure, T.P. Weber, P.W. Carr, Anal. Chem., 69 (1997) 416.
21. I.Fichan, C.Larroche, J.B. Gros, J. Chem. Eng. Data, 44 (1999) 56.
22. J. Li, S.G. Pavlostathis, R. Araujo, Environ. Int., 24 (1998) 353.
23. S. Griffin, S. Grant Wyllie, J. Markham, J. Chrom. A, 864 (1999) 221.
24. S.G. Griffin, S.Grant, J.L. Markham, D.N. Leach, Flavour Fragr. J., 14 (1999) 322.
25. V.A. Izidorov, I.G. Zenkevich, A. Slowikowski, E. Wojciujk, J. Chromatogr. A, 814 (1998) 253.
26. J.A. Platts, D. Butina, M.H. Abraham, A. Hersey, J.Chem. Inf. Comput. Sci., 39 (1999) 835.
27. M.H. Abraham, J.A. Platts, J.Org.Chem., 66 (2001) 3484.
28. J.Le, M.H.Abraham, unpublished results.
29. S. Graves, M.H. Abraham, unpublished results.
30. E.E. Stashenko, M.Cervantes, Y. Combariza, H. Fuentes, J. R. Martinez, High Resol. Chromatogr., 22 (1999) 343.
31. P.K.C. Ong, T.Acree, E. H. Lavin, J. Agric. Food Chem., 46 (1998) 611.
32. Quest International, Ashford, Kent.
33. W. Jennings, T. Shibamoto, Qualitative Analysis of Flavour and Fragrances Volatiles, Academic Press, New York, (1980).
34. W. Jennings, A. Orav, T. Kailas, M. Liiv, Chromatographia, 43 (1996) 215.
35. S. Eri, B. K. Khoo, J. Lech, T. G. Hartman, J. Agric. Food Chem., 48 (2000) 1140.
36. H. J. Woerdenbag, R. Bos, H.E. Van Meeteren, J.J.J. Baarslag, L.T.W. DeJong Van, J. Essent. Oil Research, 12 (2000) 667.
37. G.R. Mallavarapu, L. Rao, S. Ramesh, J. Essent. Oil Research, 12 (2000) 766.
38. G. Takeoka, C. Perrino, Jr., Ron Buttery, J. Agric. Food Chem., 44 (1996) 654.
39. E. Tudor, J. Chromatogr. A, 779 (1997) 287.
40. P.M. Shafi, A. Saidutty, R.A. Clary, J. Essent. Oil Research, 12 (2000) 179.
41. T.M. Malingre, H. Maarse, Phytochemistry, 13 (1974) 1531.

42. V. Ferreira, M. Ardanuy, R. Lopez, J. F. Cacho, J. Agric. Food Chem., 46 (19980 4341.
43. J. S. Elmore, M.A. Erbahadir, D.S. Mottran, J. Agric. Food Chem., 45 (1997) 2638.
44. B. Bicalho, A. S. Pereira, F.R. Aquino Neto, A.C. Pinto, C.M. Rezende, J. Agr. Food Chem., 48 (2000) 1167.
45. O.E. Schupp III, J.S. Lewis, Compilation of Gas Chromatographic Data, ASTM Data Series Publication NO. DS 25A, (1967).
46. L.L. Hinrichsen, S. B. Pedersen, J. Agric. Food Chem., 43 (1995) 2932.
47. M.R.B. Franco, T. Shibamoto, J. Agric. Food Chem., 48 (2000) 1263.
48. G.R. Takeoka, R.A. Flath, M. Guntert, W. Jennings, J. Agric. Food Chem., 36 (1988) 553.
49. F. Patte, M. Etcheto, P. Laffort, Anal. Chem., 54 (1982) 2239.
50. A.I. Allahverdiev, S. Irandoust, M. Andersson, J. Chromatogr. A, 848 (1999) 297.
51. P.K. Koukos, K.I. Papadopoulou, D.Th. Patakia, A.D. Papagiannopoulos, J. Agric. Food Chem., 48 (2000) 1266.
52. M.H. Abraham, C.F. Poole, S.K. Poole, J. Chromatogr. A, 842 (1999) 79.
53. Y. Ishihama, N. Asakawa, J. Pharm. Sci., 88 (1999) 1305.
54. M.H. Abraham, J. Andonian-Haftan, G.S. Whiting, A. Leo, R.W. Taft, J.Chem.Soc. Perkin Trans. 2, 2 (1994) 1777.
55. M.H. Abraham, H.S. Chadha, G.S. Whiting, R.C. Mitchell, J. Pharm. Sci., 83 (1994) 1085.
56. J.M.R. Gola, M.H. Abraham, Unpublished results.

Chapter 9 Advances in Prediction of Nasal Pungency Thresholds of Volatile Organic Compounds

9.0. Introduction

The ability to predict nasal pungency threshold, NPT, in humans of volatile organic compounds, VOC, would greatly improve the issues of indoor air pollution. Abraham and co-workers[1,2] are pioneers in the field of quantitative structure activity relationship, QSAR, applied to NPT of volatile organic compounds, VOCs. In 1996, they set up a predictive model, based on the Abraham solvation equation for gas to condensed phase process[3], for a set of 34 various VOCs including alcohols, acetates, ketones and a few miscellaneous compounds, see equation (9.1).[1]

$$\text{Log } (1/\text{NPT}) = -8.562 + 2.209\,\mathbf{S} + 3.417\,\mathbf{A} + 1.535\,\mathbf{B} + 0.865\,\mathbf{L} \qquad (9.1)$$

$n = 34$, $r = 0.976$, $sd = 0.27$, $F = 144$

Here and elsewhere **S**, **A**, **B** and **L** are the Abraham solvation parameters, n is the number of data points, r is the correlation coefficient, sd is the standard deviation and F the Fisher statistic. NPT values were measured by Cometto-Muniz and Cain[4-9] by means of a standardised method using a 270-cm^3 squeeze bottle. The reciprocal of NPT values was used, so that the more potent the VOC the larger is the value of log (1/NPT). A further ten NPT values were determined, and the combined set yield the upgraded solvation equation (9.2).[2]

$$\text{Log } (1/\text{NPT}) = -8.519 + 2.154\,\mathbf{S} + 3.522\,\mathbf{A} + 1.397\,\mathbf{B} + 0.860\,\mathbf{L} \qquad (9.2)$$

$n = 43$, $r^2 = 0.955$, $sd = 0.27$, $F=201$

This equation includes a wide variety of VOCs, including carboxylic acids, aldehydes, ketones, alcohols, etc., with only one outlier-acetic acid.[2]

Recently, Hau and Connel[10] developed a QSAR for describing the triggering of nasal pungency in humans, based on the partition of volatile organic compounds between the air phase and the receptor phase. The water / air partition coefficient, L^W, and the octanol / water partition coefficient, P^{oct}, were the partition parameters used in the model. The authors used the same NPT values as those in equation (9.1) with the exception of heptan-4-ol for which a reliable value of L^W is not available. Their equation is equation (9.3).

$$\text{Log } \{(\text{NPT})\, L^W\} = 7.69 - 1.16 \log P^{oct} \qquad (9.3)$$

$n = 33$, $r = 0.971$, $sd = 0.458$, $F = 520$

The standard deviation was calculated in this work in order to provide a uniform basis of comparison with equation (9.1). A much larger sd of 0.46 log units for a 33 compound training set was obtained using the Hau-Connel model. However, a better comparison could be achieved by assessing the efficiency of both equations to predict NPT values.

The aim of this chapter was first to develop a more general model based on the Abraham solvation equation including some terpenes. Then, the predictive power of the Abraham model and the Hau-Connel model was compared. Finally, a model was established to analyse a set of NPT values measured by means of a new glass bottle device.

9.1. New Solvation Equation Model for Nasal Pungency Thresholds

9.1.1. Model Development

Cometto-Muniz and Cain[4-9,11] have measured NPT values for six terpenes, listed at the bottom of Table 9.5 in section 9.5. Their corresponding descriptors were obtained as described in reference 12. NPT values were calculated using the latest solvation equation published viz. equation (9.2).[2] In Figure 9.1 are shown the observed and predicted values of log (1/NPT) for all compounds in Table 9.5. With the exception of linalool that is more potent than expected, terpenes generally conform to equation (9.1).

A new regression equation was then developed using all the VOCs given in Table 9.5, except acetic acid and linalool. The output of the MLR analysis of log (1/NPT) with the solute descriptors is summarised in Table 9.1.

$$\text{Log (1/NPT)} = -8.08\ (\pm 0.226) + 1.767\ (\pm 0.290)\ \mathbf{S} + 3.298\ (\pm 0.203)\ \mathbf{A} + 1.076\ (\pm 0.279)\ \mathbf{B} + 0.857\ (\pm 0.031)\ \mathbf{L} \qquad (9.4)$$

$n = 48$, $r^2 = 0.950$, sd = 0.270, F=211

The coefficients were calculated using MS Excel'97 software. The *e*-coefficient of the independent variable **E** was found to be statistically not significant.

Table 9.1. Output of the regression analysis of log (1/NPT) and solute descriptors

1)	**Variables**	**S**	**A**	**B**	**L**	
	A	-0.189				
	B	0.158	0.241			
	L	-0.035	-0.304	-0.142		
	Log(1/NPT)	0.088	0.301	0.179	0.754	
2)	**Variables**	**S**	**A**	**B**	**L**	**Log(1/NPT)**
	MEAN	0.531	0.147	0.418	3.519	-3.190
	SD	0.141	0.213	0.148	1.311	1.169
3)	**Variables**	***s***	***a***	***b***	***l***	***c***
	COEFFS	1.767	3.297	1.077	0.857	-8.084
	ST.DEV	0.290	0.203	0.279	0.031	0.227
	T	6.089	16.188	3.858	27.21	-35.63
	TTEST	1.000	1.000	1.000	1.000	1.000
4)	$\mathbf{r^2}$	0.950				
	sd	0.271				
	DOF	44				
	F	211.6				

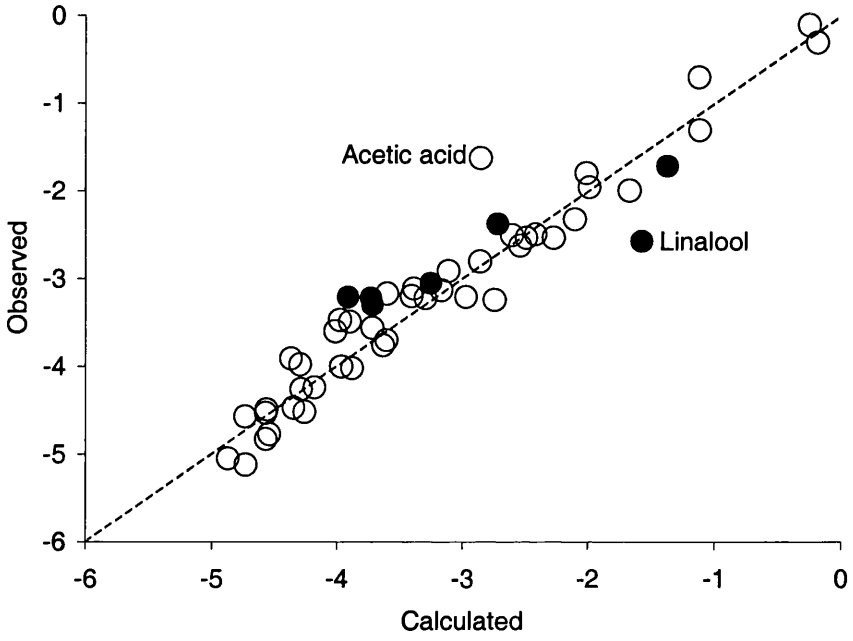


Figure 9.1. Plot of observed values of log (1/NPT) against log (1/NPT) calculated on equation (9.2). Open circles represent the original 44 VOCs, filled circles are the new terpenes. --- Identity line.

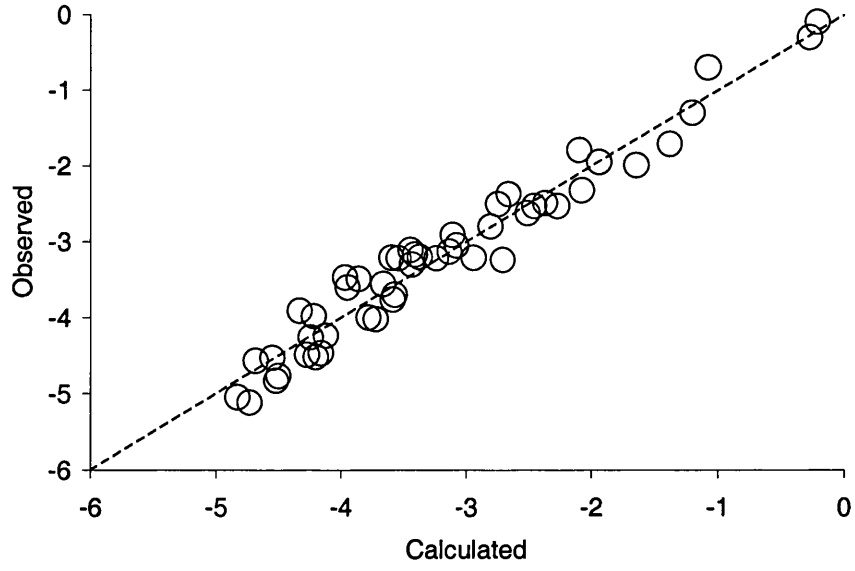


Figure 9.2. Plot of observed values of log (1/NPT) against log(1/NPT) calculated on equation (9.4). --- Identity line.

9.1.2. Discussion

The insertion of terpene compounds in the training set leads to a slight decrease in all the coefficients. Still, they are in the same range as those in equation (9.2). Hydrogen bonding acidity and basicity capacity of the receptor phase together with its dipolarity / polarisability capacity are the predominant characteristics of the phase. A plot of observed values of log (1/NPT) against those calculated from equation (9.4) shows a random scatter about the line of identity, see Figure 9.2.

The relative significance of the descriptors in equation (9.4) can be calculated from the product of the coefficient and the mean value of the corresponding descriptor. The percentage weights are then **S** (19%), **A** (10%), **B** (9%) and **L** (62%) so that the l.**L** term is by far the most important. The later is related to the size of the VOC, and very roughly to the molecule weight.

A principal component analysis, PCA, of various solvation equations for gas / solvent partition process including the one for nasal pungency threshold, was carried out to determine the solvents that would best characterise the receptor area. Here, equation (9.4) was compared to 43 models for gas / solvent partition system in order to cover a wide range of solvent physicochemical properties. Thus, alkanes, chlorinated alkanes, alcohols, aromatics, ethers, amides and other types of solvents were considered. The PCA analysis was conducted using the statistic software 'Minitab' (Minitab for Windows, Release 9.2, 1993, Minitab Inc., State College PA). A graphical display of the various processes, which shows their similarity on the basis of their physicochemical properties, was produced, see Figure 9.3. The closer the scores, the greater the similarity. Therefore, a simple visual analysis of Figure 9.3 yields the conclusion that alkanes, aromatics, ethers and chlorinated alkanes are not suitable in mimicking the receptor area. This result is however not surprising and one can use it to prove the consistency of the PCA analysis output. Water cannot be used too. Although the receptor phase is just as basic as water, it is very much less acidic[2]. Furthermore, this information shows that the binding site cannot be a very hydrophilic environment. Octan-1-ol and dimethylformamide (DMF), are very good models. N-formylmorpholine (NFM) is also a reasonable model for the receptor area, except that this amide lacks any hydrogen-bond acidity. Solvation equations for these suitable solvent models are given in Table 9.6 in section 9.5.

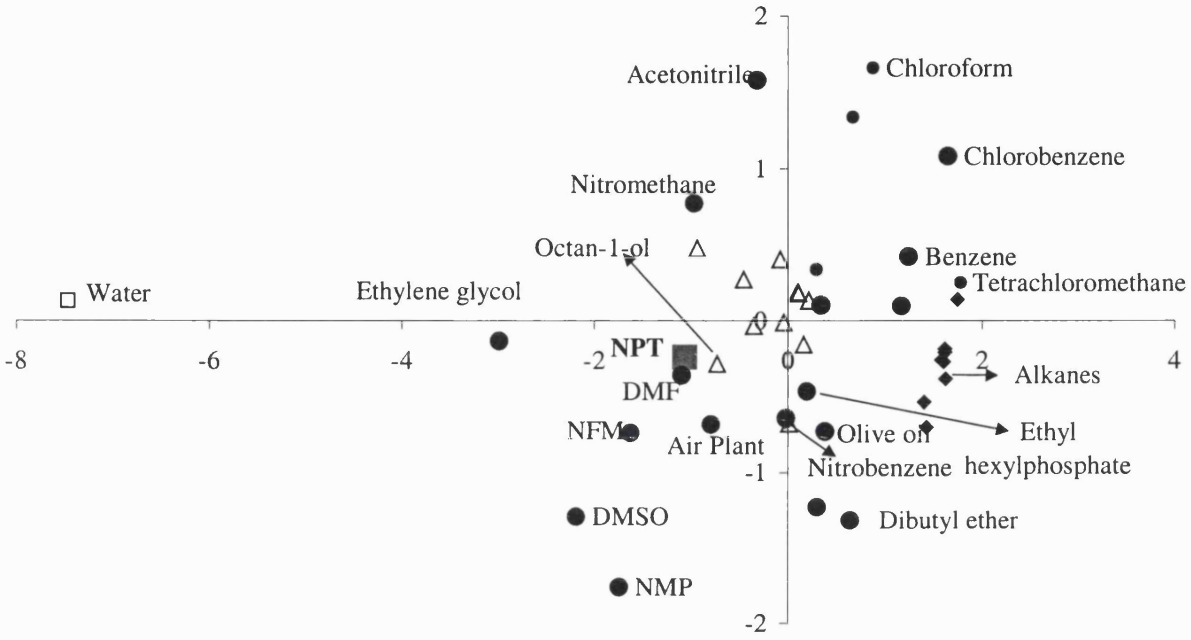


Figure 9.3. Principal component scores plot. Δ Alcohols, • Aromatics, ◆ Alkanes, ● Miscellaneous. Abbreviations: DMF, dimethylformamide, DMSO, dimethylsulfoxide, NMP, N-methylpyrrolidone, NFM, N-formylmorpholine.

9.2. Comparison between two Models for Nasal Pungency Thresholds

The predictive ability of the Abraham model[1], equation (9.1) and the Hau-Connel model[10], equation (9.3), was assessed. A series of n-carboxylic acids, aldehydes and a few terpenes was taken as the test set. Logarithms of observed NPT values, measured in ppm, are given in Table 9.2. Note that acetic acid and linalool were not considered in this study as they were clearly identified as outliers in previous analyses.

Experimental gas / water and water / octanol partition coefficient values at 298K were obtained from the MedChem 2000 database[13]. The Abraham solvation parameters are given in Table 9.5 in section 9.5. The statistical tools considered are (1) sd: standard deviation between observed and calculated NPT values, $sd = \sqrt{((obs-calc)^2/n)}$ (2) AAE:

average standard error, AAE = Σ | obs-calc | / n, and (3) AE: average error, AE = Σ(obs-calc) / n. Prediction statistic for each of the models studied is shown in Table 9.3.

The results for the test set of nine VOCs indicate that the Abraham general solvation equation is more accurate in prediction than the 'partition' model of Hau and Connel. Therefore, the model developed by Hau and Connel does not offer additional precision over that of Abraham and co-workers. Furthermore, this model does not give more insight on the nasal pungency receptor area. The authors claim that an overall slope of 1.16, see equation (9.3) indicates that the common receptor site is probably within the hydrophobic interior of the lipid bilayer membrane of the nerve ending, and that there is a slight, non-specific interaction between the VOC molecules and the receptor molecules[10]. Similar results were obtained by Abraham and co-workers.

Table 9.2. Observed log (NPT) values, NPT in ppm, and log (NPT) values calculated on equations (9.1) and (9.3)

Solute	log L^W [a]	Log P^{oct} [a]	Obs. log (1/NPT)	Calc. log (1/NPT) [b]	Calc. log (1/NPT) [c]
Butanal	2.210	0.880	4.770	-4.459	-4.471
Hexanal	2.000	1.780	3.700	-3.625	-3.531
Formic acid	5.250	-0.540	2.500	-3.066	-2.623
Butanoic acid	4.660	0.790	1.790	-2.114	-2.003
Hexanoic acid	4.560	1.920	1.300	-0.903	-1.104
Octanoic acid	4.440	3.050	0.300	0.288	-0.170
1,8-Cineole	2.250	2.820	2.370	-2.169	-2.610
Cumene	0.340	3.660	3.220	-3.104	-3.700
p-Cymene	0.500	4.100	3.050	-2.434	-3.217

[a] Values of log L^W and log Poct obtained from the MedChem database.
[b] Calculated on equation (9.3)
[c] Calculated on equation (9.1)

Table 9.3. Statistical results

	Hau-Connel Model	Abraham Model
SD	0.389 [a]	0.247 [b]
AAE	0.312 [a]	0.224 [b]
AE	0.114 [a]	0.047 [b]

[a] Calculated on equation (9.3)
[b] Calculated on equation (9.1)

9.3. Comparison between Nasal Pungency Threshold Values measured either by Squeeze Bottle System or by Glass Vessel System.

As already pointed out in chapter 2, Cometto-Muniz and Cain have measured nasal pungency thresholds in anosmics via two-stimulus delivery systems: (1) the 270-cm^3 squeeze bottle system[4-9,11], and (2) the 1,900-cm^3 glass vessels with teflon tubing and nose pieces system[14]. The first device has been widely used and allowed the measurement of NPT values for 51 VOCs, listed in Table 9.5 in section 9.5. These values were used to set up equations (9.1), (9.2) and (9.4). The second system has been recently utilised to determine the NPT values in anosmics for nine VOCs. NPT values via glass vessels are on average four times less, or 0.6 log unit less, than those reported using squeeze bottle, see Table 9.4.[13]

The nine new NPT values were used as the test set for equation (9.4). There is a rather good agreement between log (1/NPT) calculated via equation (9.4) and the observed values for the three homologous alcohols, hexyl acetate and octyl acetate, see Figure 9.4. The remaining four compounds, the three homologous ketones and butyl acetate, were not well calculated.

An indicator variable, I, was added to the regression analysis to cancel out the difference between both experiments. The best results were obtained with I as 0 for compounds whose NPT values were determined via squeeze bottle experiments and 0.6 for those studied by the new glass vessels system. The following equation was finally obtained:

$$\text{Log } (1/\text{NPT}) = -8.09\ (\pm 0.233) + 1.9381\ (\pm 0.2965)\ \mathbf{S} + 3.137\ (\pm 0.206)\ \mathbf{A} + 1.110\ (\pm 0.292)\ \mathbf{B} + 0.835\ (\pm 0.031)\ \mathbf{L} + \mathbf{I} \qquad (9.5)$$

$n = 58$, $r^2 = 0.944$, SD = 0.285, F=175

Linalool and acetic acid were not included in the analysis. A plot of equation (9.5) is shown in Figure 9.5 from which it can be seen that the incorporation of an additional parameter improves the fit.

Table 9.4. NPT values (in ppm) measured via glass vessel experiments.

VOCs	Obs. log (1/NPT)[a]	Obs. log (1/NPT)[b]	Calc. log (1/NPT)[c]	Calc. log (1/NPT)[d]
Pentan-2-one	-3.47	-2.96	-3.97	-3.30
Heptan-2-one	-2.91	-2.33	-3.11	-2.46
Nonan-2-one	-2.53	-1.42	-2.27	-1.65
Butyl acetate	-3.56	-2.73	-3.66	-3.02
Hexyl acetate	-2.80	-2.28	-2.81	-2.19
Octyl acetate	-1.95	-1.68	-1.94	-1.34
Butan-1-ol	-3.20	-2.69	-3.37	-2.81
Hexan-1-ol	-2.62	-2.29	-2.51	-1.96
Octan-1-ol	-1.99	-1.51	-1.64	-1.12

[a] Measured via squeeze bottle experiments
[b] Measured via glass vessel experiments
[c] Calculated from equation (9.4)
[d] Calculated from equation (9.5)

9.4. Conclusion

A new solvation equation for nasal pungency threshold including a series of terpenes was developed. Since this new model covers a wide spread of log (1/NPT) values, from -0.10 to 5.12, and a wide range of types of VOCs, it represents a simple way of predicting values of NPT for a large number of VOCs with a standard deviation of the predicted of 0.27. It was also pointed out that the Abraham model offers more precision and information on the nasal pungency receptor area than the model developed by Hau and Connel. Finally, an indicator variable was added to the Abraham general solvation equation in order to account for the difference between the two systems for NPT measurement devised by Cometto-Muniz and Cain.

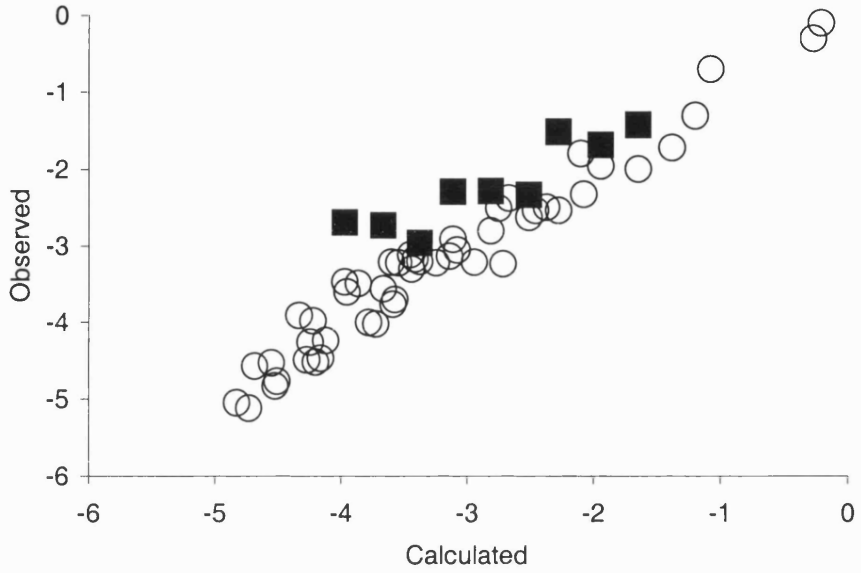


Figure 9.4. Plot Observed log (1/NPT) against log (1/NPT) calculated on equation (9.4). Open circles represent log (1/NPT) values measured from squeeze bottle experiment, filled squares are log (1/NPT) values obtained using glass bottles.

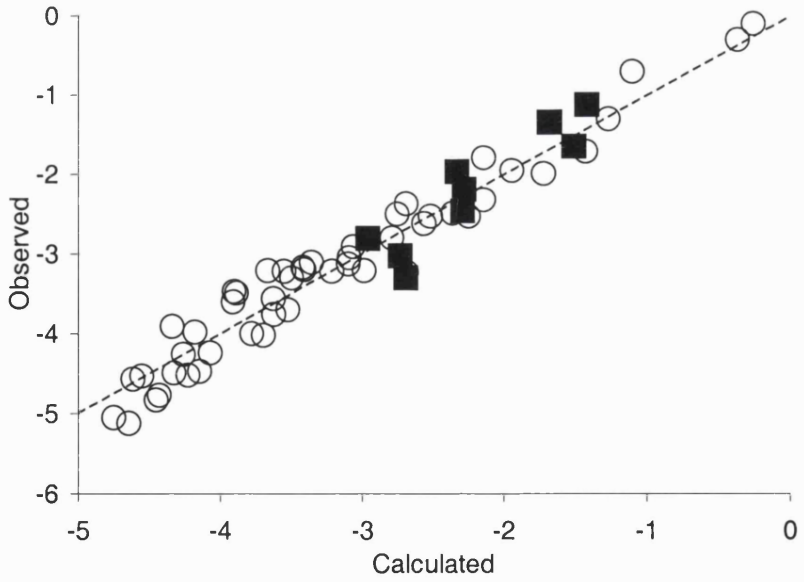


Figure 9.5. Plot Observed log (1/NPT) against log (1/NPT) calculated on equation (9.5). Open circles represent log (1/NPT) values measured from squeeze bottle experiment, filled squares are log (1/NPT) values obtained using glass bottles.

9.5. Appendix

Table 9.5. Descriptors for 51 VOCs, observed log (1/NPT) and calculated log (1/NPT) on equations (9.2) and (9.4). NPT values measured via squeeze bottle experiments. Values in ppm.

Ref.	VOCs	S	A	B	L	Obs. log (1/NPT)	Calc. log (1/NPT)[a]	Calc. log (1/NPT)[b]
7	Oct-1-yne	0.220	0.090	0.100	3.521	-4.490	-4.560	-4.269
6	Propanone	0.700	0.040	0.490	1.696	-5.120	-4.727	-4.730
6	Pentan-2-one	0.680	0.000	0.510	2.755	-3.470	-3.973	-3.968
6	Heptan-2-one	0.680	0.000	0.510	3.760	-2.910	-3.108	-3.107
6	Nonan-2-one	0.680	0.000	0.510	4.735	-2.530	-2.270	-2.271
5	Methyl acetate	0.640	0.000	0.450	1.911	-5.050	-4.868	-4.827
5	Ethyl acetate	0.620	0.000	0.450	2.314	-4.830	-4.565	-4.517
5	Propyl acetate	0.600	0.000	0.450	2.819	-4.240	-4.174	-4.119
5	Butyl acetate	0.600	0.000	0.450	3.353	-3.560	-3.714	-3.662
6	s-Butyl acetate	0.570	0.000	0.470	3.054	-3.600	-4.008	-3.949
6	t-Butyl acetate	0.540	0.000	0.470	2.802	-3.980	-4.290	-4.218
5	Pentyl acetate	0.600	0.000	0.450	3.844	-3.220	-3.292	-3.241
5	Hexyl acetate	0.600	0.000	0.450	4.351	-2.800	-2.856	-2.806
5	Heptyl acetate	0.600	0.000	0.450	4.865	-2.490	-2.414	-2.366
5	Octyl acetate	0.600	0.000	0.450	5.364	-1.950	-1.985	-1.938
5	Decyl acetate	0.600	0.000	0.450	6.373	-0.700	-1.117	-1.073
5	Dodecyl acetate	0.600	0.000	0.450	7.381	-0.100	-0.250	-0.210
4	Methanol	0.440	0.430	0.470	0.970	-4.530	-4.566	-4.547
4	Ethanol	0.420	0.370	0.480	1.485	-3.910	-4.364	-4.328
4	Propan-1-ol	0.420	0.370	0.480	2.031	-3.490	-3.894	-3.860
6	Propan-2-ol	0.360	0.330	0.560	1.764	-4.260	-4.282	-4.241
4	Butan-1-ol	0.420	0.370	0.480	2.601	-3.200	-3.404	-3.372
6	s-Butan-1-ol	0.390	0.370	0.480	2.413	-3.760	-3.630	-3.586
6	t-Butan-2-ol	0.300	0.310	0.600	1.963	-4.520	-4.255	-4.199
4	Pentan-1-ol	0.420	0.370	0.480	3.106	-3.210	-2.969	-2.939
4	Hexan-1-ol	0.420	0.370	0.480	3.610	-2.620	-2.536	-2.507
4	Heptan-1-ol	0.420	0.370	0.480	4.115	-2.320	-2.102	-2.074
6	Heptan-4-ol	0.360	0.330	0.560	3.850	-2.530	-2.488	-2.453
4	Octan-1-ol	0.420	0.370	0.480	4.619	-1.990	-1.668	-1.642
7	Toluene	0.520	0.000	0.140	3.325	-4.470	-4.344	-4.161

Ref.	VOCs	S	A	B	L	Obs. log (1/NPT)	Calc. log (1/NPT)[a]	Calc. log (1/NPT)[b]
7	Ethylbenzene	0.510	0.000	0.150	3.778	-4.000	-3.962	-3.780
7	Propylbenzene	0.500	0.000	0.150	4.230	-3.170	-3.595	-3.410
7	Chlorobenzene	0.650	0.000	0.070	3.657	-4.020	-3.876	-3.722
4	Pyridine	0.840	0.000	0.520	3.022	-3.110	-3.384	-3.446
8	Butanal	0.650	0.000	0.450	2.270	-4.770	-4.538	-4.501
8	Pentanal	0.620	0.000	0.450	2.120	-4.570	-4.732	-4.683
8	Hexanal	0.650	0.000	0.450	3.357	-3.700	-3.603	-3.570
8	Heptanal	0.650	0.000	0.450	3.865	-3.130	-3.166	-3.134
8	Octanal	0.650	0.000	0.450	4.361	-3.240	-2.740	-2.709
8	Formic acid	0.790	0.720	0.340	1.400	-2.500	-2.603	-2.744
8	Butanoic acid	0.620	0.600	0.450	2.830	-1.790	-2.008	-2.096
8	Hexanoic acid	0.600	0.600	0.450	3.920	-1.300	-1.114	-1.197
8	Octanoic acid	0.600	0.600	0.450	5.000	-0.300	-0.185	-0.271
8	Acetic acid	0.650	0.610	0.440	1.750	-1.620	-2.851	-2.946
9	Linalool	0.550	0.200	0.670	4.794	-2.570	-1.571	-1.619
4	Menthol	0.500	0.230	0.580	5.177	-1.710	-1.369	-1.377
9	1,8-Cineole	0.330	0.000	0.760	4.688	-2.370	-2.715	-2.661
9	Cumene	0.490	0.000	0.160	4.084	-3.220	-3.728	-3.542
9	p-Cymene	0.490	0.000	0.190	4.590	-3.050	-3.251	-3.076
9	Δ-3-Carene	0.220	0.000	0.100	4.649	-3.210	-3.907	-3.599
9	α-Terpinene	0.250	0.000	0.150	4.715	-3.300	-3.716	-3.436

[a] Calculated on equation (9.2)
[b] Calculated on equation (9.4)

Table 9.6. Solvation equations for gas / solvent[a] partition processes.

Solvent	SP	*c*	*e*	*s*	*a*	*b*	*l*
Plant Cuticle	Log L	-0.617	0.082	1.282	3.120	0.820	0.860
Wet Octan-1-ol	Log L	-0.198	0.002	0.709	3.519	1.429	0.858
DMF[b]	Log L	-0.161	-0.189	2.327	4.756	0.000	0.808
NFM	Log L	-0.624	0.016	2.477	3.989	0.000	0.692
NPT	Log (1/NPT)	-8.08	0.00	1.767	3.298	1.076	0.857

9.6. Reference

1. M.H. Abraham, J Andonian-Haftan, J.E Cometto-Muniz, W.S. Cain, Fund. Appl. Toxicol., 31 (1996) 71.
2. M.H. Abraham, R. Kumarsingh, J. E. Cometto-Muniz and W. S. Cain, Arch. Toxicol., 72 (1998) 227.
3. M.H. Abraham, Chem. Soc. Rev., 22 (1993) 73.
4. J.E. Cometto-Muniz and W.S. Cain, Physiol. Behav., 48 (1990) 719.
5. J.E. Cometto-Muniz and W.S. Cain, Pharmacol. Biochem. Behav., 39 (1991) 983.
6. J.E. Cometto-Muniz and W.S. Cain, Arch. Environ. Health, 48 (1993) 309.
7. J.E. Cometto-Muniz and W.S. Cain, Am. Ind. Hyg. Ass. J., 55 (1994) 811.
8. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, Exp. Brain Res., 118 (1998) 180.
9. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham and R. Kumarsingh, Pharmacol. Biochem. Behav., 60 (1998) 765.
10. K.M. Hau, D.W. Connell, B.J Richardson, Toxicological Sci., 47 (1999) 93.
11. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham and R. Kumarsingh, Ann. N-Y Acad. Sci., 855 (1998) 648.
12. M.H.Abraham, R.Kumarsingh, J.E. Cometto-Muniz, W.S. Cain, M.Roses, E. Bosh and M.L. Diaz, J.Chem.Soc., Perkin Trans 2, (1998) 2405.
13. A. Leo, The Medicinal Chemistry Project Pomona College, Claremont, CA 91711.
14. J.E. Cometto-Muniz, W.S. Cain, T.Hiraishie, M.H.Abraham and J.M.R. Gola, Chem. Senses, 25 (2000) 285.

Chapter 10 A Model for Odor Detection Thresholds

10.0. Introduction

Knowledge of the detection threshold values of odorant volatile organic compounds, VOCs, is important in understanding the level of warning that VOC odors can provide against hazardous situations[1] and in looking for insight into the olfactory mechanism. Unfortunately, odor detection threshold values, ODT, are only known for a limited number of VOCs. To overcome this problem, a number of correlations of odor detection thesholds, ODT, with various properties of odorants, has been developed; the study by Laffort and Patte[2] being one of the first to employ a physical analysis. Their investigation as well as others[3-6] has been presented in chapter 2.

Most of the studies developed so far have been based on homologous series of odorants, excluding the numerous types of important VOCs such as terpenes or inhalation anaesthetics that do not fall into any homologous series. Furthermore, although the majority of the models may be useful as empirical correlations, they yield little mechanistic information. The model of Hau and Connel[1] is significant, however, because it is the real only attempt to correlate ODT values on any mechanistic basis, see chapter 2. The partition coefficient between octan-1-ol and water, P^{oct}, and that between water and air, K^W, were used to model the partition process of VOCs into the biophase where the olfactory signal is transformed. QSARs were developed for four homologous series. The ODT values were from the AIHA compilation.[7] The model proposed for the alkyl acetates is as follows,

$$\text{Log } [\{ODT\}K_W] = -1.65 + 4.44 \log P^{oct} \qquad (10.1)$$

The interpretation of equation (10.1) was that K^W is an approximation for K^{AM}, since both refer to equilibrium between the gas phase and an aqueous condensed phase, and that K_{MB} and K_R are both functions of P^{oct}. KAM, KMB and KR represent the equilibrium constants for partition between the air phase and the mucus layer, the mucus layer and the biophase, and the biophase and the receptor, respectively.

The mammalian olfactory system is able to characterise and to discriminate a vast number and variety of different odorous volatile organic compounds, VOCs. An important element to understanding the way in which olfactory perception is processed is the relationship between odor stimuli and the receptive surface.[8] The primary event of olfactory stimulation can be described as the transport of VOCs from the air stream to the olfactory epithelium via a thin layer of mucus (10-30 μm thick).[9,10] Such diffusion may involve, at least at part, odorant binding proteins (OBPs) that can act as carriers.[11-13] The central pocket in the OBP has dimensions of 11*10*7 Å (i.e. 770Å^3) with an opening size of 6*7 Å [13], although a much larger cavity of 1100-1300 Å has been suggested[11]. When transported across the mucosal layer to a receptor area or biophase, the VOCs, or the OBP/VOC complex, can then interact with odor receptors at the surface of the cilia membrane of the olfactory neuron. The actual binding pocket in the rat OR5 receptor, however, is no less than 12 Å from the extracellular surface of the receptor[14]. The diffusion or transport process is important in the olfactory mechanism. However, is it the only factor influencing this biological process? To answer this question, it is useful to consider two types of interaction. Simple transport processes are selective, in that different VOCs will have different equilibrium constants, depending on their structure. However, small changes in structure or small positional changes of functional groups have rather small effects on such processes. On the other hand, in processes such as ligand / receptor interactions, small changes in structure can have very large effects; these processes have specific effects. A general model is presented in Figure 10.1 where are indicated which processes may be selective and which may be specific in nature. Three possible mechanisms for olfactory stimulation are hypothesised,

i. The VOC is transported by an OBP to the receptor area. The VOC/OBP complex interacts with an olfactory receptor, R, to form the complex VOC/(R+OBP).

ii. The VOC is transported by an OBP to the receptor phase. At proximity of the receptor, the VOC/OBP is disrupted. Thus only the VOC interact with the receptor, VOC/R.

iii. The VOC diffuses through the mucosal layer and then gain access to the biophase where it interacts with an olfactory receptor to form the complex VOC/R.

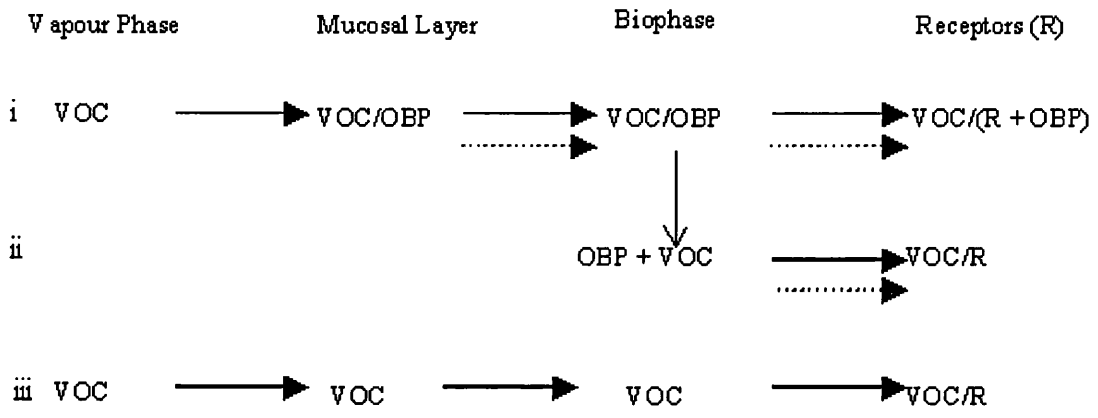


Figure 10.1. A possible model for odor thresholds. Selective processes ⟶ ; specific processes ⋯⋯▶ . VOC, volatile organic compound; OBP, odor binding protein; R, receptor phase.

Whether the VOC/OBP interactions and the VOC/R interactions are general interactions that can be modelled by a physiochemical transport processes, or whether they are more specific interactions, is a crucial point. The analysis of Hau and Connel[1] certainly supposes that the VOC/R interaction is a general interaction that can be modelled by a simple physicochemical descriptor, such as log P^{oct}.

In the present work, a model that simply reflects a passive physicochemical transport property is first used. Comparison with physicochemical transport to various solvents or to various biophases will then indicate whether or not such passive transport can model all or part of the odor detection process.

This method starts with the Abraham general solvation equation[15] for the correlation of processes, in which VOCs are transferred from the gas phase to some condensed phase,

$$\text{Log SP} = c + e\mathbf{E} + s\mathbf{S} + a\mathbf{A} + b\mathbf{B} + l\mathbf{L} \qquad (10.2)$$

The theory behind the equation has been presented in chapter 2. In chapter 9, equation (10.2) was used to correlate nasal pungency threshold values (NPT, in part per million, ppm) for 48 varied VOCs, resulting in equation (10.3),

$$\text{Log (1/NPT)} = -8.080 + 1.767\,\mathbf{S} + 3.298\,\mathbf{A} + 1.076\,\mathbf{B} + 0.857\,\mathbf{L} \quad (10.3)$$

$n = 48$, $r^2 = 0.950$, sd = 0.270, F = 211

Here and elsewhere, n is the number of data points, r is the correlation coefficient, sd is the standard deviation in the dependent variable, and F is the F-Statistic. The *e*-coefficient of the independent variable, **E**, was statistically not significant. The reciprocal of NPT values were used, so that the more potent the VOC the larger is the value of log (1/NPT). The coefficients in equation (10.3) compared well to those for various gas / solvent partitions that take place by simple transfer mechanism[16-19]. In chapter 9, a principal component analysis showed that there is a considerable similarity between the NPT equation and equations for the solubility of gaseous VOCs in octan-1-ol, dimethylformamide, and N-formylmorpholine. It is noteworthy that equation (10.3) emcopasses a wide variety of VOCs including carboxylic acids, aldehydes, ketones, alcohols, etc., with but two outliers acetic acid and linalool.

The strategy used in the present study is to apply the general equation (10.2) to ODT values, in the hope to deduce whether or not the resulting equation is consistent with simple transfer of VOCs from the gas phase to a biophase.

10.1.Results

Odor detection thresholds, ODT, for a series of 64 VOCs, including esters, aldehydes, ketones, alcohols, carboxylic acids, aromatic hydrocarbons, terpenes and a number of other VOCs, have been determined by Cometto-Muniz and Cain[20-24] and by Cometto-Muniz et al.[25,26], using a standardised protocol. The average standard deviation for all odor thresholds, expressed as log (1/ODT) is 0.63 log unit. The VOCs used in these studies are listed in Table 10.4 in section 10.4, together with log (1/ODT) values, where ODT is in parts per million, ppm.

The solute descriptors for each of the VOC in the data set listed in Table 10.4 were taken from an in-house software using SmartWare II (Informix Sofware Inc. Kansas) database. Regression analyses and statistical tests were performed with the same in-house software. Another statistical software, JMP (Version 3.2.1 SAS Institute Inc) was also used to determine the coefficient of cross-correlation, q^2. This coefficient

is a useful measure of internal self-consistency and is defined as (1-PRESS/SSY). PRESS stands for predictive residual sum squares and works by leaving out each data point in turn, building a model on the remaining points, and using that to predict the left-out one. SSY is the sum of squares of the response values or variance.

10.1.1. Comparison with Nasal Pungency Threshold

There are 48 overlapping VOCs for which a plot of observed log (1/ODT) against observed log (1/NPT) is given in Figure 10.2. Note that linalool and acetic acid were not included in the comparison graph as their observed log (1/NPT) were found out of the range, see chapter 9. In general there seems to be a robust indication of a systematic relationship between ODT and NPT. However, one group of outliers was highlighted. The aldehydes and carboxylic acids are actually more potent in terms of ODT values than would be expected from the line of fit, except for formic acid that is less potent. Moreover, since the NPT values for the aldehydes and carboxylic acids can be correlated with simple physicochemical descriptors, this strongly suggests that the difference between the ODT and NPT processes is due to some extra effect in the ODT process.

10.1.2. Development of Models

As a first step, equation (10.1) was applied to all the VOCs except the carboxylic acids and aliphatic aldehydes. The VOCs, propanone, octan-1-ol, methyl acetate and t-butyl acetate were then also revealed to be outliers, i.e. (observed log (1/ODT) - calculated log (1/ODT)) > -1 log units. This left a total of 50 VOCs for the analysis. The frequency distribution of the descriptor **L** and the frequency distribution of the variable, log (1/ODT), for all the remaining VOCs are given in Figure 10.3. Application of equation (10.1) to the 50 VOCs yielded the correlation equation,

$$\log (1/ODT) = -5.154 + 0.533\,\mathbf{E} + 1.912\,\mathbf{S} + 1.276\,\mathbf{A} + 1.559\,\mathbf{B} + 0.699\,\mathbf{L} \qquad (10.4)$$

$n = 50,\ r^2 = 0.773,\ q^2 = 0.603,\ sd = 0.579,\ F = 28.7$

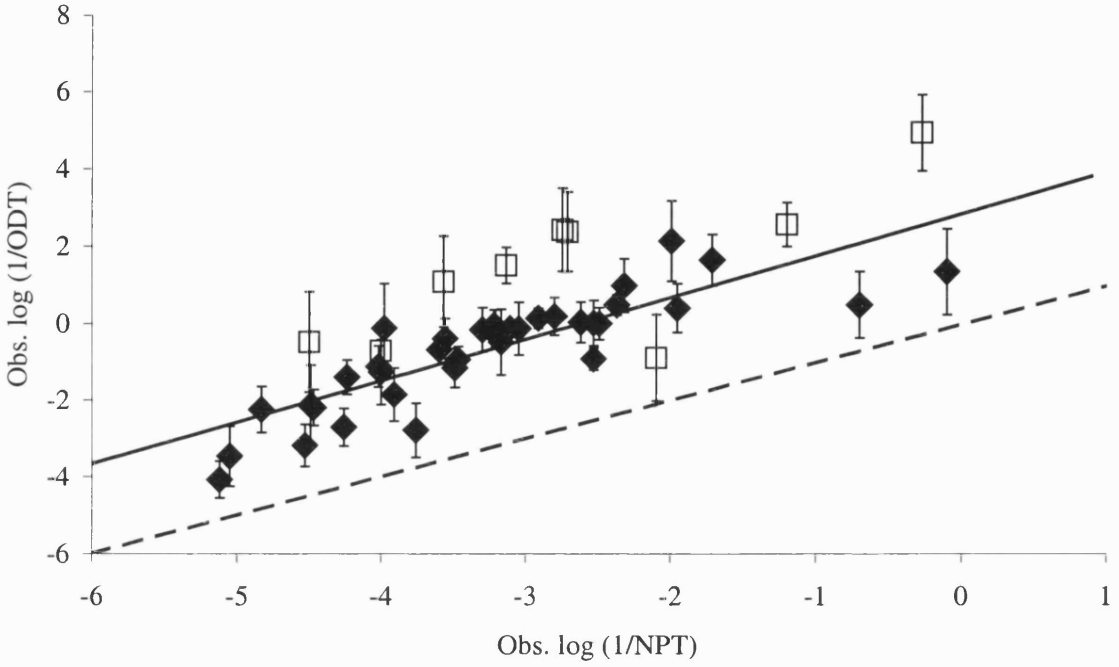


Figure 10.2. Plot of observed log (1/ODT) values against observed log (1/NPT) values. Empty squares: carboxylic acids and aldehydes, filled diamonds: remaining VOCs. ___ line of fit, ----- identity line.

A plot of observed values of log (1/ODT) against those calculated from equation (10.4) is shown in Figure 10.4. It is a great interest to compare the coefficients in the above equation with the coefficients for the other processes shown in Table 10.1. It can be seen that the coefficients in equation (10.1) are of the same sign and similar order of magnitude as those for transfer from the gas phase to organic solvents. For example, equation (10.4) compares well with the equations for gas / methanol or gas / wet octan-1-ol, as well as with the NPT equation (10.3). The relative significance of the descriptors in equation (10.4) was calculated from the product of the coefficient and the mean values of the corresponding descriptor. The percentage weights are then **E** (4%), **S** (20%), **A** (3%), **B** (12%) and **L** (61%). These values are almost identical to the ones for the nasal pungency equation. The term *l*.**L** is again the most important term. It therefore appears that simple transfer from the gas phase to a biophase must play a substantial role in the relationship of odor thresholds to the structure of VOCs, of the

order of 77% of the total effect. Inspection of Table 10.1 also leads to the conclusion that the aqueous mucus layer that covers the olfactory epithelium does not influence the process, because the equation for gas / water transfer[16] is completely different to equation (10.4). The latter equation is also in agreement with the finding that the odor receptor binding pocket, at least for the OR5 receptor, is a considerable distance away from the extracellular surface of the receptor[14].

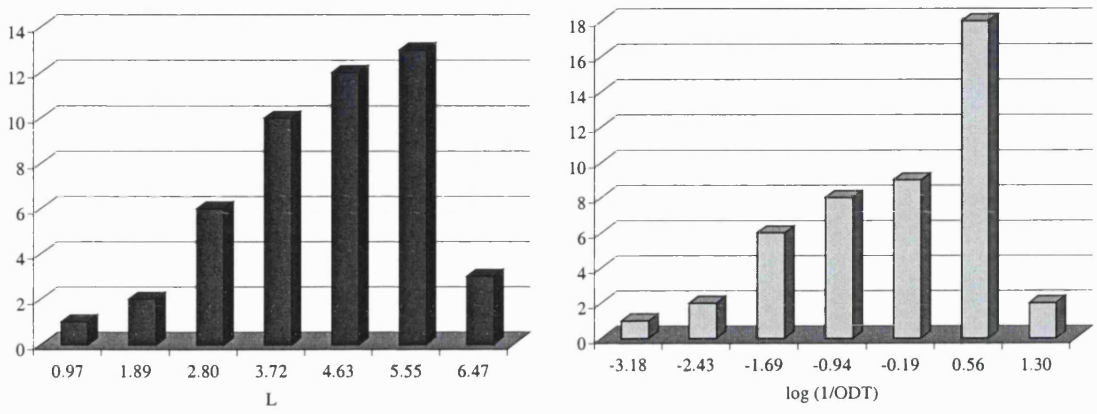


Figure 10.3. Frequency of distribution for the descriptor **L** and for the variable log (1/ODT).

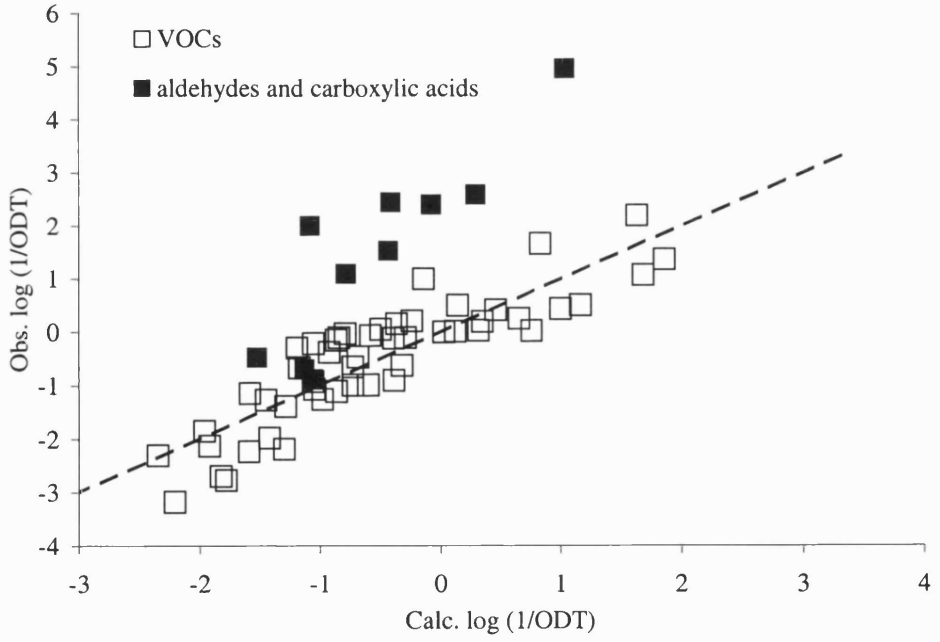


Figure 10.4. Plot of observed log (1/ODT) against calculated values on equation (10.4).

Table 10.1. Regression coefficients in equations (10.2) for gas / solvent (phase) partitions at 298 K

Phase	*e*	*s*	*a*	*b*	*l*
Wet octan-1-ol	0.002	0.709	3.519	1.429	0.859
Dry methanol	-0.215	1.173	3.701	1.432	0.769
Chloroform	-0.467	1.203	0.138	1.432	0.994
Acetone	-0.277	1.522	3.258	0.078	0.863
Dimethylformamide	-0.189	2.327	4.756	0.000	0.808
Water[a]	0.822	2.743	3.904	4.814	-0.213
Brain[a]	0.427	0.286	2.781	2.787	0.609
Muscle[a]	0.544	0.216	3.471	2.924	0.578
Nasal pungency[a]	0.000	2.154	3.522	1.397	0.860
Eye irritation[a]	-0.482	1.420	4.025	1.219	0.853
ODT, equation (10.4)[a]	0.533	1.912	1.276	1.559	0.699

[a] At 310 K.

In order to ascertain what other factors, as well as simple transport, influence the ODT values, it is instructive to plot the residuals in equation (10.4), i.e. (observed – calculated log (1/ODT)) against the **L** descriptor. Even more informative plot is of the residuals against the maximum length, **D**, of the VOC, see Figure 10.5. The latter was obtained by means of a computer-assisted molecular-modelling program called Molecular Modelling Pro ((TM) Revision 3.1 (1992) ChemSW software Inc.). The maximum value for **D** was retrieved after geometry optimisation (MM2 field). The residuals are not random, and both small VOCs and large VOCs are less potent than expected. In Figure 10.6, only the residuals for some homologous series are given, for clarity. It can be seen that the residuals follow a 'parabolic-like' curve: as molecular size increases, the residual values increases to a maximum value and then decreases. An example of maximum length determination is given in Figure 10.7.

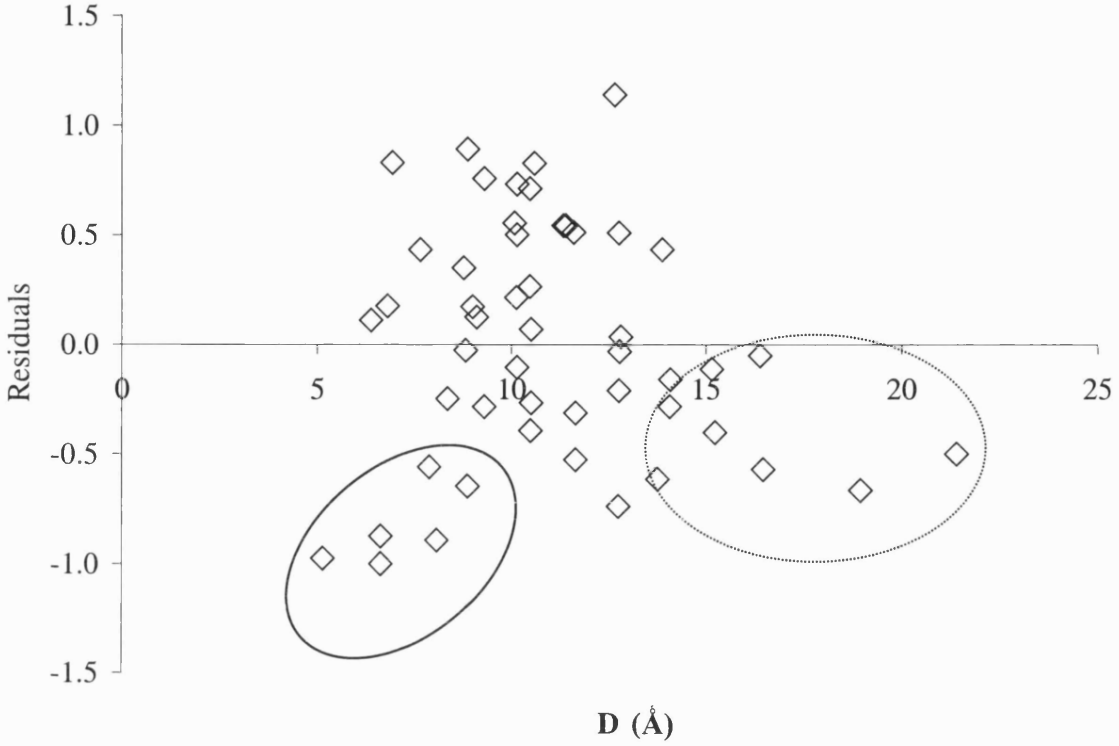


Figure 10.5. Plot of the residuals, observed log (1/ODT) – calculated Log (1/ODT) on equation 10.4 against the descriptor **L**. —— Small VOCs, ····· large VOCs.

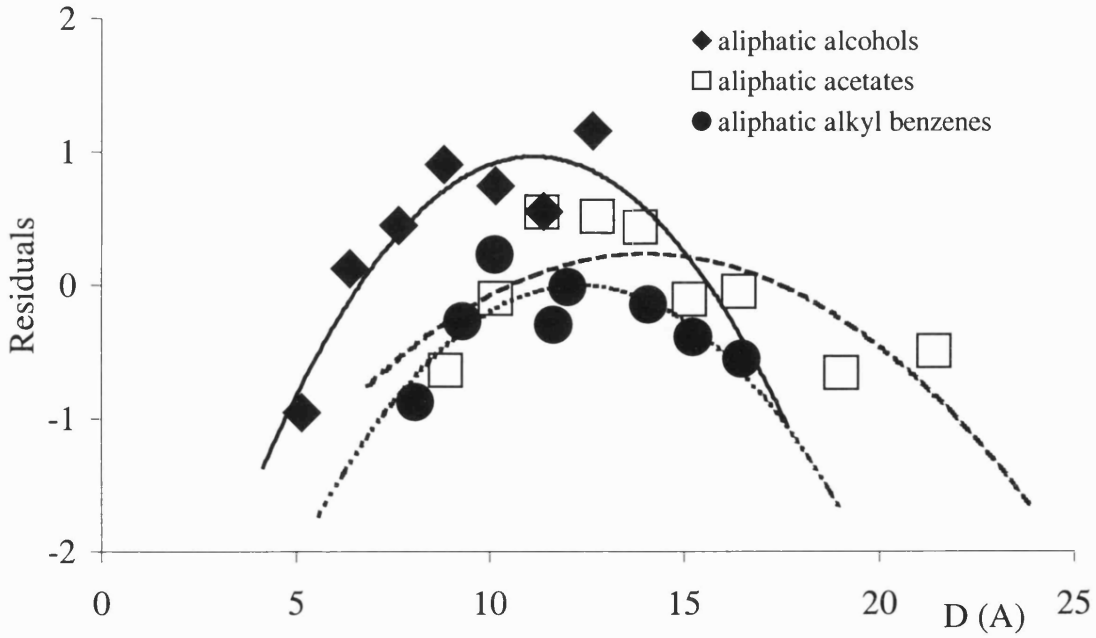


Figure 10.6. Residuals (observed - calculated log (1/ODT) values on equation (10.4) against the VOC maximum length.

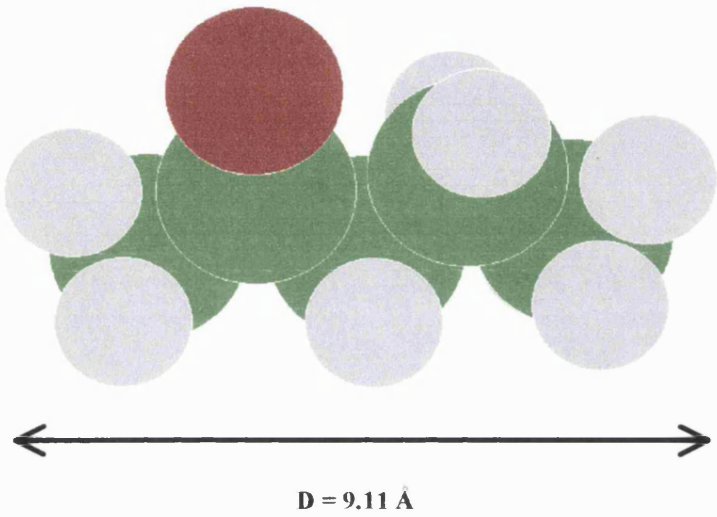


Figure 10.7. Example of maximum length determination, D, for pentan-2-one after geometry optimisation.

It was suggested that the pattern of residuals in Figure 10.6 is due to an extra effect, involving the size of the molecule, in addition to simple transfer. The effect can be quantified and incorporated into an equation for log (1/ODT) through addition of the quadratic terms **L*L** and **D*D**, see equation (10.5) and equation (10.6).

$$\text{Log (1/ODT)} = -7.109 + 0.096\,\mathbf{E} + 2.239\,\mathbf{S} + 2.221\,\mathbf{A} + 1.161\,\mathbf{B} + 1.742\,\mathbf{L} - 0.122\,\mathbf{L*L} \qquad (10.5)$$

$n = 50$, $r^2 = 0.833$, $q^2 = 0.687$, sd = 0.502, F = 44

$$\text{Log (1/ODT)} = -5.961 - 3.161\,\mathbf{E} + 4.210\,\mathbf{S} + 3.683\,\mathbf{A} - 0.844\,\mathbf{B} + 1.591\,\mathbf{L} - 0.015\,\mathbf{D*D} \quad (10.6)$$

$n = 50$, $r^2 = 0.861$, $q^2 = 0.713$, sd = 0.459, F = 44.5

By comparison with equation (10.4), the predictions are better if the quadratic term, **L*L,** is included. The *s*-, *a*-, and *b*- coefficients are close to those in equation (10.4). On the other hand, the *c*- and *l*-coefficients are larger because they depend on both the transfer of VOCs from the air stream to olfactory epithelium and some more specific interactions. It is also difficult to describe the **L*L** term. This is a squared value of a free energy term, **L**. There is no thermodynamic explanation for such situation. In a similar way, a regression analysis was carried out using the five usual solvation descriptors and the squared maximum distance, **D*D,** see equation (10.6). The statistics are reasonable, and such an equation could be used to predict ODT values of VOCs, except for carboxylic acids and aliphatic aldehydes. Although the statistical results are reasonable, equation (10.6) does not compare very well with the usual solvation equation such those in Table 10.2. The coefficients in equation 10.6 are all above their usual maximum values and the negative *b*-coefficient value is also chemically unrealistic for a transfer from the gas to the receptor phase. From these two studies, equation (10.5) and equation (10.6), it appears clearly that a quadratic 'size' term is required by the model. Equation (10.5) is more advantageous than equation (10.6) because the determination of a new parameter, such as **D**, is not required.

Another approach to the present problem was to insert a quadratic equation describing the distribution of the residuals for the 50 VOCs over the maximum length range shown in Figure 10.6. The equation is as follows,

$$\text{Residuals} = -1.600 + 0.297\,\mathbf{D} - 0.013\,\mathbf{D*D} \quad (10.7)$$

Adding equation (10.7) to the solvation equation (10.4), one obtains,

$$\text{Log (1/ODT)} = -6.757 + 0.533\,\mathbf{E} + 1.912\,\mathbf{S} + 1.276\,\mathbf{A} + 1.559\,\mathbf{B} + 0.699\,\mathbf{L} + 0.297\,\mathbf{D} - 0.013\,\mathbf{D*D} \quad (10.8)$$

$n = 50$, $r^2 = 0.82$, sd = 0.511

A plot of observed vs. calculated log (1/ODT) is given in Figure 10.8.

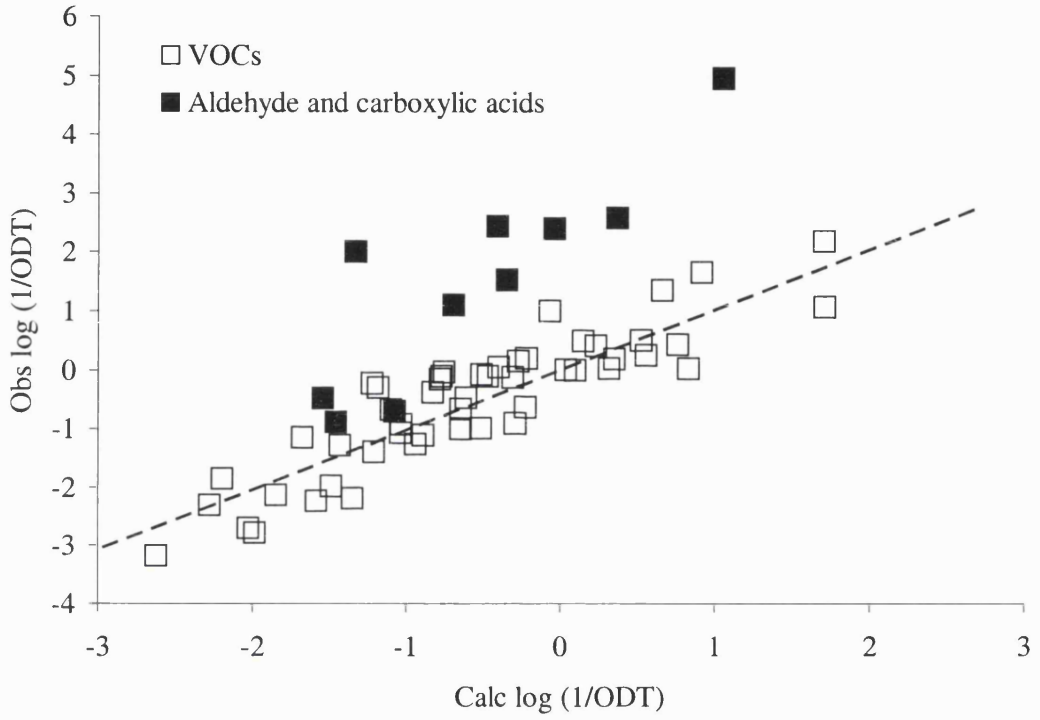


Figure 10.8. Plot of observed log (1/ODT) against calculated log (1/ODT) on equation (10.8).

The statistics of equations (10.5), (10.6) and (10.8) are reasonably good, bearing in mind the experimental error in ODT values, viz. 0.63 log units. The next step was to include the carboxylic acids and aliphatic aldehydes into the regression equation by means of an indicator variable, **H**, chosen as 2.0 for the carboxylic acids and aliphatic aldehydes and zero for all others VOCs,

$$\text{Log (1/ODT)} = -7.503 - 0.340\,\mathbf{E} + 1.652\,\mathbf{S} + 2.104\,\mathbf{A} + 1.500\,\mathbf{B} + 0.822\,\mathbf{L} + 0.381\,\mathbf{D} - 0.016\,\mathbf{D^*D} + 1.000\,\mathbf{H} \qquad (10.9)$$

$n = 50$, $r^2 = 0.84$, sd = 0.608

Equation (10.9) is a general equation for log (1/ODT) values, and could be used to predict further values to about 0.6 log units, of order of experiment error. Four compounds are again outliers to equation (10.9), viz. propanone, methyl acetate, *tert*-butyl acetate and octan-1-ol. Such an equation can be used as the basis of a model of odor detection.

On the other hand, a predictive equation can be constructed by using a parabolic term in **L**, rather than in **D**.

$$\text{Log (1/ODT)} = -7.720 - 0.060\,\mathbf{E} + 2.080\,\mathbf{S} + 2.829\,\mathbf{A} + 1.139\,\mathbf{B} + 2.028\,\mathbf{L} - 0.148\,\mathbf{L^*L} + 1.000\,\mathbf{H} \qquad (10.10)$$

$n = 50$, $r^2 = 0.85$, $q^2 = 0.536$, $sd = 0.598$

In equation (10.10) the indicator variable for aldehydes and carboxylic acids takes the value **H** = 1.6. Observed log (1/ODT) values were plotted against calculated log (1/ODT) values on equation (10.9) and on equation (10.10), see Figure 10.9 and Figure 10.10 respectively. The advantage of equation (10.10) over equation (10.9) is there is no need for further parameter determination, **L*L** is trivially measured from the descriptor **L**.

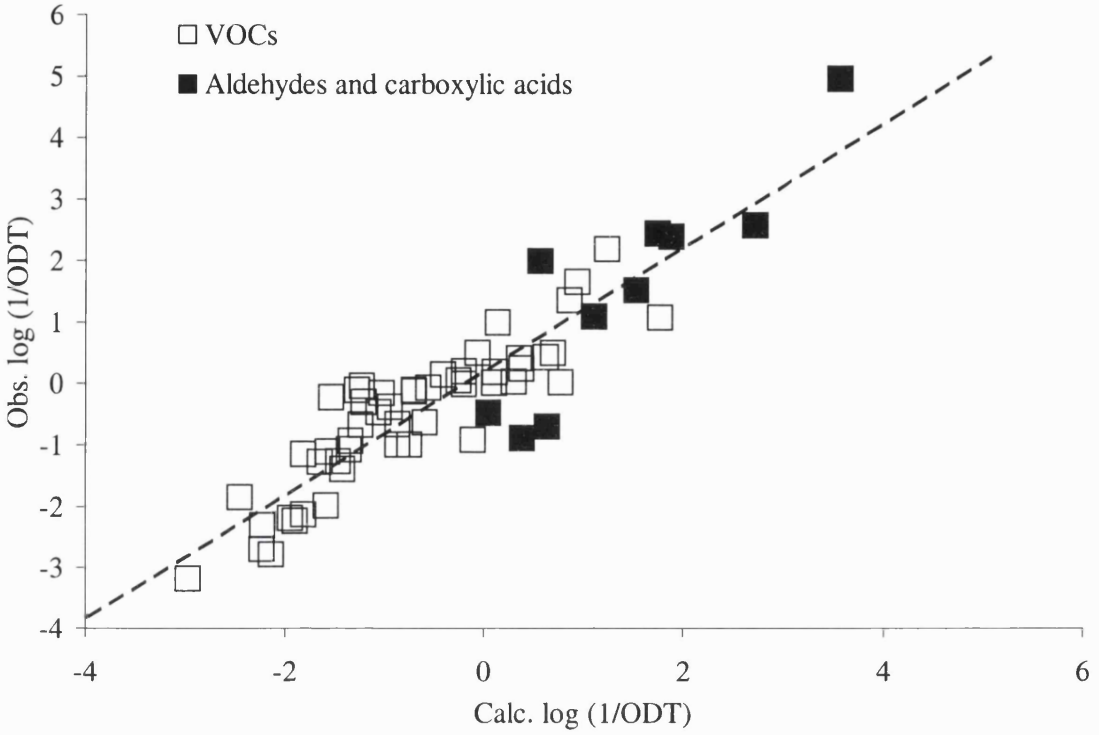


Figure 10.9. Plot of observed values of log (1/ODT) against calculated log (1/ODT) on equation (10.9).--- Regression line.

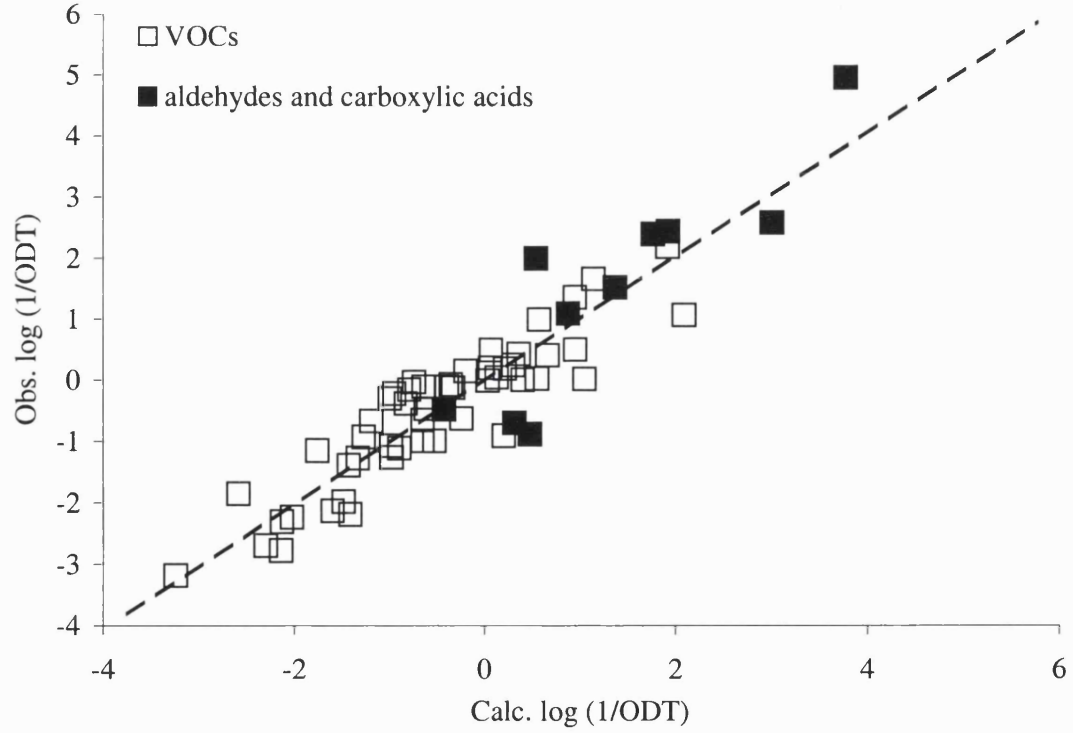


Figure 10.10. Plot of observed values of log (1/ODT) against calculated log (1/ODT) on equation (10.10).--- Regression line.

10.2. Discussion

Models for olfactory stimulation have been developed in the present work. A summary of the equations obtained here is given in Table 10.2. Equation (10.9) is not only a predictive equation, but can be considered to be compatible with the model shown in Figure 10.1. A large part of the variation in log (1/ODT) values with the structure of the VOCs is due to simple transport of the VOC from the gas phase to a biophase. In addition, there is an effect due to the size of the VOC, specifically to the maximum length. The potency of VOCs in a homologous series reaches a maximum at a length of around 11-12 Å. Now this length is almost the same as the maximum dimension of the central pocket in OBPs, viz. 11Å [13]; the alternative volume of Bianchet et al.[11] suggests a maximum length of the central pocket of 12-13 Å. Thus, one possible mechanism includes simple transfer from the gas phase to a biophase mediated by transport by OBPs, see mechanisms **i** and **ii** in Figure 10.1. The exceptions are the adehydes and carboxylic acids that are more potent than calculated by about a factor 100. In any case, it cannot be concluded that there is only OBP or even one type of OBP; there may be several types with maximum dimensions around 10-15 Å.

Of course, the above is not the only mechanism that fits our data analysis. It is possible that the OBPs have no discrimination at all, and that the maximum length effect take place on activation of the receptor. However, at least two types of interaction contribute to the overall threshold effect.

Vincent et al.[27] have recently determined complexation constants for a number of VOCs with porcine OBP. Details are in Table 10.3, with the complexation constant given as log ($1/IC_{50}$). Over the seven VOCs studies, values of log (1/IC50) vary by 0.75 log unit, whereas log (1/ODT) varies by no less than 3.99 units. It is therefore possible that the effect of OBPs is not the prime reason for the variation of log (1/ODT), but that the complexation to OBPs (or possibly the rate of complexation to OBPs) just mediates the effect of transport to, and interactions with, the receptor.

Equation (10.8) has other consequences, including the effect of homologues. Here, homologous series for acetates and alkyl benzenes are investigated. Descriptors for the higher homologues in these two series are given in Table 10.5 in section 10.4. The linear dependence of log (1/ODT) on **L**, as in Equation (10.4), would lead to a regular increase in log (1/ODT) along a homologous series, as shown in Figure 10.11 and Figure 10.12.

Table 10.2. Equations developed in the present work.

Eq.	*c*	*e*	*s*	*a*	*b*	*l*	*l*l*	*d*	*d*d*	*h*	n	r^2	q^2	sd
10.4	**-5.154**	**0.533**	**1.912**	**1.276**	**1.559**	**0.699**	**-**	**-**	**-**		**50**	**0.773**	**0.603**	**0.579**
	-0.410	*0.454*	*-0.623*	*-0.776*	*-0.731*	*-0.072*								
10.5	**-7.109**	**0.096**	**2.239**	**2.221**	**1.161**	**1.742**	**-0.122**	**-**	**-**		**50**	**0.833**	**0.687**	**0.502**
	0.609	*0.409*	*0.546*	*0.713*	*0.641*	*0.271*	*0.031*							
10.6	**-5.961**	**-3.161**	**4.210**	**3.683**	**-0.844**	**1.591**	**-**	**-**	**-0.015**	**-**	**50**	**0.860**	**0.713**	**0.459**
	0.361	*0.799*	*0.665*	*0.772*	*0.743*	*0.181*			*-0.015*					
10.8	**-6.757**	**0.533**	**1.912**	**1.276**	**1.559**	**0.699**	**-**	**2.971**	**-0.013**	**-**	**50**	**0.820**	**-**	**0.511**
	-	*0.454*	*0.623*	*0.776*	*0.731*	*0.072*		*0.118*	*0.005*					
10.9	**-7.503**	**0.340**	**1.652**	**2.104**	**1.500**	**0.822**	**-**	**0.381**	**-0.016**	**1.000**[a]	**60**	**0.840**	**-**	**0.608**
	-	*0.524*	*0.681*	*0.561*	*0.732*	*0.076*		*0.124*	*0.005*	*0.153*				
10.10	**-7.720**	**-0.060**	**2.080**	**2.829**	**1.139**	**2.028**	**-0.148**	**-**	**-**	**1.000**[b]	**60**	**0.850**	**0.536**	**0.598**
	0.647	*0.461*	*0.595*	*0.511*	*0.637*	*0.283*	*0.034*			*0.168*				

[a] With H = 2.0

[b] With H = 1.6

However, the parabolic dependence on ($\mathbf{D} - \mathbf{D}^2$), as in equation (10.8), considerably modifies the linear increase and results in the prediction shown in Figure 10.11 and Figure 10.12. The values of log (1/ODT) gradually become smaller than expected from the linear relationship, and eventually even begins to decrease. This corresponds to a chemical cut-off in potency, a prediction that is completely outside the scope of previous analyses[1,4]. This predicted cut-off effect has a very important consequence. Hau *et al.*[28] have used their partition model[1] to predict odor thresholds for VOCs found in the indoor environment. As pointed out above, these partition models do not include any cut-off effect at all, and hence higher homologues will be predicted to be more potent than on our model.

Table 10.3. Comparison of complexation of VOCs with porcine OBPs, and odor thresholds

VOC	Log (1/ODT) [a]	Log ($1/IC_{50}$) [b]
Benzyl benzoate	4.58	-0.59
Benzophenone	4.25	-0.56
Thymol	2.40	-0.40
2-Isobutyl-3-methoxy pyrazine	1.27	0.05
Undecanal	0.73	0.16
Dihydromyrcenol	0.59	0.10

[a] Equation (10.8). [b] From ref. 27

Another, very important, consequence follows from the initial Equation (10.4). The dependent variable, log (1/ODT), conceptually takes the place of the dependent variable, log K, where K is a gas / biophase equilibrium constant given by

$$K = \frac{\text{[number of molecules of VOC in the biophase]}}{\text{[number of molecules of VOC in the gas phase]}} \qquad (10.11)$$

The ODT value itself represents the number of molecules in the gas phase, so that the only way that 1/ODT can take the place of an equilibrium constant, K, is if the number of molecules of a VOC in the biophase in equilibrium with the gas phase threshold value of the VOC, is the same for each VOC. This is a more general conclusion than the supposition of Hau and Connell[1] that the minimum proportion of available receptors necessary for the detection of odors is the same for all members of a homologous series, but differs from series to series.

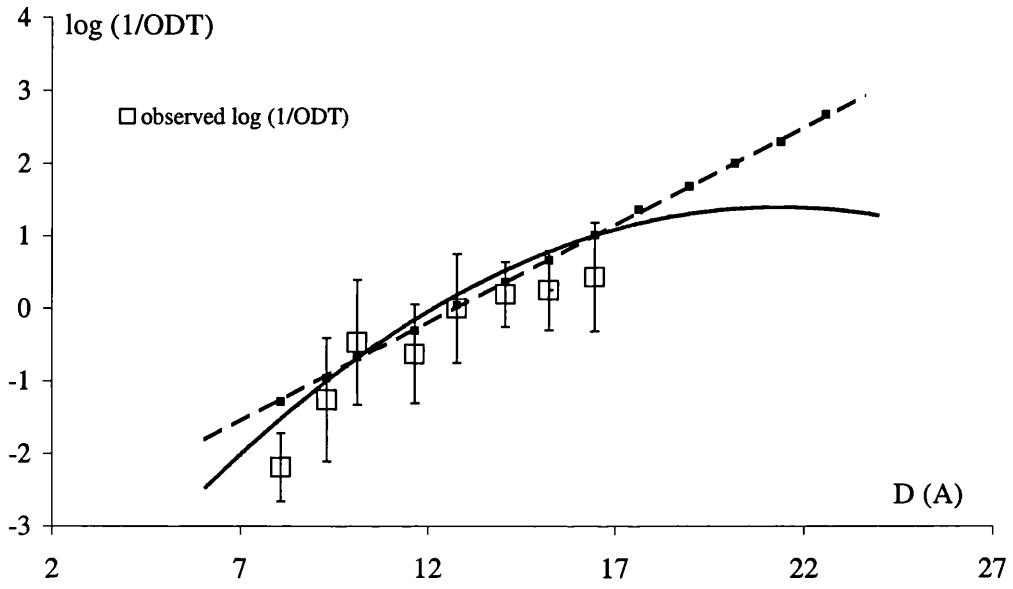


Figure 10.11. Plot of observed values of log (1/ODT) against the VOC maximum length, for the homologous series of alkyl benzenes. ---- Calculated values on equation (10.4); —— Calculated values on equation (10.9).

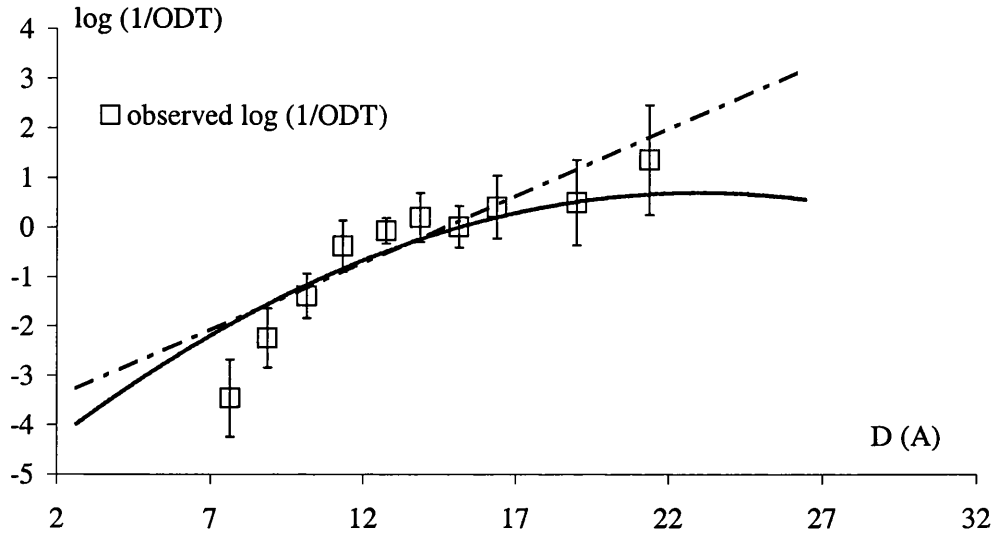


Figure 10.11. Plot of observed values of log (1/ODT) against the VOC maximum length, for the homologous series of acetates. ---- Calculated values on equation (10.4); ——Calculated values on equation (10.9)

The necessity for the use of an indicator variable for aldehydes and carboxylic acids arises because these two sets of compounds are more potent than predicted by equation (10.9). There is precedent for the extra potency of aldehydes and carboxylic acids. Alarie et al.[29] have shown that these compounds are more potent than expected in sensory irritation in mice, and suggest that they undergo some actual chemical reaction. However, aldehydes and carboxylic acids (except acetic acid) fit the general equation for nasal pungency thresholds without use of any indicator variable, see equation (10.4). There is also the problem of the four outliers, propanone, methyl acetate, t-butyl acetate and octan-1-ol. There may be extra experimental error with the first three compounds. Loss of propanone and methyl acetate due to their high volatility would result in the compounds appearing to be of lower potency. In the case of t-butyl acetate, the compound seemed to form an emulsion in some experiments, and this would result in an erroneous estimation of the ODT value. However, no explanation for the increased octan-1-ol ODT value is proposed.

Very recently, the EVA spectral descriptor has been applied to a selection of ODT values[30]. No details were given other than for 52 ODT values, q^2 was 0.57 and for

44 ODT values q^2 was 0.71; unfortunately EVA results cannot be interpreted in any chemical way so cannot lead to any mechanistic conclusions.

The odor perception of enantiomers is well known, but invariably in terms of odor quality.[31,32] Rossiter[32] and Laska et al.[33] list pairs of enantiomers that elicit different sensations of odor quality. The latter workers tested odor discrimination of 10 pairs of enantiomers and concluded that within their experimental procedure, differences in odor intensity played little or no part in discrimination of the two enantiomeric forms. Other workers have shown that ODTs for *R*(+)- and *S*(-)-nicotine are essentially the same.[34] This again suggests selective, rather than specific, transport of VOCs to the biophase.

10.2. Conclusion

Two models for olfactory stimulation have been put forward in this work. First, it was shown that a large variation in log (1/ODT) values with the structure of the VOCs is due to simple transport of the VOC from the gas phase to a biophase. So that discrimination amongst VOCs is selective but is not very specific. In addition, there is an effect that it due to the size of the VOC, as measured by the maximum length. One possible mechanism is that the complexation to OBPs mediates the effect of transport to, and interactions with, the receptor phase.

10.4 Appendix

Table 10.4. Values of log (1/ODT) with ODT in ppm and VOC descriptors used in the present work

Solute	Log (1/ODT)	E	S	A	B	L	D (Å)
Methanol	-3.180	0.278	0.440	0.430	0.470	0.970	5.150
Ethanol	-1.850	0.246	0.420	0.370	0.480	1.485	6.378
Propan-1-ol	-1.150	0.236	0.420	0.370	0.480	2.031	7.649
Propan-2-ol	-2.700	0.212	0.360	0.330	0.560	1.764	6.634
Butan-1-ol	-0.300	0.224	0.420	0.370	0.480	2.601	8.882
Butan-2-ol	-1.980	0.217	0.360	0.330	0.560	2.338	7.890
2-Methylpropan-1-ol	-2.780	0.180	0.300	0.310	0.600	1.963	6.638
Pentan-1-ol	-0.110	0.219	0.420	0.370	0.480	3.106	10.146
Hexan-1-ol	0.050	0.210	0.420	0.370	0.480	3.610	11.396
Heptan-1-ol	1.000	0.211	0.420	0.370	0.480	4.115	12.654
Heptan-4-ol	-0.910	0.180	0.360	0.330	0.560	3.850	11.650
Octan-1-ol	2.150	0.199	0.420	0.370	0.480	4.619	13.910
Pyridine	-0.110	0.631	0.840	0.000	0.520	3.022	6.814
Methyl acetate	-3.460	0.142	0.640	0.000	0.450	1.911	7.650
Ethyl acetate	-2.240	0.106	0.620	0.000	0.450	2.314	8.870
Propyl acetate	-1.390	0.092	0.600	0.000	0.450	2.819	10.154
Butyl acetate	-0.380	0.071	0.600	0.000	0.450	3.353	11.340
Pentyl acetate	-0.070	0.067	0.600	0.000	0.450	3.844	12.760
Hexyl acetate	0.200	0.056	0.600	0.000	0.450	4.351	13.880
Heptyl acetate	0.010	0.050	0.600	0.000	0.450	4.865	15.150
Octyl acetate	0.410	0.029	0.600	0.000	0.450	5.364	16.395
Decyl acetate	0.500	0.033	0.600	0.000	0.450	6.373	18.940
Dodecyl acetate	1.360	0.012	0.600	0.000	0.450	7.381	21.380
Propanone	-4.070	0.179	0.700	0.040	0.490	1.696	6.612
Pentan-2-one	-0.930	0.143	0.680	0.000	0.510	2.755	9.110
Heptan-2-one	0.150	0.123	0.680	0.000	0.510	3.760	11.610
Nonan-2-one	0.030	0.119	0.680	0.000	0.510	4.735	14.120
Toluene	-2.190	0.601	0.520	0.000	0.140	3.325	8.080
Ethylbenzene	-1.260	0.613	0.510	0.000	0.150	3.778	9.303
Propylbenzene	-0.470	0.604	0.500	0.000	0.150	4.230	10.124
Butylbenzene	-0.630	0.600	0.510	0.000	0.150	4.730	11.650
Pentylbenzene	-0.004	0.594	0.510	0.000	0.150	5.230	12.778
Hexylbenzene	0.190	0.591	0.500	0.000	0.150	5.720	14.080
Heptylbenzene	0.250	0.577	0.480	0.000	0.150	6.219	15.231
Octylbenzene	0.430	0.579	0.480	0.000	0.150	6.714	16.466
Oct-1-ene	-2.310	0.094	0.080	0.000	0.070	3.568	12.808
Oct-1-yne	-2.130	0.155	0.220	0.090	0.100	3.521	12.771
Chlorobenzene	-1.110	0.718	0.650	0.000	0.070	3.657	8.360
2-Phenylethanol	2.190	0.811	0.910	0.300	0.640	4.628	10.090
s-Butyl acetate	-0.670	0.044	0.570	0.000	0.470	3.054	10.149
t-Butyl acetate	-0.110	0.025	0.540	0.000	0.470	2.802	8.943
Butanal	-0.477	0.187	0.650	0.000	0.450	2.270	8.44

Solute	Log (1/ODT)	E	S	A	B	L	D (Å)
Pentanal	-0.699	0.163	0.650	0.000	0.450	2.851	9.690
Hexanal	1.097	0.146	0.650	0.000	0.450	3.357	10.950
Heptanal	1.523	0.140	0.650	0.000	0.450	3.865	12.200
Octanal	2.398	0.160	0.650	0.000	0.450	4.361	13.460
Formic acid	-0.886	0.300	0.790	0.720	0.340	1.400	5.260
Acetic acid	2.000	0.265	0.650	0.610	0.440	1.750	6.298
Butanoic acid	2.444	0.210	0.620	0.600	0.450	2.830	8.790
Hexanoic acid	2.585	0.174	0.600	0.600	0.450	3.920	10.290
Octanoic acid	4.959	0.150	0.600	0.600	0.450	5.000	13.800
Menthol	1.660	0.400	0.500	0.230	0.580	5.177	10.590
Cumene	-0.033	0.602	0.490	0.000	0.160	4.084	9.300
p-Cymene	-0.121	0.607	0.490	0.000	0.190	4.590	10.476
Δ-3-Carene	-0.223	0.511	0.220	0.000	0.100	4.649	6.930
Linalool	0.022	0.398	0.550	0.200	0.670	4.794	12.749
1,8-Cineole	0.495	0.383	0.330	0.000	0.760	4.688	8.788
Geraniol	1.070	0.513	0.632	0.390	0.660	5.479	13.749
α-Terpinene	-0.152	0.526	0.250	0.000	0.150	4.715	10.477
γ-Terpinene	-0.992	0.497	0.320	0.000	0.200	4.815	10.499
α-Pinene	-1.277	0.446	0.140	0.000	0.120	4.308	9.000
β-Pinene	-1.070	0.530	0.240	0.000	0.190	4.394	8.828
(R) (+) Limonene	-0.994	0.488	0.280	0.000	0.450	4.725	9.550
(S) (+) Limonene	-0.659	0.488	0.280	0.000	0.450	4.725	9.550

Table 10.5. Descriptors for higher homologous

VOC	E	S	A	B	L	D (Å)
Tridecyl acetate	0.000	0.600	0.000	0.450	7.878	22.670
Tetradecyl acetate	0.000	0.600	0.000	0.450	8.380	23.910
Pentadecyl acetate	0.000	0.600	0.000	0.450	8.883	25.180
Hexadecyl acetate	0.000	0.600	0.000	0.450	9.386	26.430
Nonylbenzene	0.578	0.480	0.000	0.150	7.212	17.640
Decylbenzene	0.579	0.470	0.000	0.150	7.708	18.980
Undecylbenzene	0.579	0.470	0.000	0.150	8.159	20.180
Dodecylbenzene	0.571	0.470	0.000	0.150	8.600	21.390
Tridecylbenzene	0.570	0.470	0.000	0.150	9.132	22.590
Tetradecylbenzene	0.570	0.470	0.000	0.150	9.619	23.950

10.5 References

1. K.M. Hau, D.W. Connell, Indoor Air 8 (1998) 23.
2. P. Laffort, F. Patte, J.Chromatogr. 406 (1987) 51.
3. M. Chastrette, SAR and QSAR in Environmental Res., 6 (1997) 215-254.
4. T. Yamanaka, Chem. Senses 29 (1995) 471.
5. M. Davos, F. Patte, J. Roualt, P. Laffort, L.J. Van Gemert, Standardized Human Olfactory Thresholds, IRL Press, Oxford (1990). M.H. Abraham, Chem.Soc.Revs., 22 (1993) 73.
6. A. Dravnieks, Ann.New York Acad.Sci., 237 (1974) 144.
7. AlHA , Odor thresholds for chemicals with established occupational health standards. Ohio, American Industrial Hygiene Association (1989).
8. T.C. Pearce, J.W. Gardner, W. Gopel, Sensors Update Vol3., Eds. Baltes,H., Gopel,W. and Hesse,J., Wiley-UCH, (1998) 73.
9. D.E. Hornung, M.M. Mozell, Biochemistry of Taste and Olfaction, Eds. Cagan,R.H and Kare,M.R, New York; London: Academic Press, (1981) 31.
10. S.H. Snyder, P.B. Sklar, J. Pevsner, J.Biol . Chem., 263 (1988) 13971.
11. M.A. Bianchet, G. Bains, P. Pelosi, J. Pevsner, S.H. Snyder, H.L. Monaco, L.M. Amzel, Nature Struct. Biol., 3 (1996) 934.
12. S. Brownlow, L. Sawyer, Nature Struct. Biol., 3 (1996) 902.
13. M. Tegoni, R. Ramoni, E. Bignetti, S. Spinelli, C. Cambillau, Nature Struct. Biol., 3 (1996) 863.
14. M.S. Singer, G.M. Shepherd, Neuroreport, 5 (1994) 1297.
15. M.H. Abraham, Chem. Soc., 22 (1993) 73.
16. M.H. Abraham, J. Andonian-Haftvan, G.S. Whiting, A. Leo, R.W. Taft, J.Chem.Soc., Perkin Trans.2, (1994)1777.
17. M.H. Abraham, G.S. Whiting, P.W. Carr, H. Ouyang, J.Chem.Soc., Perkin Trans.2, (1998) 1385.
18. M.H. Abraham, J.A. Platts, A. Hersey, A.J. Leo, R.W. Taft, J.Pharm.Sci.. 88 (1999) 670.
19. M.H. Abraham, J. Le, W.E.Jr. Acree, Collect. Czech. Chem.Comm., 64 (1999) 1748.
20. J.E. Cometto-Muniz, W.S. Cain, Physiol. Behav., 48 (1990) 719.
21. J.E. Cometto-Muniz, W.S. Cain, Pharmacol. Biochem. Behav., 39 (1991) 983.
22. J.E. Cometto-Muniz, W.S. Cain, Arch. Environ. Health, 48 (1993) 309.
23. J.E. Cometto-Muniz, W.S. Cain, Am.Ind.Hyg.Ass.J., 55 (1994) 811.
24. J.E. Cometto-Muniz, W.S. Cain, Chem. Senses, 20 (1995) 191.
25. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, Exp.Brain Res. 118 (1998) 180.

26. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, R. Kumarsingh, Pharmacol. Biochem.Behav. 60 (1998) 765.
27. F. Vincent, S. Spinelli, R. Raamolni, P. Pelosi, C. Cambillo, M. Tegoni, J.Mol.Biology 300 (2000) 127.
28. Y. Alarie, M. Schaper, G.D. Nielsen, M.H. Abraham, Arch.Toxicol., 72 (1998) 125.
29. K.M. Hau, D.W. Connell, B.J. Richardson, Regulatory Toxicology and Pharmacology, 31 (2000) 22.
30. D.B. Turner, P. Willett, Eur.J.Med.Chem., 35, (2000) 367.
31. D.H. Pybus, C.S. Sell, The chemistry of fragrances. Royal Society of Chemistry, London (1999).
32. K.J. Rossiter, Chem.Rev., 96 (1996) 3201.
33. M. Laska, A. Liesen, P. Teubner, Am.J.Physiol 277 (Regulatory Integrative Comp. Physiol. 46), (1999) R1098.
34. N. Thuerauf, M. Kaegler, R. Dietz, A. Barocka, G. Kobal, Psychpharmacol., 142 (1999) 236.

Chapter 11 Prediction of Chemosensory Effects of 'New' Volatile Organic Compounds

11.0. Introduction

Chemosensory effects of volatile organic compounds, VOCs, in humans include nasal pungency, eye irritation and odor response.[1] Quantitative structure-activity relationships, QSAR, based on the Abraham solvation equation for gas to condensed phase processes, have been developed for nasal pungency and eye irritation threshold[2], NPT and EIT respectively. A model for odor detection threshold, ODT, has also been proposed. The reciprocal NPT, EIT and ODT values were used in these models, so that the more potent the VOC the larger is the log (1/NPT), log (1/EIT) and log (1/ODT). Models for NPT and ODT values were derived from chapter 9 and 10 while the model for EIT was taken from reference 2 and are summarised below.

Nasal Pungency Threshold, NPT in ppm

$$\text{Log (1/NPT)} = -8.080 + 1.767\,\mathbf{S} + 3.298\,\mathbf{A} + 1.076\,\mathbf{B} + 0.857\,\mathbf{L} \quad (11.1)$$

$n = 48$, $r^2 = 0.950$, $sd = 0.270$

Eye irritation threshold, EIT in ppm

$$\text{Log (1/EIT)} = -7.918 - 0.482\,\mathbf{E} + 1.420\,\mathbf{S} + 4.025\,\mathbf{A} + 1.219\,\mathbf{B} + 0.853\,\mathbf{L} \quad (11.2)$$

$n = 54$, $r^2 = 0.928$, $sd = 0.360$

Odor detection threshold, ODT in ppm

$$\text{Log (1/ODT)} = -7.445 - 0.304\,\mathbf{E} + 1.652\,\mathbf{S} + 2.104\,\mathbf{A} + 1.500\,\mathbf{B} + 0.822\,\mathbf{L} + 0.369\,\mathbf{D} - 0.016\,\mathbf{D}^2 + 1.000\,\mathbf{H} \quad (11.3)$$

$n = 60$, $r^2 = 0.840$, $sd = 0.601$

Here and elsewhere, n is the number of data points, r is the correlation coefficient, sd is the standard deviation in the dependent variable. **E**, **S**, **A**, **B**, and **L** are the VOC Abraham solvation descriptors. **D** is the maximum length of the VOC. **H** is an indicator variable, chosen as 2.0 for the carboxylic acids and aldehydes and zero for all other VOCs.

The most important application of equations (11.1), (11.2) and (11.3) is that they enable the relevant chemosensory effects to be predicted by using the Abraham solvation descriptors **E, S, A, B** and **L**. Indeed the health impacts can be trivially calculated for a large variety of non reactive VOCs for which the solvation descriptors are available. There are, however, some caveats to the models. In particular, Cometto-Muniz and co-workers[3] have recently observed that the potency of higher homologous in a series becomes less than predicted. This result was attributed to a 'cut-off' effect. The cut-off is not just a manifestation of the very low saturated vapour pressure of the higher homologues, but seems to be a chemical cut-off that is related to the actual mechanism of irritation and olfaction. If higher homologous do exhibit cut-off effects, then the predicted values evoked by equation (11.1)-(11.2) will always be greater than observed, i.e. the predicted values will be lower than those observed. Note that the cut-off is encompassed in the model for olfactory stimulation, equation (10.3), as explained in chapter 10.

The predicted log (1/NPT), log (1/EIT) and log (1/ODT) values and hence the corresponding NPT, EIT and ODT give an insight into the propensity of a given VOC to generate chemsosensory effects in humans.[1] A VOC that is predicted to have a large log (1/NPT) value, and hence small NPT value, should be regarded as pungent. On the other hand VOCs generating no chemosensory effects would be characterised by small log (1/NPT) values and thus large NPT values. Three levels of nasal pungency were arbitrary chosen in this work based on the log (1/NPT) values measured by Cometto-Muniz and Cain:

- Weak nasal pungency: $\log(1/NPT) \leq -3.60$ e.g. Toluene
- Moderate nasal pungency: $-3.60 < \log(1/NPT) \leq -1.60$ e.g. Heptan-2-one
- Strong nasal pungency: $-1.60 < \log(1/NPT)$ e.g. octan-1-ol

A similar scale was chosen for the eye irritation process.

- Weak eye irritation: $\log (1/EIT) \leq -3.60$
- Moderate eye irritation: $-3.60 < \log (1/EIT) \leq -1.60$
- Strong eye irritation: $-1.60 < \log (1/EIT)$

For a typical VOC, the threshold for pungency or irritation lies above the odor threshold by as little as 1 or as many as 4 order of magnitude. The selected boundaries for odor detection thresholds were arbitrary chosen as follows.

- Weak odor response: $\log (1/ODT) \leq -1.50$
- Moderate odor response: $-1.50 < \log (1/ODT) \leq -0.50$
- Strong odor response: $-0.50 < \log (1/ODT)$

Once log (1/NPT), log (1/EIT) and log (1/ODT) values are obtained for a given VOC, the latter can be classified according to its ability to trigger chemosensory effects.

11.1. Prediction and Visualisation of Chemosensory Effects for Selected Compounds

Descriptors for eighteen refrigerants and 114 terpenes have been reported in this work. A further 250 VOCs, such as esters, lactones, diones, aliphatic alkanes and cyclic alkanes have been investigated, their descriptor values have been reported elsewhere.[4] In the present work, a selection of thirty-nine VOCs including terpenes, refrigerants, acyclic esters and lactones is investigated. An additional descriptor is needed for the calculation of ODT values. **D** is the maximum length of the VOC calculated through a computer-assisted molecular-modelling program (Molecular Modelling Pro (TM) Revision 3.1 (1992) ChemSW software Inc.). Descriptors for a selection of VOCs including are listed in Table 11.1, so that it is possible to calculate their log (1/NPT), log (1/EIT) and log (1/ODT) values. The values themselves are in Table 11.1 together with observed values for toluene, heptan-2-one and propan-1-ol.

The predicted values for each series of compounds in Table 11.1 may be plotted. Next, the calculated values for each data point is then used to colour code the respective point to represent a specific level of chemosensory effects as describes above. Block

colouring of the separation of the three categories, weak, moderate and strong effects provides a convenient way to observe the ability of a given VOC to generate chemosensory effects. Plots for log (1/NPT), log (1/EIT) and log (1/ODT) are given in Figures 11.1, 11.2 and 11.3 respectively.

It can be seen from these graphs that the terpene substances are predicted to have weak to strong pungency. The lactones are rather strong irritants and the selected acyclic esters can be regarded as moderate irritants. It appears clearly that the refrigerants are very weak irritants. Similar results are obtained for odor detection threshold but in a different concentration scale. Note that some of the compounds investigated are rather large, so that their estimated effects might be smaller than observed due to a cut-off effect.

In the next two sections, attention is focused to the chemosensory effects of terpenes and refrigerants.

11.1.1. Chemosensory Effects of Terpenes

Terpenes constitute a fairly large family of VOCs, which covers a wide spread of physicochemical properties. Hence, it is not surprising that the chemosensory effects of terpenes on humans vary considerably from weak effects to strong effects. The propensity of these effects is mainly driven by the lipophilicity of the terpene substances, as shall see now. The lower molecular weight terpene hydrocarbons, tricyclene and α-fenchene, are predicted to have small values of log (1/NPT), log (1/EIT) and log (1/ODT) and hence large NPT, EIT and ODT values. On the other hand, the large and very lipophilic terpenes, thujopsene and α-gurgujene, are predicted to have larger values of log (1/NPT), log (1/EIT) and log (1/ODT) and hence moderate chemosensory effects. This analysis shows that the greater the value of **L** for a given terpene hydrocarbon, the larger will be the chemosensory effects and therefore the stronger nasal pungency, eye irritation and odor response will the VOC have. In general, terpene compounds with polar functionality groups, such as alcohols, esters, ethers, ketone and aldehydes, will be predicted to have larger values of log (1/NPT) than the corresponding hydrocarbons. For instance, fenchol is predicted with to have larger values of log (1/NPT) than the terpene hydrocarbon α-fenchene, see Table 11.1. Similar results are obtained for prediction of EIT and ODT values. Terpene aldehydes and

ketones are quite potent, and the terpene alcohols and acetates are predicted to be very potent, i.e. to have very small pungency thresholds, e.g. citronellyl acetate.

As mentioned in chapter 8, descriptors are the same for enantiomers of a given terpene. The descriptors for R (+) limonene, as an example, will be the same for S (+) limonene. By analogy, either calculated value of log (1/NPT) or calculated value of log (1/EIT) or calculated log (1/ODT) values will be the same. This is in rather good agreement with experimental results. Cometto-Muniz and co-workers[5] have determined the ODT values in ppm for these two enantiomers. The ODT values are R (+) limonene (9.863) and S (+) limonene (4.560), so that the logarithmic values, 0.994 and 0.659 respectively, differ only by 0.335 log units. The enantiomeric differences in this particular case is not very large, and is actually less than the average standard deviation, sd, for all observed log (1/ODT), sd = 0.63 log units. There are no reports of nasal pungency thresholds and eye irritation thresholds in man for enantiomers of a given VOC. However Kasanen et al.[6] have measured the decrease in respiratory rate in mice, RD_{50}, for enantiomers of α-pinene and β-pinene. Again the enantiomeric differences was found almost negligible.

11.1.2. Chemosensory Effects of Refrigerants

Refrigerants have only a small chemosensory effect on man. Values of NPT, EIT and ODT are all larger than those for toluene that has very weak chemosensory effects. Furthermore, values of the saturated vapour pressure, in ppm at 298K [7,8] are also given in Table 11.2. For some refrigerants the saturated vapour pressure at 298K is less than the NPT and EIT values, in which case these will be no perceived effect, anyway. On the other hand, for most of the refrigerants, the saturated vapour pressure at 298K is larger than the ODT values, in which case there will be perceived effect.

Table 11.1. Estimation of chemosensory potency of the refrigerants

Code	NPT (ppm)	EIT (ppm)	ODT (ppm)	P^o (ppm)[a]
R32	8,220,000	5,200,000	105,000	16,700,000
R23	40,720,000	16,400,000	536,000	46,300,000
R14	1,673,530,000	509,000,000	16,380,000	
R11	812,000	848,000	3,700	1,050,000
R12	7,870,000	6,230,000	39,000	6,420,000
R13	95,980,000	48,530,000	660,000	35,200,000
R114	5,871,000	3,450,000	26,400	2,120,000
R115	67,100,000	28,300,000	300,000	9,000,000
R227	16,100,000	5,340,000	88,000	4,580,000
R1122	2,900,000	1,350,000	23,000	4,670,000
R1216	94,800,000	31,900,000	360,000	7,530,000
Toluene	29,500[b]	25,700[b]	155[b]	37,400
Propan-1-ol	3,090[b]	6,920[b]	14[b]	26,900
Heptan-2-one	812[b]	309[b]	1[b]	4,990
Octan-1-ol	98[b]	60[b]	0.00708[b]	99
Decyl acetate	5[b]	20[b]	0.0437[b]	30

[a] Taken from ref. [7,8]

[b] Observed values, ref. [1,9]

11.2. Conclusion

It has been shown in this chapter that it is trivial to estimate nasal pungency threshold, eye irritation threshold, and odor detection threshold values for any non reactive VOC for which descriptor values are known. Here a selection of VOCs have been investigated among the 382 VOCs whose descriptors have been determined in this work. Lactones, esters and terpenes have mostly strong potency to trigger nasal pungency, eye irritation and odor detection. However, it must be kept in mind that most of these compounds are rather big so that their predicted values might be smaller than the observed one. Finally, it was shown that refrigerants have no or very weak effects on humans.

11.3. Appendix

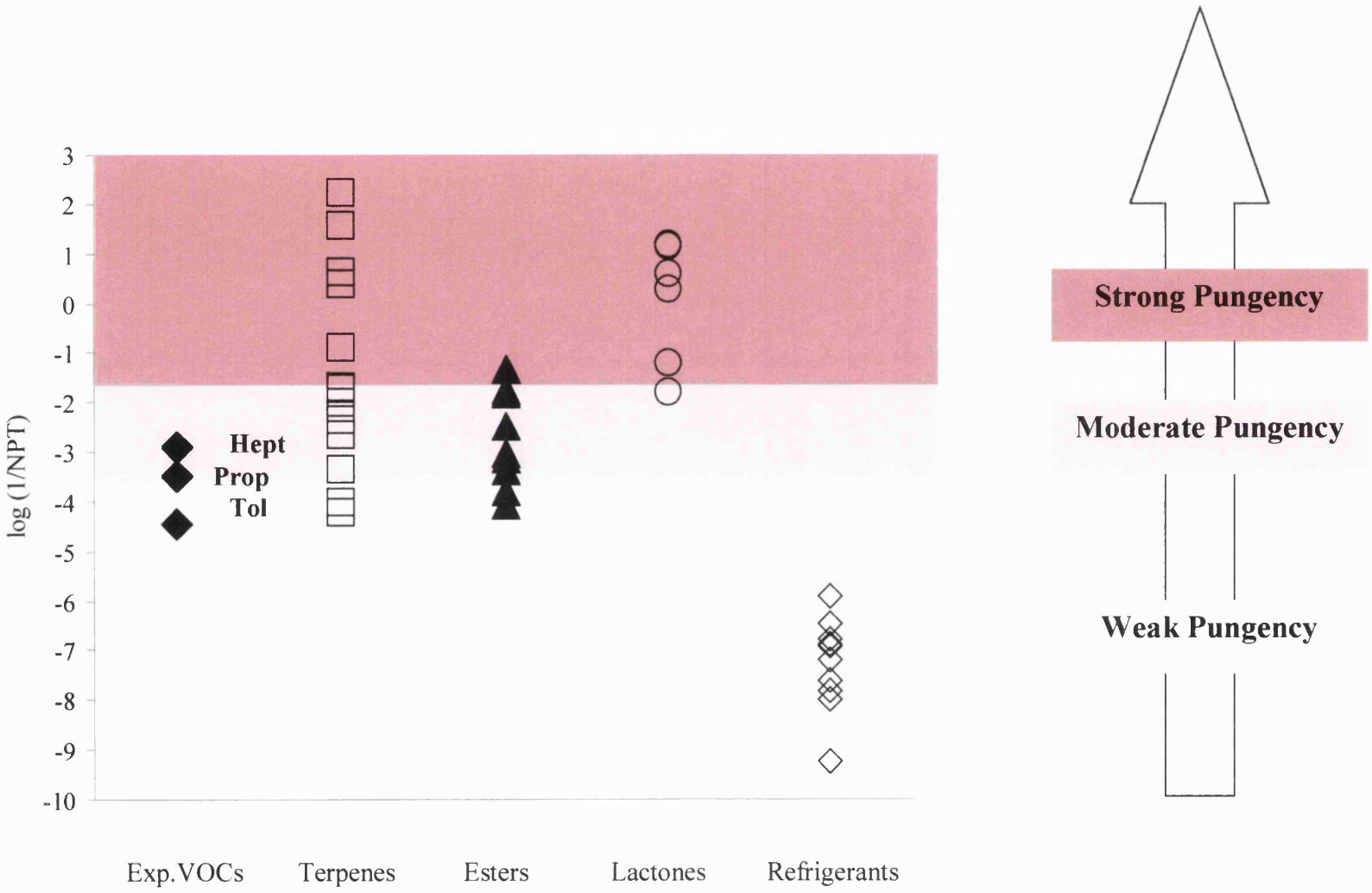


Figure 11.1. Nasal Pungency Scale for Selected VOCs. Log (1/NPT) calculated on equation (11.1)

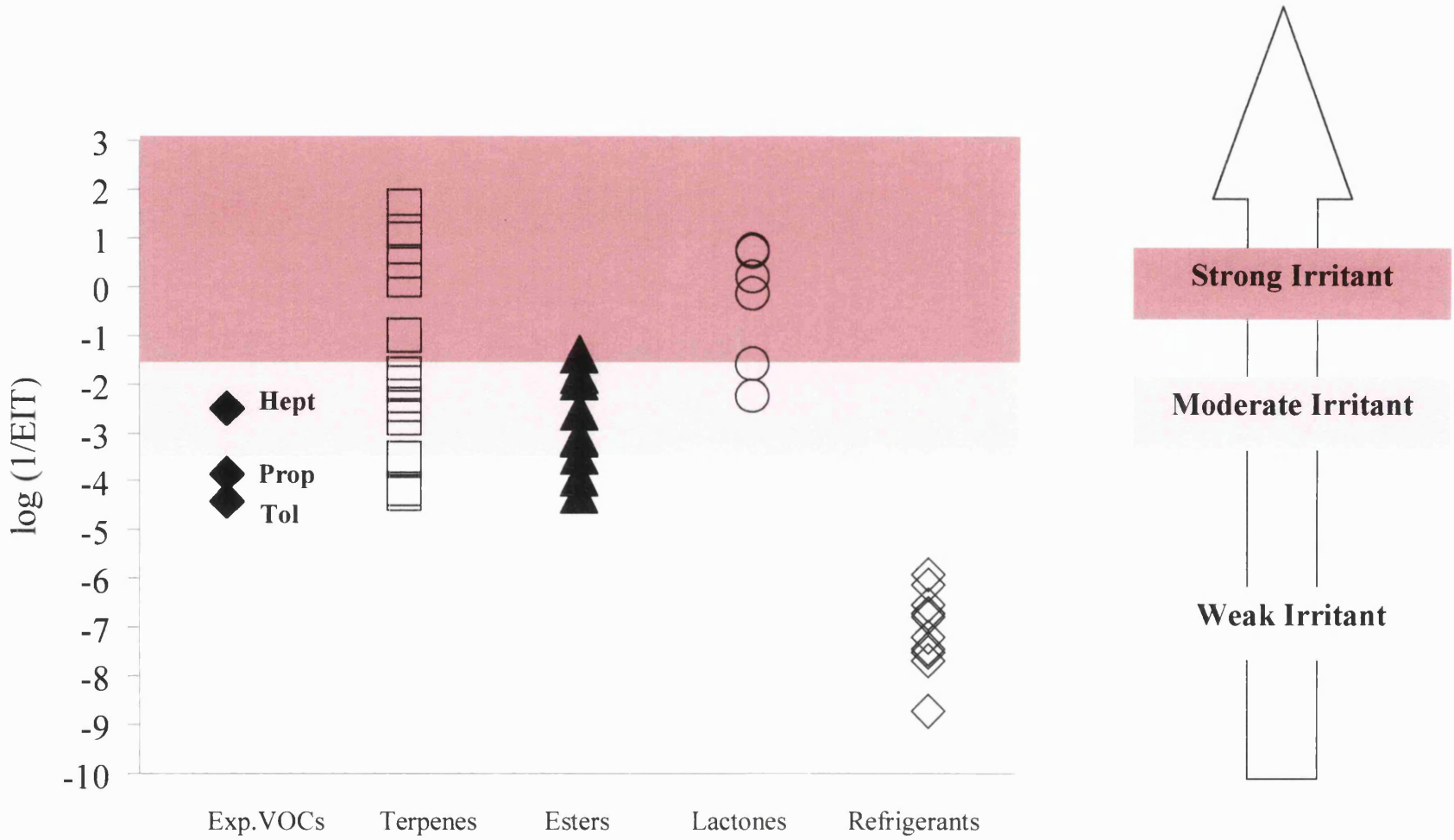


Figure 11.2. Eye irritation threshold Scale for Selected VOCs. Log (1/EIT) calculated on equation (11.2)

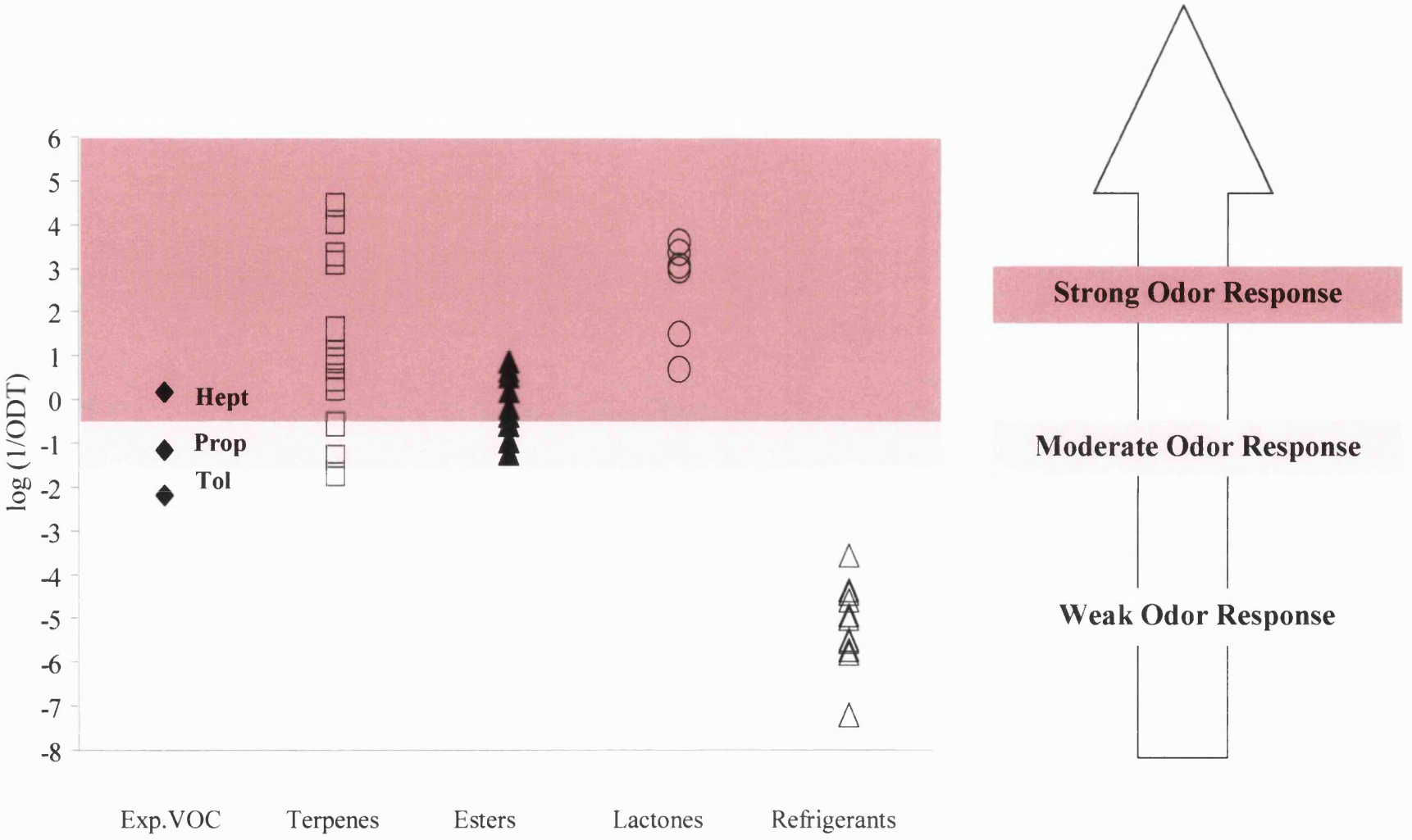


Figure 11.3. Odor Detection threshold Scale for Selected VOCs. Log (1/ODT) calculated on equation (11.3)

Table 11.2. Selection of VOCs with descriptors and predicted chemosensory effects on Humans.

VOCs	E	S	A	B	L	D[a]	log (1/NPT)[b]	log (1/EIT)[d]	log (1/ODT)[e]
Toluene	0.601	0.520	0.000	0.140	3.324	8.080	-4.470[c]	-4.463[c]	-2.190[c]
Propan-1-ol	0.236	0.420	0.370	0.480	2.031	7.649	-3.490[c]	-3.629[c]	-1.150[c]
Heptan-2-one	0.123	0.680	0.000	0.510	3.760	11.610	-2.910[c]	-3.183[c]	0.150[c]
Terpenes									
Tricyclene	0.303	0.150	0.000	0.000	4.211	7.910	-4.206	-4.259	-1.726
a-Fenchene	0.556	0.120	0.000	0.100	4.380	8.760	-4.007	-4.158	-1.323
Methyl tiglate	0.332	0.680	0.000	0.400	3.630	9.940	-3.337	-3.528	-0.550
1,8-Cineole	0.378	0.340	0.000	0.750	4.674	8.795	-2.667	-2.716	0.206
trans-Rose oxide	0.350	0.480	0.000	0.550	5.030	8.795	-2.329	-2.444	0.422
α-Gurgujene	0.737	0.100	0.000	0.100	6.762	10.070	-2.001	-2.241	0.746
Thujopsene	0.774	0.180	0.000	0.150	6.880	10.600	-1.704	-1.984	1.082
Fenchol	0.688	0.420	0.220	0.570	5.020	8.770	-1.697	-1.791	0.908
Myrtenol	0.643	0.700	0.300	0.450	5.224	10.457	-0.893	-1.022	1.616
Cedrol	0.836	0.500	0.160	0.600	7.480	10.284	0.387	0.145	3.123
α-Cadinol	0.640	0.500	0.180	0.600	7.696	10.710	0.638	0.504	3.297
Tonalid	0.698	0.920	0.000	0.700	8.482	12.070	1.568	1.140	4.432
Geranyl phenylacetate	0.876	1.040	0.000	0.550	9.210	19.510	2.242	1.663	4.044
Esters									
Methyl 3-butenoate	0.348	0.640	0.000	0.430	2.819	10.080	-4.071	-4.248	-1.226
Ethyl 3-butenoate	0.270	0.610	0.000	0.430	3.218	11.260	-3.782	-3.913	-0.939
Propyl 3-butenoate	0.223	0.600	0.000	0.430	3.719	12.550	-3.370	-3.477	-0.573
Isopropyl 2-ethylbutyrate	0.050	0.440	0.000	0.490	4.229	10.590	-3.151	-3.113	-0.378
Isopropyl 3-methyl-2-butenoate	0.230	0.540	0.000	0.500	4.162	11.330	-3.021	-3.102	-0.185
cis-3-Hexenyl propanoate	0.175	0.570	0.000	0.450	4.793	14.140	-2.481	-2.556	0.183
Hexyl 2-butenoate	0.176	0.630	0.000	0.420	5.469	16.290	-1.828	-1.931	0.540
trans 3-Hexenyl pentanoate	0.141	0.540	0.000	0.450	5.690	15.960	-1.765	-1.817	0.656
cis 3-Hexenyl hexanoate	0.128	0.560	0.000	0.450	6.161	17.200	-1.326	-1.381	0.872

VOCs	E	S	A	B	L	D[a]	log (1/NPT)[b]	log (1/EIT)[d]	log (1/ODT)[e]
Lactones									
γ-Butyrolactone	0.366	1.500	0.000	0.510	3.600	7.120	-1.796	-2.272	0.685
γ–Hexadecalactone	0.325	1.354	0.000	0.510	4.593	9.350	-1.203	-1.612	1.482
γ-Undecalactone	0.291	1.383	0.000	0.510	6.664	15.520	0.624	0.212	3.044
γ–Dodecalactone	0.298	1.440	0.000	0.510	7.169	16.770	1.157	0.720	3.371
δ-Decalactone	0.372	1.389	0.000	0.510	6.281	13.180	0.306	-0.146	2.975
δ-Dodecalactone	0.361	1.380	0.000	0.510	7.346	15.680	1.203	0.755	3.600
Refrigerants									
R32	-0.316	0.487	0.065	0.052	0.040	5.120	-6.915	-6.715	-5.019
R23	-0.427	0.183	0.110	0.034	-0.274	5.200	-7.592	-7.202	-5.729
R14	-0.550	-0.250	0.000	0.000	-0.819	5.190	-9.224	-8.707	-7.214
R11	0.207	0.240	0.000	0.070	1.950	6.430	-5.909	-5.928	-3.567
R12	0.027	0.125	0.000	0.000	1.124	6.420	-6.896	-6.795	-4.597
R13	-0.247	-0.046	0.000	0.000	0.209	5.820	-7.982	-7.686	-5.819
R114	-0.190	0.050	0.000	0.000	1.427	7.190	-6.769	-6.538	-4.421
R115	-0.360	-0.120	0.000	0.000	0.543	7.180	-7.827	-7.452	-5.482
R227	-0.557	0.012	0.070	0.030	0.688	7.690	-7.206	-6.727	-4.945
R1122	-0.340	0.285	0.150	0.000	0.723	7.080	-6.462	-6.129	-4.357
R1216	-0.500	-0.166	0.000	0.100	0.337	7.670	-7.977	-7.503	-5.555

[a] The maximum distance, D, in Å.
[b] Calculated on equation (11.1)
[c] Observed values in ppm, taken from Ref. 2,9.
[d] Calculated on equation (11.2)
[e] Calculated on equation (11.3)

11.4. References

1. W.S. Cain, J.E. Cometto-Muniz, Occup. Med., 10 (1995) 133.
2. M.H. Abraham, R. Kumarsingh, JE Cometto-Muniz, W.S. Cain, Toxicol. In Vitro, 12 (1998) 403.
3. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, Exp. Brain Res., 118 (1998) 180.
4. J.M.R. Gola, unpublished work.
5. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, R. Kumarsingh, Pharmacol. Biochem. and Behaviour, 60 (1998) 765.
6. J.-P. Kasanen, A.-L. Panasen, P. Panasen, J. Liesivuori, V.M. Kosma, Y. Alarie, Arch. Toxicol., 72 (1998) 765.
7. S. Maaβen, Dissertation (1995), TU Berlin.
8. R.M. Stephenson, S.Malanowski, Handbook of the thermodynamics of organic compounds, Elsevier Science Publishing Co., (1987).
9. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muñiz and W.S. Cain, Arch. Toxicol., 72 (1998) 227-232.

Chapter 12 Association Between Two Volatile Organic Compounds at the Proximity of Receptor Areas

12.0. Introduction

The chemosensory discomfort experienced in any indoor place may in principal come from a single volatile organic compound, VOC, but more commonly comes from a mixture of VOCs. Hence, in order to understand and predict chemosensory impact, both single VOCs and mixtures require study. Odor response and chemesthetic responses, at suprathreshold level and threshold level, to mixtures of VOCs, have been the subject of a number of investigations on humans; those are presented in detail in chapter 2. In this chapter, attention is drawn to investigations carried out by Cometto-Muniz and Cain on mixtures at threshold levels.[1-4]

Cometto-Muniz et al[2] have found that the constituents of mixtures of up to nine components exhibited varying degrees of agonism, i.e. dose additivity as regards odor, and sensory irritation, in the evocation of threshold. Agonism was larger for sensory irritation than for odor thresholds. Also agonism tended to increase with number of components and their lipophilicity. In a recent report on the chemosensory detectability of butan-1-ol and heptan-2-one singly and as a mixture[3], it was pointed out that the binary mixture exhibits chemosensory agonism regarding odor, and sensory irritation. Similar results were obtained for toluene and heptyl acetate binary mixture.[4] However, departures from simple additivity may be observed because of physicochemical interactions between the components of the mixture during their transport process from the vapour phase to a receptor phase. Such interactions could occur in the vapor of the mixture, either before administration or in the nasal airways, or in the mucus layer of the nose, or at or near to the site of physiological action. Therefore, it is worth assessing any possible interaction in order to understand the rules governing the sensory impact of mixtures.

12.0.1. Interactions between Volatile Organic Compounds

Because VOCs are all non-ionic, the two main types of interaction that can take place between two VOCs are either dipole / dipole or hydrogen bonding. Of these, the hydrogen bonding interaction occupies a special position, being strong enough to have a profound effect on the reactivity of the molecules engaged in it, yet weak enough to prevent their permanent conversion into other compounds. Hydrogen bond interaction has, therefore, occupied a key position in the interpretation of biophysical processes. Because of its importance in interaction between molecules and between nearby functions in the same molecules, the hydrogen bond offers special promise for the study of such interactions.[5]

The hydrogen bonding is by far the largest interaction in cases where a reasonably strong hydrogen bond acid, e.g. an alcohol, can interact with a reasonably strong hydrogen bond base, e.g. a ketone, as shown in Figure 12.1. Now for an equilibrium in which a hydrogen bond acid A, and a hydrogen base, B, form a complex C, a specific equilibrium constant for hydrogen bonding, K, can be defined as follows,

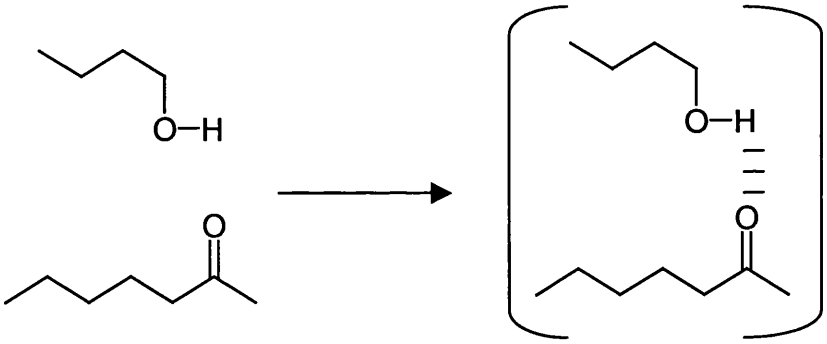


Figure 12.1. Hydrogen bond interaction between VOCs

$$A + B \underset{}{\overset{K}{\rightleftharpoons}} C \qquad (12.1)$$

	A	B	C
at t = 0	ai	bi	0
at t = t	ai – cf	bi - cf	cf

and

$$K = cf / ((ai-cf)(bi - cf)) \qquad (12.2)$$

Where ai is the initial concentration of component A, bi is the concentration of VOC, and B and cf represents the amount reacted at the equilibrium time t.

Now if K, ai and bi are known, a quadratic equation for the amount reacted, cf, can be set up. The percentage of complexation between A and B is determined according to equation (12.3). This calculation is performed using ai and bi depending on which is the smallest.

Percentage of complexation = 100.[C] / [A] (12.3)

The extent of complexation is an indication as to whether or not the interaction between the two VOCs is large enough to interfere with the chemosensory process. The extent of complexation depends highly on the ability of both components A and B to participate to hydrogen bond interaction but also on the polarity of the bulk solvent. An increase in complexation is generally observed with increasing stability of the proton donor, A, and increasing basicity of the proton acceptor, B. Furthermore, the percentage of association decreases as the solvent polarity increases. Hence, large percentage values will be obtained in inert solvents such as alkanes. Conversely, low percentage values will be determined in polar solvents.

12.0.2. Experimental Methods

A method for examining interaction between VOCs is to determine complexation constants, K. A number of experimental techniques is available to achieve this purpose. Some of the methods utilised are UV/visible spectroscopy [6], nuclear magnetic resonance, ^{1}H NMR [7] and ^{19}F NMR [8], microcalorimetry[9] and optical activity[10] measurements[11-15]. However, vibrational spectroscopy is possibly the most widely used experimental technique for equilibrium constant determination. This technique refers to the observation of the stretching frequency of a covalently bond hydrogen atom that is perturbed in a characteristic way when that atom participates in hydrogen bonding to an acceptor group.[5] In recent years, the development of Fourier transform infrared spectroscopy (FTIR) has vastly increased its accuracy and reliability[16]. An example of complexation constant determination by vibrational spectroscopy among a large quantity of similar studies, is the work carried out by Laurence and co-workers[17] on sulfonyl bases. The authors observed 1:1 complex formation between thirteen sulfonyl

bases and 4-fluorophenol and methanol in tetrachloromethane by monitoring the infrared shits, on complexation, of the ν(OH) vibrations of the two hydrogen bond acids with the sulfonyl bases.[17]

Interactions between hydrogen bond acids and hydrogen bond bases have nearly always been investigated in nonpolar or slightly polar solvents because larger complexation constant values will be measured in such environments, e.g. tetrachloromethane, cyclohexane, carbon disulfide, tetrachloroethylene, dichloromethane, 1,2-dichloroethane. Experiments have also been made in benzene, chlorobenzene and 1,2-dichlorobenzene. Noteworthy, only few investigations have been carried out to measure the extent of interaction between hydrogen bond acids and hydrogen bases in aqueous solvent. For instance, Nakano and Higushi[10]determined complexation constants of triptophan with some amides at 298K in water by means of triptophan optical activity measurements. Complexation constant values varied from 0 dm^3 mol^{-1} (triptophan / tetramethyl succinamide system) to 30 dm^3 mol^{-1} (triptophan / caffeine). However the complexes formed are not simple !:1 hydrogen bond complexes. A few examples of 1:1 hydrogen bond complexation involve charged species: the phenolate anion and a monoprotonated amine [6] The largest complexation constant value, 0.79 dm^3 mol^{-1}, was obtained for the system phenolate ion / ammonium ion. Simple 1:1 complexation between uncharged species in polar solvents such as water and alcohols seems not to have been observed. Note that hydrogen bond equilibrium constants will always be given in units of dm^3 mol^{-1}.

12.0.3. Abraham Approach

There is a considerable literature on equilibrium constants for 1:1 hydrogen-bond complex formation, $K_{1:1}$, in nonpolar solvents, especially in the particular solvent tetrachloromethane. For hydrogen-bonding interaction in tetrachloromethane, $K_{1:1}$, values are very accurately given[18] by the equation,

$$\text{Log } K_{1:1} = 7.354\ (\alpha_2^H.\beta_2^H) - 1.094 \qquad (12.4)$$

$n = 1312$, $r = 0.9956$, $sd = 0.093$, $F = 147880$

Here and elsewhere, n is the number of data points, r is the correlation coefficient, sd is the standard deviation, and F is the F-statistic. The parameters that define the VOCs are α_2^H , the VOC 1:1 hydrogen bond acidity, and β_2^H the VOC 1:1 hydrogen basicity. Quite recently, similar equations to equation (12.4) have been constructed for hydrogen-bond formation in the gas phase[19], 1,1,1-trichloroethane[20], 1,2-dichloroethane[21] and chlorobenzene[21]. A general linear free-energy relationship for hydrogen bond complexation can be put forward as follows.

$$\text{Log } K_{1:1} = m\,(\alpha_2^H.\beta_2^H) + c \qquad (12.5)$$

The constants, m and c, are characteristic of the solvent. Values of the constants m and c for the solvents investigated so far are given in Table 12.1. Equations (12.6)-(12.9) are important in that it shows that the bilinear formalism that takes place in tetrachloromethane solution also applies to various solvents. The intercept, c, of equations in Table 12.1 are quite close to that in equation (12.4). According to Abboud et al.[22], this strongly suggests that the 'magic point', -1.1, found by Abraham [see equation (12.4)] is endowed with a deep physical meaning. The slope, m, decreases as the polarity of the solvent increases, indicating a modest attenuation of the hydrogen bond interactions.

Table 12.1. Values of constants m and c for various solvents.

Eq.	Solvent	m	c
(12.6)	Gas Phase	9.130	-0.870
(12.7)	1,1,1-Trichloroethane	6.856	-1.144
(12.8)	Chlorobenzene	6.822	-1.086
(12.9)	1,1-Dichloroethane	6.024	-1.202

12.0.4. Empirical Solvent Polarity Scales

Empirical solvent polarity scales are intended to provide quantitative measures of solvent / solute interactions, viz. hydrogen bonding and dipole / dipole interactions.[23] A number of theoretical treatments have been developed that quantitatively link the interaction to properties of the bulk solvents. These have been reviewed in a book.[24] In this work, attention was drawn to the Dimroth and Reichard solvent polarity parameter[25], $E_T(30)$, and the Kamlet-Taft parameters[26-30], α, β and π^*. The π^* scale of solvent polarity measures the effect of dipolarity / polarizability, the α scale is concerned with the ability of the solvent to accept an electron pair and the β scale, the ability of the solvent to donate a share of an electron pair in hydrogen bond interaction. [26-30] The choice of these parameters was driven by the fact that they are known for a large number of solvents and have successfully characterised solvent effects on organic reactivity and on absorption spectra.[26] These solvatochromatic parameters were discussed in chapter 3.

Dimroth and Reichardt proposed a solvent parameter, $E_T(30)$, that is based on the transition energy for the longest-wavelength solvatochromic absorption band of the pyridinium-N-phenoxide betaine dye.[27] Owing to an exceptionally displacement of the solvatochromic absorption band, the $E_T(30)$-values provide an excellent and very sensitive characterisation of the polarity of the solvents, high $E_T(30)$-values corresponding to high solvent polarity.[24] $E_T(30)$ values have been determined for more than 300 pure organic solvents, and for a great number of binary solvent mixtures. (See references 66-72, 124, 174-192 in Ref. 24) Unfortunately, $E_T(30)$-values have by definition the dimension of kcal.mol^{-1}.[24] Thus, the use of the so-called normalised E_T^N-values has been recommended.[24] They are defined according to equation (12.10), using water and tetramethysilane, TMS, as extreme reference solvents.

$$E_T^N = (E_T(\text{solvent}) - E_T(\text{TMS})) / (E_T(\text{water}) - E_T(\text{TMS})) \qquad (12.10)$$

Hence, the corresponding E_T^N-scale ranges from 0.000 for TMS, the least polar solvent, to 1.000 for water, the most polar solvent. Further, the solvents can be roughly divided into three groups according to their $E_T(30)$ or E_T^N values depending on their specific

solvent / solute interactions: (i) protic solvent ($0.5 < E_T^N < 1.0$), (ii) dipolar non protic ($0.3 < E_T^N < 0.5$), and (iii) apolar non protic ($0.0 < E_T^N < 0.3$).[24]

Of particular interest is the correlation between $E_T(30)$-values and Kamlet-Taft solvatochromic parameters.[26, 31] Equation (12.11) suggests that $E_T(30)$-values measure not only a blend of solvent dipolarity and polarisability but also the solvent hydrogen bond acidity of protic solvents.

$$E_T(30)\ (\text{kcal.mol}^{-1}) = 14.6\ (\pi^* - 0.23\ \delta) + 16.5\ \alpha + 30.31 \qquad (12.11)$$

Where δ, a polarisability correction term, is 0.0 for nonchlorinated aliphatic solvents, 0.5 for polychlorinated aliphatic solvents, and 1.0 for aromatic solvents.

In the present work, the normalised parameter, E_T^N is preferred to $E_T(30)$ because has the advantage of being dimensionless and in the range of 0 and 1, hence commensurate with the α-scale and β-scale.

Of course, the receptor phase may be more polar than any of the above phases. It was shown in chapter 9, that octan-1-ol and dimethylformamide, DMF, are very good models for the nasal pungency receptor area. This result can be applied to odor threshold area and eye irritation. Unfortunately, there are no literature values for hydrogen bond complex formation in octan-1-ol and DMF. However, log $K_{1:1}$ values between two given VOCs are generally reduced as the condensed phase becomes more polar, so that the largest calculated log K_H value in solvents such as cyclohexane, tetrachloromethane, trichloroethane will represent the maximum possible interaction between two given VOCs. Thus if the calculated interaction is too small to affect nasal pungency potency, the actual interaction in a more polar environment will certainly be too small to do so.

One aim of the present work was to establish equations similar to equation (12.4) for hydrogen bonding interaction in solvents such as octan-1-ol and dimethylformamide. Hence, an headspace gas chromatographic method was devised to measure complexation constants, $K_{1:1}$, for a series of acids and bases in octan-1-ol. Next, the relationship between the values for the constants c and m and solvent polarity was studied. This approach required the use of empirical solvent polarity scales. Note that the term polarity refers to the action of all possible, specific and non-specific, intermolecular forces between solvent and solute molecules.

Table 12.2. Empirical solvent parameters for selected solvents.[a]

Solvent	$E_T(30)$	$E_T^{N\,c}$	β	α	π^*	δ
Gas phase	25.2[b]	-0.170	0.000	0.000	-1.100	0.000
Cyclohexane	30.8	0.006	0.000	0.000	0.000	0.000
Tetrachloroethylene	31.8	0.138	0.000	0.000	0.280	0.500
Tetrachloromethane	32.4	0.052	0.000	0.000	0.280	0.500
Benzene	34.5	0.111	0.100	0.000	0.590	1.000
1,1,1-Trichloroethane	36.2	0.170	0.000	0.000	0.490	0.500
Chlorobenzene	36.8	0.188	0.070	0.000	0.710	1.000
1,2-Dichloroethane	41.3	0.327	0.000	0.000	0.810	0.500
Octan-1-ol	48.1	0.543	0.860[d]	0.77	0.410	0.000

[a] From ref. 24 otherwise notified.
[b] This work, see appendix 2.
[c] Calculated from equation (12.10)
[d] From ref. 32.

12.1. Results

12.1.1 Hydrogen Bond Interactions in Octan-1-ol

A new method was devised to determine $K_{1:1}$ hydrogen bond interaction constants in octan-1-ol. The technique, based on a headspace gas chromatographic method, is presented in section 12.2. Twenty-six $K_{1:1}$ values among the twenty-nine measured were used to set up the following equation,

$$\log K_{1:1} = 2.950\ \alpha_2^H.\beta_2^H - 0.741 \qquad (12.12)$$

$n = 26,\ r^2 = 0.950,\ sd = 0.092,\ F = 458$

This is the first time that such an equation is developed for the estimation of 1:1 hydrogen bonding interaction in a polar and protic solvent such as octan-1-ol. Values of intercept, c, and slope, m, differ from those listed in Table 12.1. This result is however not surprising, because it reflects the difference of physicochemical properties between octan-1-ol and the solvents investigated so far. The slope is some 69 % smaller than in

the gas phase, indicating a rather large attenuation of the hydrogen bonding in solution. The intercept value is worth considering. In octan-1-ol, the value of c, -0.741 (± 0.070), is more positive than those listed in Table 12.1. However, this difference may be due to the lack of small values for the constant $K_{1:1}$ in octan-1-ol.

Now, values of α_2^H are known for about 200 compounds and values of β_2^H for some 500 compounds so that log $K_{1:1}$ values can be calculated for 100,000 complexations of the type shown in equation (12.12). A simple calculation through equation (12.12) will give a quantitative measure of the interaction between two VOCs in octan-1-ol as the condensed phase.

12.1.2 Hydrogen Bond Interactions in Selected Solvents

Log K1:1 values for 1:1 interactions were retrieved from the literature. Three solvents were investigated, viz. cyclohexane, tetrachloroethylene, and benzene. Equations similar to equation (12.12) were developed.

Cyclohexane

$$\log K_{1:1} = 7.674\ \alpha_2^H.\beta_2^H - 0.954 \qquad (12.13)$$

$n = 430$, $r^2 = 0.975$, $sd = 0.174$

Tetrachloroethylene

$$\log K_{1:1} = 7.382\ \alpha_2^H.\beta_2^H - 1.087 \qquad (12.14)$$

$n = 79$, $r^2 = 0.993$, $sd = 0.107$

Benzene

$$\log K_{1:1} = 6.744\ \alpha_2^H.\beta_2^H - 1.080 \qquad (12.15)$$

$n = 33$, $r^2 = 0.979$, $sd = 0.123$

As mentioned above, the values of c and m are highly dependent upon the physicochemical nature of the investigated solvent. Based on this apparent relationship,

empirical solvent scales have been considered in order to predict values of c and m and expand the bilinear formalism to a larger set of solvents. This study is presented in the following section.

12.1.3 Application of The Abraham Method to Selected Solvents

It was hypothesised that the constant m is linearly related to one or to a combination of solvents parameters so that the following correlation expression should be applicable.

(i). $m = \text{intercept} + e \,.\, E_T^N$

(ii). $m = \text{intercept} + e \,.\, E_T^N + b.\,\beta$

(iii). $m = \text{intercept} + a.\alpha + b.\beta + p.\pi^*$

(iv). $m = \text{intercept} + a.\alpha + b.\beta + p.\pi^* + d.\delta$

Similar correlations were considered with the constant c as the dependent variable.

A MS Excel'97 with the function 'Data Analysis' was used to calculate the regression coefficients, e, a, b, p and d. Values of those are given in Table 12.3 and Table 12.4 together with the statistical results.

The regression of the slope m over E_T^N gives reasonable statistical results (r^2 = 0.924, sd 0.499, F = 85.17). However, a statistical improvement of the fit (r^2 = 0.984, sd = 0.242, F = 192.83) is obtained when the hydrogen bond basicity properties of the solvent, expressed by β is included in the regression. The ratio of the coefficients of E_T^N and β is 2.65, hence the sensitivity of the slope c to E_T^N is larger than β. As shown in equation (12.11), E_T^N is a blend of hydrogen bond acidity and dipolarity / polarisability of the solvent so that these two properties contribute principally in the value of the slope m. Although there are slight improvements sd and r^2 in equation (12.18) and equation (12.19) relative to equation (12.17), the F statistic is somewhat worse, so that the improvement is only marginal.

It appears clearly from Table 12.4 that there is no correlation between the solvent polarity parameter, E_T^N and the intercept m. However, the statistical results improved significantly when the hydrogen bond basicity parameter was considered in the regression analysis. The ratio of the E_T^N and β is 0.85, hence the sensitivity of the intercept m to β is slightly larger than E_T^N. Again a marginal improvement of the fit was obtained when values of m were correlated to the Kamlet-Taft parameters.

It emerges from this study that a combination of E_T^N and β is a good approach for estimating the values of the slope m and the intercept c in equation (12.5) for any solvent whose solvent polarity parameters are available. In Table 12.5, are given some examples of values for m and c for selected solvents. . Of course, the predictive equations (12.17) and (12.21) have been developed using a set of only nine solvents, among which are eight aprotic solvents. Hence, it is highly possible that the results are biased towards octan-1-ol.

In order to see the effect of solvents on complexation constants, values of $\alpha_2^H.\beta_2^H$ in equation (12.5) were chosen arbitrarily between 0.05, e.g. very weak hydrogen bonding interaction, and 0.8, e.g. large hydrogen bond interaction. Then the hydrogen bond complexation constant, $K_{1:1}$, can be calculated for all $\alpha_2^H.\beta_2^H$ values in any of the given solvents in Table 12.5. A plot of log $K_{1:1}$ values against $\alpha_2^H.\beta_2^H$ is shown in Figure 12.1.

Figure 12.2 illustrates well the common trend of a decreasing value of log $K_{1:1}$ as the polarity of the solvent increases, for any given acid / base pair. These results also highlight the potential difficulties that would be encountered should an attempt be made to measure the hydrogen bond interaction between acids and bases in water.

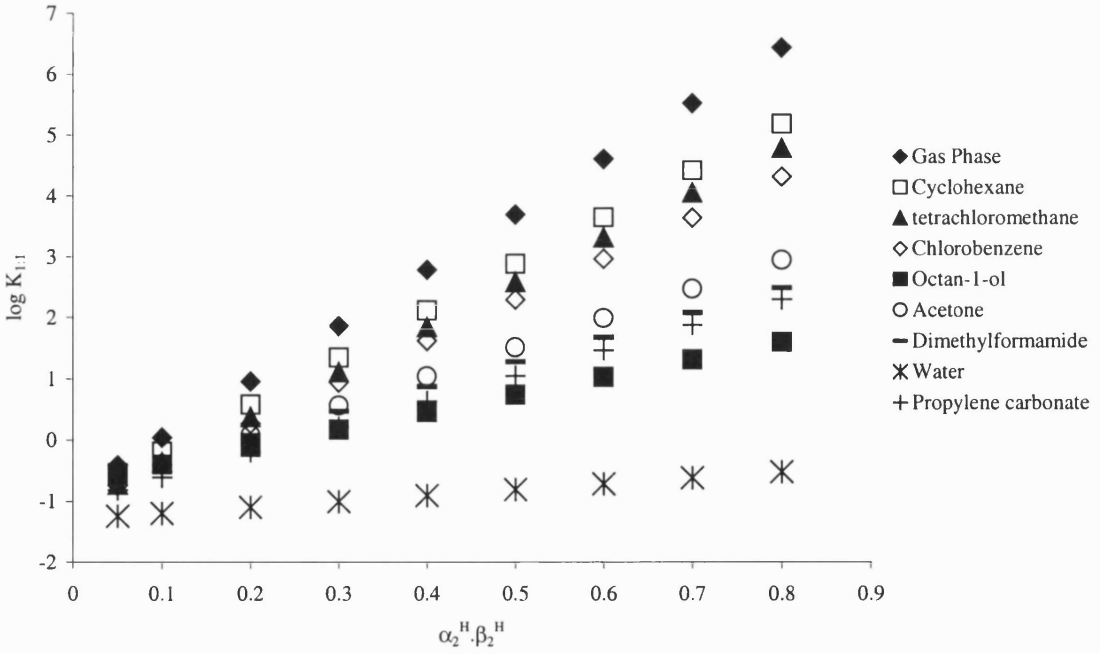


Figure 12.2 Plot of log $K_{1:1}$ values against $\alpha_2^H.\beta_2^H$ in selected solvents.

Table 12.3. Relationship between the slope, m and various empirical solvent polarity parameters

Eq.	Regression Coefficients							Statistical Results		
	Intercept	e^n	b	p	a	d	n	r^2	sd	F
(12.16)	**7.993**	**-8.126**	**-**	**-**	**-**	**-**	**9**	**0.924**	**0.499**	**85.17**
	0.213	0.880	-	-	-	-				
(12.17)	**7.897**	**-5.835**	**-2.200**	**-**	**-**	**-**	**9**	**0.984**	**0.242**	**192.83**
	0.105	0.635	0.451	-	-	-				
(12.18)	**7.603**	**-**	**1.534**	**-1.504**	**-7.084**	**-**	**9**	**0.988**	**0.231**	**140.79**
	0.096	-	2.360	0.156	2.580	-				
(12.19)	**7.479**	**-**	**-0.828**	**-1.637**	**-4.208**	**0.417**	**9**	**0.989**	**0.248**	**92.38**
	0.227	-	4.638	0.276	5.473	0.684				

Table 12.4. Relationship between the intercept, c and various empirical solvent polarity parameters

Eq.	Regression Coefficients							Statistical Results		
	Intercept	e^n	b	p	a	d	n	r^2	sd	F
(12.20)	**-1.044**	**0.149**	**-**	**-**	**-**	**-**	**9**	**0.035**	**0.168**	**0.254**
	0.071	0.296	-	-	-	-				
(12.21)	**-1.010**	**-0.683**	**0.800**	**-**	**-**	**-**	**9**	**0.932**	**0.048**	**41.44**
	0.020	0.126	0.089	-	-	-				
(12.22)	**-1.035**	**-**	**0.727**	**-0.173**	**-0.263**	**-**	**9**	**0.955**	**0.043**	**35.49**
	0.017	-	0.437	0.028	0.477	-				
(12.23)	**-0.980**	**-**	**1.778**	**-0.114**	**-1.540**	**-0.185**	**9**	**0.977**	**0.034**	**42.55**
	0.031	-	0.641	0.038	0.757	0.094				

Table 12.5. Calculated values of the slope m and intercept c for a series of solvents.[a]

Solvent	E_T^N	β	m	c
Carbon disulfide	0.065	0.000	7.518	-1.054
Dibutyl ether	0.071	0.460	6.471	-0.690
1,4-Doxane	0.164	0.370	6.126	-0.826
Bromobenzene	0.182	0.060	6.703	-1.086
Tetrahydrofuran	0.207	0.550	5.479	-0.711
Dichloromethane	0.309	0.000	6.096	-1.221
Acetone	0.355	0.480	4.770	-0.868
DMF [b]	0.404	0.690	4.022	-0.734
DMSO [b]	0.444	0.760	3.634	-0.705
Acetonitrile	0.460	0.310	4.531	-1.076
Propylene carbonate	0.491	0.400	4.152	-1.025
Water	1.000	0.500	0.962	-1.293

[a] Values for E^T_N and β are taken from ref. 24.
[b] Abbreviations: DMF dimethylformamide, DMSO Dimethylsulfoxide.

12.1.4 Percentage of Association of Hydrogen Bond Complexation Butan-1-ol and Heptan-2-one in Octan-1-ol and Dimethylformamide

The maximum vapor concentrations used for eye irritation and nasal pungency threshold studies were almost the same, and were orders of magnitude greater than those used in the odor detection studies. Four particular mixtures of butan-1-ol and heptan-2-one used by Cometto-Muniz and Cain to study nasal pungency threshold in humans were selected together with four mixtures of VOCs in the study of odor detection threshold. The maximum vapor phase concentrations in ppm of butan-1-ol and heptan-2-one in the presented vapor are given in Table 12.6. These mixtures are denoted as (i)-(iv); all mixtures are composed with same amount of heptan-2-one but with varying concentrations of butan-1-ol. Note that in Table 12.6, hept denotes heptan-2-one and BuOH refers to butan-1-ol. Also in Table 12.6, the corresponding concentrations converted to $mol.dm^{-3}$ are displayed, ($mol.dm^{-3}$ = ppm / (100000*24.46)).

Table 12.6 Maximum concentrations in the vapor mixtures[a]

	Hept (ppm)	BuOH (ppm)	Hept ($mol.dm^{-3}$)	BuOH ($mol.dm^{-3}$)
NPT Studies				
(i)	89.1	371	$3.6\ 10^{-6}$	$1.5\ 10^{-5}$
(ii)	89.1	631	$3.6\ 10^{-6}$	$2.6\ 10^{-5}$
(iii)	89.1	1259	$3.6\ 10^{-6}$	$5.1\ 10^{-5}$
(iv)	89.1	2137	$3.6\ 10^{-6}$	$8.7\ 10^{-5}$
ODT Studies				
(i)	0.03	0.24	$1.2\ 10^{-9}$	$1.1\ 10^{-8}$
(ii)	0.03	0.55	$1.2\ 10^{-9}$	$2.2\ 10^{-8}$
(iii)	0.03	1.26	$1.2\ 10^{-9}$	$5.1\ 10^{-8}$
(iv)	0.03	2.40	$1.2\ 10^{-9}$	$9.8\ 10^{-8}$

[a] Studies on eye irritation thresholds were conducted using the same concentration as those for nasal pungency thresholds.

In general, molar concentrations are greatly increased on transfer from the gas phase to any solvent phase. A quantitative measure is the gas-liquid partition coefficient or Ostwald solubility coefficient[33], L, defined through equation (12.24):

$$L = [\text{molar conc. in solution}]/[\text{molar conc. in gas phase}] \qquad (12.24)$$

The Abraham general solvation equation, equation (12.25), can be used to correlate all manner of gas-to-condensed phase processes, including log L values.[34]

$$\text{Log } L = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L} \qquad (12.25)$$

The above equation has been presented in detail in chapter 3. As previously mentioned, octan-1-ol and dimethylformamide are reasonable models for the receptor area. Then log L for butan-1-ol and heptan-2-one in the gas phase / wet octanol system can be used as the concentration factors for transfer of these VOCs from the gas phase to the receptor area. Experimental values[35] are in Table 12.7, together with log L values for dimethylformamide[36]. The latter were calculated from the Abraham general equation for gas / dimethylformamide transfer process. The maximum concentrations of the two VOCs at the receptor area can be calculated taking octan-1-ol and dimethylformamide as the models, see Table 12.8.

Table 12.7. Log L values for butan-1-ol and heptan-2-one

Solvent:	Dimethylformamide[a]	Wet octanol
Butan-1-ol	4.635	4.34
Heptan-2-one	4.436	4.21

[a] Calculated on equation (12.25).

Table 12.8. Maximum concentrations, in $mol.dm^{-3}$, at the chemosensory receptor area for butan-1-ol and heptan-2-one, taking dimethylformamide and octan-1-ol as models.

	Dimethylformamide		Octan-1-ol	
Mixture	Hept	BuOH	Hept	BuOH
NPT Studies				
(i)	$9.9\ 10^{-2}$	$6.5\ 10^{-1}$	$5.9\ 10^{-2}$	$3.3\ 10^{-1}$
(ii)	$9.9\ 10^{-2}$	$11.1\ 10^{-1}$	$5.9\ 10^{-2}$	$5.6\ 10^{-1}$
(iii)	$9.9\ 10^{-2}$	$22.2\ 10^{-1}$	$5.9\ 10^{-2}$	$11.2\ 10^{-1}$
(iv)	$9.9\ 10^{-2}$	$37.7\ 10^{-1}$	$5.9\ 10^{-2}$	$19.1\ 10^{-1}$
ODT Studies				
(i)	$3.4\ 10^{-5}$	$4.3\ 10^{-4}$	$2.0\ 10^{-5}$	$2.2\ 10^{-4}$
(ii)	$3.4\ 10^{-5}$	$9.7\ 10^{-4}$	$2.0\ 10^{-5}$	$4.9\ 10^{-4}$
(iii)	$3.4\ 10^{-5}$	$2.2\ 10^{-4}$	$2.0\ 10^{-5}$	$1.1\ 10^{-3}$
(iv)	$3.4\ 10^{-5}$	$4.2\ 10^{-3}$	$2.0\ 10^{-5}$	$2.1\ 10^{-3}$

[a] Studies on eye irritation thresholds were conducted using the same concentration as those for nasal pungency thresholds.

In order to calculate interactions between the two VOCs, use can be made of equations developed in this work for the estimation of hydrogen bond equilibrium constants, $K_{1:1}$, in octan-1-ol and dimethylformamide.

Octan-1-ol

$$\log K_{1:1} = 2.950\ \alpha_2^H.\beta_2^H - 0.741 \qquad (12.12)$$

Dimethylformamide

$$\log K_{1:1} = 4.022\ \alpha_2^H.\beta_2^H - 0.734 \qquad (12.26)$$

For butan-1-ol, the value of α_2^H is 0.330, and a value of 0.500 for β_2^H has been obtained for heptan-2-one from the data of Massat et al.[37] These parameters together with equations (12.12) and (12.26) then yield values for hydrogen bond complexation in octan-1-ol (0.6 dm^3 mol^{-1}) and dimethylformamide (0.8 dm^3 mol^{-1}).

ai and bi, as given by the concentration of butan-1-ol and heptan-2-one in the selected solvent, respectively, are listed in Table 12.8. The percentage of interaction at the nasal pungency, and odor thresholds was determined according to equation (12.3). The calculation is performed using ai or bi depending on which is the smallest. Results are given in Table 12.9.

Table 12.9. Determination of the amount reacted and the percentage of interaction between butan-1-ol and heptan-2-one at the receptor area.

	Dimethylformamide		Octan-1-ol	
Mixture	cf	%[a]	cf	%[a]
NPT Studies				
(i)	$2.6\ 10^{-2}$	12.1	$1.2\ 10^{-2}$	10.5
(ii)	$4.5\ 10^{-2}$	18.6	$2.0\ 10^{-2}$	15.4
(iii)	$8.9\ 10^{-2}$	30.2	$4.0\ 10^{-2}$	24.5
(iv)	$1.5\ 10^{-1}$	40.7	$6.8\ 10^{-2}$	32.5
ODT Studies				
(i)	$3.5\ 10^{-9}$	0.010	$2.6\ 10^{-9}$	0.008
(ii)	$1.9\ 10^{-9}$	0.022	$5.9\ 10^{-9}$	0.018
(iii)	$1.8\ 10^{-8}$	0.050	$1.3\ 10^{-9}$	0.042
(iv)	$3.4\ 10^{-8}$	0.095	$2.6\ 10^{-9}$	0.080

[a] Calculated on equation (12.3)

12.1.5. Implications of the Results Obtained

These results are based on the assumption that octan-1-ol and dimethylformamide are good models for the chemosensory receptor area. The vapour concentrations used in the determination of ODT values were very low, therefore at the proximity of the odor receptor the percentage of interaction is insignificant. This result accounts for the fact that dose additivity is found for odor detection.

More interestingly however, with regards to the chemesthetic impact of VOCs, the results indicate there is some degree of complexation between the components of the mixture at the nasal pungency and eye irritation receptor area. The concentration of

VOCs used for experiments on NPTs however were much higher than those used for the odor studies. Hence, there is a greater amount of interaction between the two components. However the fact that dose additivity is observed for sensory irritation thresholds suggests that the interaction observed does not reduce the chemosensory potency of these VOCs.

12.2. Determination of Complexation Constants using a Headspace Gas Chromatographic Method.

A headspace gas chromatographic method, HS-GC, with internal standard was devised to obtain the complexation constant between a hydrogen bond donor and a hydrogen bond acceptor in solution in polar and protic solvents such as octan-1-ol.

12.2.1. Complexation between two VOCs

Assume a dilute solution of a particular VOC A and an inert standard substance D in a given solvent set up in a closed vial as shown in Figure 12.3, part **1**, so that A and D will distribute between the liquid and gas phase above the solvent. As explained in chapter 4 (section 4.2.3), the concentrations of A and D in the liquid phase, a_L and d_L are related to their gas chromatographic peak areas ai' and di',

$$a_i/d_i = k \,.\, a_i'/d_i' \qquad (12.27)$$

where k is the constant of proportionality including the following variables; the sample preparation, the headspace sampling, the chromatographic conditions and the detector sensitivity.

Now if a less-volatile component, B, is added to the system and interacts with the component A to form an adduct C according to equations (12.1) and (12.2), a new system is envisaged, as shown in Figure 12.3, part **2**.

Because of the complex formation between A and B, the concentration of A in solution will decrease and hence, in accordance with equation (12.24), the gas phase concentration will also decrease. The decrease in the concentration of free A species in solution will lead to a diminution of the peak area as shown in Figure 12.4. Thus, changes in the concentration of free A in the liquid can be monitored by gas chromatography. D is an inert component. Then, the concentration of D will nominally be the same before and after the addition of B.

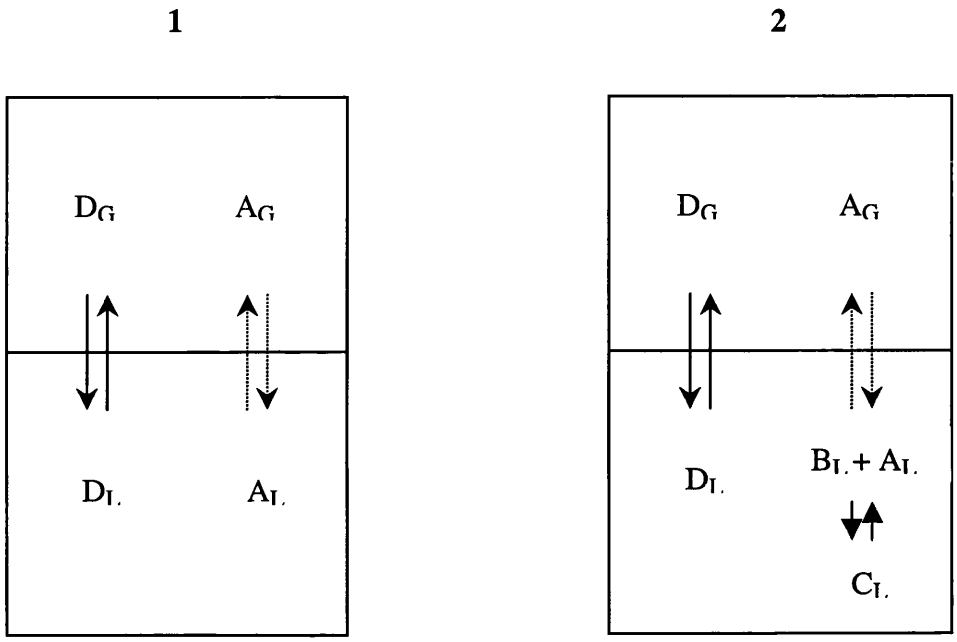


Figure 12.3. The system before and after addition of the component B.

However, the apparent concentration in the gas phase can change because of sampling or analytical errors, and hence the peak areas may differ before and after the addition of B. After the addition of compound B, equation (12.28) becomes:

$$a_f / d_i = k \,.\, a_f' / d_f' \qquad (12.28)$$

where af is the final concentration of A and a_f' and a_f' are the peak of areas of compounds A and D. k is the proportionality constant and is assumed not to be affected by the addition of B to the matrix system. Hence, the k-value will be the same in equations (12.27) and (12.28). Then, combining these equations,

$$(a_f / d_f) / (a_i / d_i) = a_f / a_i = (a_f' / d_f')/ (a_i' / d_i') = R \qquad (12.29)$$

with R as the ratio of peak areas after and before the addition of the second component, B. It is not necessary to introduce any gas-solution partition from B, because this is chosen as an involatile compound.

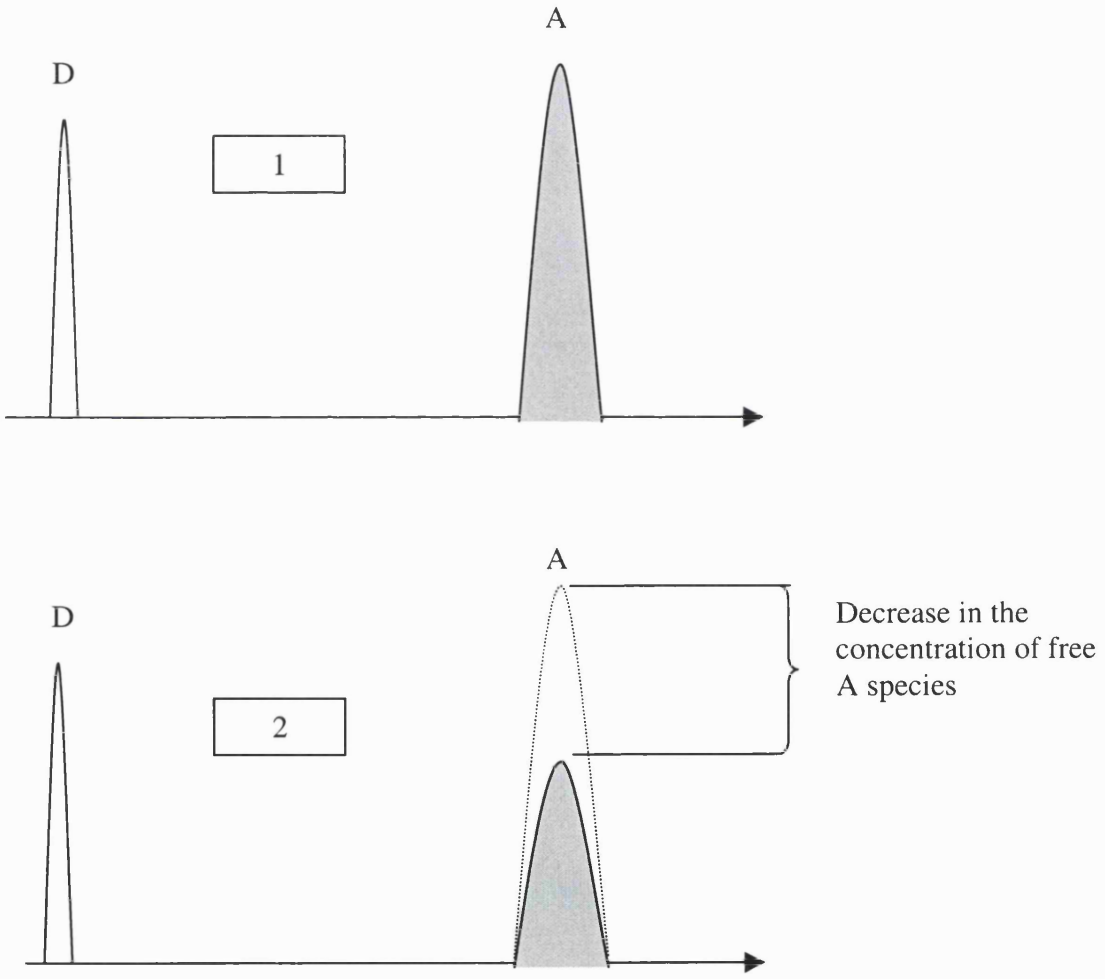


Figure 12.4. Schematic of gas chromatograms of the species A and D before (1) and after (2) addition of B.

Substituting c_f by (a_i-a_f) in equation (12.2), one obtains:

$$K = (a_i - a_f) /\{a_f [b_i - (a_i - a_f)]\} \qquad (12.30)$$

Inserting R in equation (12.30), then:

$$K = (1-R) / R [b_i - (a_i - a_f)] \quad (12.31)$$

$$K = (1-R) / [R. b_i - a_i. R. (1-R)] \quad (12.32)$$

The above assumes that the initial concentration of B is at least equal to that of A. Then the overall complexation constant, K, can be obtained from the gas chromatographic ratio, R, and the initial concentrations of A and B from the final equation (12.32).

12.2.2 Theory of Alcohol Self-Association

Interestingly, the HS-GC method can be used to study alcohol self-association. Here, an alcohol A plays the role of both hydrogen bond acid and hydrogen bond base. Consider n molecules of an alcohol A interacting to form a complex Cn as follows:

$$nA \rightleftharpoons Cn \quad (12.33)$$

at t = 0	ai	0
at t = t	af	cf_n

where ai and af are the initial and final concentration of free species of alcohol in solution respectively, and cf_n is the concentration of complex formed. Concentration ai is related to af and cf:

$$ai = af + ncf_n \quad (12.34)$$

The equilibrium constant, K_n, refers to the total interactions between n molecules of alcohol and is obtained from equation (12.35).

$$K_n = cf_n / (af)^n \quad (12.35)$$

The equilibrium constant value is then obtained from the gradient of the plot cf_n vs.$(af)^n$.

Now, as already pointed out, the initial and final concentration of free pentan-1-ol species in solution, ai and af, are related by the ratio of peak areas after and before formation of complex Cn,

$$R' = ai / af = R'obs / R'calc \qquad (12.36)$$

R'obs is the ratio of observed peak areas, ai' and di'. Thus, R'obs is expected be linearly correlated with the ratio ai'/di'. A departure from linearity would imply that a mixture of free alcohols and self-associated alcohols is in solution. It is then necessary to calculate the ratio of peak areas from the linear portion of the plot, R' calc. At low concentration of the investigated alcohol, it is assumed that only free species are in solution. Thus, in such a range of concentration, R'calc values have to be proportional to the ratio of concentrations, ai / di,

$$R'calc = k' (ai / di) \qquad (12.37)$$

Where k' is the constant of proportionality. Once R'obs, R'calc are known, R can be calculated. Then af, cf_n and K_n can be calculated from to the appropriate equation.

12.2.3 Experimental Part

The main procedures of the HS-GC method are:

i) to prepare a solution of the required compounds, solute A and standard D, in an appropriate solvent,
ii) to allow time for mixture equilibration,
iii) to measure the peak areas of components A and D, ai' and di' respectively,
iv) to add the complexing agent, B,
v) to allow time for mixture equilibration,
vi) to measure the peak areas of components A and D, af' and df' respectively, and hence, by comparison with the initial peak areas, calculate the complexation constant, K.

12.2.3.1. Solvent Selection

Two solvents were used in this work. n-hexadecane and octan-1-ol were chosen because of their low volatility and molecular properties. First, various reasons led to the choice of n-hexadecane. n-Hexadecane is an inert non-volatile alkane. Only weak dispersion interactions with solutes can occur, specific interactions are assumed to be largely absent. Therefore, the value of K should be larger in this solvent than for any other solvent and hence a measure of the direct complexation between an acid and a base should be more easily obtained. Secondly, Octan-1-ol is a highly protic and polar solvent. It was then speculated that complexation between two VOCs was significantly reduced than in n-hexadecane because of the dipole-dipole interactions and hydrogen bonding interaction occurring between the solutes and the solvent.

12.2.3.2. Solute selection and Solute Concentration

As the aim of this work was to measure the equilibrium constant for a 1:1 complexation, all acids and bases chosen must possess one acidic or basic site and hence undergo the desired interaction, as illustrated in Figure 12.4.

In n-hexadecane solvent, experiments were carried out on pentan-1-ol, a slightly strong hydrogen bond donor, and a series of non-volatile bases (nonan-2-one, nitrobenzene, benzonitrile, acetophenone, dimethyl sulfoxide, triethylphosphate and hexamethylphosphorus triamide). Pentan-1-ol self-association in n-hexadecane was also investigated. These above bases were selected in order to obtain a range of values for $\alpha_2^H.\beta_2^H$ and for the experimentally determined log K values, thus facilitating a direct comparison with a similar range of literature values. The concentration of pentan-1-ol was kept close to 0.05 $mol.dm^{-3}$ in order to avoid the risk of self-association. The concentration of base used for each experiment was determined according to its value of β_2^H. Also, in accordance with equation (12.32), the concentration of base must be equal to, or larger than, the concentration of acid. For weak bases, it was more judicious to use a large excess of base in order to ensure that the complexation can be easily detected by the reduction of the gas chromatographic peak. For strong bases, the complexation with the acid is much more readily detected by gas chromatographic measurements, hence a large excess of base was not needed.

In octan-1-ol, stronger acids and bases were selected in order to provide workable results. A series of fluorinated alcohols were chosen for these experiments due to their good volatility and their relatively strong acidity. The corresponding bases, were chosen to cover a range of basicity and to provide a measurable complexation. The concentration of acid in octan-1-ol needed to be rather large in order to give the best possible gas chromatographic data. The chosen concentration varied according to the investigated alcohol. Because of the use of high concentration, it was necessary to ensure that the alcohols did not self-associate and that the results remained within the headspace linearity. The concentration of the base used for each experiment was determined according to its value of β_2^H, as explained above.

Names of acids and bases together with some of their properties are listed in Tables 12.10.

12.2.3.3. Gas Chromatographic Apparatus

Gas chromatographic measurements were obtained using a Perkin Elmer F-33 gas chromatograph equipped with a flame ionisation detector (F.I.D) under the conditions specified in Table 12.11

Table 12.11. Gas chromatographic conditions

Chromatographic Parameters	Polar compounds to be analysed
Stationary phase	Carbowax 20 M
Amount of stationary phase (%)	12
Support	Chromosorb W
Column length	70 cm
Column temperature*	70-120 °C
Carrier gas	Nitrogen
Carrier gas flow rate*	25-30 $cm^3.min^{-1}$
Sensitivity	Range 1
Sample volume*	1-3 cm^3

* Depending on the system under investigation

Table 12.10 List of hydrogen bond acids and bases used in the present work.

VOCs		Formula	MW	d	$LogL^{oct}$	α_2^H	β_2^H	Source
Pentan-1-ol	PentOH	CH2-CH2-CH2-CH2-CH2-OH	88.15	0.811		0.33	0.50	ALDRICH
2-Fluoroethanol	FEOH	CH_2F-CH_2-OH	64.06	1.108	-0.76	0.40		FLUKA
2,2,2-Trifluoroethanol	TFE	CF_3-CH_2-OH	100.04	1.373	0.41	0.57		ALDRICH
1,1,1,3,3,3-Hexafluoro-2-methyl	HFMP	CF_3-C(-CH_3)(-OH)-CF_3	182.07	1.302		0.66		ACROS
1,1,1,3,3,3-Hexafluoropronan-1-ol	HFIP	CF_3-CH(-OH)-CF_3	168.04	1.596	1.66	0.77		APOLLO
Perfluoro-tert-Butanol	PTB	$(CF_3)_3$C-OH	236.04	1.693		0.87		APOLLO
Nonan-2-one	NON	CH_3-C(=O)-$(CH_2)_6$-CH_3	142.24	0.832	3.14	0.00	0.51	ALDRICH
Acetophenone	ACE	CH_3-C(=O)-C_6H_5	120.15	1.030	1.58	0.00	0.51	ALDRICH
N,N-Dimethylformamide	DMF	$(CH_3)_2$N-CH=O	73.10	0.944	-1.01	0.00	0.66	ALDRICH
N,N-Dimethylacetamide	DMA	$(CH_3)_2$N-CH(=O)-CH_3	87.12	0.937	-0.77	0.00	0.73	ACROS
Dimethyl sulfoxide	DMSO	CH_3-S(=O)-CH_3	78.13	1.101	-1.35	0.00	0.78	ALDRICH
Triethylphosphate	TEP	$(CH_3CH_2O)_3$P=O	182.16	1.072	0.80	0.00	0.79	ACROS
1,1,3,3-Trimethylguanidine	TMG	$(CH_3)_2$N-C(=NH)-N$(CH_3)_2$	115.18	0.918	0.41	0.00	0.92	ALDRICH
Hexamethylphosphoramide	HMPA	$((CH_3)_2N)_3$P	179.20	1.030	0.28	0.00	1.00	ALDRICH
Decane	Dec	CH_3-$(CH_2)_8$-CH_3	142.29	0.730		0.00	0.00	ALDRICH

Chromatograms were recorded and treated by the means of a computer program called PIC3. Solute retention times and peak areas were automatically retrieved.

12.2.3.4. Choice of the Inert Standard

An inert compound was used in the method to compensate for sample losses that occurred during sample preparation and final gas chromatographic analysis. The standard must elute quite close to the analyte of interest and their corresponding peak areas must be obtained on the same scale of sensitivity in order to facilitate their comparison.

A series of qualitative gas chromatographic analyses were conducted in order to determine which substance would be the most suitable as a standard. Studies focused on several inert n-alkanes (pentane to decane). Upper n-alkanes were not used because of their tendency to give broad peaks.

0.02 mm^3 of pure n-alkane (n-pentane 99+% aldrich, n-hexane, BDH, n-heptane +99%, aldrich, n-octane +99%, aldrich, n-nonane ~99%, sigma, decane 99+%, aldrich) was measured using a 0.05 mm^3 Hamilton syringe and injected into the chromatograph. Conditions of analyses were as follows:

Instrument:	Perkin Elmer F-33
Detector:	F.I.D
Column:	Carbowax 20 M
Carrier Gas:	Nitrogen at 30 $cm^3.min^{-1}$
Injection volume:	0.02 mm^3
Injection Temperature:	100°C
GC oven Temperature:	80°C

The average retention time value for each alkane analysed twice by gas chromatography are given below. The lower alkanes were found to elute too quickly and therefore n-decane was chosen as the standard.

Table 12.12. Retention times for a series of alkanes

Solute	Retention time (min.)
n-Pentane	0.32
n-Hexane	0.50
n-Heptane	0.85
n-Octane	1.34
n-Nonane	3.00
n-Decane	5.58

In general, the concentration of the standard was chosen so that the decane peak area was smaller than the solute peak area. Therefore, the amount of decane needed in the HS-GC experiments depends upon the system investigated. For instance, a concentration of 0.08 $mol.dm^{-3}$ was used to study the interaction between perfluoro-tert-butan-2-ol and a series of bases in octan-1-ol although a concentration of 0.140 $mol.dm^{-3}$ was used to measure the complexation constant between pentan-1-ol and several bases in n-hexadecane.

12.2.3.5 Headspace Linearity Determination

Headspace linearity refers to the linear relationship between the original concentration of the analyte in the sample, a_o and its concentration in the gas phase, a_G or between a_o and A_G^A, the peak area obtained when analysing an aliquot of the headspace. Experiments were carried out to determine the headspace linearity of the entire set of hydrogen bond donors, A, used in this work.

$10cm^3$ of the solvent were poured into flask **1** by using a $10cm^3$ pipette (Class A). The two solutes were inserted directly into the solvent using Hamilton syringes ($50mm^3$, $100mm^3$, $500mm^3$ and 1 cm^3). The amount of solute introduced was determined by weighting it. Then flask **1,** capped with a rubber septum, was delicately shaken to allow the components to mix. Next, $5cm^3$ of the stock solution were withdrawn and inserted into flask **2**. Two $5cm^3$ samples containing the same amount of solutes were then obtained. Septa were replaced and then pierced with a small needle to avoid pressure build-up. Finally, flasks **1** and **2** were placed into a water bath thermostated at 25°C and occasionally shaken manually for at least 30 minutes.

Then, 3 cm^3 of the vapour phase in each flask was removed by the means of a 5 cm^3 gas-tight syringe (Hamilton) and injected on to a F-33 Perkin Elmer GC-FID equipped with a Carbowax 20 M column. The vapour phase in each flask was analysed twice. An equilibrium time of at least 30 minutes was allowed between each headspace

removal. A defined mass of hydrogen bond acid was added to flasks **1** and **2**. Then, the flasks were sealed with new septum, shaken manually for few minutes, replaced in the water bath and let to equilibrate at 25°C for at least 30 minutes. Two GC analyses of the headspace in flasks **1** and **2** were conducted as previously described. Solutions with increasing concentration of hydrogen bonding acid were prepared and analysed in an identical way.

The correlation between the concentration and the GC measurements allows the headspace linearity range of the investigated hydrogen bond acid to be studied. The concentration of the hydrogen bond acid in solution was plotted against the peak area ratio, ai'/di'. The headspace linearity of fluoroethanol was investigated as explained above. It appears clearly from Figure 12.5 that the peak area ratio is linearly correlated with fluoroethanol concentration over the concentration range of 0.03-0.4 mol.dm^{-3}. Similar studies were carried out for the remaining hydrogen bond donors.

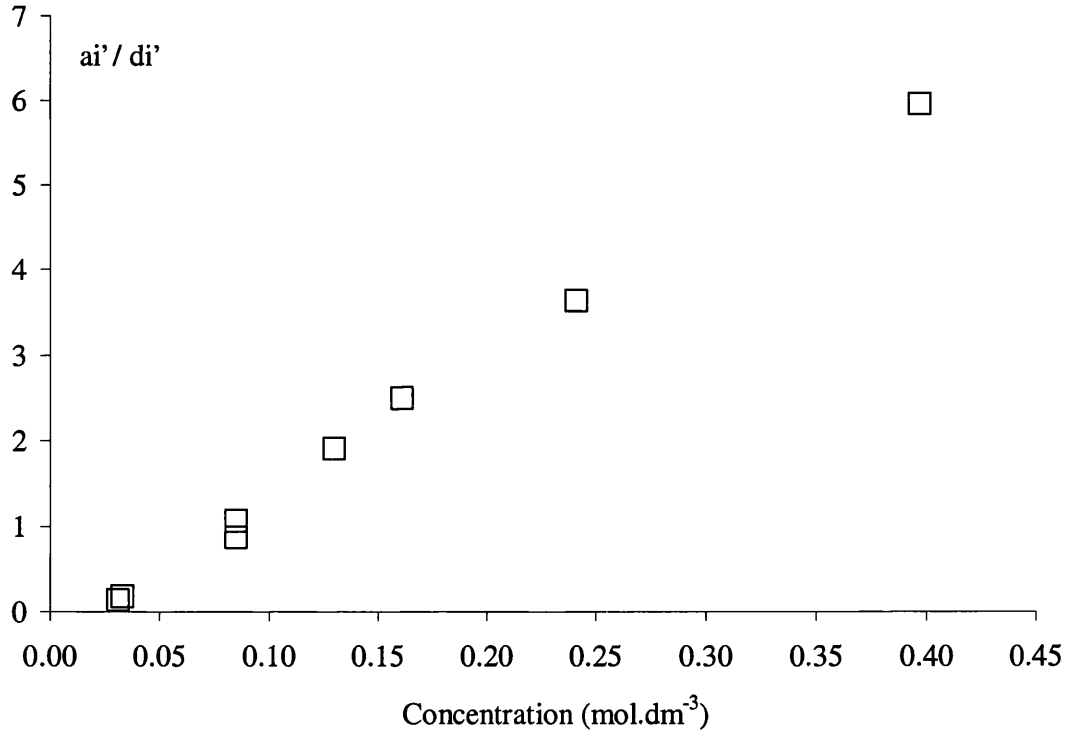


Figure 12.5 Plot of ratio of peak area (ai' / di') against the concentration (mol.dm^{-3}) for FEOH in octan-1-ol.

12.2.3.6 Complexation Constant Determination

Experiments were conducted in duplicate using two 155 cm^3 glass flasks, flask **1** and **2**. The liquid sample was composed of an inert standard, decane, a volatile hydrogen bond acid in dilute solution in a non-volatile solvent.

10 cm^3 of the solvent was poured into flask **1** by using a 10 cm^3 pipette (Class A). The two solutes were inserted directly into the solvent using Hamilton syringes (50mm^3, 100mm^3, 500mm^3 and 1 cm^3). The amount of solute introduced was determined by weighting it. Then flask **1** capped with a rubber septum, was delicately shaken to allow the components to mix. Next, 5 cm^3 of the stock solution were withdrawn and inserted into flask **2**. Two 5 cm^3 samples containing the same amount of solutes were then obtained. Septa were replaced and then pierced with a small needle to avoid pressure build-up. Finally, flasks **1** and **2** were placed into a water bath thermostated at 25°C and occasionally shaken manually for at least 30 minutes. Then, 3 cm^3 of the vapour phase in each flask was removed by the means of a 5 cm^3 gas-tight syringe (Hamilton) and injected on to a F-33 Perkin Elmer GC-FID equipped with a carbowax 20 M column. The vapour phase in each flask was analysed twice. An equilibrium time of at least 30 minutes was allowed between each headspace removal. The complexing agent or hydrogen bond base was introduced by means of a syringe of appropriate volume into the headspace glass flasks. Again the amount of solute added was determining by weighting it. The septa were changed after addition of the base. Then flask **1** and **2** were delicately shaken to allow the components to mix. They were replaced into the water bath thermostated at 25°C and let to equilibrate for at least one hour. Headspace in flasks **1** and **2** was analysed by GC as explained above.

Solutes peak areas before and after addition of the hydrogen bond base were retrieved by the means of a computer-assisted program called PIC 3.

12.2.3.7. Experimental note

All experiments were conducted in duplicate. All glassware and equipment used in this work was washed by a standard procedure, to minimise contamination. Firstly, the equipment was washed in hot water. Secondly, the equipment was thoroughly rinsed with reagent grade acetone. The equipment was finally allowed to dry on a nitrogen gas line.

12.2.3.8 Experimental Difficulties

Even though the HS-GC method is very simple experimental technique, its performance can be adversely influenced by several variables, such as sample preparation, headspace sampling and chromatographic analysis. Therefore, precautions had to be taken through the entire process in order to obtain reliable quantitative results.

a- Sample Preparation

Components had to be of high purity. Rubber septa, which were used as stoppers, were frequently changed to minimise the adsorption of lipophilic components onto the rubber surface as pointed by Drozd and Novak [38] in their review on headspace gas analysis by chromatography. Enough time had to be given to the volatile components to equilibrate between the sample and the gas phase. The so-called equilibration time depends on the ability of the components to distribute between the two phases, but also on the sample solution volume and upon the amount of aliquot withdrawn from the gas phase. Recently, Chai and Zhu [39] showed that the use of larger volume ratio (V^G/V^L) could reduce the equilibrium time for the experiment significantly. For instance, when volume ratio of 399, 199 and 1 were used, the system required less than five minute, five minutes and more than thirty minutes to equilibrate, respectively. In the present work, 155 cm^3 vials were filled with 5 to 6 cm^3 of solution depending on the system investigated, leading then to a phase ratio of 31. Thus, it was believed that the system would be fully equilibrated after 30 minutes.

b- Headspace Sampling

The aliquot of the headspace was taken manually by means of a Luer lock gas – tight syringe (Hamilton). This type of syringe insures that no loss of the aliquot can occur between removal from the vial and injection into the gas chromatogram. However, when the headspace sample is removed using gas-tight syringe, the syringe temperature must be equal or higher than that of the headspace. If it is not, the vapour could condense thus altering the concentration of the analysed components in the gas phase. The concentration of the components could be altered due to adsorption onto the

walls of the syringe. To prevent such effects the gas-tight syringe was kept on the top of the water bath control box, at temperature roughly equivalent to the water bath temperature (298K). In addition, the syringe was cleared after each sampling by flushing through nitrogen gas. Finally, the syringe was cleaned before storage according to the manufacturer's instructions.

A major uncertainty in analysis by HS-GC is due to the loss of headspace sample when it was introduced into the injector of the gas chromatogram. To prevent such leakage, injector seals had to be replaced often.

The results of the experiments carried out in n-hexadecane and octan-1-ol solvents are detailed in the following section.

12.2.4. Results and Discussion

12.2.4.1 Experiments in n-Hexadecane Solvent

Interactions between pentan-1-ol and 5 various bases: nitrobenzene, nonan-2-one, acetophenone, DMSO and triethylphosphate, in n-hexadecane solvent were investigated as described above. The concentrations of solutes in n-hexadecane varied between 0.03 and 0.05 mol.dm^{-3} for pentan-1-ol and between 0.12 and 0.13 mol.dm^{-3} for n-decane, see Table 12.15 in section 12.4. Oven temperature was set at 353K. Complexation constant values for pentan-1-ol and various bases in n-hexadecane solvents were calculated using equation (12.33). The complexation constant values are listed in Table 12.15 together with the concentration of pentan-1-ol, decane and bases; decane and pentan-1-ol peak area values before and after addition of the base.

a- Pentan-1-ol Self-Association in n-Hexadecane Solvent

Monomer concentration of pentan-1-ol in n-hexadecane

The peak area values, R'obs, were plotted against the concentration ratio, ai / di, see Figure 12.6. A curve was fit to the experimental data. The best relationship between the two parameters is given in equation (12.38)

$$R'obs = 14.61\ (ai/di)^4 - 20.02\ (ai/di)^3 - 9.73\ (ai/di)^2 + 25.32\ (ai/di) \quad (12.38)$$

$n = 15,\ r^2 = 0.983$

The linear relationship between Robs and ai/di, defined in equation (12.37), did not fit with the experimental results. This departure from linearity implied that pentan-1-ol was present as complexed species as well as free species in solution. R'calc could be deduced from the tangent of the curve at the origin. After differentiation of equation (12.38), one obtains,

$$R'calc = 25.32\ (ai/di) \quad (12.39)$$

R'calc values were calculated on equation (12.39) and are displayed in Table 12.16. R The ratio R' could then be estimated from R'obs and R'calc values. The concentration of free species in n-hexadecane, af, was obtained, see Table 12.16 in section 12.4.

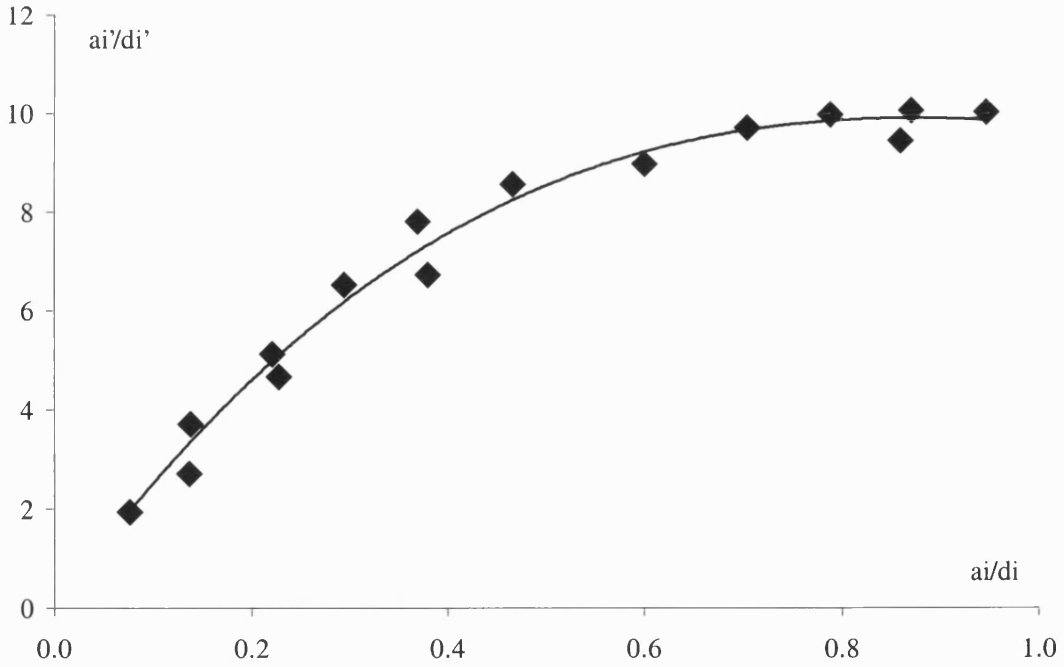


Figure 12.6. Plot of the ratio of observed peak areas, ai'/di', against the ratio of concentration, ai/di.

b- Nature of the pentan-1-ol self-association

The nature of alcohol self-association in various nonpolar solvents has been widely investigated.[40-42] The experimental data obtained through number of studies have been interpreted with a variety of models, e.g. one associated species in solution [monomer-trimer, monomer-dimer...] or several associated species in solution [monomer-trimer-octamer].[40-42] In this work, attention was essentially drawn to cases where only one complexed species was in solution. Three monomer-complexed species equilibrium were then investigated:

- Monomer-dimer species (n = 2)
- Monomer-trimer species (n = 3)
- Monomer-tetramer species (n = 4)

where n is the number of molecules complexing in accord with equation (12.34).

Monomer-dimer equilibrium

The concentration of dimer in solution, cf_2, was deduced from equation (12.34). Next, according to equation (12.35), cf was plotted against $(af)^2$ in order to obtain the equilibrium constant K_2. A linear equation was fitted to the experimental data with the intercept set to the origin. The gradient of this line will give the value of K_2. The K_2-value and the value of the correlation coefficient, r^2, are displayed in Table 12.13.

Monomer-trimer equilibrium

The concentration of trimer in solution, cf_3 and the value of the equilibrium constant K_3 were determined as in the previous example with now n equal to three. Here, cf_3 was plotted against $(af)^3$. Values of K_3 as well as the correlation coefficient of the fit, r^2, are given in Table 12.13.

Monomer-tetramer equilibrium

The concentration of tetramer in solution, cf_4, was obtained as described in the monomer-dimer equilibrium example. The value of the equilibrium constant, K_4, was deduced from the plot cf vs $(af)^4$. Results are given in Tables 12.13.

Table 12.13. K_n and r^2 values

Pentan-1-ol			Methanol		
n	K_n	r^2	n	K_n	r^2
2	9.43	0.80	2	13.03	0.79
3	25.21	0.94	3	195.37	0.91
4	1835.40	0.96	4	3174.1	0.97

Since a poor correlation coefficient was obtained, it was not reasonable to assume the presence of dimer species in n-hexadecane. The results of the investigation mainly showed that there is no significant difference between the monomer-trimer and monomer-tetramer equilibrium. However, the monomer-tetramer fit was slightly better than the monomer-trimer fit indicating that an assumed tetramer rather than trimer is more nearly consistent with the data. For this reason, it was decided to carry on the investigation on the monomer-tetramer equilibrium.

Penta-1-ol self-association in n-decane solvent was investigated by Sjöblom and co-workers [41] by measurements of the pentan-1-ol vapour pressure and of the NMR chemical shift of the hydroxyl protons. In their study, the authors considered several models for association equilibrium in the solutions mainly based on cases where two alcoholic species are in solutions: the alcohol monomer and the associated species containing n associated species. Employing a least-square computer program, they determined the more suitable model fitting with the experimental data. For pentan-1-ol in n-decane the monomer-tetramer and monomer-pentamer equilibrium lead to the best fittings. In the present study on pentan-1-ol, the most likely monomer-associated species mixture in solution in n-hexadecane was found to be the monomer-tetramer one. This result agrees well with the one obtained by Sjöblom and co-workers in n-decane.

Comparison with Tucker et al. Data

The following investigation on methanol self-association in n-hexadecane was undertaken using literature data in order to determine whether an association model can be reasonably predicted from the results of the gas chromatographic headspace analysis. In previous studies on association of alcohol in non-polar solvents, Tucker and co-

workers [40,42] measured a number of vapour pressures, P, and corresponding concentrations of methanol in n-hexadecane, ai, at 298K. Then, the monomeric concentration in the vapour phase, a_G, could be obtained as follows:

$$a_G = P / RT \qquad (12.40)$$

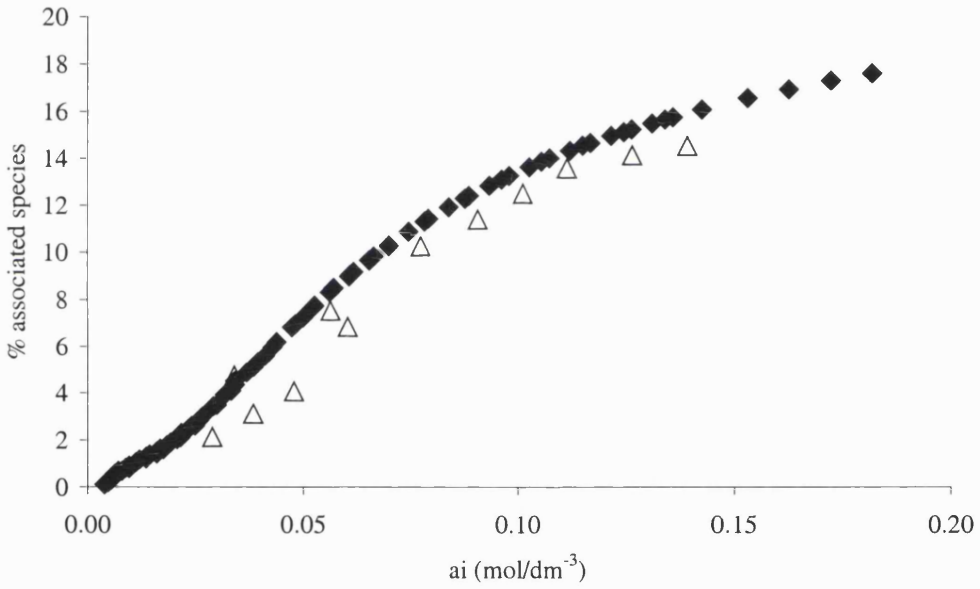
where VP is the vapour pressure of methanol (atm), R is the gas constant (R=8.20575E-02 $dm^3.atm.K^{-1}.mol^{-1}$) and T the temperature (K). Then, from the monomer concentration in the gas phase and the Ostwald solubility coefficient n-hexadecane, L[16], the monomer concentration in the liquid phase, af, was obtained according to equation (12.24). The distribution constant value, L, was determined by Tucker and co-workers[42] and was found to be 8.249 at 298K.

Cases, in which only one associated species was present in solution, were investigated, viz. monomer-dimer species, monomer-trimer species and monomer-tetramer species. For each situation, the amount of self-associated methanol, cf_n, as well as the value of the equilibrium constant, K_n, were determined as described above. K_2, K_3 and K_4-values were given by the gradient of the plots cf_n vs af^n, with n equal to 2, 3, 4, respectively. The calculations for methanol association have been reported elsewhere. In Table 12.13 are displayed the equilibrium constant values, K_n, and the coefficient of the correlation between cf_n and $(af)^n$, r^2, for methanol association. Due to the poor r^2 value, the monomer-dimer species equilibrium in n-hexadecane seems to be unreasonable. The best fit was found to be the monomer-tetramer one, indicating that an assumed tetramer is consistent with the data.

In conclusion, tetramer species are the predominant associated species in pentan-1-ol and methanol solution in non-polar solvents. On a same graph, was plotted the percentage of pentan-1-ol and methanol tetramer species in n-hexadecane solvent ($100(cf_4/ai)$) against the initial concentration of alcohol in solution, ai, see Figure 12.7. The plots are rather similar. It was concluded that the amount of self-associated species does not depend on the nature of the alcohol. From the experimental point of view, it is true that vapour pressure data provide a more severe test of an associated model [42] than do gas chromatographic headspace analysis data. However, data obtained in this work and these derived from the Tucker and co-workers study agree pretty well, proving that the present method can lead to consistent and reliable results.

12.2.4.2 Validation of Experimental Procedure

Since HS-GC method to measure 1:1 interaction was developed in this work, it was very important to validate the present experimental technique. The usual way to validate a technique is to compare experimental data to literature data. Unfortunately no log K values for 1:1 association in n-hexadecane or in octan-1-ol have been measured so far. However, 430 log K values in cyclohexane have been used to establish equation



(12.13).

Figure 12.7 Plot of the percentage of tetramer species in n-hexadecane against the concentration in alcohol in solution, ai. Empty triangles: pentan-1-ol, filled diamonds: methanol.

This equation (12.13) can be used to estimate log K value for any 1:1 interaction as long as the α_2^H and β_2^H values are known. Therefore the first experiment conducted was to measure log K values in n-hexadecane whose properties are similar to those for cyclohexane. Then, an experimental values in n-hexadecane could be compared with those estimated on equation (12.13). The results are tabulated in Table 12.14.

Table 12.14. Validation of experimental method by comparing observed log K values in n-hexadecane with calculated log K values in cyclohexane.

Acid	Base	α_2^H	β_2^H	$\alpha_2^H.\beta_2^H$	Obs. Log K[a]	Calc log K[b]	Residual[c]
Pentan-1-ol	Nitrobenzene	0.328	0.340	0.112	-0.103	-0.098	-0.004
Pentan-1-ol	Nonan-2-one	0.328	0.510	0.167	0.470	0.330	0.140
Pentan-1-ol	Acetophenone	0.328	0.510	0.167	0.286	0.330	-0.044
Pentan-1-ol	DMSO	0.328	0.780	0.256	1.186	1.009	0.177
Pentan-1-ol	TEP	0.328	0.790	0.259	1.450	1.034	0.416

[a] Experimental log K values obtained with an error of ± 0.05 log units.
[b] Calculated on equation (12.13) with an error of ± 0.174 log units.
[c] Residual = Obs. log K - Calc. log K

The method developed in the present work was able to produce compatible results to those calculated on equation (12.13), at least for the first four systems listed in Table 12.14. This investigation and the study carried out on pentan-1-ol self-association in n-hexadecane validated the current experimental procedure.

12.2.4.3. Experiment in Octan-1-ol

Interactions between five hydrogen bond donors, FEOH, TFE, HFMP, HFIP and PTB, and five hydrogen bond acceptors, DMF, DMA, DMSO, TEP, TMG and HMPA, have been investigated. Experiments were carried out as explained above. GC Oven temperature varied according to the hydrogen donor studied: FEOH (353K), TFE (K), HFMP (343K), HFIP (393K) and PTB (363K). Note that no experiments were carried out to determine the complexation constant between PTB and TMG due to instant reaction between the two components; fumes appeared in the vial. Hence, twenty-nine complexation constants were measured, as listed in Table 12.17 in section 12.4.

12.2.4.4. Analysis of Error

Having two sets of data available for each measurement is useful, since it allows some error range to be assessed. The relative standard deviation, RSD, (RSD = (sd*100)/mean value) for peak area ratio before and after addition of the base, ai'/di' and af'df', in Table 12.15 varies from 0.1 to 34%. The average error in log K values in n-hexadecane was also estimated and was found to be around 0.050, ranging from 0.023

log units to about 0.110 log units. The largest discrepancy was obtained for the system pentan-1-ol / nonan-2-one. The average error in log K values in octan-1-ol was calculated to be around 0.08 log units, ranging from 0.004 log units (FEOH / DMF) to about 0.18 log units (HFMP / TEP).

12.2.4.5. Relationship Between Acidity and Basicity of VOCs and the Extent of Complexation in Octan-1-ol

The twenty-nine experimentally determined log K values were plotted against the corresponding $\alpha_2^H.\beta_2^H$ values. The α_2^H and β_2^H values were taken from the literature. A common trend for the complexation of various VOCs in octan-1-ol was clearly evident in Figure 12.8. Three outliers were, however, highlighted; they were the pairs HFIP/TEG , PTB/ DMSO and PTB/ HMPA.

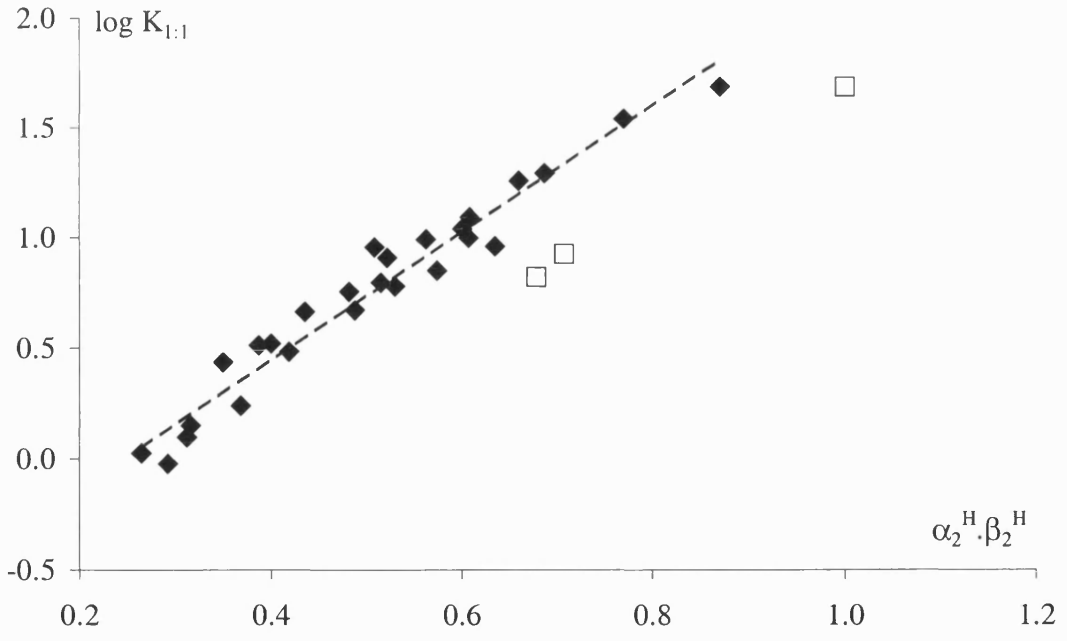


Figure 12.8. Plot of log $K_{1:1}$ values for twenty-nine binary mixtures against the term $\alpha_2^H.\beta_2^H$ for various acids and bases in octan-1-ol. --- Outliers

If the log K for overall complexation is presumed to be essentially the same as the 1:1 hydrogen bonding interaction then, the following equation can be constructed.

$$\log K_{1:1} = 2.950\ (\pm 0.138)\ \alpha_2^H.\beta_2^H - 0.741\ (\pm 0.070) \qquad (12.12)$$

$n = 26$, $r^2 = 0.950$, sd = 0.092, F = 458

A plot of observed log $K_{1:1}$ values against log $K_{1:1}$ values calculated on equation (12.15) may be done, see Figure 12.9.

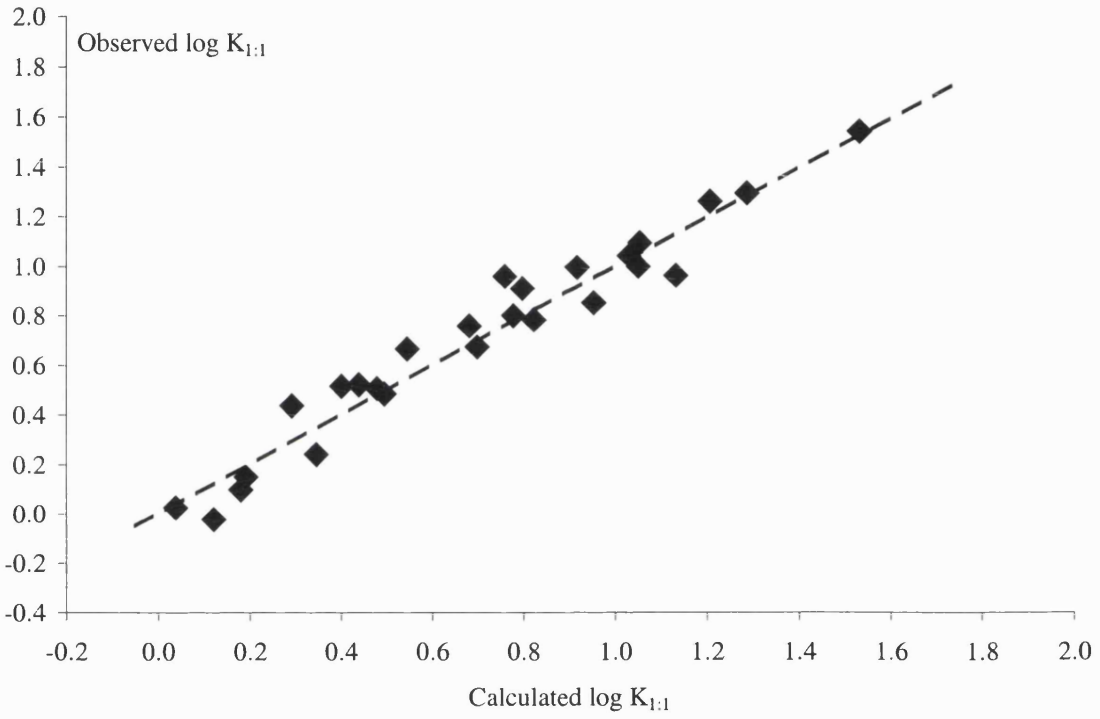


Figure 12.9. Plot of observed log $K_{1:1}$ in octan-1-ol against calculated log $K_{1:1}$ on equation (12.15). ---- Identity line

12.3. Conclusion

In this research the complexation between various acids and bases in octan-1-ol was measured by the headspace gas chromatographic method of analysis. From the data obtained, the complexation constants for the 1:1 interaction between a binary mixture of butan-1-ol and heptan-2-one was calculated. A model was proposed to estimate complexation constants in other solvents such as dimethylformamide. From the experimental and calculated complexation constants, the percentage of interaction

between the two VOCs at the receptor area was estimated, with octan-1-ol and dimethylformamide as receptor area models.

The results obtained indicate that at the olfactory receptor area, the interaction between the VOCs is insignificant, thus accounting for the observation of dose of additivity with regards to odor thresholds.

In addition, the results revealed that the interactions between VOCs at the nasal pungency receptor area are minimal. This finding is in agreement with the observation of dose additivity for sensory irritation thresholds. Therefore one can substantiate the hypothesis that the chemesthetic potency of VOCs is largely governed by the physicochemical transport processes.

As octan-1-ol and/or dimethylformamide, and the receptor area are polar and acidic, the extent of interaction between these phases and the VOCs is so large that there is very little complexation between the volatile components. Such an observation is in line with the theory that the extent of complexation between VOCs at the receptor area is either insignificant or minimal. Therefore one conclude that the chemosensory potency of a mixture of VOCs is not significantly reduced by interactions between the components.

This research involved a novel procedure, the hydrogen bond complexation constants for uncharged species are the first to be determined in such an acidic and polar medium as octan-1-ol. Consequently it is impossible to make comparisons with other previous research. However, the results obtained illustrate the success of this new method and hence provide the opportunity for achieving a better understanding of the chemosensory impact of mixtures of VOCs in the not so distant future.

12.4. Appendix

12.4.1. Appendix 1

Table 12.15. Headspace analysis results in n-hexadecane solvent

Base	ai [a]	di [a]	ai'/di' [b]	sd [c]	af [d]	bf [d]	af'/df' [b]	sd [c]	R [e]	K [f]	logK	sd [c]
Nitrobenzene	0.028	0.138	3.377	0.041	0.025	0.426	2.514	0.078	0.744	0.790	-0.103	0.023
Nitrobenzene	0.028	0.139	4.012	0.182	0.026	0.506	2.909	0.059	0.725			
Nonan-2-one	0.045	0.132	5.847	0.537	0.042	0.585	2.776	0.238	0.475	2.949	0.470	0.034
Nonan-2-one	0.050	0.145	4.716	0.713	0.047	0.579	2.312	0.114	0.490			
Acetophenone	0.057	0.133	6.598	0.216	0.054	0.455	3.934	0.169	0.592	1.933	0.286	0.109
Acetophenone	0.058	0.135	6.676	0.159	0.054	0.498	3.496	0.304	0.483			
DMSO	0.044	0.124	5.585	0.268	0.044	0.081	2.971	0.125	0.532	15.348	1.186	0.031
DMSO	0.047	0.131	6.067	0.313	0.054	0.083	3.169	0.038	0.522			
TEP	0.056	0.122	6.615	0.140	0.055	0.086	2.746	0.217	0.415	28.193	1.450	0.044
TEP	0.059	0.128	7.228	0.153	0.058	0.084	2.916	0.268	0.403			

[a] Initial concentration of pentan-1-ol (ai) and Decane (di), $mol.dm^{-3}$.
[b] Average ratio of two measurements of peak areas, ai' and di', before addition of the base.
[c] Standard deviation based on two ratio values.
[d] Final concentration of pentan-1-ol (af) and concentration of the base (bf), $mol.dm^{-3}$.
[d] Average ratio of two peak areas, ai' and di', measurements after addition of the base.
[e] Calculated on equation (12.29).
[f] Average value obtained from two K values calculated on equation (12.32).

Table 12.16. Initial concentration in pentan-1-ol and decane in n-hexadecane and ratio of peak area values.

ai ($mol.dm^{-3}$)	ai/di	R'obs [a]	R'calc[b]
0.029	0.221	5.12	5.60
0.034	0.228	4.67	5.76
0.038	0.294	6.52	7.45
0.048	0.369	7.82	9.35
0.056	0.379	6.72	9.61
0.060	0.466	8.57	11.80
0.077	0.600	8.98	15.21
0.091	0.704	9.71	17.83
0.101	0.788	9.98	19.95
0.111	0.869	10.06	22.01
0.126	0.858	9.44	21.74
0.139	0.946	10.03	23.96

[a] Average values of two measurements.
[b] Calculated on equation (12.39)

Table 12.17. Headspace analysis result in octan-1-ol solvent

Base used	ai [a]	di [a]	ai'/di' [b]	sd [c]	af [d]	bf [d]	af'/df' [b]	sd [c]	R [e]	K [f]	logK	sd [c]
						FEOH						
DMF	0.096	0.052	2.742	0.073	0.090	0.833	1.457	0.019	0.531	1.061	0.026	0.032
DMF	0.096	0.052	2.749	0.250	0.090	0.801	1.556	0.005	0.566			
DMA	0.092	0.046	2.898	0.109	0.084	0.851	1.639	0.023	0.566	0.949	-0.023	0.004
DMA	0.092	0.046	2.839	0.334	0.085	0.807	1.636	0.044	0.576			
DMSO	0.089	0.050	3.428	0.183	0.086	0.586	2.011	0.140	0.537	1.255	0.099	0.161
DMSO	0.089	0.050	2.916	0.317	0.086	0.600	1.841	0.119	0.653			
TEP	0.092	0.047	3.011	0.048	0.081	0.757	2.199	0.169	0.730	1.410	0.150	0.131
TEP	0.092	0.047	3.022	0.175	0.081	0.713	1.978	0.072	0.655			
TMG	0.105	0.049	3.332	0.330	0.103	0.152	2.683	0.088	0.846	1.738	0.240	0.144
TMG	0.105	0.049	3.019	0.167	0.103	0.132	2.595	0.091	0.807			
HMPA	0.098	0.049	3.050	0.154	0.097	0.104	2.308	0.726	0.757	3.315	0.520	0.127
HMPA	0.098	0.049	2.859	0.118	0.097	0.102	2.338	0.120	0.818			
						TFE						
DMF	0.622	0.236	3.127	0.707	0.547	1.557	0.807	0.131	0.258	2.731	0.436	0.053
DMF	0.622	0.236	3.375	0.612	0.543	1.570	0.765	0.136	0.227			
DMA	0.528	0.192	3.361	1.002	0.485	0.932	1.156	0.078	0.344	3.268	0.514	0.030
DMA	0.529	0.192	3.395	0.673	0.482	0.991	1.044	0.109	0.308			

base used	ai [a]	di [a]	ai'/di' [b]	sd [c]	af [d]	bf [d]	af'/df' [b]	sd [c]	R [e]	K [f]	logK	sd [c]
DMSO	0.465	0.223	1.937	0.084	0.432	1.020	0.651	0.063	0.370	3.199	0.505	0.183
DMSO	0.553	0.224	2.087	0.028	0.426	1.056	0.522	0.000	0.248			
TEP	0.601	0.240	2.905	0.283	0.505	0.937	1.032	0.297	0.355	3.045	0.484	0.016
TEP	0.468	0.240	2.823	0.266	0.392	0.954	0.898	0.020	0.318			
TMG	0.395	0.097	3.464	0.369	0.330	1.314	0.590	0.019	0.170	4.715	0.673	0.004
TMG	0.395	0.097	3.358	0.024	0.331	1.286	0.578	0.027	0.172			
HMPA	0.469	0.229	1.917	0.200	0.420	0.598	0.647	0.076	0.337	6.027	0.780	0.012
HMPA	0.470	0.229	1.813	0.156	0.366	0.628	0.563	0.002	0.311			
						HFMP						
DMF	0.627	0.088	8.210	0.649	0.557	1.389	1.468	0.478	0.179	4.618	0.664	0.041
DMF	0.627	0.088	8.228	0.456	0.557	1.470	1.530	0.026	0.186			
DMA	0.599	0.110	7.808	1.018	0.557	0.782	2.835	0.788	0.363	5.709	0.757	0.177
DMA	0.599	0.110	6.882	0.782	0.556	0.779	1.849	0.425	0.269			
DMSO	0.599	0.093	9.399	0.482	0.556	1.080	2.121	0.157	0.229	6.275	0.798	0.110
DMSO	0.599	0.093	9.477	0.471	0.550	1.161	1.692	0.145	0.162			
TEP	0.653	0.124	7.484	1.746	0.538	1.035	1.635	0.296	0.218	8.107	0.909	0.178
TEP	0.573	0.124	8.108	1.507	0.470	1.053	1.055	0.059	0.130			
TMG	0.230	0.120	1.139	0.011	0.217	0.451	0.301	0.007	0.265	10.201	1.009	0.040
TMG	0.230	0.120	1.132	0.031	0.235	0.480	0.266	0.027	0.235			
HMPA	0.625	0.130	8.156	0.034	0.561	0.608	1.989	0.456	0.244	18.257	1.261	0.048
HMPA	0.627	0.130	6.832	1.099	0.563	0.592	1.622	0.462	0.237			

base used	ai [a]	di [a]	ai'/di' [b]	sd [c]	af [d]	bf [d]	af'/df' [b]	sd [c]	R [e]	K [f]	logK	sd [c]
						HFIP						
DMF	0.713	0.090	1.138	0.119	0.647	1.200	0.138	0.003	0.121	9.061	0.957	0.170
DMF	0.714	0.090	1.075	0.050	0.632	1.450	0.153	0.002	0.143			
DMA	0.587	0.098	0.703	0.013	0.546	0.807	0.148	0.018	0.210	9.858	0.994	0.007
DMA	0.588	0.098	0.750	0.045	0.476	0.844	0.138	0.002	0.184			
DMSO	0.814	0.101	0.980	0.084	0.751	1.064	0.188	0.019	0.168	11.017	1.042	0.014
DMSO	0.814	0.101	0.957	0.019	0.747	1.141	0.167	0.024	0.154			
TEP	0.653	0.124	1.319	0.040	0.737	0.839	0.335	0.046	0.254	12.437	1.095	0.114
TEP	0.573	0.124	1.220	0.013	0.736	0.852	0.248	0.018	0.204			
TMG	0.497	0.040	5.072	0.724	0.442	0.872	0.813	0.041	0.160	8.469	0.928	0.148
TMG	0.497	0.040	3.976	0.212	0.443	0.868	0.903	0.028	0.227			
HMPA	0.735	0.091	1.062	0.044	0.658	0.694	0.192	0.021	0.181	34.999	1.544	0.101
HMPA	0.736	0.091	1.001	0.004	0.646	0.708	0.139	0.008	0.139			
						PTB						
DMF	0.622	0.008	3.127	0.251	0.570	1.080	0.591	0.021	0.189	7.102	0.851	0.013
DMF	0.622	0.008	3.110	0.070	0.568	1.116	0.546	0.081	0.176			
DMA	0.622	0.008	3.105	0.358	0.554	1.178	0.425	0.041	0.137	9.169	0.962	0.011
DMA	0.622	0.008	3.036	0.129	0.553	1.190	0.398	0.030	0.131			
DMSO	0.596	0.009	3.065	0.099	0.567	0.682	0.988	0.001	0.330	6.674	0.824	0.006

base used	ai [a]	di [a]	ai'/di' [b]	sd [c]	af [d]	bf [d]	af'/df' [b]	sd [c]	R [e]	K [f]	logK	sd [c]
DMSO	0.596	0.009	2.804	0.069	0.568	0.676	0.931	0.039	0.336			
TEP	0.687	0.010	4.375	0.123	0.611	0.645	0.892	0.025	0.204	19.739	1.295	0.156
TEP	0.687	0.010	3.516	0.248	0.608	0.676	0.843	0.080	0.240			
HMPA	0.622	0.008	3.142	0.193	0.563	0.573	0.494	0.009	0.157	49.090	1.691	0.067
HMPA	0.622	0.008	3.176	0.138	0.562	0.574	0.549	0.017	0.173			

[a] Initial concentration of pentan-1-ol (ai) and decane (di) in $mol.dm^{-3}$.
[b] Average ratio of two measurements of peak areas, ai' and di', before addition of the base.
[c] Standard deviation based on two ratio values.
[d] Final concentration of pentan-1-ol (af) and concentration of the base (bf), in $mol.dm^{-3}$.
[d] Average ratio of two peak areas, ai' and di', measurements after addition of the base.
[e] Calculated on equation (12.29).
[f] Average value obtained from two K values calculated on equation (12.32).

12.4.2. Appendix 2 – E_T (30) value in the gas phase

In Table 12.2 contains a gas-phase $E_T(30)$ value of 25.2 kcal.mol-1, E_T^N = -0.170, which derived from empirical linear correlations between $E_T(30)$ and thermodynamic data of the solvent dependent equilibrium between configurational isomers. Three correlations of the form $E_T(30) = M.X + C$ were developed with x being either $-\Delta G^o$ for equilibrium between the 1,2-dibromo-4-t-butylcyclohexanes[43] or DE, the rotamer energy difference for 1,2-dichloroethane and 1-fluoro-2-chloroethane in various solvents[44]. Summary of the correlations is displayed in Table 12.18. Values for gas-phase $E_T(30)$ were calculated from equations (12.42)-(12.43) in Table 12.18 and were then average to give the $E_T(30)$ values used throughout the present work, 25.2 ± 0.8. This calculated $E_T(30)$ value is in remarkable agreement with another gas-phase $E_T(30)$[24] values derived from linear correlations between measured and calculated maximum wave length values of the betaine dye.

Table 12.18. Relationship between $E_T(30)$ values and thermodynamic data.

Ref	Eq.	X	M	C	n	r^2	Calc $E_T(30)$	sd
43	(12.41)	-ΔGo	0.017	25.400	6	0.961	25.400	1.266
44	(12.42)	ΔE	-16.900	46.200	8	0.957	25.920	0.981
44	(12.43)	ΔE	-10.800	34.000	8	0.987	24.287	0.759

12.5 References

1. J.E. Cometto-Muniz, W.S. Cain, H.K. Hudnell, Percept. & Psychophys., (1997).
2. J.E. Cometto-Muniz, M.R. Garcia-Medina, A.M. Calvino, Chem. Senses, 14 (1989) 163.
3. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham and J.M.R. Gola,, Physiology and behavior, 67 (1999) 269.
4. J.E. Cometto-Muniz, W.S. Cain, M.H. Abraham, J.M.R. Gola, Toxicol. Sci. (2001).
5. E.M. Arnett, L. Joris, E. Mitchell, T.S.S.R. Murty, T.M. Gorrie, P.V.R. Schleyer, J.Am.Chem.Soc., 92 (1970) 2365.
6. D.J.W. Grant, T. Higushi, Techniques of Chemistry, Vol. XXI, John Wilhey & Son, 1990.
7. M. Nakano, N.I. Nakano, T. Higushi, J. Phys. Chem., 71 (1967) 3954.
8. D. Gurka, R.W. Taft, J. Am. Chem. Soc., 91 (1969) 4794.
9. E.M. Arnett, L. Joris, E. Mitchell, T.S.S.R. Murty, T.M. Gorrie, P.v.R. Schleyer, J. Am.Chem.Soc., 92 (1970) 2365.
10. M. Nakano, T. Higushi, J. Pharm. Sci., (1968) 1865.
11. G.C. Pimentel, A.L. McClellan, The Hydrogen Bond, Freeman, New York, (1960).
12. J. Yarwood, Spectroscopy and Structure of Molecular Complexes, Plenum Press, London, (1973).
13. P. Schuster, G. Zundel, C. Sandorfy, The Hydrogen Bond, Vol. I to III, North Holland, Amsterdam, (1976).
14. M. Ratajaczack, W.J. Orville-Thomas, Molecular Interactions, Vol. I to III, Wiley, New York, (1980).
15. P. Schuster, Hydrogen Bond –Topics in Current Chemistry, Vol. 120, Springer, Berlin, (1984).
16. P.R. Griffiths, J.A. de Haseth, Fourier Transform Infrared Spectrometry – Chemical Analysis, Vol. 83, Wiley, New York, (1986).
17. A. Chardin, C. Laurence, M. Berthelot, D.G. Morris, J.Chem.Soc., Perkin Trans. 2, (1996) 1047.
18. M.H. Abraham, P.L. Grellier, D.V. Prior, R.W. Taft, J.J. Morris, P.J. Taylor, C. Laurence, M. Berthelot, R.M. Doherty, M.J. Kamlet, J.-L. M. Abboud, K. Sraidi, J. Am. Chem. Soc., 110 (1988) 8534.
19. J.Marco, J.M.Orga, R.Notario and J.L.M.Abboud, J.Am.Chem.Soc. ,116 (1994) 8841.
20. M.H.Abraham, M.berthelot, C.Laurence and P.J.Taylor, J.Chem.Soc.Perkin Trans 2, 1998, 187.
21. M.H.Abraham, unpublished results.

22. J.-L. Abboud, R. Notario, V. Botella, in Quantitative Treatments of Solute / Solvents Interactions – Theoritical and Computational Chemistry Vol. 1, Eds. P. Politzer, J.S.Murray, Elsevier Science B.V., (1994) 135.
23. T. Abe, J.-L. M. Abboud, F. Belio, E.Bosch, J.I. Garcia, J.A. Mayoral, R. Notario, J.Ortega, M. Roses, J. Phys. Org. Chem., 11 (1998) 193.
24. C. Reichardt, Solvents and Solvent Effects in Organic Chemistry, 2nd Ed., VCH, Weinheim, (1990).
25. K. Dimroth, C. Reichardt, T. Siepmann, F. Bohlmann, Liebigs Ann. Chem., 661 (1993) 1.
26. M.J. Kamlet and R.W. Taft, J.Am.Chem.Soc., 98 (1976) 2886.
27. M.J. Kamlet, J.L. Abboud, R.W. Taft, J.Am.Chem. Soc., 99 (1977) 6027; ibid. 99 (1977) 8325.
28. J.-L. M. Abboud, M.J. Kamlet, R.W. Taft, Progr.Phys.Org.Chem., 13 (1981) 485.
29. R.W. Taft, J.-L. M. Abboud, M.J. Kamlet, M.H.Abraham, J.Solution Chem, 14 (1985) 153.
30. M.J. Kamlet, R.W. Taft, Acta Chem. Scand., Part B 40 (1986) 619.
31. M.J. Kamlet, J.-L. M. Abboud, M.H. Abraham, R.W. Taft, J.Org. Chem., 48 (1983) 2877.
32. Y. Marcus, J. Solut. Chem., 20 (1990) 929.
33. M.H.Abraham, P.L.Grellier and R.A. McGill, J.Chem.Soc.Perkin Trans2, (1987) 797.
34. M.H.Abraham, Chem.Soc.Rev., 22 (1993) 72.
35. M.H.Abraham and J.Le, unpublished work.
36. M.H. Abraham, unpublished work.
37. A.Massat, A.Cosse-Barbi and J.P.Doucet, J.Mol.Struct., 13 (1989) 212m.
38. J.Drodz and J.Novak, J.Chromatograph., 165 (1979) 141.
39. Chai
40. E.E.Tucker, E.D.Becker, J.Phys.Chem, 1973, 77, 1783.
41. E.Sjoblom, U.Henriksson and P.Stenius, Finn.Chem.Lett, 1982, 114.
42. E.E.Tucker, S.B. farnham and S.D. Christian, J.Phys.Chem, 1969, 73, 3820.
43. M.H. Abraham, L.E. Xodo, M.J. Cook, R. Cruz, J. Chem. Soc. Perkin Trans II, (1982) 1503.
44. R.J. Abraham, E. Bretschneider, Medium Effects on Rotational and Conformational Equilibria.

Chapter 13 Future Work

Future work will no doubt involve determination of Abraham descriptors for more indoor air pollutants, but attention should be drawn to the analysis of cut off phenomenon, as outlined below.

13.0. Resolution of physical versus biological cut offs

As mentioned under preliminary studies, previous work on human chemosensory detection of members of homologous chemical series uncovered the existence of cut offs for chemesthetic potency. However, this work was not designed to focus on the phenomenon. Such cut offs can, in principle, reflect a physical or a biological mechanism. Under a purely physical mechanism the maximum available concentration of stimulus in the vapor phase (at, for example, room temperature) simply falls below the concentration necessary to evoke the chemosensory response. Under a biological mechanism the stimulus molecules lack a different kind of key property (e.g., size, shape, maximum length) to trigger transduction.

Although both mechanisms offer valuable information as to the conditions in which to expect chemesthetic effects from a VOC, only those stimuli reaching a cut off via a biological mechanism constitute an effective tool to probe the molecular boundaries for evoking human trigeminal chemesthesis at the receptor biophase. Elucidation of this issue in the future could follow a two-step approach.

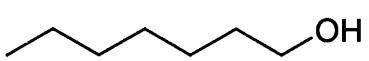
(i). A solid assessment of whether a cut off stems from a physical or a biological mechanism can be made by examining the relationships between observed chemesthetic thresholds[1], predicted chemesthetic thresholds[2,3] and saturated vapor concentrations (at 296 K, i.e., room temperature) along each homologous series targeted for study. If a projection of predicted thresholds for homologs within (and, even, beyond) the cut off range shows the thresholds to be clearly below saturated vapor concentration at 296 K, a biological mechanism is likely to be responsible for the cut off.

(ii). Further insight into the type of cut off mechanism involved could be obtained by testing whether the chemesthetic potency of an inactive (i.e., cut off) homolog can be induced by increasing the temperature of the vapor source (for example, raising it to 310 K). This procedure will raise the saturated vapor concentration achieved by the stimulus, increasing even further the gap with predicted threshold and providing ample room for the chemesthetic response to be evoked if, indeed, a concentration restriction is the problem. If the homolog still fails to evoke a response under these conditions, the possibility of a physical cut off can be safely ruled out.

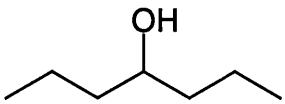
These experiments to further rule-out a physical cut off could be performed on the first (i.e., lowest) homolog reaching a cut-off within each series (for nasal pungency and for eye irritation). If, for any of the series, the loss of chemesthetic impact rests on a physical cut off, the first homolog to show such loss will be the first to overcome it when an increased temperature raises the saturated vapor concentration. Chemesthetic responses for these stimuli under a raised temperature could be tested at two sites: the nasal mucosa (to probe nasal pungency) and the ocular mucosa (to probe eye irritation).

13.0.1. Calculation of cut-off molecular determinants

In the preliminary analysis of odor thresholds, the maximum length (i.e., the unfolded length) of a VOC was used. The calculations of maximum length were carried out using the molecular modeling package Molecular Modeling Pro. As a first step, the maximum length analysis could be applied to trigeminal chemosensory results for homologous series of general structure (C_nH_{2n+1})-X where X is a functional group. However, it would be extremely instructive if analogous VOCs were studied where the functional group was in the middle of the carbon chain, rather than at the end. Of all the functional group possibilities, the alcohols seem the most interesting, because the configuration of the carbon backbone remains the same, all the carbon atoms having a tetrahedral disposition of valences,



A: Heptan-1-ol



B: Heptan-4-ol

Examination of these alcohols, A and B, and also other compounds such as ethers, R-O-R, and ketones, R-C(=O)-R, would enable one to deduce if the 'maximum length' applies to the overall maximum length irrespective of functional group, or to the maximum length to the functional group. The overall maximum length of A and B are about the same, but the length to the functional group is much larger in A than in B.

Most of the molecules in the homologous series that have been investigated, as well as those suggested above, have been acyclic aliphatic compounds that are flexible enough to adopt various conformations. It would be of considerable interest to examine more rigid molecules in order to test the hypothesis of 'maximum unfolded length' in more detail, or, indeed, to develop other hypotheses. There are a few acyclic compounds that exist as single conformers for which exact dimensions can be calculated; two such examples are t-butyl methyl ether (t-Bu-O-Me) and pinacolone (t-Bu-CO-Me). However, most examples of conformationally fixed molecules are cyclic compounds. As ethers there are 1,8-cineole, α-pinene oxide and β-pinene oxide. Very rigid ketones include norcamphor, camphor, 1-adamantyl methyl ketone and 3-chloro-2-norbornanone. Ketones with limited flexibility are cyclopentanone and cyclohexanone (and their alkyl substituted derivatives). Rigid esters include bornyl acetate and exo-2-norbornyl formate. There are a large number of alcohols that are either rigid or are of limited flexibility, for example adamantanol, borneol, the cis- and trans-methylcyclopentanols and the cis- and trans-methylcyclohexanols, as well as cyclopentanol and cyclohexanol themselves.

Although there may be problems over volatility, aromatic compounds give rise to numerous series that are conformationally restricted. The p-substituted ethers of type R-C_6H_4-OMe where R = H, Methyl, Ethyl, iso-Propyl and tert-Butyl form such a series, as do ketones, R-C_6H_4-COMe, and many other series of the general form R-C_6H_4-X

where R is as before and X can be F, Cl, Br, CN or even OH although volatility problems may preclude the last (OH) series.

In any examination of a chemical cut-off effect in a homologous series, it is useful to be able to estimate a 'base line' or 'normal' effect. However, when dealing with compounds that do not form a homologous series, for example the more rigid molecules discussed above, it is essential to be able to estimate a 'base line' effect - this is the only way of evaluating quantitatively any diminution of potency. Such an estimation requires the construction of a general QSAR that will predict 'base line' potencies, hence the absolute necessity for equations like those given as equations for nasal pungency and odour thresholds. Once extra information on the potency of flexible and rigid molecules is available, the 'base line' equations can be modified to include parameters related to the cut-off effect, as presented in chapter 10.

Of course, the length parameter is hardly the only cut-off parameter one could calculate and test. The breadth / length of a compound could be investigated, or the surface area / length or volume / length might be useful descriptors. In addition, there may be functional group effects that modify or alter any size descriptor. Again, these effects can most easily be evaluated through QSAR equations.

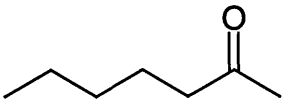
It is clearly not viable for all the above possibilities to be examined in every homologous series, and so some prior selection is needed, based on our theoretical understanding so far. Three possible sets of experiments, as follows, could lead to the maximum information with the minimum number of experiments.

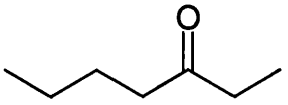
(i). As regards the maximum length effect, a suitable series to test the maximum length effect would be the series: heptan-2-one, heptan-3-one and heptan-4-one. These ketones are commercially available, are not too involatile, and could be used to test whether the maximum length is the overall maximum, or the maximum to the functional group.

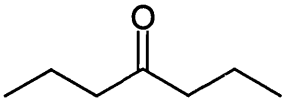
(ii). A descriptor that has been used in the past (although in other contexts) is the breadth -to-length dimensional ratio in a molecule. Now one could take heptan-2-one or hexan-2-one as the standard, and investigate cyclohexanone and t-butyl methyl ketone (pinacolone, readily available) as more bulky compounds.

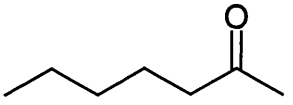
(iii). Remaining in the ketone series, the requirement for a single point of attachment through one functional group could be studied through the commercially available (and, again, not too involatile) compounds, butan-2,3-dione, pentan-

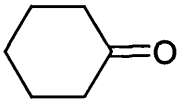
2,4-dione and hexan-2,5-dione, with the additional possibility of hexan-2,3-dione or hexan,3,4-dione.

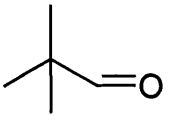
a-  Heptan-2-one

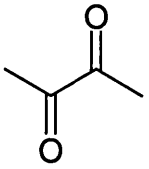
 Heptan-3-one

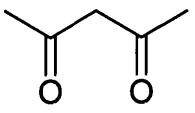
 Heptan-4-one

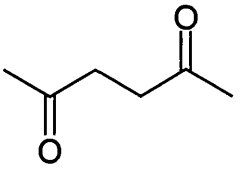
b- Heptan-2-one 

Cylohexanone 

tert-Butyl ketone 

c-  Buta-2,3-dione

 Penta-2,4-dione

 Hexa-2,5-dione

These series, a, b, and c, would be priority series. If they yielded relatively unambiguous answers, the entire experimentation would be streamlined, through experimental design based on theory.

13.0.2 Potential difficulties and limitations

An important feature of the proposal is the rigorous monitoring of all vapor stimuli through gas chromatography (GC). This includes weekly follow-up of chemicals in glass vessels to ensure stability. It is not very common for chemosensory studies on airborne chemicals to rely on GC of vapors, and even less common to include repetitive GC measurements as a follow-up procedure. Nevertheless, we consider that these laborious and time-consuming measurements are an essential tool for the quantification of the stimulus, a fundamental step for development of our QSARs.

Further studies could explore trigeminal responses to acute (1-3 sec) presentations of the chemicals. This constitutes a limitation when trying to extrapolate the present findings to residential or occupational exposures that typically linger for much longer periods (up to months in some cases) and, in addition, involve probably much lower concentrations of VOCs. Time of exposure constitutes, no doubt, a crucial parameter in the overall understanding of trigeminal chemosensory responses, although such work is likely to be in the more distant future.

13.1. References

1. Cometto-Muñiz, J.E., in Indoor Air Quality Handbook, Eds. J.D. Spengler, J.M. Samet, J.F. McCarthy, McGraw-Hill, New York, (2001) 20.1.
2. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muñiz, W.S. Cain, Arch. Toxicol., 72 (1998a) 227.
3. M.H. Abraham, R. Kumarsingh, J.E. Cometto-Muñiz, W.S. Cain, (1998b).Toxicol. in Vitro, 12 (1998b) 403.