UNIVERSITY OF SOUTHAMPTON
FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS
SCHOOL OF CHEMISTRY

Prediction of Retention for Pharmaceutical Molecules in Supercritical Fluid Chromatography. The Synthesis and Analysis of a Library of Sulfonamides.

by

Amaury Cazenave Gassiot

Supervisor: Dr. G. John Langley
Advisor: Dr. Richard C. D. Brown

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

November 2007

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS
SCHOOL OF CHEMISTRY
Doctor of Philosophy

PREDICTION OF RETENTION FOR PHARMACEUTICAL MOLECULES IN SUPERCRITICAL FLUID CHROMATOGRAPHY. THE SYNTHESIS AND ANALYSIS OF A LIBRARY OF SULFONAMIDES.

by Amaury Cazenave Gassiot

A library of thirty-two sulfonamides was designed for analysis by supercritical fluid chromatography (SFC) with the aim of highlighting the relationship between properties of the test analytes and their chromatographic retention.

The pharmaceutical setting of the present project prompted to focus on the design of a library of drug-like analytes. Sulfonamides were chosen for their ease of synthesis and because the sulfonamide functionality is widely spread amongst drug molecules. The design of the library was also undertaken by taking into account such concepts as privileged structures and Lipinski's Rule of Five in order to maximise the drug-likeness of the test analytes. The thirty-two sulfonamides were subsequently synthesised for the purpose of the study.

During a pilot study involving a restricted test set and using an isocratic approach with 20 % methanol (MeOH) in carbon dioxide (CO_2) as a mobile phase and three different stationary phases (2-ethyl-pyridyl, cyano and diol bonded silica), simple trends linking the retention time of the analytes to their structural features were identified.

Following this pilot study, a more systematic approach was used to study properties-retention relationships of the test compounds. Polycratic studies were undertaken on the test library using CO_2-MeOH mobile phases (in the presence or absence of additive) and a 2-ethyl-pyridyl (2-EP) column. Taking a restricted range of retention factor, k, ($1 < k < 10$) and keeping the proportion of modifier in the mobile phase, φ, above 10 %, it was shown that log k varies linearly with φ ($R^2 > 0.98$), although the relationship is not linear at mobile phase compositions below 10 % modifier. From these relationships, different retention characteristics of the analytes were calculated. Most quantitative structure-retention relationships (QSRR) found in the literature deal with gas-chromatography (GC) or high-performance liquid chromatography (HPLC). In the latter case, it has been shown by Kaliszan and co-workers that retention characteristics can be correlated with three simple molecular descriptors to derive equations predicting the retention behaviour of new compounds: total dipole moment, μ, molecular surface area, A, and the electronic charge on the most negatively charged atom, δ_{min}. The values of these three descriptors were calculated for the test compounds using a molecular modelling package and this work shows that they are correlated with measured retention characteristics of the test analytes. The correlation of chromatographic measurements with calculated molecular descriptors may allow the prediction of the retention behaviour for an unknown compound provided its properties are known.

Results obtained during the polycratic study showed a net improvement of peak shapes and a decrease in retention time when ammonium acetate was used as an additive in the mobile phase. The effects of increasing concentrations of ammonium acetate additive in supercritical fluid chromatography were studied on silica (Si), 2-ethyl-pyridine (2-EP) and endcapped 2-ethyl-pyridine (2-EP-C) stationary phases. The study involved the addition of increasing concentrations of ammonium acetate either in the mobile phase modifier (methanol) or in the sample solvent. The effects of ammonium acetate on the retention and the peak shape of the analytes were evaluated. Compounds that exhibited satisfactory chromatographic behaviour in the absence of the additive were virtually unaffected by its presence in mobile phase or sample solvent. Nevertheless, compounds that exhibited late elution and strongly tailing peak shapes when pure methanol was used showed dramatically improved chromatographic behaviour in the presence of the additive. Shorter retention was observed, not only when the modifier was introduced in the mobile phase but also when it was present in the sample solvent.

à mes parents

pour m'avoir toujours laissé libre de mes choix
et m'avoir permis de poursuivre de si longues études,
ce travail vous est dédié.

Contents

Figures

Tables

Schemes

Acknowledgements

First of all I would like to acknowledge the support of my supervisor. I am very grateful to John for giving me, when most needed, the opportunity to take on this PhD. John possesses this rare and valuable quality that consists in being always there when you need him without ever imposing his presence upon you. His gentle guidance kept me on the tracks while his trust provided me with independence.

This work was supported by the Combinatorial Centre for Excellence Consortium at the School of Chemistry, University of Southampton which involved the following partners: Evotec, GlaxoSmithKline, Mettler Toledo, Pfizer and Princeton Chromatography. It could not have been brought to this point without all the partners involved. I would like to thank Bob, Jeff, Stephen, Paul, Clare, Frank, Mark and Craig for their help and support.

I cannot imagine what would have been the last three years without my team-mate. Working everyday with Mohini was always a pleasure. I would like to thank her for her kindness and her everlasting support.

Thanks also to the other members of the group: Julie, Louisa and Stephen for the friendly environment and for patiently dealing with my outrageous accent.

Despite what some might think, life is not all about working in the lab... Moving to the United Kingdom was a big change and I am thankful to Marion and Nicolas for our distance-proof friendship and for their comforting support in the hardest days. A new country also means new friends. The list is long of people I would like to thank for all the great moments I had during these three years. I will keep it short: thanks so much to Justine, Niall and Julien for taking me by their side.

I would like to thank the Domzalskis, whose hospitality provided me with a peaceful retreat where I could put most of this work together.

Finally, the last word is for you, Staycie. Because you are my everyday inspiration, this work would not be the same without you. Merci.

Abbreviations

2-EP	2-ethylpyridine
2-EP-EC	endcapped 2-ethylpyridine
2-PPU	2-pyridyl-propyl urea
ACD	available chemical database
ADMET	absorption-distribution-metobalism-excretion-toxicity
APCI	atmospheric pressure chemical ionisation
BBB	blood-brain barrier
CA	carbonate anhydrase
CLND	chemiluminescent nitrogen detection
CMC	comprehensive medicinal chemistry
CN	cyanopropyl
CSP	chiral stationary phase
DEA	diethylamine
DMF	dimethylformamide
EDMA	ethyldimethylamine
ELSD	evaporative light scattering detection
eq.	equivalent
ES	electrospray
ESI	electrospray ionisation
FID	flame ionisation detector
Fmoc	*9H*-fluorenylmethyl carbamate
GC	gas chromatography
GC-MS	gas chromatography-mass spectrometry
GPCR	G-protein coupled receptor
HR-MS	high resolution mass spectrometry
HTA	high-throughput analysis
HTS	high-throughput screening
I.D.	internal diametre
IPA	isopropylamine
IR	infrared

IUPAC	International Union of Pure and Applied Chemistry
LC	liquid chromatography
LC-MS	liquid chromatography-mass spectrometry
LR-MS	low resolution mass spectrometry
MDDR	MDL drug data report
min	minutes
mmol	millimole
MS	mass spectrometry
MWT	molecular weight
PK	pharmacokinetic
pSFC	packed column supercritical fluid chromatography
RMM	melative molecular mass
rp-HPLC	reversed phase high performance liquid chromatography
r.t.	room temperature
SAR	structure-activity relationships
Si	bare silica
SFC	supercritical fluid chromatography
TEA	triethylamine
THF	tetrahydrofuran
TFA	trifluoroacetic acid
TLC	thin layer chromatography
t_R	retention time
UV	ultra violet
v/v	volume/volume
WDI	world drug index

Chapter 1.

Introduction

1.1. Chromatographic analysis in a pharmaceutical setting.

1.1.1. Applications of chromatography coupled to mass spectrometry in pharmaceutical sciences.

In the past 15 years, drug discovery has asked more and more from analytical sciences. When new challenges arise and existing analytical techniques cannot tackle them, one possible way forward, rather than investigate and develop an entirely new concept, is to hyphenate pre-exisiting techniques. The first successful hyphenated technique was gas chromatography coupled to mass spectrometry (GC-MS) in the 1960s. At the time the success of this technique was huge, since GC-MS could tackle most of the problems encountered at the time, noticeably by the petroleum industry.

However, hyphenation can be hampered by the lack of compatibility between the two techniques of interest. In the case of liquid chromatography-mass spectrometry (LC-MS), the hyphenation was problematic because MS is inherently a gas phase technique and the transition from the liquid to the gas phase was a major issue. Eventually, the advent of electrospray ionisation (ESI) and atmospheric pressure chemical ionisation (APCI) provided scientists with an easy way to hyphen MS with LC. ESI and APCI are the two main sources of ionization in LC-MS in pharmaceutical analysis. They are referred to as atmospheric pressure ionisations (API) due to the fact that ionisation occurs at atmospheric pressure through a combination of high voltage and heat. In ESI, charge droplets are

formed at the tip of the inlet capillary under the effect of a high voltage (typically 3-5 kV). Droplets shrink as they approach the analyser until individual ions evaporate and enter the analyser. In APCI, the eluent is vaporised under heating. A corona discharge then ionises the solvent that, in turn, ionises the analytes through chemical ionisation mechanisms.

In the last ten years, following the development of ESI and APCI, LC-MS has become an analytical tool of prime importance in all steps of drug development. The drug discovery process involves the screening of compound libraries to select leads that will be modified and optimised into a compound suitable for development as a drug candidate. During this process, the identification and quantification of compounds of interest and impurities is mostly undertaken through LC-MS, although other technique like nuclear magnetic resonance (NMR) and infrared (IR) or ultraviolet (UV) spectroscopy are also of importance.

In the early steps of drug discovery, *i.e.* at the stage of library synthesis, LC-MS is used for identification and purity assessment of newly synthesized chemical entities. Very often, LC-MS systems at that stage are highly automated and open-access. They are also used for purification purposes, with mass triggered collection.

The selection of chemical entities for further development will be based on the affinity of the synthesized compounds for the biological target of interest. High-throughput screening (HTS) of *in vitro* biological activity mostly involves fluorescence techniques. However, LC-MS methods exist that can be used in activity assays.[1]

Another stage of drug discovery where LC-MS is definitely of prime importance is absorption-distribution-metabolism-excretion (ADME) and pharmacokinetic (PK) studies. The critical importance of early ADME-PK studies will be discussed in Chapter 2. The assessment of the metabolic stability of a given compound and the identification of its metabolites is one of the main aspects of drug discovery, since it will give insight into the half-life and the toxicity of the potential drug. Most of the applications of LC-MS methods in the drug discovery process are *in vitro* and *in vivo* assays of ADME-PK properties. LC-MS-MS is also use extensively in that field.[2]

In summary, LC-MS has become the analytical technique of choice in all stages of the pharmaceutical research and development process, from library purity

assessment to metabolite identification, from HTS to ADME studies. However, pharmaceutical analysis must, more and more, be undertaken in a high-throughput fashion. In HPLC, long analysis times and low column efficiency, due to the low solute diffusion in the mobile phase, are often observed. High-throughput analysis (HTA) aims at the rapid analysis of vast numbers of compounds. HTA is driven by the need for analytical support for new biological targets, by the need to analyse large combinatorial libraries and also by the requirement to shorten analysis and method development times. Therefore analytical chemists could use an alternative method affording chromatographic separation comparable to (or better than) HPLC separation while decreasing the analysis time. Speeding LC-MS methods has been attempted to increase the throughput, noticeably by using shorter columns or increasing mobile phase flow rate. Another possibility is to use supercritical fluid chromatography.

1.1.2. Why supercritical fluid chromatography ?

To achieve the quality and safety requirements expected for new drug compounds, analytical chemists are faced with the challenge of developing new analytical methods capable of quick, highly efficient separations for the characterization of all compounds and impurities. As seen previously, until recently HPLC-MS has been preferentially used for this purpose. However, with the aim of maximising the information gathered in a given time, packed column supercritical fluid chromatography coupled to mass spectrometry (pSFC-MS) appears more and more as a complementary technique for HTA. Interestingly, pSFC is suitable for most of the tasks undertaken by LC-MS, whether it be structure analysis (pSFC-MS-MS), quantification (pSFC-CLND/ELSD-MS), purification (preparative scale pSFC), chiral separations (pSFC-MS) or purity assessment (pSFC-UV-MS).

For all these applications, high chromatographic resolution is required. This means that the need is for a technique providing chromatographic efficiency (*i.e.* a high number of theoretical plates), speed of analysis, sensitivity and selectivity. From this point of view pSFC presents many advantages for pharmaceutical analysis. It has a high separation efficiency, it is suitable for separation of isomers or structurally similar analytes, the selectivity can be adjusted by varying several

parameters (mobile phase, stationary phase, temperature, pressure, *vide infra*), it is fast (three to five times faster than LC, with reduced column equilibration times), it is cost effective and generates less toxic waste than LC. Indeed, SFC generally uses CO_2 as mobile phase main component, although other substances can be used (*vide infra*). For that purpose, the CO_2 is collected as a by-product of other chemical reactions or collected directly from the atmosphere and, therefore, contributes no new CO_2 to the environment. The "green" potential of SFC, noticeably on a preparative scale, has been demonstrated.[3]

In terms of covered chemical space, any analyte soluble in methanol or a less polar solvent is suitable for SFC analysis. The technique is therefore suitable for non-polar analytes. The addition of an organic modifier in the mobile phase (possibly with the addition of a third component at low concentration, *vide infra*) even affords elution of polar compounds such as organic acids and bases and their salts. The elution of peptides using SFC has also been reported.[4] The feasibility of pharmaceutical analysis using SFC has also been demonstrated. Pinkston and co-workers, for instance, analysed a large and diverse library of pharmaceutical compounds and found that SFC was suitable for the analysis of 75 % of the analytes (as compared with 79 % for HPLC).[5] Zhao and co-workers came to the conclusion that pSFC can be implemented in all steps of drug discovery and is a valuable complement to LC. However, they also concluded that dedicated academic research and hardware development are still needed to reach that stage, especially because in-depth knowledge of retention mechanism in SFC is still missing.[6]

It is with the aim of getting insight into how pSFC works for pharmaceutically relevant compounds that the current project was initiated.

1.1.3. Advantages of a predictive model.

In a HTA context, the time spent on method development and the analysis time itself are crucial. The fast analysis time and the fast column equilibration time in pSFC can dramatically reduce the overall analysis time. However, pSFC has not yet, and will probably never, totally replace HPLC, especially because the two techniques are complementary rather than competitive. In analytical laboratories

equipped with both types of instrument then comes the problem of deciding which technique to choose for a given batch of samples. To avoid time-consuming double analysis of test samples, to decide which technique will be more suitable for which analytes, prior knowledge of the two techniques is required to make an educated guess on the best way forward. In HPLC, chemometric studies have afforded software packages that greatly help method development by predicting the retention behaviour of an analyte when its retention in given conditions are known.[7] As useful as they are, these tools are not yet able to predict the retention of analytes from mere molecular properties without analysing the compounds of interest in well-defined conditions. Numerous quantitative structure-retention relationships (QSRR) studies (discussed in Chapter 3) have attempted to predict the HPLC retention of analytes from numerical values of properties describing their structures.[8] Nevertheless, such studies, despite their potentially fruitful applications, are still scarce in SFC. With the aim of evaluating the potential application to SFC of an existing model predicting retention of analytes in HPLC, this project was focused on a QSRR study of a small library of pharmaceutical compounds in SFC.

1.2. Overview of supercritical fluid chromatography.

1.2.1. Supercritical fluids.

The critical temperature of a substance is the temperature above which that substance can no longer exist as a liquid, no matter how much the pressure is increased. In the same way, the critical pressure is the pressure above which the substance can no longer exist as a gas, no matter how high the temperature is. In a phase diagram, these pressure and temperature values define the critical point. (Figure 1)

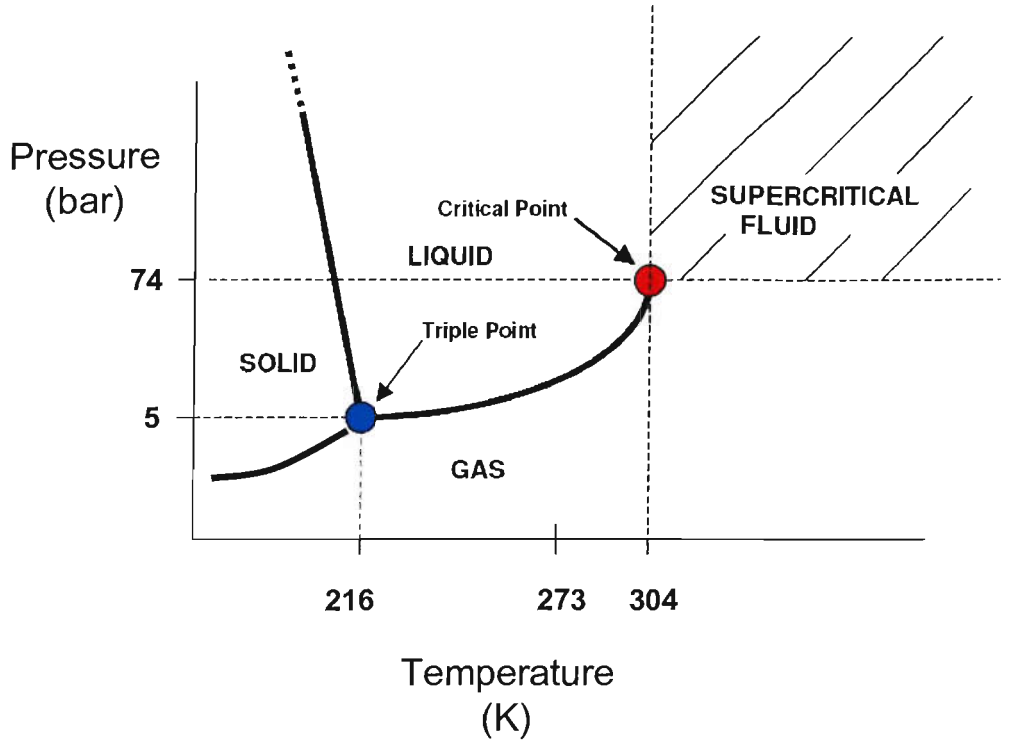


Figure 1. Phase diagram of a pure substance. Shown temperature and pressure values are those of pure carbon dioxide.

Supercritical fluids are obtained either by heating a gas above its critical temperature or by compressing a liquid at a higher pressure than its critical pressure. It is impossible to draw a clear line between supercritical fluids and liquids or gases, since the transition from liquid to supercritical fluid by raising the temperature at constant pressure or from gas to supercritical fluid by increasing the pressure at constant temperature is continuous. Critical temperatures, pressures and densities for a number of pure substances are shown in Table 1.

Table 1. Critical pressure (P_c), temperature (T_c) and density (ϱ_c) of pure substances.

Substance	P_c (bar)	T_c (°C)	ϱ_c (g mL^{-1})
CO_2	72.9	31.3	0.47
N_2O	72.5	36.5	0.45
SF_6	37.1	45.5	0.74
Xe	58.4	16.6	1.10
CH_3OH	78.9	240.5	0.27
$CH_3CH(OH)CH_3$	47.0	235.3	0.27
H_2O	218	374	0.32

Under supercritical conditions, the properties of the substance (*e.g.* density, viscosity, diffusion coefficient...) are intermediate between those of liquid and gas. The density is typically of the order of magnitude of liquid density (from 0.1 to 0.8 g cm^3). The solvating power of supercritical fluids is also very similar to the one of many conventional organic solvents and much higher than in gases. Conversely, the diffusion coefficient and viscosity of supercritical fluids are about 5 to 50 times higher than in liquids. This is illustrated in Table 2.

Table 2. Order of magnitude of physical properties density, diffusion and viscosity for gaseous, supercritical and liquid states.

	Density (g cm^3)	Diffusion (cm^2 s^{-1})	Viscosity (g cm^{-1} s^{-1})
Gas	10^{-3}	10^{-1}	10^{-4}
Supercritical fluid	10^{-1} – 1 *Liquid-like*	10^{-4} – 10^{-3} *Liquid-like*	10^{-4} – 10^{-3} *Gas-like*
Liquid	1	< 10^{-5}	10^{-2}

The properties of supercritical fluids, intermediate between the properties of gas and liquids, make them interesting for use as chromatographic mobile phases. Gas chromatography (GC) allows high resolution separation of complex mixtures. However, GC is limited to thermally stable compounds that are volatile and of low molecular mass. Reversed phase high performance liquid chromatography (RP-HPLC) is recognized to be the most convenient separation technique for a wide range of compounds, including substances of high molecular mass and with thermal lability. Nevertheless, long analysis times and low column efficiency due to the low solute diffusion in the mobile phase are often observed. According to the intermediate properties of supercritical fluids between those of a gas and a liquid, supercritical fluids appear to be a good solution to avoid problems of both HPLC and GC.

Further, supercritical fluids used as mobile phase introduce the possibility of influencing retention by varying temperature and pressure. In a supercritical state, the density of the mobile phase changes significantly as the pressure and the temperature vary, which is not the gas in liquid or gaseous states. As a consequence, an increase in pressure at constant temperature results in an increase in density and, therefore, in a greater solvation power. As a result, a solute becomes more soluble in the mobile phase and the retention decreases. At a temperature

near the critical point, density of the mobile phase falls dramatically as temperature increases, reducing the solvating effect to a greater extent than it is compensated by the rise in vapour pressure, once again influencing retention of an analyte.

1.2.2. Supercritical fluid chromatography.

1.2.2.1. The ups and downs of SFC.

Using a supercritical fluid as a mobile phase to perform chromatographic separations was suggested more than forty years ago.[9] The first packed column SFC system was marketed in 1982 by Hewlett Packard (HP), followed by a capillary SFC instrument in 1986 by Less Scientific SFC. Analytical chemists had been long awaiting the possibilities offered by SFC. The fact that, in SFC, one can develop chromatographic methods not only by varying mobile phase composition but also temperature and pressure was very much anticipated. At the time, many people even thought that SFC would replace HPLC altogether.[10] This, obviously, did not happen. The poor reproducibility obtained with capillary SFC in the early 1990s, combined with the lack of user-friendliness and the high cost of early SFC systems resulted in the technique being disregarded as inefficient and too expensive. SFC is still in many analysts' minds, at best, a substitute for normal phase LC. Another factor that might have influenced the poor acceptance of SFC is its market situation. The SFC market is composed of small companies that cannot compete with the main HPLC manufacturers in terms of marketing and sales staff. Several companies gave SFC a try and then decided to give up. In 1995, HP sold its SFC section to Terry Berger who founded Berger Instrument, which was subsequently bought by Mettler-Toledo in 2000, before being ceded again, to Thar, in 2007. This is a major concern for potential buyers that do not want to be left with an instrument of which nobody is going to ensure maintenance.

Despite all this, there has been recently a resurgence of interest in SFC, noticeably in the pharmaceutical industry. As seen in § 1.1.2, the many advantages of SFC (in terms of speed of analysis, resolution, solvent consumption, *etc.*..) make it a technique of choice for high-throughput pharmaceutical analysis and SFC, at

both analytical and preparative scales is used routinely in an increasing number of pharmaceutical companies (Pfizer in La Jolla, Eli Lilly, GSK...). Interest is also growing in the petroleum industry with SFC being very suitable for the analysis of biofuels and in environmental analysis for the detection of pesticides, for instance.

1.2.2.2. SFC *vs.* HPLC

According to the intermediate properties of supercritical fluids between those of a gas and a liquid (*vide supra*), supercritical fluids appear to be a good solution to avoid problems of both HPLC and GC. Supercritical fluid mobile phases have much greater solubilizing power than gaseous ones and can, therefore, be used for the separation of involatile and high-molecular-mass samples unsuited to GC. Although the typical solute diffusion coefficient in supercritical fluids is intermediate between those of gas and liquid, it is noteworthy that the diffusion coefficient is an order of magnitude greater than in liquid. This fact has important chromatographic implications concerning separation time and column efficiency. SFC is theoretically up to ten times faster than HPLC,[11] because of the lower viscosity and higher diffusivity in the mobile phase, SFC columns typically provide a three- to five-fold reduction in analysis time over HPLC.[12] Moreover, column equilibration times are far shorter with SFC compared with HPLC,[10] reducing once again overall analysis time. The typical minimum values for height equivalent to a theoretical plate for packed-column SFC (pSFC) and HPLC are very similar, the most important difference, however, is that the minimum value in SFC is achieved at linear velocities three to five times greater than for HPLC.[13, 14] Figure 2 shows that the optimal height equivalent to a theoretical plate (HETP) is similar in HPLC and SFC (typically 12 μm) but, in SFC, it is reached at much higher linear velocity (typically 0.1 cm s^{-1} for HPLC and 0.4-0.5 cm s^{-1} for SFC). Further, the low viscosity of supercritical fluids results in lower pressure drops along the column, thus up to 10 columns can be assembled, serially, to afford up to 200,000 theoretical plates.

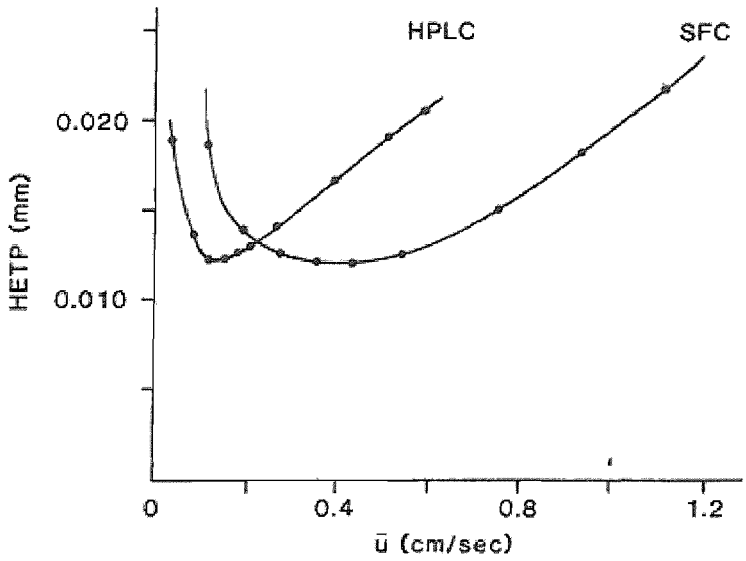


Figure 2. Van Deemter plots for the analysis of pyrene by HPLC and SFC. HETP is the height equivalent to a theoretical plate, ū is the linear velocity. (reproduced from Gere[13]).

Another important advantage of SFC compared with HPLC is that SFC provides rapid separations without the use of large volumes of organic solvents. With the desire for environmentally conscious technology, the use of organic chemicals, as used in HPLC, could be reduced with the use of SFC. This is especially true when it comes to preparative scale where large quantities of solvents are involved.

To summarise, SFC possesses a number of advantages when compared to HPLC: shorter analysis time, higher efficiency, fast column equilibration, less harmful and more cost-effective mobile phases, easy to hyphenate with many detectors (*cf.* 1.2.2.4) and easy to scale-up from analytical to preparative scale.

However, SFC and HPLC should not be seen as competitive techniques. HPLC will handle samples that are not suitable for SFC, and the converse is true. The two techniques are actually complementary and being equipped with both types of instruments will provide the analyst with the means to tackle more analytical challenges.

1.2.2.3. Supercritical mobile phases.

The moderate critical conditions of carbon dioxide, 304 K and 74 bar, make it favourable for the analysis of the thermally unstable compounds. Taking

into account that CO_2 is barely toxic, chemically inert, non-flammable, non-explosive and presents a low response in most detection systems, it is a solvent of choice for application in SFC.[15] Moreover, CO_2 is readily available in high purity (therefore suitable for obtaining clean chromatographic baseline) and is miscible with most organic solvents.

Nevertheless, CO_2 presents the important disadvantage of being very non-polar, even though the solvation power in a supercritical state depends on the density of the fluid. At a density of 0.25 g cm^{-3}, CO_2 has a similar solvent strength to perfluorinated alkanes and, at a density of 0.98 g cm^{-3}, it is slightly more polar than hexane.[13] CO_2 is therefore unable to elute polar compounds. As a consequence, the mobile phase is very often a binary or even ternary mobile phase: organic modifiers are added to CO_2 in order to increase polarity. A wide range of modifiers can be used: *e.g.* methanol, ethanol, isopropanol, butanol, acetonitrile, water, tetrahydrofuran and dimethylsulfoxide.[16] The most commonly used modifier is methanol. The effects of the modifier on the retention and selectivity of the mobile phase are complex. They are not only due to an increase in polarity of the mobile phase. Modification of the density of the mobile phase and interactions with the stationary phase by modification of the conformation of the bonded phase (in the case of chiral stationary phase) and by deactivation of active sites are also involved.[15] In the case of ternary mobile phases, the third component is often an acidic or basic additive typically added in a small amount (<1 % v/v) in order to increase chromatographic efficiency and obtain symmetrical, well-shaped peaks.[17] Additives can be chosen according to the nature of the analyzed compounds: elution of carboxylic acids, for example, will be better with trifluoroacetic acid (TFA) or citric acid, while elution of bases will be improved by using aliphatic amines, *e.g.* isopropylamine (IPA), diethylamine (DEA), dimethylethylamine (DMEA) and triethylamine (TEA).[18] The use of volatile ammonium salts (*e.g.* ammonium acetate) has also been reported as discussed in Chapter 5.

As in HPLC, the composition of the mobile phase can be programmed from 0 to 100 % of modifier, using isocratic, gradient or step gradient. Increasing the proportion of modifier makes the fluidity of the mobile phase decrease, increasing the analysis time, however SFC still remains faster than HPLC since chromatograms can easily be collected with flow rates of more than 5 mL min^{-1}.[19,20]

It is noteworthy that increasing the concentration of modifier can lead to a binary mobile phase which is not actually in a supercritical state in the conditions of the experiment. Since there is no discontinuity between sub- and super-critical properties of the fluid, this fact is not of major importance for the separation of the solutes and it makes some authors to prefer use the term "unified chromatography" regardless of the exact state of the mobile phase: very high temperature liquid, condensed gas, sub- or super-critical fluid.[21]

1.2.2.4. Stationary phases.

Both packed and open-tubular capillary columns can be used in SFC. According to their higher efficiency, open capillary columns are preferred for the resolution of complex mixtures. They are typically between 1 and 35 m in length, 0.025 to 0.1 mm I.D., with a coating thickness of 0.1 to 3 μm. Less efficient, packed columns are used for less complex mixtures as they allow shorter analysis times and much higher loadability. Typical column characteristics are 30 to 250 mm length, 2.0 to 4.6 mm I.D., 5-6 μm particle size. Nowadays, SFC is mostly used with packed columns.

A wide range of achiral and chiral stationary phases (CSP) can be used in pSFC, most of which are silica-based, though polysaccharide, zirconia, polystyrene, divinylbenzene[22, 23] and porous graphitic carbon[24] based packings also exist. Most columns used in pSFC were first designed for HPLC. One drawback of using HPLC columns in SFC is not related to chromatographic separation but rather to experimental conditions. For instance, if one has a column designed to only withstand HPLC-like temperatures, around 30°C, and then uses it in a 60°C SFC method, adverse consequences are likely. For that matter, one might wish that manufacturer will design more SFC-dedicated columns.

Although SFC is often considered a normal-phase technique, *i.e.* involving a polar stationary phase and a non-polar mobile phase, non-polar HPLC stationary phases such as the very popular octadecylsilane packing (C_{18}) have been used in SFC. The choice of stationary phases is huge and listing all of them is not of interest here. As examples of achiral polar stationary phases one can cite some of the most commonly used: bare silica, diol, cyano-propyl (CN), amino-propyl, 2-

pyridyl-propyl urea and one that has been especially developed for SFC: 2-ethyl-pyridine (2-EP). Structures of these stationary phases are shown in Figure 3.

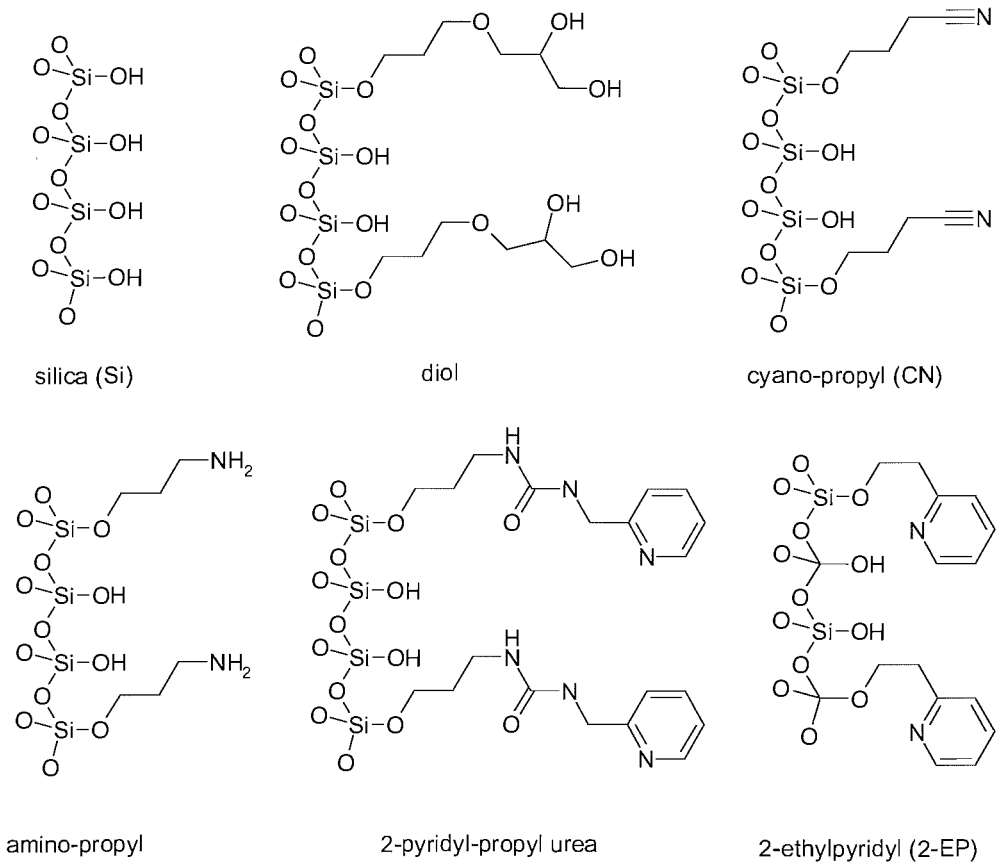


Figure 3. Structure of six achiral stationary phases.

For chiral separations, HPLC-designed CSP are again used in SFC. As an example one can cite the polysaccaride-based CSP Chiralcel OD (cellulose tris[3,5-dimethyl- phenylcarbamate]), Chiralcel OJ (cellulose tris[4-methylbenzoate]) and Chiralpak AD (amylose tris[3,5-dimethylphenylcarbamate]), which are all efficient in both SFC[12,25] and HPLC.[26]

1.2.2.5. Detection.

Theoretically, SFC displays the widest possible choice of detection techniques, being compatible with most of LC and GC detectors. However, most of them are not commercially available in combination with a SFC system. The most common detector used in SFC is the UV absorbance detector, because of its

sensitivity, its wide dynamic range and also because SFC mobile phases are generally UV-transparent (CO_2 is UV-transparent below the cut-off wavelength of most UV detectors). Other non-informative detection techniques can be used such as evaporative light scattering detection (ELSD),[27] chemiluminescent nitrogen detection (CLND) or flame ionization detection (FID).[28] Recently, Brunelli and co-workers have described the hyphenation of pSFC with corona-charged aerosol detection.[29] However, it is more and more desirable for high-throughput analysis methods to use informative detection techniques, *i.e.* techniques that provide structural information about the eluted compounds.

As with LC, mass spectrometry (MS) is an obvious choice of informative technique. pSFC-MS has been known in the literature for more than twenty years[30] and a substantial amount of research has been carried out demonstrating the efficiency of the technique using different ionization conditions such as positive or negative ion atmospheric pressure chemical ionization (APCI) or electrospray ionization (ESI). SFC-NMR and SFC-IR have also been studied.[31] In this project, a hyphenated pSFC-UV-MS using a quadrupole mass spectrometer with an electrospray ionisation source was used.

Chapter 2.

The SOTLIB Library of Sulfonamides

2.1. Rational design of combinatorial libraries.

2.1.1. The advent of combinatorial chemistry and the limitations of "irrational design".

The advent of high throughput techniques and combinatorial chemistry, over the past 15-20 years, provided medicinal chemists with the possibility of synthesizing ever larger and more diverse libraries of potential new drugs. This was not only driven by the mere possibility of synthesising libraries of several tens, or even hundreds, of thousands of compounds; but also by the advances of genomics that will provide scientists with possibly thousands of new biological targets for which pharmaceutical companies will be willing to find ligands.[32] High throughput technologies also provide the means to actually screen vast numbers of compounds against those new targets.

In the early days, the trend in drug development had rapidly evolved towards the design of vast (>100,000 compounds) diversity-based libraries. These libraries were focused mainly on synthetic accessibility and therapeutic interests (*i.e.* the finding of ligands for one or several targets). Other matters, of relevance to industrial scale production, such as development and production costs, target population or modes of administration were happily overlooked.

Combinatorial chemistry rapidly became popular and widely used in academic and industrial environments.[33] And, in spite of sometimes being nicknamed "irrational design", combinatorial chemistry has been in many respects

successful, most of the hit compounds nowadays generated by the pharmaceutical companies coming from screening of combinatorial libraries.[34] Nevertheless, the increase in research productivity, expected from the tremendous increase in the number of screened compounds, did not occur. While the average number of compounds screened by pharmaceutical companies was multiplied a hundred times between the early 1990s and 2000,[35] the number of new chemical entities reaching the market each year remained stationary (oscillating between 27 and 43).[36] All in all, an estimated 7,000,000 compounds are screened for each single new marketed drug.

This failure to deliver the expected results has inevitably led to a move towards a new paradigm of library design where structure-based, rational design and combinatorial chemistry are more and more integrated. Multidisciplinary teams nowadays tend to design smaller libraries, often targeted to a single receptor and carefully-designed by considering as early as possible parameters influencing diversity, activity, selectivity, bioavailability, toxicity, chemical tractability, patent possibilities, *etc.*

2.1.2. Rational design of combinatorial libraries.

2.1.2.1. Requirements for medicinal chemistry libraries.

In order to maximise the chances of success, the design of an optimised combinatorial library must take into account many different requirements. First of all, for the final compounds to have a chance to be developed as drugs, they must exhibit to some extent "drug-like" properties (more detailed information about the concept of drug-likeness is given in §2.1.2.3.). Another matter of interest is the structural novelty of the new entities, which is a requisite for patent purposes. Other criteria like coverage of chemical space, diversity, toxicity and chemical tractability are also of prime importance in library design. Moreover, an important realisation has been that these requirements will vary with regards to which point of drug development is concerned. Drug development is indeed a multi-step process and a distinction has to be made between hits, leads and drugs. A hit is defined as a compound of confirmed structure displaying an activity (usually better

than a predefined threshold) against the biological targets. Hits will subsequently be developed into leads. The latter should already display a variety of properties such as basic structure-activity relationships (SAR), known mechanism of action, activity on cells, *etc*... Leads series will be further optimised into drug candidates. A summary of the characteristic features of hits, leads and early drug candidates is shown in Figure 4.

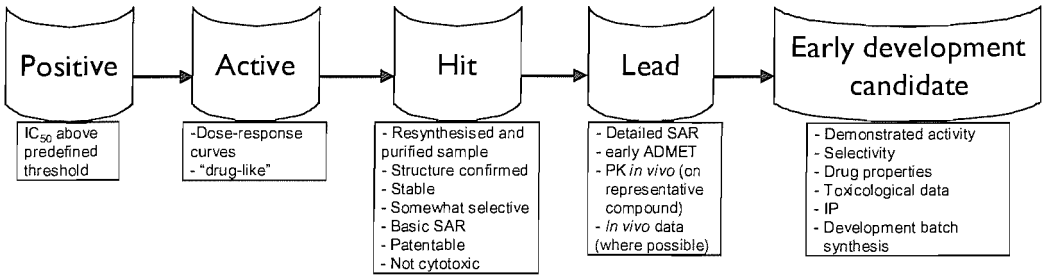


Figure 4. The drug discovery process.[34]

These differences have obvious repercussions on the way a library is conceived. For instance, high-throughput screening libraries aiming at hit-finding will generally be designed so as to display maximal diversity and to fill gaps in the represented chemical space; whilst lead-optimisation libraries will display small diversity and make optimal use of the information known about the target.[33] In the same manner, because drugs are developed from leads, they are *de facto* structurally closer to leads than they are to hits. This similarity is, however, relative and the optimisation from lead to drug involves the addition of potentially many substituents that alter the properties of the initial lead scaffold. For that reason, it has been observed that leads are generally smaller, less hydrophobic that drugs. Thus, whenever the conception of a lead series is involved, and whatever the criteria chosen to define drug-likeness are, those criteria must be kept different enough to their limit value so that further alteration of the structure remains possible.

2.1.2.2. Computer-aided library design.

Because of the numerous parameters involved in the conception of a combinatorial library, modern drug design makes regular use of computational methodologies to assist medicinal chemists in drug development.[37]

The lowest level of sophistication in computational chemistry is sometimes referred to as "ocular design".[38] It consists of the simple visualisation of the ligand-receptor complex using display software. As rudimentary as it may seem, it can lead to quick and useful developments when combined with the knowledge of an experienced medicinal chemist. When the three-dimensional shape of the binding site is known, docking (or virtual screening or *in silico* screening) is a more advanced computational tool that allows the determination of binding energies of a potentially large number of structures and conformations with the target receptor. This tool can also be used to select not only structures as a whole but also fragments that exhibit affinity for the active site and are therefore of interest for inclusion in combinatorial libraries. Structural similarities within a family of biological targets (G-protein coupled receptors (GPCRs), ion channels, nuclear receptors) can be used to design potential ligands of more than one specific target. Docking methodologies have been to date time-consuming and limited in scope, especially as the 3D structure of the receptor is not always available. Nevertheless, a number of other computational tools exist to assist the medicinal chemist in the conception of a new combinatorial library. Statistical methods for library optimisation are based on a variety of approaches. One can cite calculation of molecular descriptors (similarity, diversity, predicted activity for instance), selection by reagents availability, optimised combination of available fragments, *etc.*[39] However, an exhaustive listings of these methodologies would be beyond the scope of this work; reviews have been published that summarise these techniques.[33]

2.1.2.3. Early-ADMET and Lipinski's Rule of Five.

Drug-likeness and the evaluation of ADMET properties.

Another field where computational chemistry is growing in importance is the prediction of ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) properties, blood-brain barrier (BBB) permeation and solubility characteristics of potential new drugs. Indeed, to be developed as a drug, a new structure must not only be active and selective towards its target, but must also exhibit physico-chemical properties that will allow for its absorption and distribution in the patient's body. Moreover it must also not be excreted before reaching its target, and it must affect as few other, non-pathological pathways as possible. Poor pharmacokinetic properties and toxicity are one of the main causes of the high attrition rate observed in drug discovery; with 40 % of new chemical entities being rejected due to their inability to reach their target receptor.[34] As a consequence, evaluation of "drug-likeness" has become the centre of interest of many research groups. The term "drug-like" to qualify a molecule is usually understood as meaning "containing functional group and/or having physical properties consistent with the majority of drugs".[40] As seen previously, it may, on some occasions, be wiser to devise "lead-like" rather than "drug-like" compounds. The underlying concept however remains the same; that is to define criteria on which to decide as early as possible whether or not a given compound is likely to be developed into an efficient drug. The rigorous determination of pharmacokinetic properties necessitates complex *in vivo* measurements that make their implementation practically impossible in a high-throughput fashion. For instance, turbidimetric measurements of solubility give only an estimate of the actual thermodynamic solubility, the measurement of which is impractical in a drug-discovery setting.[41] To facilitate quick decision-making when selecting candidate structures for library conception, a number of structure-based rules have been devised that allow chemists to rule out compounds that are likely to exhibit properties deleterious to pharmacokinetics. These models must be used carefully as their reliability is debated; opponents deem them too restrictive. The major argument is that, as the selection criteria are based on the analysis of sets of known drugs, they tend to rule out any innovative structure that would lead to a drug that would have suitable potency but does not relate to known drugs or binds to an

insofar unidentified target.[39] However, they have proved useful in the selection of structures to include in libraries.

Basic selection rules can be used with regards to metabolism and toxicity. They are mainly based on filtering functional groups that are known to be reactive and/or toxic. As a rule, nitro groups, Michael acceptors, aldehydes, sulfonyl halides, primary alkyl halides, epoxides or aziridines, sulfonate esters, phosphonate esters, long aliphatic chains, peroxides, 1,2-dicarbonyls and acid halides should be avoided.[34, 40] Metabolism prediction has also been attempted by identifying substructures recognized by cytochromes P450 or having known metabolic pathways, and by reactivity calculations. However, metabolism and toxicity assessments are still unreliable and emphasis has mainly been put on absorption properties. Indeed, oral bioavailability has been shown to correlate with simple physico-chemical properties such as lipophilicity, solubility and pK_a. Lipophilicity (measured through log P and log D) influences membrane permeation, absorption, distribution and metabolism. Solubility is especially important for oral administration and can compensate for poor absorption of an orally administered compound. pK_a affects lipophilicity and solubility and therefore any phenomenon influenced by them.[39] Other studies have highlighted the importance of molecular weight, number of hydrogen bonding groups, polar surface area, flexibility, number of exocyclic methylene groups, number of halogen atoms, number of rotatable bonds or molar refractivity.[34]

Lipinski's "Rule of Five".

The most popular selection rule is probably the well-known "Lipinski's Rule of Five". The base hypothesis postulated by Lipinski and co-workers[41] was that drug candidates with poor absorption properties would go no further than pre-clinical or possibly "phase I" (*i.e.* tests on healthy volunteers) drug development phases. From the 50,000 compounds in the World Drug Index (WDI) database, they selected the 8,545 structures bearing either or both denominations customarily given to candidates reaching "phase II" (tests on patient volunteers), *i.e.* International Non-proprietary Name (INN) and United States Adopted Name (USAN). They subsequently refined the subset to eliminate compounds without "indications and usage" and compounds such as polymers, peptides, quaternary salts or phosphorous structures. They were then left with a

database of 2,245 compounds supposedly displaying physico-chemical properties compatible with good absorption. The aim of the study was then to highlight a set of properties-based rules allowing an audience of medicinal chemists to easily identify structures presenting properties suitable for acceptable absorption. Because of the targeted audience, an emphasis was put on finding descriptors that would be of straightforward interpretation for chemists, *i.e.* descriptors identifiable with a chemist's skills.

The first chosen descriptor was relative molecular mass (RMM). High RMM had been linked to poor intestinal and BBB permeation. Actually, only 11 % of the compounds in the studied library had a RMM over 500, as compared to 22 % in the WDI database.

Lipophilicity had often been taken into consideration in absorption studies and was therefore chosen as the second descriptor. In addition to the ACD calculated log P values (Clog P), Lipinski and co-workers designed their own algorithm for lipophilicity calculations, based on Moriguchi's work (Mlog P). They found that only 10 % of the test library had a Clog P greater than 5 (corresponding to a Mlog P greater than 4.15).

The third and fourth selected descriptors were the numbers of hydrogen bond donors and acceptors, due to the fact that a high hydrogen-bonding capacity can be deleterious for cell membrane permeation. As with log P calculations, scales of hydrogen bonds donor and acceptor capacity had been previously implemented, based on solvatochromic parameters.[42, 43] Lipinski and co-workers, however, decided to use an easier way to quantify hydrogen-bonding capacity. Hydrogen bond donor character was evaluated by simply counting the number of NH and OH bonds in the structure of interest; whilst an estimate of the hydrogen bond accepting character was obtained by counting the number of oxygen and nitrogen atoms in the molecule. Although this counting method, as confessed by Lipinski himself, is not nearly as good as the solvatochromic scales, it was effectively found that only 8 % of the compounds in the test library contained more than five NH and OH bonds; and than only 12 % had more than ten oxygen and nitrogen atoms.

These observations led to the statement of the "Rule of Five" as expressed in Lipinki's paper:

"*Poor absorption or permeation are more likely when:*

- *There are more than 5 H-bond donors (expressed as the sum of OHs and NHs);*
- *The MWT is over 500;*
- *The Log P is over 5 (or MLogP is over 4.15);*
- *There are more than 10 H-bond acceptors (expressed as the sum of Ns and Os)*

Compound classes that are substrates for biological transporters are exceptions to the rule."

Only 10 % of all compounds in the studied library had a combination of two parameters outside the range.

The rule has been further refined by researchers like Ghose,[44] who extended Lipinski's work and found that 80 % of the drugs in their test set had a log P between -0.4 and 5.6, a molar refractivity between 40 and 130, a RMM between 160 and 480 and between 20 and 70 atoms.

Lipinski's Rule of Five can be generalised across all classes of drugs, and has proved popular among medicinal chemists. It is, of course, a broad selection rule and well-known exceptions are often cited to undermine its applicability, *e.g.* vitamins, fungicides and antibiotics like macrolides, tetracyclins and rifamycins. Some of these exceptions can nevertheless be dismissed as irrelevant. In his 1997 paper, Lipinski takes the example of the antibiotic azithromycin, which happens to simultaneously not comply with the Rule of Five (RMM = 749) and to be active orally. The authors noted that this compound actually exhibits a very poor intestinal permeation, as expected from its failure to pass the Rule of Five. Its efficiency is most probably due to its very high aqueous solubility (50 mg mL^{-1}) that compensates for its poor permeation. Beside, one can question the relevance of citing antibiotics as exceptions to the Rule of Five. Indeed, these compounds are not targeted to the host's cells but to bacteria and, thus, do not necessarily need to display good absorption properties. In the case of the treatment of an intestinal infection for instance, poor intestinal permeation would be synonymous with increased concentration of the drug at the site of infection and therefore, in that case, a failure to pass Lipinski's rule would actually be desirable. Lipinski himself stated that such broad filters giving rough estimates of ADME properties might not be acceptable in a development stage, but there are of great interest in the discovery stage to allow screening library designers to conceive compounds more

likely to reach drug needs. This rule provides an overall, simple guideline for what is absorbed and what is not, and deviation from it in the discovery stage should only be motivated by specific reasons.

2.1.2.4. Privileged substructures.

So far, only methods rating drug-likeness according to projected ADMET properties of the candidates have been presented. But, the idea of analysing drug databases to identify features linked to good pharmacokinetics can obviously be extended to the search of structural features that regularly appear in drug molecules and can be related to the ability of these molecules to bind to a biological target. Muegge and co-workers rated drug-likeness according to the presence of characteristics pharmacophoric elements in the compounds. A certain level of drug-likeness would then be achieved by molecules featuring between 2 and 7 of the following structural groups: amine, amide, alcohol, ketone, sulfone, sulfonamide, carboxylic acid, carbamate, guanidine, amidine, urea and ester. [40]

The idea that molecular substructures could be identified presenting binding capability with various receptors was introduced in 1988 by Evans.[45] In his seminal paper, he described the design and synthesis of potent and orally active cholecystokinine antagonists. He developed an approach which consists in designing new compounds by modifying structures already known for their binding affinity for diverse receptors (1,4-benzodiazepin-2-ones) and coined the term "privileged structures" to refer to such scaffolds. Although the alternative denomination "privileged substructures" is sometimes preferred with regards to the fact that they are structural elements rather than molecules in their own right.[32]

The concept of privileged (sub)structure is potentially of great interest for the design of screening libraries and has been successfully put into practice many times since Evans' work. Opponents to privileged structures argue that their promiscuous nature leads to a lack of selectivity. However, studies have shown that selective compounds can be build by further modifications of privileged structures. For instance, the biphenyl structure is present in 4.3 % of all known drugs. If privileged structures are the scaffold that brings the features necessary for binding,

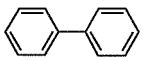
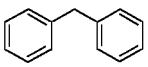
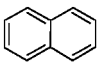
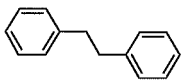
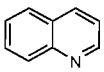
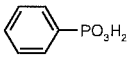
the adjunction of appropriate substituents allow for the design of selective compounds.

The explanation of why privileged structures actually exist is a matter of debate. Obviously, they must possess key physico-chemical properties that allow for their promiscuous binding to receptors. It has been proposed that some of them can mimic elements of protein secondary structure like β- and γ-turns. Benzodiazepines (the first privileged structures presented by Evans) are indeed β-turn mimetics. The concept of privileged structure suggests that, although chemical space is virtually infinite, compounds with biological activity are structurally related to privileged structures. The abundance of privileged structures in natural products could be explained by evolutionary pressure: organisms synthesizing compounds with biological activity would have been favoured through natural selection. It could also derive from the fact that biosynthesis is actually carried out by enzymes; therefore the resulting products must possess substructures with ability to bind to these proteins. The existence of a binding pocket, conserved amongst distinct receptors and in which privileged structures could bind, has also been postulated. As a matter of fact, mutagenesis experiments and sequence analysis have shown that, if GPCRs exhibit high variability in some parts of their binding sites, explaining the recognition of a high diversity of ligands, a conserved pocket common to class A (and to some of class B) GPCRs also exists. Bondensgaard and co-workers showed that ligands possessing privileged structures bind into this pocket.[46] It is mainly constituted of aromatic residues, therefore resulting in non polar interactions. Aromatic-aromatic interactions are favourable due to the entropically advantageous desolvation of non polar surfaces but also due to the enthalpic contribution of π-π interactions. As a result, aromatic-aromatic interactions can be as strong as hydrogen bonds and are also directional in character. The conserved pocket only partly accommodates the privileged structures, parts of them being in the variable part. That would explain why a given privileged structure binds preferentially only to a subset of receptors and not to any GPCR. A universal privileged structure therefore seems unlikely to exist.

Several groups have been trying to identify privileged structures. Hajduk and co-workers[47] have decomposed 10,080 compounds from an NMR database into 104 fragments. The database contained NMR-based information about whether or not each compound had binding affinity for 11 proteins. Twelve

fragments were identified as being highly represented in compounds binding to these proteins; most of these being statistically preferred for binding to one particular protein. However, COOH and biphenyl substructures were found to be preferred for binding to 50 % of the proteins. Diphenylmethyl also proved to be preferred for binding to three proteins. The identified privileged structures are listed in Table 3; including these structures in screening libraries could increase the chances of finding active compounds.

Table 3. Privileged substructures identified in an analysis of NMR-derived binding data (reproduced from Hajduk's article[47]).

Name	Structure	Protein Targets[a]	Δ[b]	β[c]	% Library[d]	% Drugs[e]	% ACD[f]
Carboxylic acid	COOH	6 (55%)	6.1-44.7	0.59-2.03	12.6	19.4	8.1
Biphenyl		5 (45%)	5.5-99.7	2.15-2.40	1.9	4.3	2.7
Diphenyl-methyl		3 (27%)	12.1-13.8	1.00-1.31	5.2	8.6	6.4
Naphtyl		1 (9%)	6.4	1.01	1.7	3.3	3.3
Phenyl		1 (9%)	3.9	0.75	68.7	73.3	73.2
Cyclohexyl		1 (9%)	3.4	1.51	4.5	12.2	6.3
Bibenzyl		1 (9%)	6.5	0.73	3.0	5.9	6.0
Benzimidazole		1 (9%)	17.5	1.61	1.0	0.8	0.3
Quinoline		1 (9%)	34.3	2.13	0.9	4.2	1.5
Triazine		1 (9%)	22.5	2.09	0.8	0.2	0.4
Benzofurane		1 (9%)	22.5	0.83	0.5	0.8	0.3
Phenyl-phosphonate		1 (9%)	98.8	3.52	0.2	0.03	0.04

[a] The number of projects (out of 11) for which the substructure was represented significantly greater than chance in the active compounds; [b] Increase over chance of the substructure's representation in the active compounds; [c] Regression coefficient from the logistic regression model. Where the substructure was preferred against multiple targets, the range for β against these targets is given. [d] Percentage of compounds (of 10,080) used in the analysis that contain this substructure. [e] Percentage of 154,000 compounds from the WDI and MDDR that contain this substructure. [f] Percentage of 177,000 compounds from ACD that contain this substructure.

In a fashion similar to that of Lipinski's, Bemis and co-workers have refined the Comprehensive Medicinal Chemistry (CMC) database to eliminate veterinary drugs, diagnostics aids, cosmetics, anesthetics, and other irrelevant compounds. They have analysed the 5,120 remainings structures in terms of atomic properties (atom type, hybridisation, charge) and graph properties, through which molecules are viewed as assemblies of ring systems, linkers, frameworks (that is rings and linker together) and side chains (Figure 5).[48]

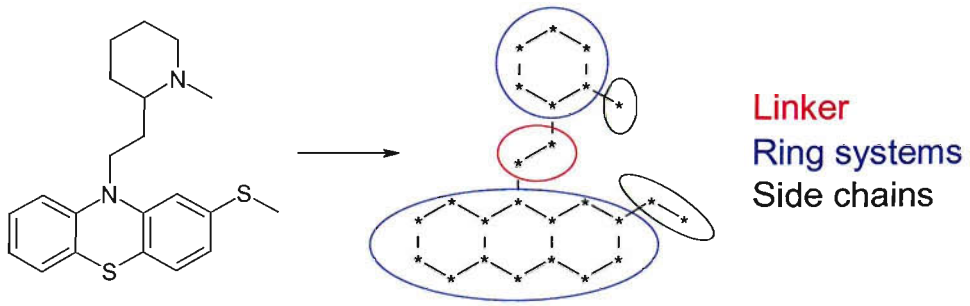


Figure 5. Conversion from a compound structure to a molecular framework, as described in Bemis *et al.*[48]

In their first approach, the authors only took the graph properties into account. This led to the definition of 1,179 frameworks (only 6 % of the structures did not lead to any framework, *i.e.* were acyclic). 783 (66 %) frameworks were unique (*i.e.* present in only one structure). However, amongst the remaining frameworks, 32 were represented 20 times or more in the structures and accounted for 50 % of the whole set. Secondly, atomic properties (element type, hybridisation and bond order) were included in the analysis, leading to the identification of 2,506 atomic frameworks, only 41 of which accounted for 24 % of the database. The 20 most commonly occurring frameworks are shown in Figure 6.

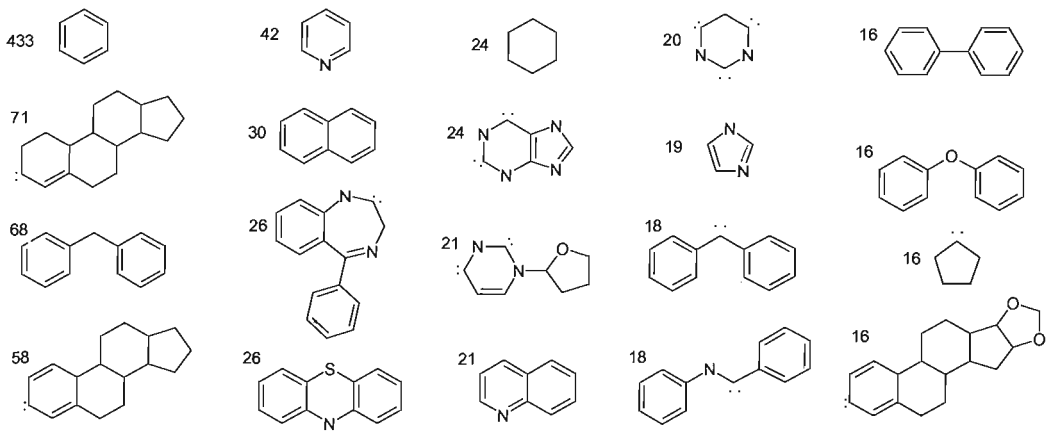


Figure 6. The 20 most commonly occurring atomic frameworks identified by Bemis and co-workers within the CMC database, along with their number of occurrences.[48]

The authors noted that the set of obtained frameworks may not be exhaustive, since they were derived from a restricted database. Beside, observed trends might be due to restrictions such as patent considerations or availability of building blocks rather than to identification of the inherent properties of drugs. However, these frameworks could be interesting to include in a screening library. Interestingly, diphenylmethane and indole groups were already present in the benzodiazepines described by Evans in 1988 (Figure 7).

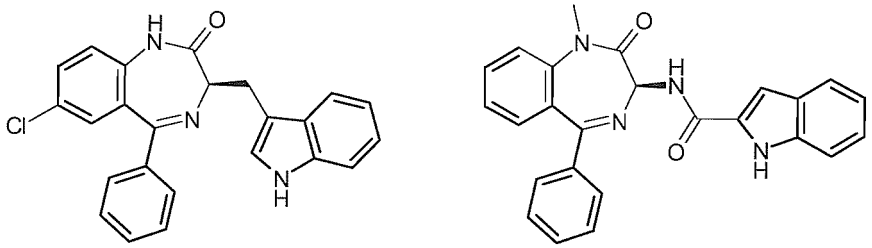


Figure 7. Two examples of benzodiazepine designed by Evans. Both already feature the diphenylmethane group later identified as privileged structures in their own right.

In a subsequent article,[49] Bemis focused his analysis on side chains. The analysis revealed that 92 % of the studied structures had a side chains. 18,664 side chains were identified, leading to an average of four side chains per molecular scaffold. The most frequently occurring number of side chains is two, with the majority of the compounds having between one and five side chains. The most commonly occurring side chains are shown in Figure 8.

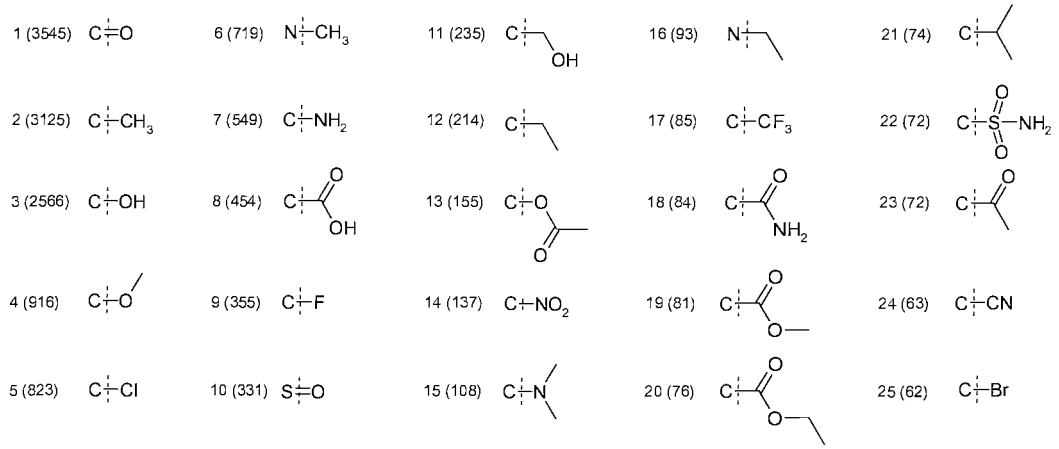


Figure 8. Most frequently occurring side chains in the CMC database (first number: rank; number in brackets: occurrences in the database).[49]

In a comprehensive review,[32] Horton enumerates the main described privileged structures to date as well as the drugs in which they appear, their

pharmaceutical applications and the typical synthetic pathways through which they can be obtained. The detailed privileged structures include: 1,4-benzodiazepin-2-one, biphenyl, 1,4-dihydropyridine, benzopyran, pyranocoumarin, 2,6-dichloro-9-thiabicyclo[3.3.1]nonane, isoxazole, 3,5-linked pyrrolin-4-ones, dihydro-β-agarofuran sesquiterpenes, spiroindoline sulfonamide, spiroindanyl piperidine, β-glucose and monosaccharides in general, benzazepinone, diphenylmethane, biphenyltetrazole, spiropiperidine, 4-substituted piperidine, indole and benzylpiperidine.

2.2. Design of the SOTLIB library.

2.2.1. Requirements of the SOTLIB library.

The compounds to be included in the SOTLIB library were not aimed at any given biological target. Instead, they were merely designed to be studied in terms of their retention behaviour. The absence of a given biological target implied that there was no specific pharmacophoric structure to comply with, nor was there a known lead compound, the structure of which was to be optimized to improve, for instance, binding affinity. The set of molecules to be synthesized was therefore to be chosen freely, although keeping in mind that the compounds should exhibit properties as compatible as possible with pharmaceutical applications. Moreover, from a practical point of view, because the very aim of the project was the chromatographic analysis of the compounds and not their synthesis *per se*, the designed structures had to be obtained as quickly as possible and in good yields. It was therefore decided to design the compounds according to the following criteria:

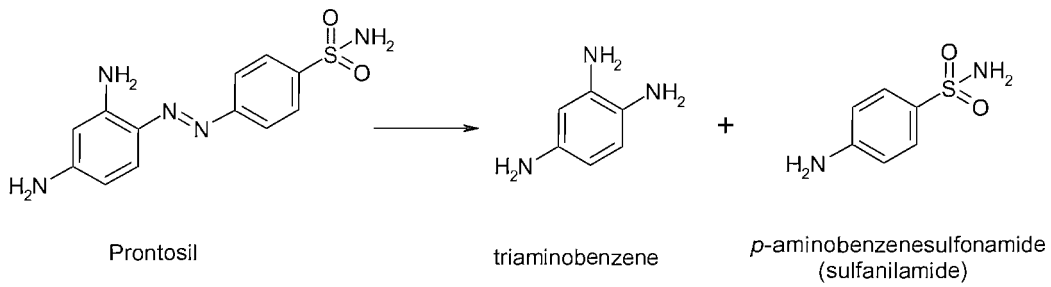
- Common scaffold of relevance to the pharmaceutical industry
- Inclusion of privileged substructures
- Compliance with Lipinski's Rule of Five
- Ease of synthesis
- Availability of starting materials
- Inclusion of neutral and basic compounds

2.2.2. The choice of a library of sulfonamides.

2.2.2.1. The use of sulfonamides in medicinal chemistry.

The sulfonamide functionality is ubiquitous in natural products and synthetic drugs. They are used, for instance, as antibiotics, anti-convulsivant, anti-hypertensive, hypoglycaemic and herbicide. Moreover, the sulfonamide group is amongst the structural group scoring for drug-likeness in Muegge's work[40], and the list of most occurring side chains in drugs compiled by Bemis also features the primary sulfonamide group.[49]

The history of the medicinal use of synthetic sulfonamides starts in 1935, with the discovery by Domagk[50] of the antibiotic properties of a sulfonamide-containing diazoic dye: the sulfamidochrysoidine or Prontosil.[i] The same year, Tréfouël and co-workers[51] demonstrated that *p*-aminobenzenesulfonamide (or sulfanilamide) presents broader antibiotic properties than Prontosil and speculated that sulfanilamide is the actual active principle of Prontosil, released *in vivo* by reduction of the diazo bond, hypothesis that would later be proved true. (Scheme 1)



Scheme 1. *In vivo* cleavage of Prontosil into triaminobenzene and antibiotic sulfanilamide.

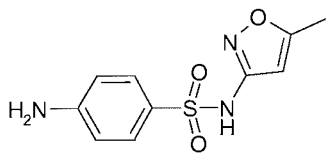
Sulfanilamide then constituted the starting point for the development a whole class of antibiotics, amongst which are sulfadiazine and sulfamethoxazole. Sulfonamides are still widely used as antibacterial agents, *e.g.* to prevent bacterial infections and to promote growth and treat disease in livestock.[52]

Other applications of sulfonamides soon followed, especially with the discovery in 1940 of sulfanilamide's capacity to inhibit the enzyme carbonic

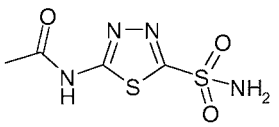
[i] This discovery won Domagk the 1939 Nobel Prize in Physiology or Medicine.

anhydrase (CA).[53] This prompted its use as a diuretic, followed by the development of more potent inhibitors of CA, such as acetazolamide, now used in the treatment of glaucoma. Other sulfonamides have been developed as diuretics, acting on different molecular targets, for instance the widely used thiazides (*e.g.* hydrochlorothiazide - molecular target unknown) and the high-ceiling diuretics (*e.g.* furosemide - inhibition of Na/K/2Cl symport). Sulfonamide diuretics are used in the treatment of glaucoma, hypertension and oedema. Further modifications of the original structures and new clinical observation[54] led to the development of new derivatives with diverse applications: glaucoma, hypertension, diabetes (*e.g.* the sulfonylurea tolbutamide), cancer, viral infections. Nowadays, sulfonamides are still of interest for pharmaceutical companies and are often integrated in new structures.[55] Cyclic sulfonamides (β-sultams) are functional analogues of β-lactams and can be interesting to include in the design of new antibiotics, although this concept has not yet yielded any efficient compounds.[56, 57] Acyclic sulfonamides are isosteres of the transition state of the hydrolysis of the amide bond[58] and also exhibit zinc-chelating properties.[55] As a result, they have been widely used as a substitute for the amide bond in peptidomimetic structures aiming at the reversible inhibition of proteases. Such an application has wide applicability in pharmaceutical research. Proteases are enzymes hydrolysing amid bond in peptides and are involved in various physiological pathways. Protease inhibitors could find application in the treatment of parasitic, fungal and viral (HIV) infections, hepatitis, cancer and inflammatory, immunological, respiratory and cardiovascular disorders. Compounds containing the sulfonamide group have been described as inhibitors of matrix metalloproteases (possible applications: atherosclerosis, rheumatoid arthritis and osteoarthritis, cancer), HIV-1 protease (anti-HIV drug amprenavir), renin (anti-hypertensive drug zankiren) and thrombin (argatroban, used against peripheral arterial occlusive diseases).[59, 60] Examples of marketed drugs containing a sulfonamide group are shown in Figure 9.

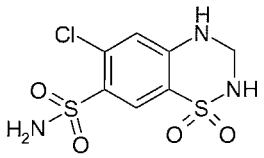
The widespread use of sulfonamides and their broad range of applications in pharmaceutical sciences made them ideal candidates for the design of a library aimed at drug-like compounds.



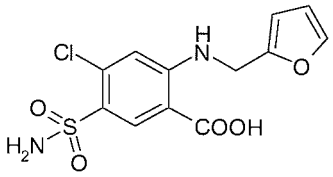
Sulfamethoxazole, antibiotic



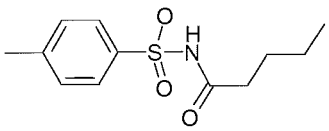
Acetazolamide, anti-glaucoma (CA inhibitor)



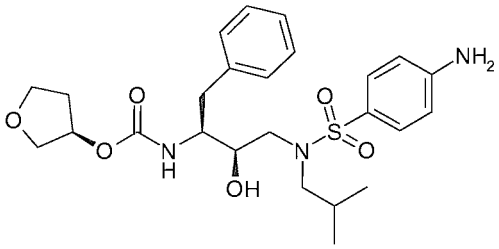
Hydrochlorothazide, diuretic



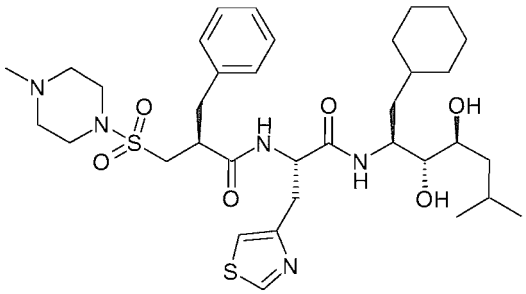
Furosemide, diuretic



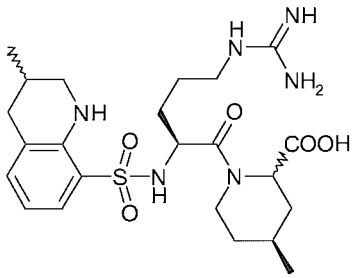
Tolbutamide, hypoglycemic



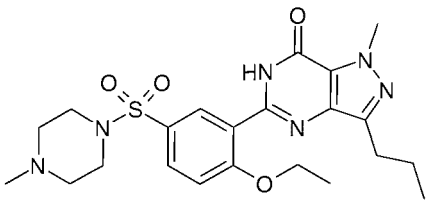
Amprenavir, anti-HIV



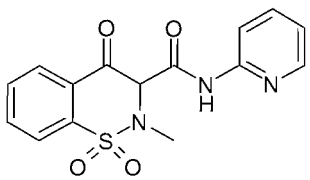
Zankiren, anti-hypertensive



Argatroban, anti-thrombin



Sildenafil, PDE5 inhibitor

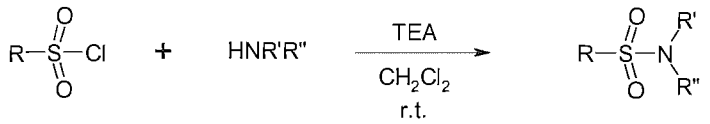


Piroxicam, NSAID

Figure 9. Examples of sulfonamide-containing drugs and their applications.

2.2.2.2. Ease of synthesis.

Sulfonamides are readily obtained through a nucleophilic substitution reaction between a sulfonyl chloride and an secondary amine,[61, 62] as shown in Scheme 2. This method allows for the synthesis of secondary or tertiary sulfonamides depending on whether a primary or secondary amine, respectively, is used. Experimental conditions will be detailed later.



Scheme 2. Nucleophilic substitution between a sulfonyl chloride and a secondary amine.

This ease of synthesis was consistent with the requirement that the compounds should be obtained quickly and in good yields. This, combined with the occurrence of sulfonamides in the drugs described above, prompted the selection of a library of sulfonamides for the purpose of the study.

2.2.3. The selection of the substituents and the final library.

The selection of substitutents was conditioned by the previously mentioned requirement: inclusion of privileged structures and consistency with the Rule of Five (to ensure drug-likeness of the analytes), presence of neutral and basic compounds (basic compounds constitute the majority of known drugs), availability of starting materials. Four sulfonyl chlorides were selected for the library: benzenesulfonyl chloride, *p*-trifluoromethylbenzenesulfonyl chloride, biphenyl-4-sulfonyl chloride and *p*-aminobenzenesulfonyl chloride (Figure 10). The latter was not commercially available, due to obvious reactivity issues. It was nonetheless selected because the presence of the ubiquitous amino side chain was deemed interesting and an alternative synthetic route could be designed still involving only three synthetic steps.

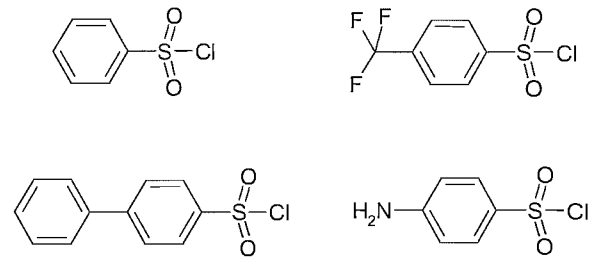


Figure 10. Sulfonyl chlorides chosen for the synthesis of the library.

According to the same principles, eight secondary amines were chosen for the synthesis of the library: *N*-ethylcyclohexylamine, *N*-(diphenylmethyl)methylamine, benzimidazole, 1-cyclohexyl-piperazine, 1-(diphenylmethyl)piperazine, 4-(ethylaminomethyl)pyridine, *N*,*N*,*N*′,*N*′-tetraethyldiethylenetriamine, *N*′-benzyl-*N*,*N*-dimethylethylenediamine. They are shown in Figure 11.

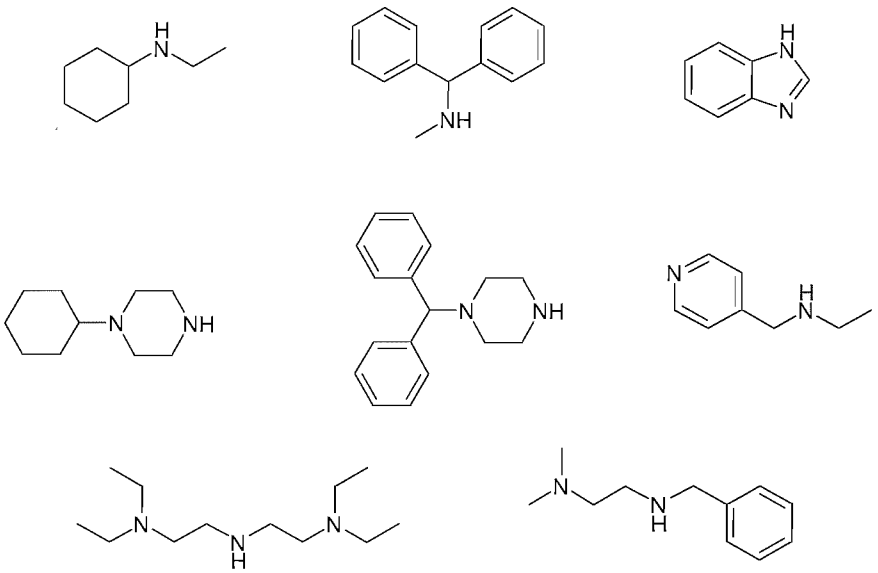


Figure 11. Amines chosen for the synthesis of the library.

The coupling of the four sulfonyl chlorides with the eight amines yielded thirty-two sulfonamides that constitute the test library. The test compounds are shown in Figure 12.

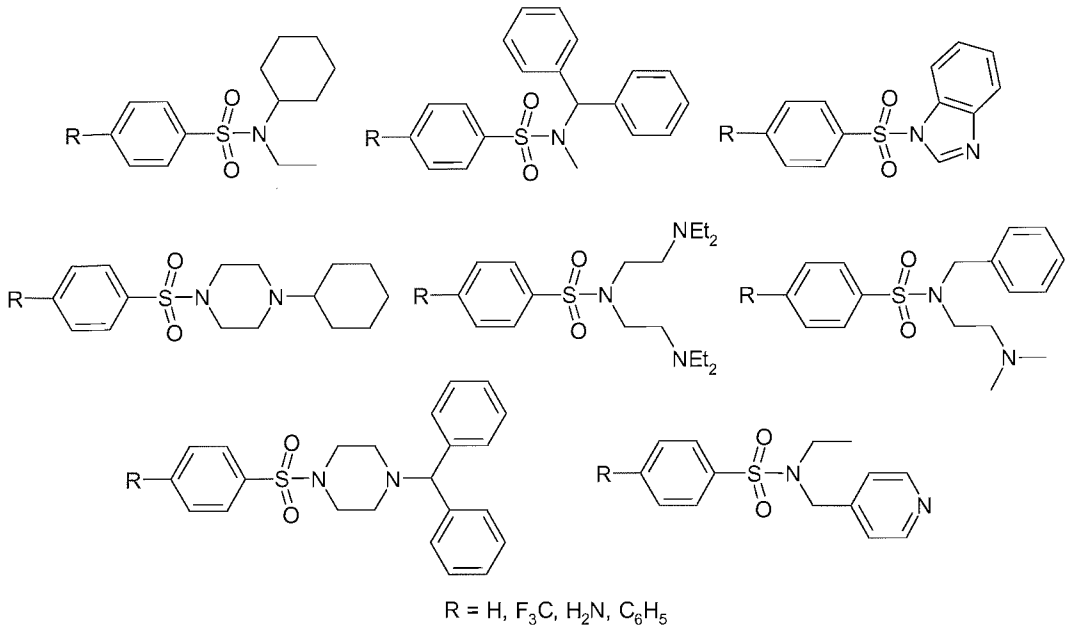


Figure 12. Test compounds.

2.2.4. Consistency of the library compounds with Lipinski's Rule of Five.

The number of hydrogen bond donors (*i.e.* OHs and NHs) and acceptors (*i.e.* Os and Ns), Clog P and molecular weight (MW) of the test compounds have been calculated to assess whether or not they display properties consistent with Lipinski's Rule of Five. All 32 sulfonamides meet the Lipinski's criteria of MW and number of hydrogen bond donors and acceptors. As far as Clog P is concerned, ten analytes (compounds **2**, **4**, **5**, **6**, **19**, **21**, **23**, **24**, **26** and **27**) are too much hydrophobic according to the Rule of Five. However, as mentioned in §2.1.2.3, only 10 % of the compounds in Lipinski's database had a combination of two or more criteria outside the required range. As a consequence, the usual practice is to not exclude compounds that do not meet one only of the criteria. According to this guideline, all thirty-two test sulfonamides are in accordance with the rule. Beside, if one considers the range defined by Ghose ($-0.4 < \text{Clog P} < 5.6$), only three compounds (compounds **5**, **8** and **26**) are too hydrophobic. Moreover, as one of the aims of the study is the design of a general SFC method suitable for a wide range of compounds, the broadening of the range of Clog P values is not detrimental to the current work. The calculated physico-chemical properties are

shown in Table 4. Clog P, Clog D at pH 7 and pK_a have been calculated using ACD software.[ii]

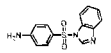
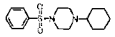
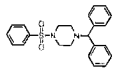
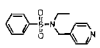
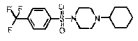
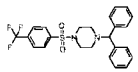
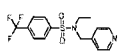
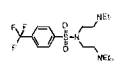
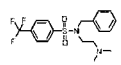
Table 4. Compounds structures and their physico-chemical poperties.

Compounds	Structures	H-bond donors	H-bond acceptors	Clog P[a]	RMM	Consistency with Rule of 5[b]	pK_a[a]	Clog D pH 7[a]
1		0	3	3.491 ±0.281	267.40	✓	-5.80 ±0.20	3.94
2		0	3	5.066 ±0.364	337.44	✓	-7.18 ±0.20	5.07
3		0	4	2.588 ±0.647	258.3	✓	3.72 ±0.18	-0.53
4		0	3	4.913 ± 0.362	335.39	✓	-6.53 ±0.20	4.91
5		0	3	6.038 ±0.430	405.44	✓	-8.53 ±0.20	6.04
6		0	4	3.559 ±0.666	326.3	✓	3.53 ±0.18	0.63
7		0	3	5.588 ±0.336	343.49	✓	-5.87 ±0.20	5.59
8		0	3	6.714 ±0.409	413.54	✓	-7.30 ±0.20	6.71
9		0	4	4.235 ±0.660	334.4	✓	3.69 ±0.18	1.14
10		1	4	3.261 ±0.309	282.41	✓	-6.66 ±0.20; 1.92 ±0.20	1.92
11		1	4	4.386 ±0.386	352.46	✓	-8.67 ±0.20; 1.68 ±0.20	3.27

[a] Clog P, pK_a and Clog D at pH 7 were calculated using ACD software.
[b] Structures are deemed consistent with Lipinski's rule when no more than one criterion is outside the range.

[ii] Thanks are due to Dr. Mark Taylor, from Pfizer, for the calculation of physico-chemical properties.

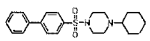
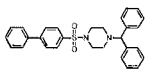
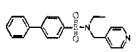
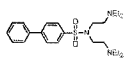
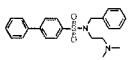
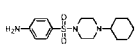
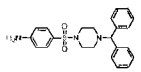
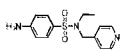
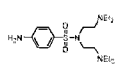
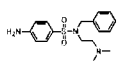
Table 4 (continued). Compounds structures and their physico-chemical poperties.

Compounds	Structures	H-bond donors	H-bond acceptors	Clog P[a]	RMM	Consistency with Rule of 5[b]	pK$_a$[a]	Clog D pH 7[a]
12		1	5	1.606 ±0652	273.32	✓	-0.15 ±0.20; 3.85 ±0.18	-1.72
13		0	4	3.559 ±0.390	308.45	✓	4.58 ±0.20; -9.63 ±0.20	-0.43
14		0	4	4.588 ±0.515	392.52	✓	2.83 ±0.20; 9.89 ±0.20	2.35
15		0	4	4.246 ±0.324	276.36	✓	4.16 ±0.20; -10.33 ±0.20	-1.29
16		0	5	4.246 ±0.444	257.36	✓	9.90 ±0.20; 9.24 ±0.20; -9.67 ±0.20	-13.7
17		0	4	3.813 ±0.403	318.44	✓	8.66 ±0.20; -12.34 ±0.20	-4.25
18		0	4	4.531 ±0.452	376.44	✓	4.17 ±0.20; -10.98 ±0.20	0.96
19		0	4	5.560 ±0.563	460.52	✓	2.42 ±0.20; -11.23 ±0.20	3.73
20		0	4	3.242 ±0.397	344.36	✓	4.11 ±0.20; -11.68 ±0.20	-0.27
21		0	5	5.218 ±0.499	311.33	✓	9.81 ±0.20; 9.15 ±0.20; -11.02 ±0.20	-12.54
22		0	4	4.785 ±0.464	386.44	✓	8.56 ±0.20; -13.68 ±0.20	-3.18

[a] Clog P, pKa and Clog D at pH 7 were calculated using ACD software.

[b] Structures are deemed consistent with Lipinski's rule when no more than one criterion is outside the range.

Table 4 (continued). Compounds structures and their physico-chemical poperties.

Compounds	Structures	H-bond donors	H-bond acceptors	Clog P[a]	RMM	Consistency with Rule of 5[b]	pK_a[a]	Clog D pH 7[a]
23		0	4	5.206 ±0.432	384.54	✓	4.46 ±0.20; -9.75 ±0.20	1.34
24		0	4	6.235 ±0.547	468.62	✓	2.71 ±0.20; -10.00 ±0.20	4.12
25		0	4	3.918 ±0.374	352.46	✓	4.14 ±0.20; -10.45 ±0.20	0.37
26		0	5	5.893 ±0.481	431.65	✓	9.86 ±0.20; 9.20 ±0.20; -9.79 ±0.20	-11.96
27		0	4	5.461 ±0.444	394.54	✓	8.62 ±0.20; -12.45 ±0.20	-2.65
28		1	5	2.879 ±0.411	323.46	✓	4.89 ±0.50; 2.34 ±0.10; -11.12 ±0.40	-3.15
29		1	5	3.908 ±0.531	407.54	✓	3.14 ±0.50; 2.34 ±0.10; -11.38 ±0.40	-0.38
30		1	5	1.591 ±0.349	291.37	✓	4.14 ±0.19; 1.23 ±0.10; -11.82 ±0.50	-2.67
31		1	6	3.566 ±0.462	258.34	✓	9.97 ±0.25; 9.32 ±0.25; -11.16 ±0.50	-15.32
32		1	5	3.133 ±0.423	333.46	✓	8.75 ±0.28; 1.00 ±0.10; -13.83 ±0.50	-5.56

[a] Clog P, pKa and Clog D at pH 7 were calculated using ACD software.
[b] Structures are deemed consistent with Lipinski's rule when no more than one criterion is outside the range.

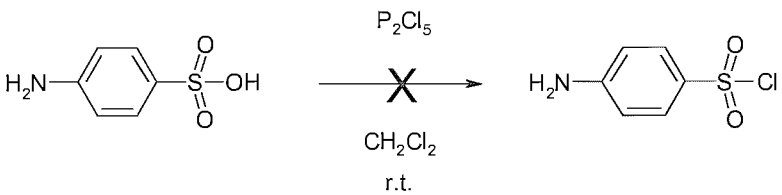
In addition to consistency with Lipinski's Rule of Five, the test library was aimed at neutral and basic compounds, hence the calculation of pK_a and Clog D (*cf.* Table 4). The presence of bases was even more desirable, since 67.5 % of all compounds from the WDI database are basic. For compounds **1**, **2**, **4**, **5**, **7** and 8, Clog P and Clog D at pH 7 values are very close, which indicates that these compounds will not be ionized at physiological pH, while benzimidazole and aniline groups display pK_{aH} of *ca.* 3.8 and 2.0, respectively, implying that these products must be ionized at physiological pH, but should not behave as bases. Other compounds exhibit differences between their Clog P and Clog D values and range from slightly to very basic. The most basic compounds are *N*,*N*,*N'*,*N'*-tetramethyldiethylenetriamine and *N'*-benzyl-*N*-dimethylethylenediamine derivatives (compounds **16**, **17**, **21**, **22**, **26**, **27**, **31** and **32**) with pK_{aH} above 8.

2.3. Synthesis of the library.

All benzenesulfonyl, trifluorobenzenesulfonyl and biphenylsulfonyl derivatives have been synthesized according to the procedure described in Scheme 2.

As 4-amino-benzenesulfonyl chloride was not readily available, syntheses of compounds **10** to **12** and **28** to **32** required its prior synthesis. Several procedures have been evaluated to reach this product, all starting from sulfanilic acid (4-aminobenzenesulfonic acid).

The first attempt involved sulfanilic acid and phosphorus pentachloride (PCl_5) in CH_2Cl_2 at room temperature, as described in Scheme 3.

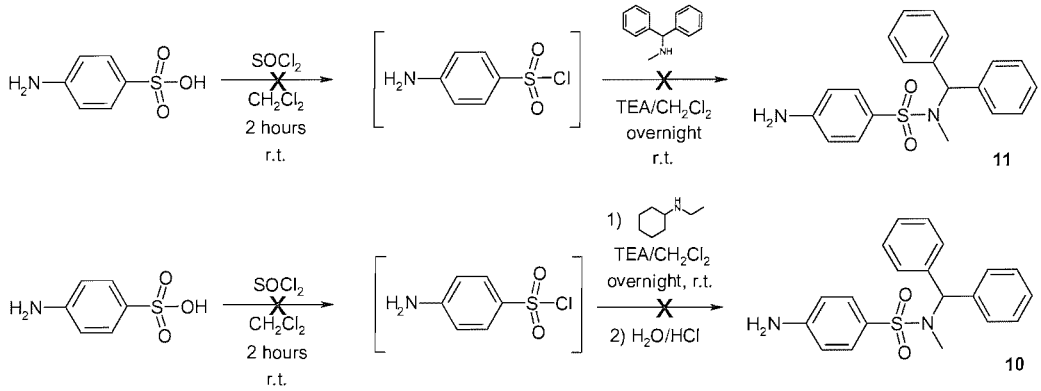


Scheme 3. Possible synthesis of 4-amino-benzenesulfonyl chloride.[63]

The reagents were not soluble in CH_2Cl_2 and, after stirring overnight, HPLC analysis showed that starting materials had not reacted.

The two next attempts were undertaken using DMF as solvent and thionyl chloride ($SOCl_2$) as chlorination agent. The reaction mixture was refluxed for several hours.[64, 65] Unexpectedly, the two reactions gave different results. The first time, the reaction mixture turned pink, while in the second attempt, the reaction mixture turned yellow. However, in both cases, the expected product could not be identified by MS.

Because of these difficulties, the next strategy was not to isolate the sulfonyl chloride but to synthesize it *in situ* and to make it react directly with an amine. (Scheme 4)



Scheme 4. Synthesis of 4-amino-*N*-benzhydryl-*N*-methyl-benzene sulfonamide and 4-amino-*N*-cyclohexyl-*N*-ethyl benzenesulfonamide.

Here again, the expected compounds were not obtained, after overnight stirring at room temperature the amine remained unreacted.

One possible explanation for the failure of these reactions is the poor solubility of sulfanilic acid in any of the organic solvents which were used during the reactions. This poor solubility could be explained by the fact that the molecule is likely to exist in its zwitterionic form. (Figure 13)

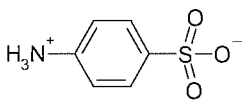
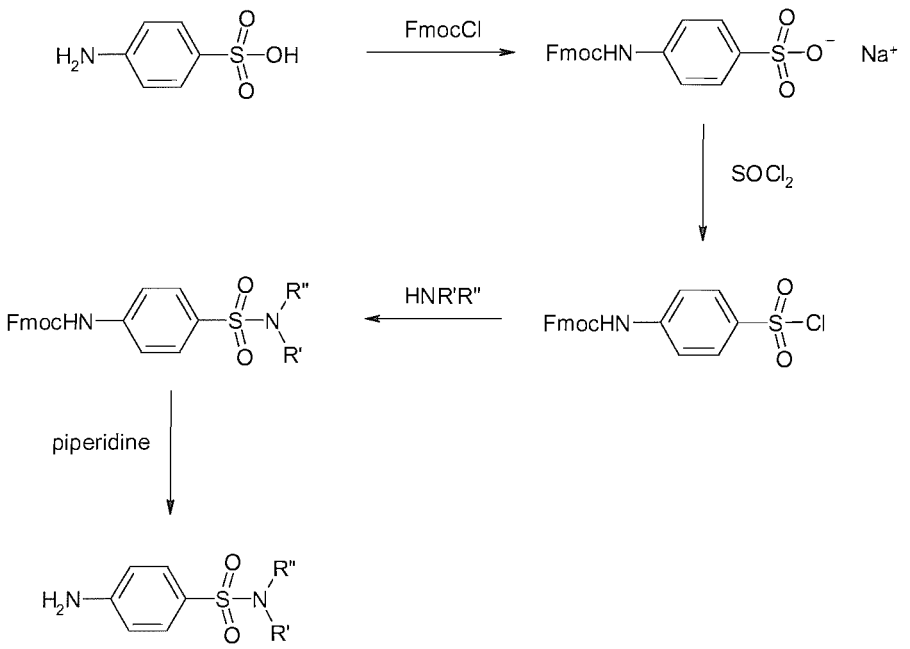


Figure 13. Zwitterionic form of sulfanilic acid.

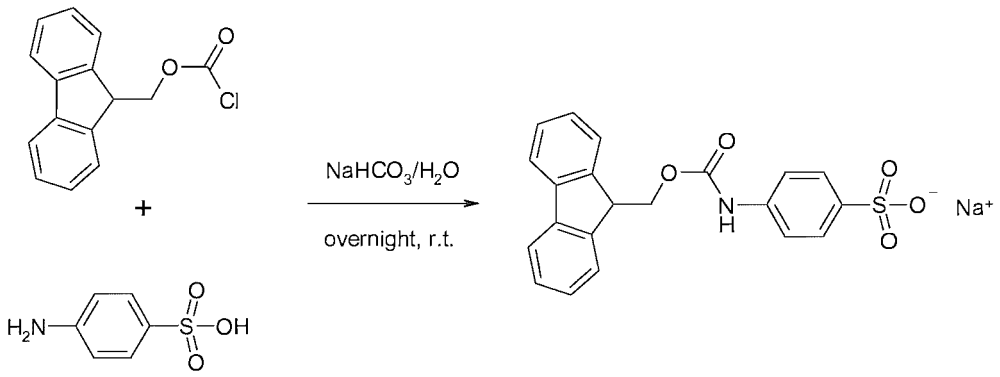
To avoid this problem, it was decided to protect the amino functionality of the sulfanilic acid. The protection group had to be stable to the chlorination conditions, *i.e.* in acidic conditions. The 9-fluorenylmethyl carbamate (Fmoc) protecting group was chosen due to its stability in acidic conditions and because it is readily cleaved, non-hydrolytically, by simple amines, and the protected amine is

liberated as its free base.[66] The new synthetic route leading to the compounds involved two more steps: the protection of the amino functionality and its deprotection in the final step, as described in Scheme 5.[67,68]



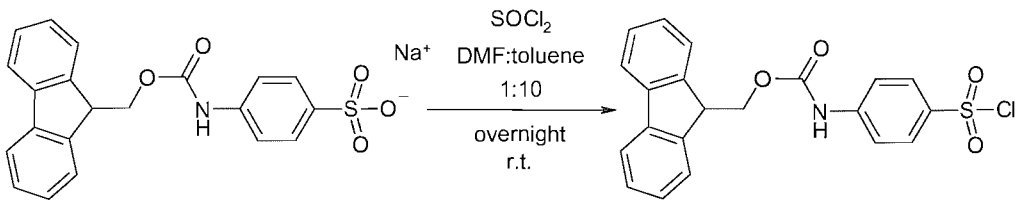
Scheme 5. Synthetic route to 4-amino-benzenesulfonamide derivatives.[35,36]

The Fmoc-protection is undertaken by reacting sulfanilic acid and 9-fluorenylmethyl chloroformate in saturated aqueous $NaHCO_3$ at room temperature overnight as described in Scheme 6.



Scheme 6. Synthesis of sodium 4-(9*H*-fluoren-9-ylmethoxycarbonylamino)-benzenesulfonate.[35,36]

The subsequent synthesis of the sulfonyl chloride was undertaken by using thionyl chloride as a chlorination agent and a mixture DMF:toluene 1:10 as a solvent. (Scheme 7)

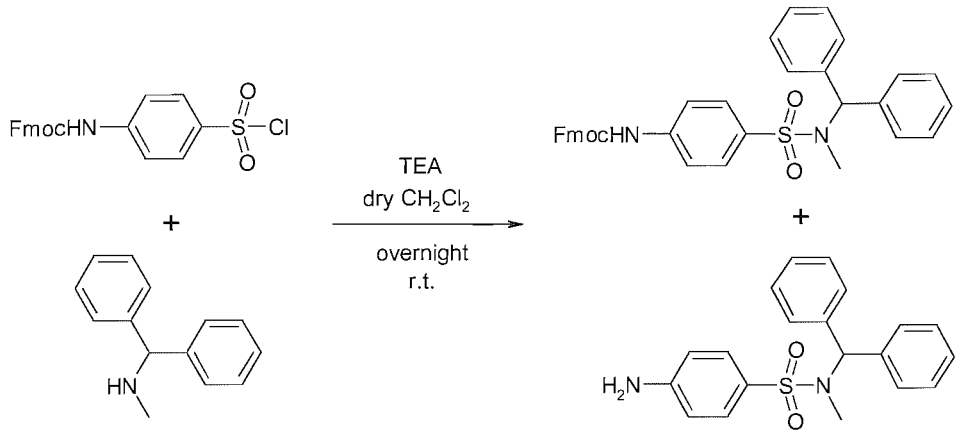


Scheme 7. Synthesis of (4-chlorosulfonyl-phenyl)-carbamic acid 9*H*-fluoren-9-ylmethyl ester. [35,36]

Sodium 4-(*9H*-fluoren-9-ylmethoxycarbonylamino)-benzenesulfonate was only partially soluble in such a mixture. The failure of the first attempt was thought to be due to this limited solubility and other mixtures were tried to achieve complete dissolution of the starting material, without achieving better results. It appeared that the failure was actually due to the presence of water in the solvent since the simple fact of drying the solvents on molecular sieve allowed the reaction to occur, affording the expected sulfonyl chloride in 60 % yield.

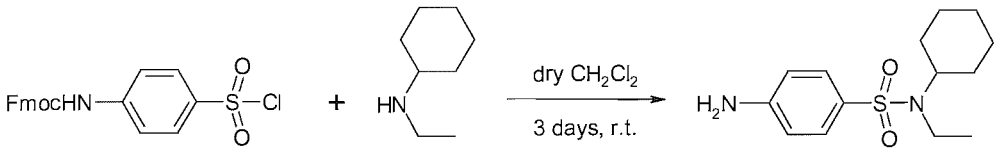
The final compounds were expected to be synthesized in subsequent steps, first substitution of the chlorine atom in CH_2Cl_2 or THF in the presence of triethylamine (TEA) as done for previous compounds and then deprotection of the amine functionality by action of piperidine in CH_2Cl_2.

Two observations led to slight modifications of this procedure. First, the hydrolysis of the sulfonyl chloride, which was not observed in previous syntheses, occurred. This issue was solved by undertaking the reactions in dried solvents. Secondly, during the synthesis of 4-amino-*N*-benzhydryl-*N*-methyl-benzene sulfonamide **11** using TEA in dry CH_2Cl_2, it was observed that the presence of amine and TEA allowed strong enough basic conditions to carry out the deprotection of the Fmoc-group readily in the substitution reaction mixture. (Scheme 8)



Scheme 8. Reaction between 4-amino-*N*-benzhydryl-*N*-methyl-benzene sulfonylchloride and *N*-(diphenylmethyl)methylamine.

One equivalent of piperidine was added to effect complete deprotection and the expected compound **11** was afforded in 37 % yield after filtration and purification. As a consequence, the synthesis of 4-amino-*N*-benzhydryl-*N*-methyl-benzene sulfonamide **10** was tried using excess appropriate amine in dry CH_2Cl_2 without TEA. (Scheme 9)



Scheme 9. Synthesis of 4-Amino-*N*-benzhydryl-*N*-methyl-benzene sulfonamide **10**.

Compound **10** was afforded after filtration of reaction mixture and purification in 82 % yield.

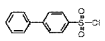
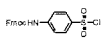
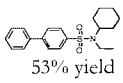
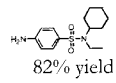
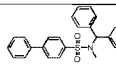
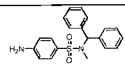
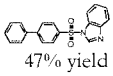
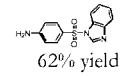
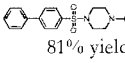
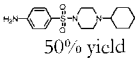
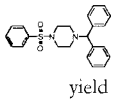
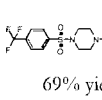
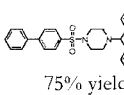
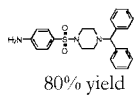
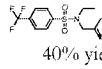
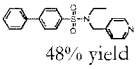
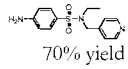
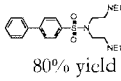
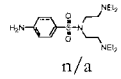
2.4. Conclusion.

A library of drug-like compounds was designed with the aim of studying the retention behaviour of pharmaceutical compounds in SFC. The design was underpinned by a number of requirements with which the final SOTLIB library had to be consistent: drug-likeness of the analytes, presence of neutral and basic compounds and ease of synthesis. Drug-likeness of the SOTLIB compounds was maximised by using such concepts as privileged substructures and Lipinski's Rule

of Five and by choosing a chemical functionality that is widely represented amongst drug classes: the sulfonamide group. This functionality also presented the advantage of being readily synthesised by coupling a sulfonyl chloride with an amine. This, combined with the selection of readily available reagents featuring privileged substructures, led to the design of thirty-two sulfonamides displaying drug-like properties.

The thirty-two SOTLIB compounds were subsequently synthesised in good yields for analysis by SFC. Compounds structures and numbers are shown in Table 5.

Table 5. SOTLIB compounds structures and numbers.

Sulfonyl chlorides / Amines				
	 1 59% yield	 4 73% yield	 7 53% yield	 10 82% yield
	 2 54% yield	 5 70% yield	 8 30% yield	 11 37% yield
	 3 36% yield	 6 52% yield	 9 47% yield	 12 62% yield
	 13 64% yield	 18 63% yield	 23 81% yield	 28 50% yield
	 14 69% yield	 19 69% yield	 24 75% yield	 29 80% yield
	 15 49% yield	 20 40% yield	 25 48% yield	 30 70% yield
	 16 65% yield	 21 79% yield	 26 80% yield	 31 n/a
	17 70% yield	22 58% yield	27 55% yield	32 75% yield

Chapter 3.

Thermodynamics of Chromatography and Quantitative Structure-Retention Relationships

3.1. Thermodynamics of chromatography.

Before starting with quantitative structure-retention relationships (QSRR) and their applications, it is of interest to explain the fundamental bases on which they are built or, rather, what are the limitations of a rigorous thermodynamic treatment of the chromatographic process that make sensible the use of QSRR.

Chromatographic separation consists in the differential migration of analytes carried through a stationary phase by a flow of mobile phase. Thus, chromatographic separation is driven by the molecular equilibrium of the analytes between the two phases. The basic thermodynamic principles underlying this process have been well described,[69] herein they are briefly explained.

3.1.1. Retention factor *k* and distribution coefficient *K*.

The retention of an analyte in a chromatographic column is most simply characterised by the retention time, t_R, *i.e.* the time needed by the analyte to travel from the injection point *via* the column to the detection point. The retention time can be adjusted to take into account the fact that part of the retention time is spent in (ideally) unretaining tubing. The corrected retention time, t'_R, is obtained as follows:

$$t'_R = t_R - t_0 \qquad (3.1)$$

where t_0 is the dead time of the column in the conditions used, *i.e.* the retention time of an unretained solute. The corrected retention time therefore corresponds to the time that the analyte spends in the stationary phase.

A convenient way of normalizing retention is to calculate the retention factor k that is defined as the ratio of the corrected retention time t'_R to the dead time t_0:

$$k = \frac{t'_R}{t_0} \qquad (3.2)$$

By definition, an unretained solute does not interact at all with the stationary phase. Thus, t_0 can be seen as the time that any analyte spends in the mobile phase. Since any analyte spends the same time as an unretained compound in the mobile phase, t'_R represents the time that the analyte spends in the stationary phase and, therefore, k represents the ratio of the time that any analyte spends in the stationary phase, t_{stat}, to the time it spends in the mobile phase, t_{mob}:[7]

$$k = \frac{t_{stat}}{t_{mob}} \qquad (3.3)$$

In isocratic conditions, the time spent by a solute in a given phase is proportional to the number of moles of this solute in that phase at any time. If we call n^{stat} and n^{mob} the number of moles of analyte in the stationary and mobile phases respectively, equation 3.3 becomes:

$$k = \frac{n^{stat}}{n^{mob}} \qquad (3.4)$$

In partition chromatography, the quantity of a given analyte in a phase is simply given by the product of the volume of the phase by the concentration of the analyte in that phase, and therefore:

$$k = \frac{c_i^{stat} V^{stat}}{c_i^{mob} V^{mob}} \quad (3.5)$$

where c_i are the concentrations of the analyte i in the stationary and mobile phases, and V are the volumes of the phases. The ratio of the volumes of the chromatographic phases is called the phase ratio β. And the ratio of the concentrations is, by definition, the partition coefficient K, thus:

$$k = \beta K \quad (3.6)$$

In adsorption chromatography, the expression of β and K are slightly different since the concentration of the analyte in the stationary phase is calculated with regard to surface area rather than volume, but the principles remain the same.

The retention factor k is therefore directly related to the distribution coefficient K of the solute between the two chromatographic phases. The chromatographic separation of two analytes is the final result of unequal distributions of the analytes between the phases arising from their differences in K. But what are the parameters influencing K?

3.1.2. Relation between Gibbs free energy *G*, chemical potential μ and distribution coefficient *K*.

Let us first consider a closed system, *i.e.* a system that can exchange heat and work, but no matter, with its surroundings. In that case, the first law of thermodynamics states that the variation of the energy of the system equals the sum of the work and heat applied to it:

$$dE = q + w \quad (3.7)$$

where q is the heat given to the system and w the work applied to it. If one considers that the only work applied is related to the pressure p in the system, then:

$$w = -pdV \qquad (3.8)$$

where dV is the variation of the volume V of the system (a positive work involves a contraction, therefore a negative value of dV)

The second law of thermodynamics relates q to the entropy S and the temperature T of the system:

$$q \leq TdS \qquad (3.9)$$

This leads to reformulating equation 3.7 into:

$$dE \leq TdS - pdV \qquad (3.10)$$

By definition, the Gibbs free energy G equals:

$$G = H - TS = E + pV - TS \qquad (3.11)$$

Hence, by differentiation of equation 3.11 followed by substitution using equation 3.10:

$$dG = dE + pdV + Vdp - TdS - SdT \qquad (3.12)$$

$$dG \leq Vdp - SdT \qquad (3.13)$$

Equation 3.13 is the mathematical translation of the known fact that at constant temperature and pressure, a naturally occurring process (*e.g.* diffusion) must have a negative dG. Moreover, at equilibrium, any transformation at constant T and p is characterised by $dG = 0$.

Let us now consider an open system, *i.e.* a system that exchanges work, heat and also matter with its environment. Any molecule taken in or out of the system will contribute to changes in the Gibbs free energy G. The variation of G with the coming in or out of a component i is proportional to the variation of the number of moles of i in the system, dn_i. The rate of variation of G with n_i ($\delta G/\delta n_i$) is defined as the chemical potential μ_i of component i. In the presence of multiple

components entering ($dn_i > 0$) or leaving ($dn_i < 0$) the system, individual contributions sum up in the expression of G, therefore:

$$dG = -SdT + pdV + \sum_i \mu_i dn_i \qquad (3.14)$$

If temperature and pressure are kept constant, equation 3.14 is simplified into:

$$dG = \sum_i \mu_i dn_i \qquad (3.15)$$

The chemical potential μ_i of component i in the system can be expressed as follows:

$$\mu_i = \mu_i^0 + RT \ln c_i \qquad (3.16)$$

where R is the gas constant and μ_i^0 is the standard chemical potential that characterises the affinity of i for the system, *i.e.* how much favourable it is for i to find itself in the given system. μ_i^0 depends on the energy of interactions between the solute and the system and to a lesser extent to entropic contributions. The second term, $RT \ln c_i$, is related to the entropy of dilution of i and therefore to its activity a_i. In the present case only highly diluted components (solutes) are considered and therefore the activity is considered equal to the concentration c_i of component i in the system.

Now let us consider a solute i in a chromatographic system, that is constituted of two phases (mobile and stationary) between which the solute will distribute. At equilibrium, and considering constant T and p, one has:

$$\Delta G = \Delta \mu_i = \mu_{i,eq}^{stat} - \mu_{i,eq}^{mob} = 0 \qquad (3.17)$$

It comes therefore that:

$$\mu_{i,eq}^{stat} = \mu_{i,eq}^{mob} \qquad (3.18)$$

And, by substitution using equation 3.16:

$$\mu_i^{0,stat} + RT\ln c_{i,eq}^{stat} = \mu_{i,eq}^{0,mob} + RT\ln c_{i,eq}^{mob} \Rightarrow \exp(\frac{-\Delta\mu_i^o}{RT}) = \frac{c_{i,eq}^{stat}}{c_{i,eq}^{mob}} = K \quad (3.19)$$

As seen in §3.1.1, K is the distribution (or partition) coefficient and is directly related to the retention factor k of the analyte i. Therefore, the retention of the analyte depends on its distribution between the two phases, *i.e.* on K, which itself depends on the difference between the standard chemical potentials of the analyte in the two phases. When equilibrium is reached in the column, the chemical potentials μ_i of the analytes in the two phases are equal (equations 3.17 and 3.18). According to equation 3.16, this equality of μ_i means a smaller μ_i^0 in one of the two phases is compensated for by an increase in the concentration of i in that same phase.

In summary, the differential distribution of i between the two phases, that is the cause of the retention phenomenon, is linked to the difference in the affinity of i for the mobile and stationary phases.

3.1.3. Forces influencing retention.

The thermodynamic principles exposed above can be applied to any type of chromatographic system. The difference of standard chemical potentials of the analyte, $\Delta\mu_i^0$, will drive its retention. Because μ_i^0 is defined as the increase of G per mole of added solute i in standard conditions, $\Delta\mu_i^0$ can be written:

$$\Delta\mu_i^0 = \Delta H_i^0 - T\Delta S_i^0 \quad (3.20)$$

where H_i^0 and S_i^0 are the standard partial molar enthalpy and entropy of solute i. Using equation 3.19, this can be reformulated:

$$\ln K = -\frac{\Delta H_i^0}{RT} + \frac{\Delta S_i^0}{T} \quad (3.21)$$

The retention is thus influenced by enthalpic and entropic terms, although enthalpic terms are usually predominant.[69] Enthalpic factors are related to intermolecular forces between the analyte and the phases, while entropic forces relate to the degrees of freedom of the analyte.

Intermolecular forces influencing retention are forces that act without modifications of the molecules involved (*e.g.* formation or cleavage of covalent bonds), with ion-exchange chromatography being a somewhat special case. They are of different types.

3.1.3.1. Van der Waals forces.

Van der Waals interactions are chemically non-specific attractive interactions that result from the electric field produced by non-ionized molecules. They operate at close range, their intensity being proportional to r^6, where r is the distance between the two molecules involved. Van der Waals forces are of different types.

Dipole-dipole interactions (sometimes called orientation interactions) operate between two molecules having a permanent dipole. Their intensity is proportional to the square of the product of the two dipoles.

When a molecule does not possess a permanent dipole, a so-called "induced dipole" can be created by the proximity of a dipolar molecular, giving rise to another type of van der Waals force, dipole-induced dipole (or inductive) interactions, the intensity of which depends on the polarizability of the apolar molecule.

The third kind of van der Waals interactions is the dispersive interactions that occur between apolar molecules. In that latter case, intermolecular attraction is explained by the fact that, even though the molecules do not have a permanent dipole moment, fluctuations in the electronic distribution in the molecules generate, at any time, a small dipole.

3.1.3.2. Ionic interactions.

Still related to electric field, ionic interactions can also be of importance for chromatographic separations (they rule the separation in ion-exchange chromatography for instance). Dipolar molecules interact with ions through coulombic forces called ion-dipole interactions.

3.1.3.3. Chemically specific interactions.

Chemically specific interactions, *i.e.* interactions that are induced by the presence of chemical element or features rather than by electro-magnetic properties, also play an important role in chromatographic separations. Hydrogen bonds involve a donor (that provides a hydrogen atom bonded to an electronegative atom such as oxygen or nitrogen) and an acceptor (that is an electron rich group such as oxygen in alcohols, ethers, carbonyls; nitrogen in amines; or π-systems).

Another type of interaction, electron pair donor-electron pair acceptor interaction, occurs between molecules that have a pair of electrons to share (lone pairs of nitrogen or oxygen atoms for instance) and molecules with electron deficiency (*e.g.* metals, aromatic polynitro compounds).

3.1.4. Limitations of the thermodynamic approach.

Since the thermodynamic principles of chromatography are known, the idea of predicting the retention of an analyte from thermodynamic properties seems reasonable. However, the fact is that the calculation of μ_i^0 is impossible from the molecular features of an analyte. Specific equations of state must be especially devised to best describe the systems being investigated. Authors have attempted to design models of chromatographic separation explaining and predicting solute retention. The most successful model has been developed by Martire and gives an

unified thermodynamic model for GC, LC and SFC.[70, 71] Schoenmakers developed a thermodynamic model for SFC based on the Lee and Kesler equation of state.[72] The model was in good agreement with existing experimental data but cannot be used for prediction. Bartle and co-workers have described the retention of four aromatic hydrocarbons using the Peng-Robinson equation of state.[73] All these models are effective at explaining already existing chromatographic data, but are nonetheless impractical because they necessitate the knowledge of parameters that are either not readily available or altogether impossible to calculate from molecular structures.

The impracticality of a purely theoretical description of the chromatographic process makes necessary to resort to empirical equations obtained by means of correlations of experimental data with molecular properties that can be computed for any kind of analyte.

3.2. Quantitative structure-retention relationships (QSRR).

3.2.1. Introduction.

Alternatively to using a purely thermodynamic model, the characteristic of interest can be described using an empirical model. For that purpose, the characteristic is measured experimentally for a series of analytes that are described by means of empirical or calculated properties describing their molecular structure. Practically the values of the characteristic (or "dependent variable") are analysed to determine whether or not they are mathematically related to the descriptors, also called "independent variables" because they must not be inter-correlated to ensure the statistical validity of the model. The outcome of the procedure is, for instance, an equation describing the characteristic as a function of the descriptors. The same methodology is used, for instance, to understand how a series of drugs interact with their biological target: quantitative structure-activity relationships (QSAR).

When applied to chromatography, this method is referred to as quantitative structure-retention relationships (QSRR). In that case, the dependent variable can be, for instance, the retention factor, while the descriptors are properties that encode the molecular structures of the studied analytes and relate to the chromatographic process. Ideally, the successful QSRR models make the prediction of the retention of compounds that have not yet been analysed, possible. Since their introduction in the 1970s, QSRR studies have been widely used, not only for retention prediction but also for retention mechanism elucidation, drug compounds classification, meaningful molecular descriptors identification or stationary phases classification. A review summarising QSSR works for the period 1996-2006 has been published by Hébeger,[8] although it does not include a review of the research in SFC. West and Lesellier used linear solvation energy relationships (LSER, *cf.* §3.2.2) to characterise an extensive set of stationary phases in subcritical fluid chromatography.[74-77] Other models have been used to predict retention factors. Alvarez and Baumann used first principles and a model of electrostatic interactions between weakly polar solutes and the dielectric continuum of the CO_2 mobile phase, and predicted the retention times of seven pesticides.[78] Fatemi and Baher proved the suitability of artificial neural network techniques to predict capacity factors of organic compounds.[79]

3.2.2. Molecular descriptors.

The choice of a set of descriptors for derivation of a QSRR equation is a difficult problem. The number of different descriptors that can be calculated for a given analyte is virtually infinite. When choosing molecular descriptors to mathematically model a given process, one should always try to focus on descriptors that can be physically interpreted in term of this process. For instance, in QSRR, bulkiness descriptors like van der Waals area or molecular mass are often useful because they can be interpreted in terms of the ability of the analyte to be involved in dispersive interactions. Similarly, total dipole moment or orbital energies of HOMO and LUMO are descriptors encoding the polarity of the molecule and therefore its ability to interact with mobile and stationary phases through dipole-dipole or hydrogen bonding interactions.

The most widely used QSRR model is based on linear solvation energy relationship (LSER) descriptors, introduced and developed by Carr and Abraham; this uses the logarithm of the retention factor, log k, as retention characteristic.[80, 81] In this model, reversed phase HPLC log k is described as a sum of terms accounting for interactions of the probe analytes with the chromatographic system, leading to an equation of the following type:

$$\log k = eE + sS + aA + bB + vV + c \qquad (3.22)$$

where lower case letters are constants characterising the chromatographic system used and upper case letters are molecular descriptors of the probe analytes. E is the excess molar refraction (calculated from the refractive index of the molecule) and models polarizability contributions from n and π electrons; S is the solute dipolarity/polarizability; A and B are the solute overall hydrogen-bond acidity and basicity; V is the McGowan characteristic volume.

The main advantage of this model is that all the parameters can be interpreted in terms molecular interactions, giving insight into the retention mechanism. Its major drawback is that polarizability and overall hydrogen-bond acidity and basicity are empirical parameters. Although these parameters are available for an increasing number of compounds, the model is impractical when dealing with completely new entities. For that reason, one should use non-empirical descriptors, *i.e.* descriptors that can be obtained from computational chemistry calculations. Examples of calculable descriptors are shown in Table 6, classified by type of property encoded.

Table 6. Examples of non-empirical molecular descriptors (adapted from Kaliszan[69])

Property encoded	Descriptors
Molecular-bulkiness	Carbon number, molecular mass, refractivity, polarizability, van der Waals area, solvent accessible surface area, total energy
Molecular polarity	Dipole moment, partial atomic charge, orbital energy of HOMO and LUMO
Molecular geometry	Length-to-breadth ratio, moment of inertia
Molecular topology (graph-derived)	Connectivity indices, topological electronic indices

Kaliszan and co-workers used a general QSRR model to compare retention properties of diverse HPLC columns.[82-84] Their model is based on the following molecular descriptors: (a) total dipole moment, μ (not to be confused with the chemical potential also conventionally noted μ), that accounts for the dipole–dipole and dipole-induced dipole attractive interactions; (b) electron excess charge of the most negatively charged atom, δ_{min}, that models ability of probe analytes to participate in polar and hydrogen-bonding type interactions; (c) water-accessible molecular surface area, A_{H2O}, that accounts for dispersive interactions. These descriptors are used to derive, through multiple regression analysis, an equation allowing for the calculation of the chosen retention characteristics (in that case $\log k_w$, *cf.* §4.2):

$$\log k_w = a\mu^2 + bA_{H2O} + c\delta_{\min} + d \qquad (3.23)$$

where a, b, c and d are characteristics of the system.

3.3. Calculations of molecular descriptors of the test compounds.

3.3.1. Spartan '02.

Spartan '02 for Windows is a molecular modelling package developed by Wavefunction, Inc. (Irvine, CA, USA). The software has been designed to be used by experimental chemists who do not necessarily have an extended knowledge of molecular mechanics and quantum chemical calculations. Through a graphical interface permitting the drawing of organic and inorganic molecules, it allows the user to calculate a variety of molecular properties using various calculation methods.

Empirical molecular mechanical models are the simplest computational methods. They are usually applied to determine equilibrium conformations. The

models used in Spartan '02 are based on SYBYL and Merck Molecular Force Field (MMFF94) force fields.

Semi-empirical molecular orbital models provide conformational information but also equilibrium and transition state geometries as well as thermodynamic and kinetic data for molecules of up to two hundred atoms. The models available in Spartan '02 are the modified neglect of diatomic overlap (MNDO) model and its variants, Austin model 1 (AM1) and parameterised model 3 (PM3).

Spartan '02 also provides molecular orbital methods such as the Hartree-Fock (HF) model that may be used for determination of equilibrium and transition state geometries, thermodynamic or kinetic data, and vibrational frequencies. These models can handle molecules up to one hundred atoms but cannot be used when the structure includes transition metals.

For the latter case, correlated methods are required, *e.g.* Møller-Plesset (MP) models, although they are much more computationally costly and are not suited for routine calculations.

The molecular descriptors calculated in the present study, that are described below, have been calculated by means of molecular mechanics geometry optimisation followed by semi-empirical molecular orbital calculations (AM1) and HF molecular orbital calculations. The geometry of the molecules was first optimized to a minimum using an empirical molecular mechanics based on MMFF94. The use of MMFF94, along with the careful drawing of the structures using Spartan's structure drawing interface, ensured the reproducibility of the obtained optimised conformers. This provided the surface area, A, values (*vide infra*). Other descriptors were calculated for single point energy at ground state, using a HF method with a 3-21G* basis set, starting from MMFF conformer and AM1 geometry.

3.3.2. Molecular surface area, *A*.

The surface area, A, calculated using Spartan '02 corresponds to the van der Waals surface area of the molecule. However, in his paper, Kaliszan made use of the solvent accessible surface (SAS) area.[85] The concept of SAS originates from

the work of Lee and Richards.[86] Practically the solvent-accessible area is calculated by moving a spherical probe representing the solvent molecule over the van der Waals surface of the molecule of interest. The trajectory of the probe defines the SAS, the area of which can be computed. Because he studied HPLC systems, Kaliszan had chosen the water-accessible surface area (A_{H2O}) as a descriptor.

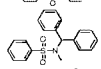
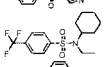
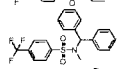
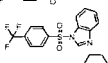
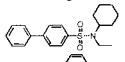
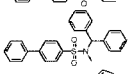
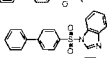
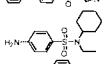
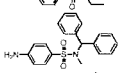
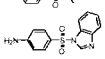
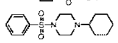
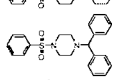
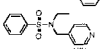
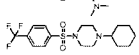
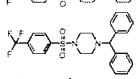
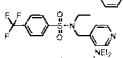
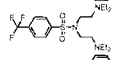
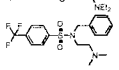
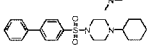
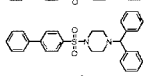
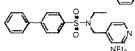
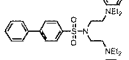
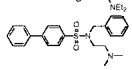
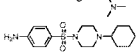
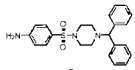
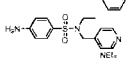
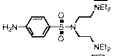
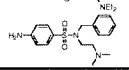
Spartan '02 does not provide the user with the possibility of calculating SAS areas. Kaliszan and co-workers used another chemical modelling package: Hyperchem (Hyper-Cube, Waterloo, Canada). An evaluation version of this software was obtained in order to calculate SAS areas.

The critical parameter in the calculations of SAS areas is the radius of the probe. A literature search showed that the commonly accepted radius for a water probe is 1.4 Å. The origin of that value can be traced back to the seminal article by Lee and Richards in 1971.[86] Interestingly, the value of 1.4 Å was actually "assumed" by the authors, although it seems that the assumption originates from the van der Waals radii values compiled by Bondi.[87, 88] In the present case, the mobile phase is mainly constituted of MeOH and CO_2. It seems therefore more appropriate to use MeOH or CO_2 accessible surface area (A_{MeOH} and A_{CO2} respectively). The probe radius for MeOH was chosen at 1.7 Å.[89] For the CO_2 molecule, the radius was set at 1.8 Å.

Using Hyperchem, the geometry of the molecules was first optimized to a minimum using an empirical molecular mechanics based on MM+ force field. This step was followed by semi-empirical calculations based on the AM1 model. The surface area, A, calculated by the software corresponds to the van der Waals surface of the molecule. The solvent-accessible surface areas were calculated for the different solvents: water, MeOH and CO_2.

The results of molecular surface area A calculations using Spartan and Hyperchem and SAS areas calculations for the three solvents using Hyperchem are shown in Table 7.

Table 7. Molecular surface area calculations.

Cpds	Structures	Spartan calculations	Hyperchem calculations			
		A (Å^2)	A (Å^2)	A_{H2O} (Å^2)	A_{MeOH} (Å^2)	A_{CO2} (Å^2)
1		295.43	292.16	469.05	516.09	532.43
2		358.71	343.83	542.89	593.90	611.42
3		261.61	251.74	433.85	481.20	496.89
4		332.51	326.03	514.96	566.39	583.98
5		395.66	376.65	585.26	635.98	653.93
6		298.42	281.18	480.19	529.92	545.93
7		375.03	375.52	579.52	634.35	651.35
8		438.44	429.98	672.43	733.46	755.14
9		341.34	324.67	545.16	601.60	620.22
10		308.86	315.40	490.68	538.72	555.09
11		369.68	358.33	561.19	614.29	631.41
12		275.06	263.80	451.31	499.01	514.76
13		334.76	336.70	540.05	592.58	610.28
14		416.46	410.88	639.05	698.06	717.56
15		297.50	291.65	471.94	520.85	537.23
16		429.00	430.57	638.10	692.27	711.30
17		356.86	346.87	544.41	596.95	615.33
18		371.89	361.59	582.44	636.97	656.51
19		453.52	443.30	687.36	748.32	770.29
20		334.52	325.01	498.71	546.97	563.27
21		464.68	460.36	667.24	724.62	744.83
22		394.18	375.11	551.15	601.58	618.66
23		414.38	411.34	655.26	715.16	735.73
24		492.08	484.89	756.07	822.38	845.64
25		372.40	372.44	588.67	645.46	664.34
26		510.02	506.45	750.13	810.94	832.10
27		436.94	423.53	649.41	708.52	729.26
28		348.07	351.97	561.52	615.48	633.73
29		429.83	427.62	664.27	724.65	745.68
30		301.56	313.12	495.26	544.86	561.72
31		443.57	447.03	647.68	703.46	721.06
32		370.42	365.74	564.19	617.06	634.04

Van der Waals surface areas, obtained using Spartan, are very similar to those calculated with Hyperchem, as expected. A plot of Spartan against Hyperchem values gives a linear distribution, with a slope close to unity (0.999), a small intercept (-5.969 Å^2), a squared correlation coefficient R^2 of 0.986 and a standard error of estimate of 7.636 (Figure 14).

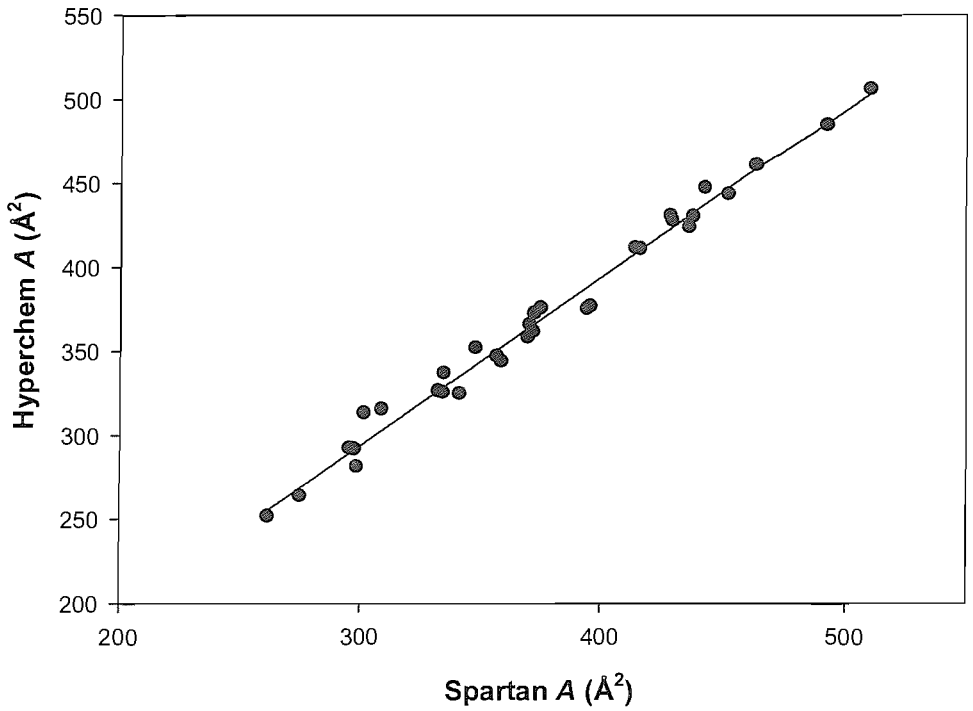


Figure 14. Spartan calculated van der Waals surface areas *A vs.* Hyperchem calculated values. Intercept = -5.97 Å^2, slope = 1.00, R^2 = 0.99, standard error of estimate = 7.64.

More interestingly, the SAS areas calculated with Hyperchem are linearly correlated with the van der Waals surface areas obtained with the same software. For instance the plot of Hyperchem *A vs.* Hyperchem A_{wat} is shown in Figure 15. Linear regression analysis gives a slope of 1.29, an intercept of 101.12 Å^2 and a R^2 of 0.96.

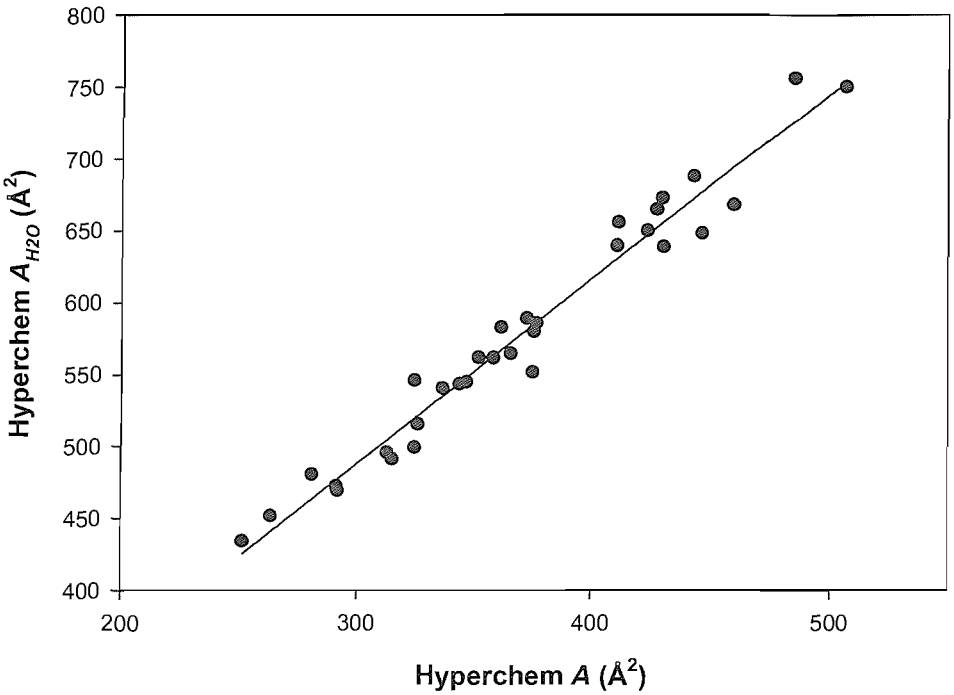


Figure 15. Hyperchem calculated van der Waals surface areas A *vs.* Hyperchem calculated water-accessible surface areas. Intercept = 101 Å^2, slope = 1.3, R^2 = 0.96.

Similar correlations are observed between A and A_{MeOH} and between A and A_{CO2} (data not shown). The fact that van der Waals and SAS areas are linearly correlated implies that the use of one or another in a statistical model is valid. Indeed, whatever the studied characteristic, if shown to be correlated with one descriptor, say A_{H2O}, the above observed correlation implies that it will also be correlated with A, A_{MeOH} and A_{CO2}. The use of one calculated area instead of another would merely change the numeric value of the coefficients of the yielded correlation equation, without affecting the statistical significance.

For this reason, and because the other descriptors used in this work were calculated using Spartan '02 (*vide infra*), Spartan'02 A values were chosen for the derivation of the QSRR equations described in the next chapter.

3.3.3. Total dipole moment, μ.

The total dipole moment gives an indication of the charge distribution in the molecule. It can be seen as a measure of the polarity of a chemical entity. Its

intensity can be measured experimentally, although its sign and direction are much more difficult to determine through experiments. These properties can be calculated with reasonable accuracy using Hartree-Fock models. Total dipole moments of the test analytes were calculated using a HF method with a 3-21G* basis set, starting from MMFF conformer and AM1 geometry. The calculated values are shown in Table 8.

3.3.4. Partial atomic charges.

Atomic charges cannot be individually defined, since they are the result of the overall repartition of electrons in the molecule. More importantly, they cannot be measured experimentally. Different methods exist to evaluate partial atomic charges. The so-called Mulliken procedure starts by integrating the electron density function to determine the total number of electrons in a molecule. Then, electrons are partitioned among individual basis functions (*i.e.* atoms) to compute partial electronic charges on each atom. An alternative method makes use of the electrostatic potential that represents the energy of interaction of a single positive charge with the molecule, due to attractive interactions with the electronic cloud and repulsive interactions with the nuclei. Alternatively, the electrostatic potential can be seen as resulting from the interactions of the single positive charge with partially charged atoms. Hence, atomic partial charges can be calculated as follows:[90]

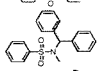
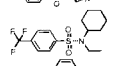
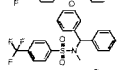
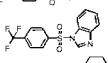
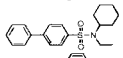
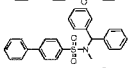
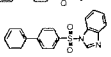
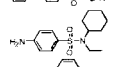
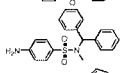
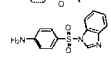
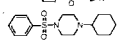
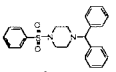
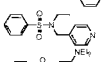
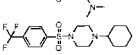
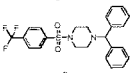
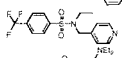
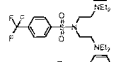
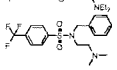
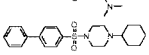
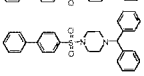
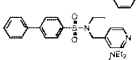
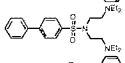
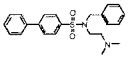
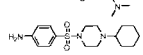
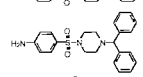
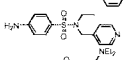
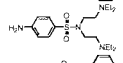
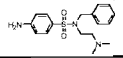
- calculate wavefunction of the molecule of interest;
- define grid points around the molecule;
- calculate electrostatic potential for each of the grid points;
- fit the calculated electrostatic potential with a model potential based on partial atomic charges (that are treated as parameters to be determined with the constraint that the sum of partial charges equals the total charge of the molecule)[iii]

[iii] It is important to note that the grid points are defined arbitrarily and that the calculated charges are dependent on the settings.

According to Kaliszan,[85] the electronic charge on the most negatively charged atom, was chosen as a descriptor for the QSRR study described in the next chapter. Atomic charges were calculated using Spartan'02, in which the electrostatic potential method described above is implemented. As for total dipole moment, calculations were based on a Hartree-Fock method with a 3-21G* basis set, starting from MMFF conformer and AM1 geometry.

Calculation of δ_{min} showed that, in most cases, the most negative partial charge is located on one of the two sulfonamide oxygen atoms, although the specific most electronegative atom varies with the structure. For instance, in the case of amino-substituted benzene rings, the most electronegative is the nitrogen atom of the $-NH_2$ group. The values of δ_{min} for the test compounds are shown in Table 8.

Table 8. Calculated dipole moments, μ, and electronic charges on the most negatively charged atom, δ_{min}.

Cpds	Structures	μ (debyes)	δ_{min} (electrons)
1		5.5589	-0.696576
2		5.4793	-0.692614
3		5.4776	-0.681072
4		4.9617	-0.703724
5		5.0594	-0.684566
6		2.1631	-0.673554
7		5.8939	-0.708109
8		5.7657	-0.691448
9		6.3054	-0.682336
10		6.4013	-1.085524
11		6.6950	-1.092456
12		8.0977	-1.082750
13		4.7297	-0.664081
14		4.9130	-0.665910
15		7.0769	-0.777623
16		4.0004	-0.774443
17		4.3599	-0.670165
18		5.0541	-0.663996
19		5.1688	-0.654347
20		4.1545	-0.817220
21		4.1476	-0.747352
22		4.2278	-0.680205
23		4.9515	-0.685576
24		6.9363	-0.771575
25		4.0683	-0.746009
26		4.3237	-0.747402
27		4.7138	-0.698782
28		5.4079	-1.090688
29		5.6084	-1.088727
30		4.7766	-1.072884
31		4.9464	-1.086398
32		5.5313	-1.111102

3.4. Conclusion.

Quantitative structure-retention relationships (QSRR) are statistical relationships linking chromatographic characteristics of a series of analytes, to measured or calculated quantities related to their structural differences. The interest for this type of empirically derived relationships results from the impracticality of applying basic thermodynamic principles to describe the retention process, due to its complexity. QSRR are widely used for retention prediction, retention mechanism elucidation, meaningful molecular descriptors identification or stationary phase classification.

Modern computational methods provide the analytical chemists with a vast diversity of descriptors to encode the structures of the analytes of interest. The best descriptors are those that can be physically interpreted rather than being merely abstract quantities. The retention of drug-like compounds in rp-HPLC have been described by Kaliszan and co-workers using three molecular descriptors: total dipole moment, μ, water-accessible surface area, A_{H2O}, and electron excess charge of the most negatively charged atom, δ_{min}.

μ and δ_{min} have been calculated for the thirty-two SOTLIB compounds using the molecular modelling package Spartan'02 for windows. Spartan'02 does not provide means to calculate solvent-accessible surface (SAS) areas, but only van der Waals surface areas, A. SAS areas were calculated using another modelling package, Hyperchem 7.52. The SAS areas were observed to be highly correlated to van der Waals areas, resulting in the statistical validity of using one *in lieu* of the other. For consistency purpose, μ and δ_{min} being calculated using Spartan'02, the van der Waals surface area, A, was chosen as a descriptor rather than SAS area.

These descriptors have been used to derive a QSRR equation describing the retention of SOTLIB compounds in SFC, as described in the next chapter.

Chapter 4.

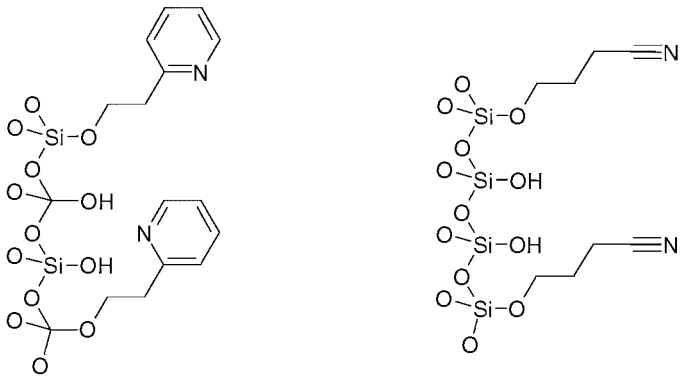
SFC Analysis of the SOTLIB Library of Sulfonamides

4.1. Preliminary study.

4.1.1. An isocratic approach.

SOTLIB compounds were synthesised *de novo* as described in Chapter 2. As a consequence, and although the suitability of SFC for diverse drug-like compounds had been demonstrated, the chromatographic behaviour of the test compounds was initially unknown. For that reason, the study was started using standard conditions recommended by the manufacturer. The outlet pressure was set at 100 bar, oven temperature 35°C, flow rate 4 mL min^{-1} with a 4 μL injection volume. Moreover, an emphasis was put on studying the effect of the stationary phase by keeping a single mobile phase while varying the column packing.

Initial experiments were undertaken using three silica-based stationary phases, 2-ethylpyridine (2-EP), cyanopropyl (CN) and diol. Structures of the stationary phase are shown in Figure 16. All columns were 250 mm in length by 4.6 mm I.D., 6 μm particle size and 60 Å pore size.



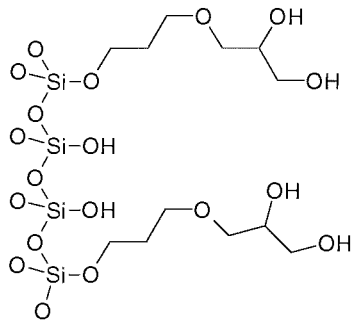


Figure 16. Structures of stationary phases.

With regard to the mobile phase, and as seen in Chapter 1, the moderate critical conditions of carbon dioxide (CO_2), combined with some other favourable properties make it the solvent of choice for SFC.[15] However, the fact that CO_2 is very non-polar results typically in the use of organic modifiers to increase the polarity of the mobile phase and make it suitable for the elution of polar analytes. For the purpose of this study, the most commonly used modifier was selected: methanol (MeOH). As mentioned above, experiments were undertaken using a single mobile phase under isocratic conditions (20 % v/v of MeOH in CO_2).

In order to investigate potential relationships between the analytes' structures and retention, a preliminary study was undertaken on a partial test set consisting of twelve sulfonamides (compounds **1** to **12**). Their retention was characterised using their retention time (t_R) under the applied conditions.

4.1.2. First results.

The analysis of these twelve compounds was undertaken with the aim of highlighting correlations between the analytes' physico-chemical properties and their retention behaviour.

For clarity purposes, in the following paragraphs the term "sulfonyl part" of the test analytes will refer to that substructure of the analyte related to the sulfonyl chloride used in the synthesis reaction, while "amine part" will refer to that substructure of the analyte related to the amine used during the synthesis. Consequently, "benzenesulfonyl derivatives", for instance, refer to those compounds yielded by the reaction of benzenesulfonyl chloride with all the amines, while "*N*-ethylcyclohexylamine derivatives" refer to those compounds obtained by coupling all sulfonyl chlorides with *N*-ethylcyclohexylamine. (Figure 17)

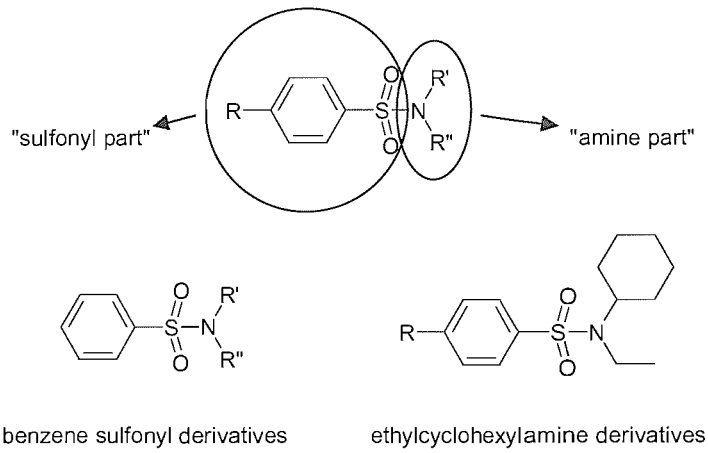


Figure 17. "Sulfonyl part" and "amine part" in the test analytes.

Under the conditions described above, using 2-EP and CN stationary phases, simple trends were identified linking the structures of the analytes to their retention time. With regard to the sulfonyl part of the compounds, 4-(trifluoromethyl)benzenesulfonyl derivatives are less retained than the non-substituted benzenesulfonyl derivatives which, in turn, exhibit shorter retention times than the biphenyl-4-sulfonyl derivatives. 4-aminobenzenesulfonyl derivatives are eluted last. As far as the amine part of the molecule is concerned, retention time increases from *N*-ethylcyclohexylamine derivatives to *N*-(diphenylmethyl)methylamine derivatives; benzimidazole derivatives showing intermediate retention (Figure 18).

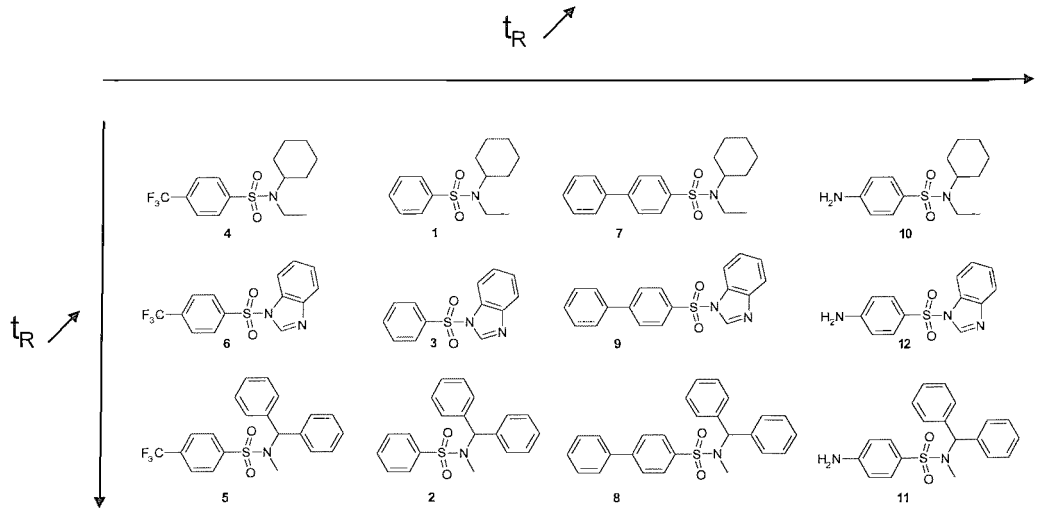


Figure 18. Evolution of retention times according to structural features. 2-EP and CN columns (4.6 x 250 mm, pore size 60 Å, particle size 6 μm), isocratic elution 20 % MeOH in CO_2, 35°C, 100 bar, 4 mL min^{-1}.

The late elution of *N*-(diphenylmethyl)methylamine derivatives using 2-EP and CN columns could be explained by π-π interactions between the two benzene rings of the analytes and the pyridine ring/nitrile group of the stationary phase. This would be consistent with the early elution of trifluoromethylbenzenesulfonyl derivatives (in which π-electrons are pulled from the benzene ring by the electron-withdrawing CF_3 group) and the later elution of biphenyl-4-sulfonyl derivatives with regard to unsubstituted benzenesulfonyl derivatives. In the case of 4-aminobenzene derivatives, the action of the NH_2 group, both as electron-donating group enriching the benzene ring in electrons and as hydrogen bond donor and acceptor, would explain the increased retention. These effects could have been further investigated by synthesizing 4-methoxybenzene sulfonyl derivatives, however this could not be undertaken within timescale of the project.

When using the same chromatographic conditions with the diol stationary phase, the same trend was observed with regards to the sulfonyl part of the structures. However, an inversion of the order of elution was observed for benzimidazole and *N*-(diphenylmethyl)methylamine derivatives: the former being eluted last.[iv] (Figure 19)

[iv] The exact same trends were subsequently observed when using a bare silica (Si) stationary phase.

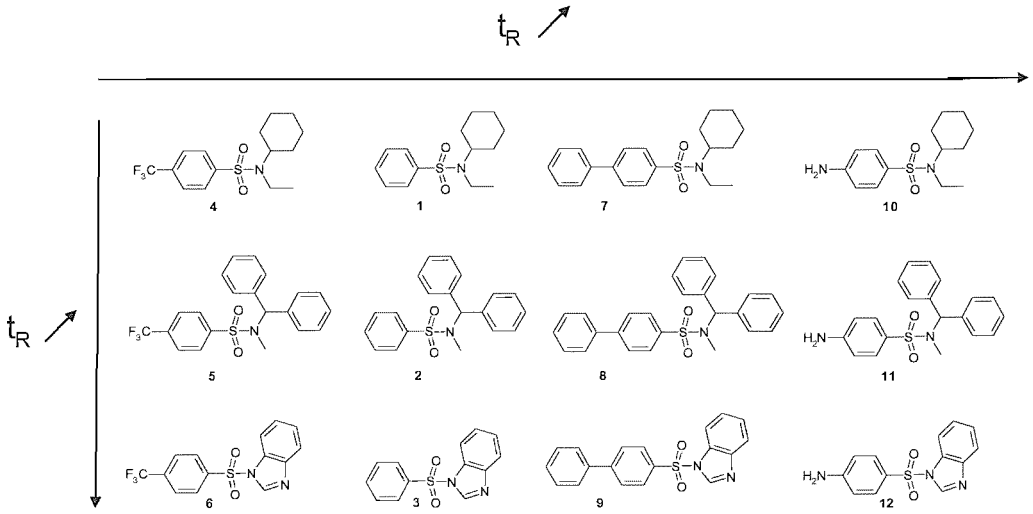


Figure 19. Evolution of retention times according to structural features. Diol column (4.6 x 250 mm, pore size 60 Å, particle size 6 μm), isocratic elution 20 % MeOH in CO_2, 35°C, 100 bar, 4mL min^{-1}.

The presence of numerous hydroxyl groups in the diol stationary phase would be expected to make hydrogen bonding an important factor in the retention mechanism. This would explain the late elution of benzimidazole derivatives; these can form hydrogen bonds with the stationary phase through the extra nitrogen atom of the benzimidazole ring. Coincidently, the absence of aromatic or other π-electron-rich features in the stationary phase deprives the *N*-(diphenylmethyl)methylamine derivatives of an important retentive interaction, resulting is their shorter retention than the benzimidazole derivatives. Nevertheless, this balance between aromatic and hydrogen bonding interactions does not explain the conservation of the relative elution order of trifluoromethylbenzenesulfonyl, benzenesulfonyl and biphenyl-4-sulfonyl derivatives. [v]

A comparison of the retention times of the twelve test analytes using 2-EP, CN and diol stationary phases is shown in Figure 20. The graph also features the retention times obtained with a bare silica (Si) stationary phase that were subsequently measured.

[v] The same arguments could be used to explain the same trends later observed using the Si stationary phase.

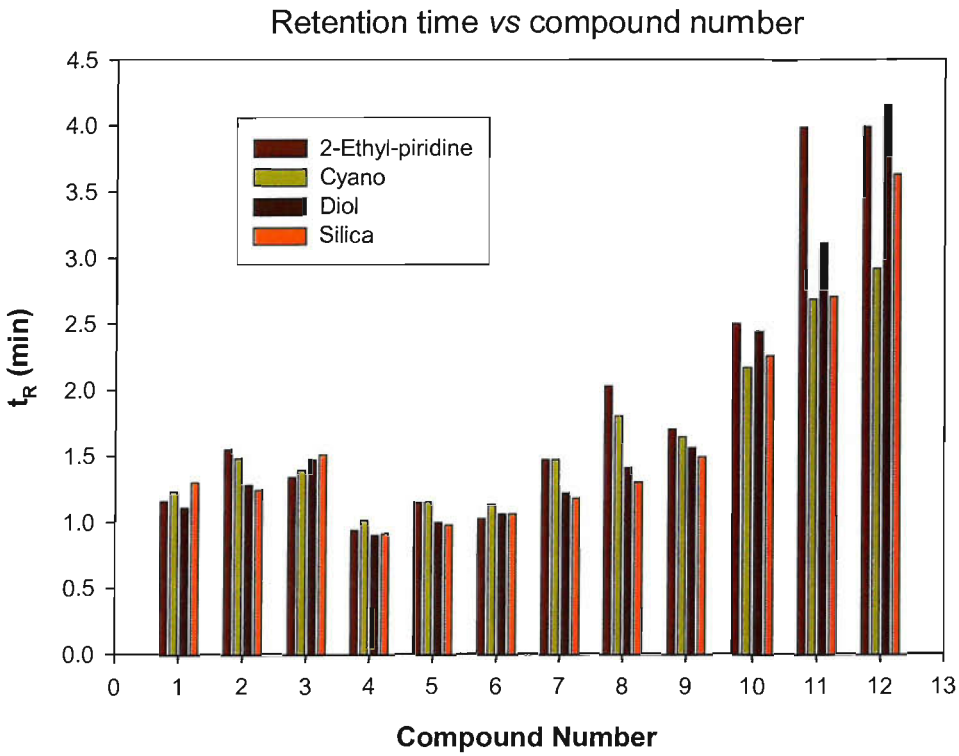


Figure 20. Retention time *vs.* compounds numbers, 2-EP, CN, diol and Si stationary phases (4.6 x 250 mm, pore size 60 Å, particle size 6 μm), isocratic elution 20 % MeOH in CO_2, 35°C, 100 bar, 4mL min^{-1}.

Compounds **8**, **11** and **12** excepted, retention times differ only slightly from one column to another. None of the columns retains the analysed compounds more than the others: some compounds are eluted faster using the 2-EP column while some are less retained on CN, diol or Si stationary phases. The four columns seem to be equally efficient for the analysis of this set of twelve test compounds.

4.1.3. Extension to more compounds.

With the advancement of the syntheses, the isocratic approach was extended to the whole SOTLIB library. The twenty new analytes (compounds **13** to **32**) are more basic than the twelve initially synthesized. Although their elution was possible in the aforementioned conditions using the 2-EP stationary phase, they were much more strongly retained by the CN and diol columns. This prompted the use of a basic additive in the modifier. The effect of additives, added in small quantity into the mobile phase modifier to improve peak shape and decrease retention is discussed in Chapter 5. In the present case, triethylamine (TEA) and

diethylamine (DEA) were evaluated at a concentration of 0.1 % v/v in the modifier (MeOH), both leading to improved peak shapes for tailing compounds and elution of analytes that were not eluted from CN and diol columns in the absence of additive (data not shown). A comparison of the retention times of the twenty test analytes using 2-EP, CN and diol stationary phases is shown in Figure 21.

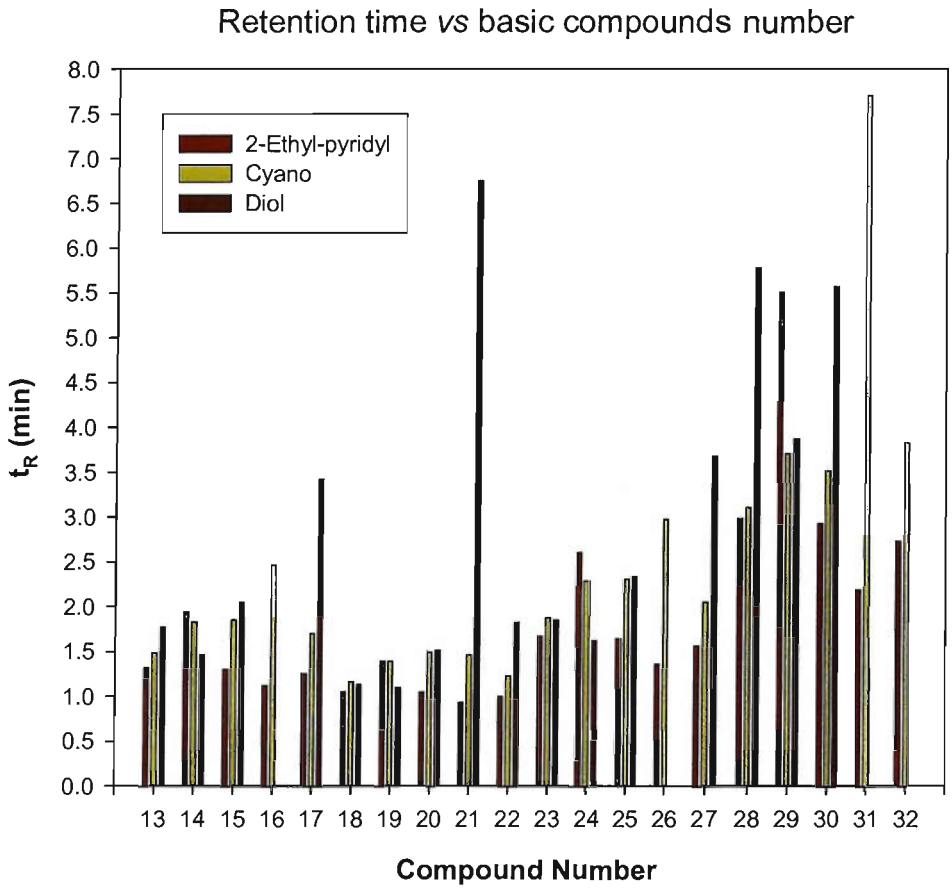


Figure 21. Retention time *vs.* compounds numbers, 2-EP, CN and diol stationary phases (4.6 x 250 mm, pore size 60 Å, particle size 6 µm), isocratic elution 20 % modifier in CO_2, 35°C, 100 bar, 4mL min^{-1}. Modifier: MeOH + 0.1 % v/v TEA.

As previously with the initial pilot study, the retention times of these twenty compounds were investigated with the aim of highlighting trends relating structure and retention. With regard to the sulfonyl part of the analytes, the previously observed trend was also identified on all three stationary phases (Figure 22).

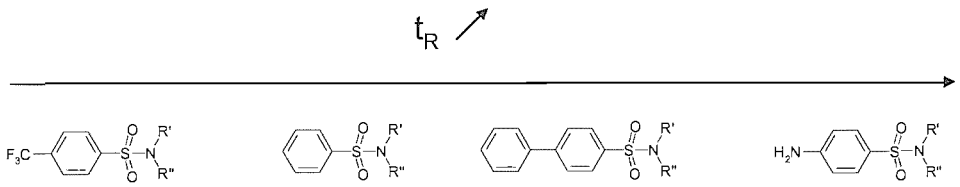


Figure 22. Evolution of retention times according to structural features. 2-EP, CN and diol columns (4.6 x 250 mm, pore size 60 Å, particle size 6 µm), isocratic elution 20 % modifier in CO_2, 35°C, 100 bar, 4mL min^{-1}. Modifier: 0.1 % v/v TEA in MeOH.

In contrast, the relative elution of the compounds could not be ranked with respect to the amine part. The introduction of five more amines afforded a larger diversity of possible interactions between the solutes and the stationary phases, making the simple interpretation of experimental data in terms of molecular features more difficult.

4.1.4. Limitations of the isocratic approach.

Even though these preliminary results were somewhat encouraging, because they highlighted that there must be a relationship of some kind between the properties of the studied compounds and their structural features, they were also insufficient as much that they did not allow identification of any specific relationship. At this point of the study, the only known properties of the test compounds were those that had been calculated to assess the drug-likeness of the SOTLIB compounds, Clog P, RMM and pK_a. The retention times of the analytes did not show any direct correlation with these properties. For instance, the trends identified previously were somewhat weakened by the fact that, using 2-EP stationary phase, compounds **11** and **12** have exactly the same retention time, although their physico-chemical properties are different. Co-elution of those two compounds using the same conditions led to only one observed peak, as shown in Figure 23. Similarly, using the Si column, compounds **3** and **9** have very close retention times (1.51 and 1.49 min respectively) under the conditions used, despite having distinct differences in Clog P and pK_a.

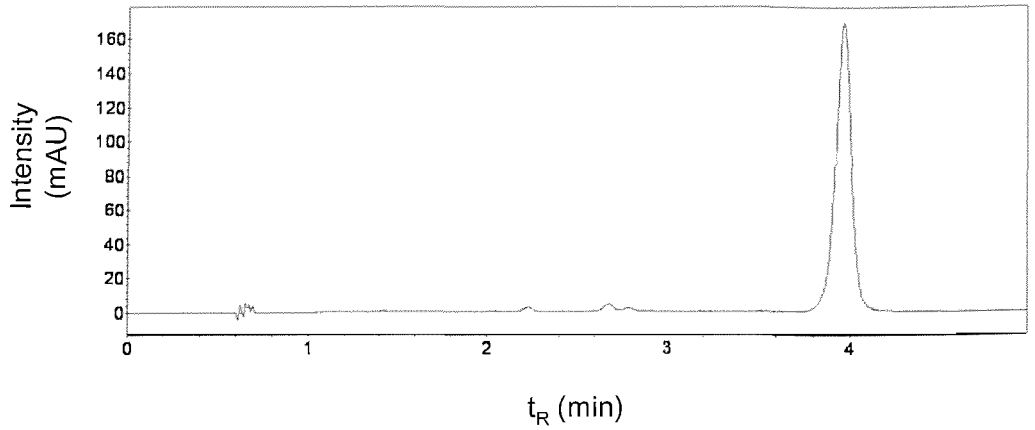


Figure 23. Co-elution of compounds **11** and **12**. 2-EP column (4.6x250 mm, pore size 60 Å, particle size 6 μm), isocratic elution 20 % MeOH in CO_2, 35°C, 100 bar, 4mL min^{-1}.

The identification of a relation between molecular substructures and retention time suggested that a relationship could be established between descriptors of the analyte and their retention time, whilst the observation of co-elution of compounds of distinct physico-chemical properties implied that these properties were not directly related to the retention mechanism. The calculation of more significant molecular descriptors was therefore required. These descriptors had to be suitable for description of the retention process and calculable using available molecular modelling package, hence the choice of Kaliszan's descriptors, described in Chapter 3.

Moreover, the composition of the mobile phase had been arbitrarily chosen to be 20 % modifier (pure MeOH or MeOH and additive) in CO_2. The analysis of samples of diverse structures requires variability in the mobile phase composition. Most often in a pharmaceutical setting, the presence of compounds of very different nature in one sample requires the use of gradient elution, where the proportion of modifier in the mobile phase varies during the analysis.

The use of the retention time as a characteristic of retention can also be considered too simple an approach, as t_R can hardly be related to the thermodynamics of chromatography.

For these reasons, a more systematic approach was adopted for further investigation. SOTLIB compounds were studied at different mobile phase compositions using a "polycratic approach" that allow for the definition of more meaningful retention characteristics.

4.2. Polycratic study.

4.2.1. Definition.

A better parameter was required to characterise the retention of the test analytes. As seen in Chapter 2, the retention factor, k, of a given analyte is directly related to its partition coefficient between mobile and stationary phases, K, which can be in turn related to the interaction forces that are the bases of the retention mechanism. The retention factor, k, was therefore an obvious choice for the characterization of the retention of an analyte.

Successful models of retention have been developed by Snyder[91] and Soczewinski.[92] They showed that the logarithm of retention factor varies linearly with the proportion of modifier, φ, in the mobile phase in reversed phase systems (octadecylsilane (ODS) stationary phase with H_2O/MeOH mobile phase) and over a limited range of mobile phase composition:

$$\log k = \log k_0 + S\varphi \qquad (4.1)$$

where log k_0 is the intercept of the curve log k = f(φ) with the y axis and S is the slope of the curve. Log k_0 represents the extrapolation of the logarithm of retention factor to a hypothetical 0 % modifier in the mobile phase. The physical meaning of this extrapolation is debatable and cannot be considered as the actual value that log k would take if elution of the analyte were to be undertaken in the absence of modifier in the mobile phase. However, log k_0 represents a convenient way of characterising the retention of an analyte and is considered as more reliable than an arbitrarily chosen isocratic log k.[69] Another parameter of interest can be calculated: φ_0, namely the value of φ for which log k equals 0. This parameter can be easily calculated through:

$$\varphi_0 = -\frac{\log k_0}{S} \qquad (4.2)$$

φ_0 is a convenient parameter because it represents a tangible value (the mobile composition for which the retention time t_R of the analyte equals two times the dead time t_0 of the column), whilst log k_0 is an abstract quantity, the value of which is more difficult to interpret directly.

Although SFC is often referred to as a normal phase technique and strong deviation from linearity has been observed in normal phase systems, the linearity of the log k *vs.* φ relationship in SFC was studied to characterise the retention of the test analytes.

The evaluation of the linearity of log $k = f(\varphi)$ was undertaken by means of a polycratic study of the test library. The term "polycratic" means that each analyte is analysed under isocratic conditions, the analysis being undertaken at a number of different mobile compositions so that the curve log $k = f(\varphi)$ can be plotted and, linearity permitting, the retention parameters log k_0, S and φ_0 calculated.

4.2.2. Experimental conditions.

The observations made in the preliminary study described in §4.1 led to slight modifications of the experimental setting.

Firstly, the better results obtained with the 2-EP stationary phase prompted its selection for the polycratic study. The most basic analytes had failed to elute from CN and diol column even in the presence of an additive and so these two stationary phases were deemed unsuitable. Further, the column length was reduced from 250 mm to 50 mm to shorten analysis time and the overall duration of the study. The 50 mm column has a I.D. of 4.6 mm, particle size 5 µm and pore size 60 Å.

Secondly, having restricted the study to one only stationary phase, slightly more diversity was introduced in the mobile phase. Polycratic studies of each analyte were undertaken using three different mobile phases. CO_2 was kept as the main component of the mobile phase and MeOH as the modifier but the latter was used either pure or modified by an additive. Ethyl-dimethyl-amine (EDMA) at a concentration of 0.1 % v/v in MeOH and ammonium acetate (NH_4OAc) at a concentration of 0.6 mM in MeOH were used as additives to improve the peak shape of some of the analytes.

Finally, the linearity of $\log k = f(\varphi)$ was studied over a restricted range of $\log k$, according to practical interest, *i.e.* $0 < \log k < 1$.[93] A negative $\log k$ (*i.e.* $k < 1$) is not desirable because the analyte is almost not retained on the stationary phase, which is of no interest in terms of chromatographic separation. To keep analysis time short and avoid peak broadening, a $\log k$ less than 1 is preferable. Moreover, late elution decreases the sensitivity because the signal is weakened by dilution.

4.2.3. Linear regression analysis.

Experimental values of k for each analyte were acquired in three replicate experiments to ensure repeatability of the data. The mean values of $\log k$ for each percentage φ were calculated. The linearity of the $\log k = f(\varphi)$ relationship for each analyte was assessed by linear regression analysis using SigmaPlot version 10.0 (Systat Software, Inc., Richmond, USA). Regression equations were derived giving values of $\log k_0$ and S. The statistical validity of the results was assessed by calculation of squared correlation coefficients (r^2), standard errors of estimates, significance levels of each term of the equations (P), and values of the t-test of significance (t). φ_0 values were subsequently determined.

4.2.4. Results.

4.2.4.1. Elution within the range of $\log k$.

Initially, it was necessary to assess whether all test compounds could be eluted within the required range of $\log k$. Compounds exhibiting $\log k$ greater than 1 in the conditions initially used (typically $\varphi = 20$ %) could be eluted sooner by increasing φ up to 50 %. However, three analytes (compounds **4**, **6** and **21**,) were omitted from the study because of a negative $\log k$ value in all conditions. Reducing φ to 1 % did not allow for sufficient retention of these analytes. Compounds **18** and **22** (when no additive was used) and compounds **18**, **20** and **22**

(when EDMA or NH_4OAc were used as the additive) could be eluted within the required range only by decreasing φ below 10 %. However, the plots of log k *vs.* φ for those compounds were not linear. Figure 24 shows an example of a non-linear plot as compared with a linear one.

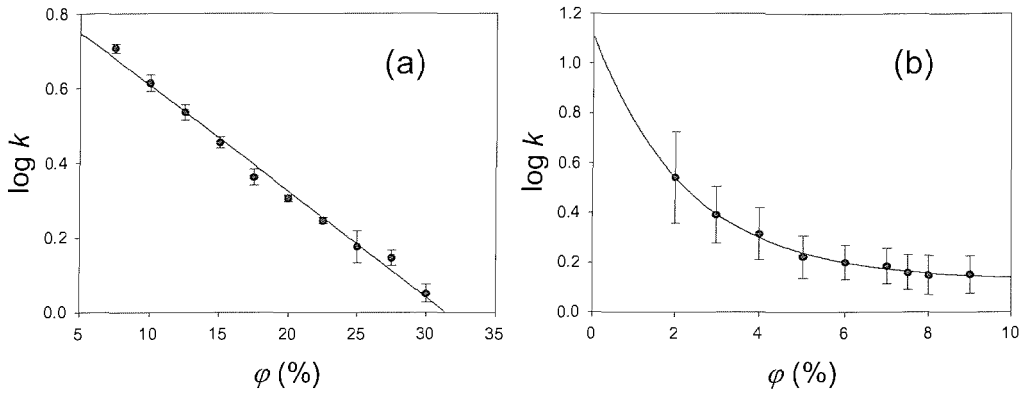


Figure 24. Examples of log k = f(φ) plots. (a) Linear plot, compound **9**, modifier: pure MeOH; (b) Non-linear plot, compound **18**, modifier: pure MeOH.

It appears that using a mobile phase composition lower than 10 % is detrimental to the linearity of log k = f(φ). This can be explained by the dramatic change in the polarity of the mobile phase on the addition of a small quantity of MeOH to the CO_2. All these early-eluting analytes were subsequently eliminated from the study. Such compounds should either be analysed using a less polar modifier (*e.g.* acetonitrile or a chloroform/methanol mix) or under different temperature and pressure conditions, parameters that are known to be of great influence in SFC retention. The remaining compounds of the library exhibited $0 < \log k < 1$ at $\varphi > 10$ %.

4.2.4.2. Peak tailing and peak splitting.

With pure MeOH modifier, and to a lesser extent in the presence of additives, a number of test analytes showed tailing and/or splitting peaks,[vi] examples are shown in Figure 25.

[vi] In the case of peak splitting, the identity of the peaks was confirmed by MS.

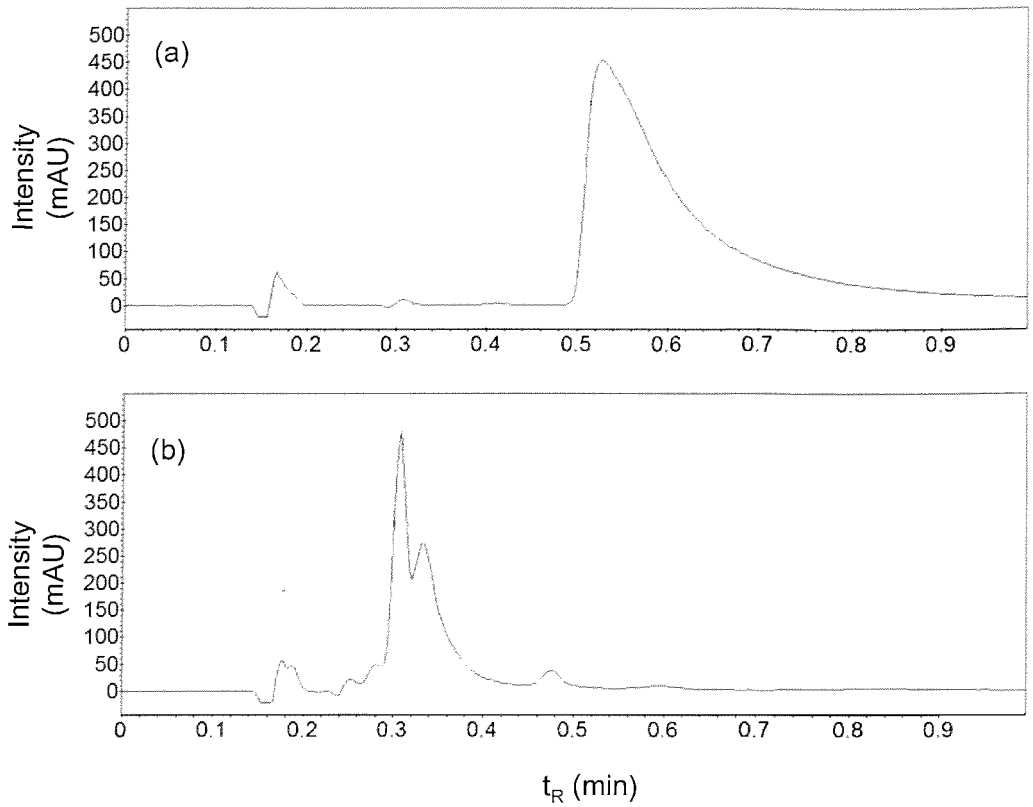


Figure 25. Examples of peak shapes. (a) tailing peak; (b) splitting peak.

Such chromatographic behaviour did not systematically affect the linearity of the variation of log k with φ, but gave rise to greater deviations of the measured values of log k (Figure 26), leading to uncertainty in the calculation of retention characteristics. Moreover, peak splitting may be due to the co-existence of competing retention mechanisms, which could interfere with the aim of establishing a coherent retention model for all compounds of the library. Compounds exhibiting peak tailing and/or splitting were therefore not included in the multiple regression analysis described below.

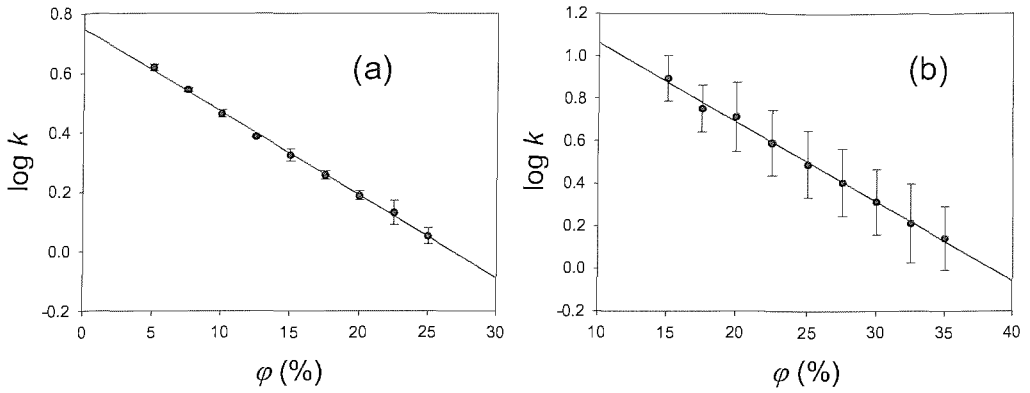


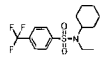
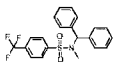
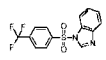
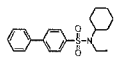
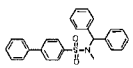
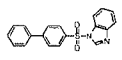
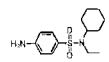
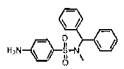
Figure 26. Comparison of standards deviations on measured log *k*. (a) Small standard deviations, compound **2**, symmetric peak shape without additive; (b) large standard deviations, compound **28**, tailing peak without additive.

Removal of several compounds from the original data set is of concern with regards to the method being applied to a wider range of analytes. However, experimental data have shown that working at a higher concentration of additive (typically 5 mM or more of NH_4OAc in MeOH) allows for a dramatic improvement of the peak shapes of tailing compounds without affecting the retention of analytes with symmetric peak shape. This is discussed in Chapter 5.

4.2.4.3. Retention characteristics.

Only sixteen out of the thirty-two test compounds were studied when pure MeOH was used as a modifier. Their retention characteristics, along with the statistical parameters of the regression analysis, are shown in Table 9. Linearity of $\log k = f(\varphi)$ was observed for the sixteen analytes, with $r^2 > 0.98$ and good significance level for the coefficients ($P < 1\times10^{-4}$).

Table 9. Results of polycratic studies of test compounds using pure methanol as modifier.

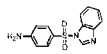
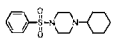
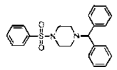
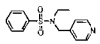
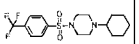
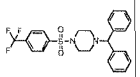
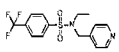
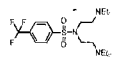
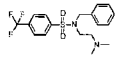
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
1		10	0.2200 (±0.0060) (37)[b]	-0.0116 (±0.0005) (-22)[b]	0.9905	0.0106
2		9	0.7469 (±0.0090) (83)[b]	-0.0279 (±0.006) (-50)[b]	0.9977	0.0186
3		10	0.6693 (±0.0089) (75)[b]	-0.0261 (±0.0006) (-47)[b]	0.9969	0.0188
4[a]		n/a	n/a	n/a	n/a	n/a
5[a]		9	0.3141 (±0.0142) (22)[b]	-0.0189 (±0.0014) (-14)[b]	0.9799	0.0205
6[a]		n/a	n/a	n/a	n/a	n/a
7		9	0.7325 (±0.0077) (95)[b]	-0.0307 (±0.0005) (-59)[b]	0.9984	0.0147
8		9	0.9669 (±0.0118) (82)[b]	-0.0297 (±0.0005) (-58)[b]	0.9977	0.0192
9		10	0.8908 (±0.0140) (64)[b]	-0.0283 (±0.0007) (-41)[b]	0.9959	0.0274
10		9	1.3427 (±0.0189) (71)[b]	-0.0396 (±0.0008) (-49)[b]	0.9975	0.0271
11		9	1.4974 (±0.0257) (58)[b]	-0.0355 (±0.0009) (-39)[b]	0.9973	0.0306

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 9 (continued). Results of polycratic studies of test compounds using pure methanol as modifier.

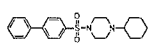
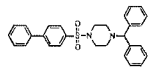
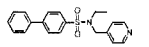
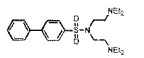
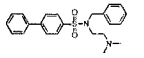
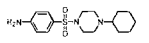
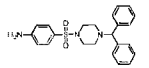
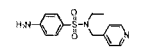
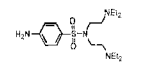
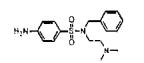
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
12		10	1.5786 (±0.0260) (61)[b]	-0.0362 (±0.0009) (-41)[b]	0.9969	0.0345
13[a]		10	0.7231 (±0.0412) (18)[b]	-0.0297 (±0.0026) (-12)[b]	0.9459	0.0804
14		10	0.8814 (±0.0179) (49)[b]	-0.0280 (±0.0008) (-35)[b]	0.9925	0.0314
15		8	0.7354 (±0.0201) (37)[b]	-0.0350 (±0.0013) (-28)[b]	0.9929	0.0195
16[a]		6	1.5608 (±0.1424) (11)[b]	-0.0736 (±0.0102) (-7)[b]	0.9583	0.1339
17[a]		10	1.0015 (±0.1274) (8)[b]	-0.0359 (±0.0086) (-4, 3×10^{-4})[b]	0.9045	0.1688
18[a]		n/a	n/a	n/a	n/a	n/a
19[a]		8	0.6008 (±0.0313) (19)[b]	-0.0258 (±0.0018) (-15)[b]	0.9692	0.0353
20		8	0.4784 (±0.0111) (43)[b]	-0.0321 (±0.0013) (-25)[b]	0.9948	0.0178
21[a]		n/a	n/a	n/a	n/a	n/a
22[a]		n/a	n/a	n/a	n/a	n/a

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 9 (continued). Results of polycratic studies of test compounds using pure methanol as modifier.

Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
23[a]		10	0.9508 (±0.0578) (16)[b]	-0.0310 (±0.0028) (-11)[b]	0.9523	0.087
24		10	1.1619 (±0.0161) (72)[b]	-0.0329 (±0.0007) (-46)[b]	0.9972	0.0283
25		9	0.9909 (±0.0135) (74)[b]	-0.0384 (±0.0007) (-57)[b]	0.9969	0.0191
26[a]		9	3.2533 (±0.1641) (20)[b]	-0.0765 (v0.0045) (-17)[b]	0.9882	0.0610
27[a]		8	1.1097 (±0.2204) (5)[b]	-0.0286 (±0.0109) (3, 1.52×10^{-2})	0.8786	0.2149
28[a]		9	1.4387 (±0.0994) (14)[b]	-0.0374 (±0.0038) (-10)[b]	0.9528	0.1291
29		10	1.4537 (±0.0279) (52)[b]	-0.0331 (±0.0009) (-38)[b]	0.9953	0.0343
30		8	1.4100 (±0.0246) (57)[b]	-0.0428 (±0.0010) (-43)[b]	0.9966	0.0282
31[a,c]		n/a	n/a	n/a	n/a	n/a
32[a]		5	1.7457 (±0.3402) (5, 2×10^{-4})[b]	-0.0421 (±0.0131) (-3, 6.6×10^{-3})[b]	0.8882	0.27

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

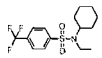
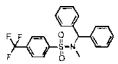
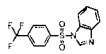
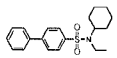
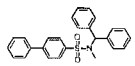
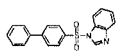
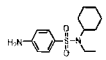
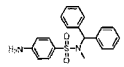
[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

[c] Compound **31** exhibited strong tailing preventing reliable retention measurement.

The introduction of an additive in the modifier greatly improved the peak shapes of several of the tailing/splitting analytes. As a consequence, in the presence of EDMA and NH_4OAc, twenty-three and twenty-one of the thirty-two compounds respectively were subjected to regression analysis and their retention parameters calculated. Results of the polycratic studies with EDMA and NH_4OAc are shown in Table 10 and 11 respectively. Again $\log k = f(\varphi)$ proved linear with high r^2 and good significance levels P.

Table 10. Results of polycratic studies of test compounds using EDMA as an additive in the modifier.

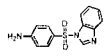
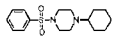
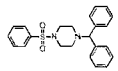
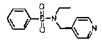
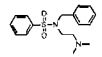
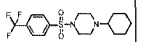
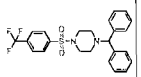
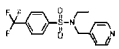
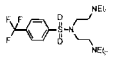
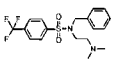
Compounds	Structures	n	log k_θ (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
1[a]		10	0.3509 (±0.0111) (32)[b]	-0.0174 (±0.0011) (-15)[b]	0.9948	0.0145
2		9	0.7270 (±0.0081) (89)[b]	-0.0259 (±0.005) (-57)[b]	0.9978	0.0177
3		10	0.6344 (±0.0076) (83)[b]	-0.0247 (±0.0005) (-53)[b]	0.9978	0.0155
4[a]		n/a	n/a	n/a	n/a	n/a
5[a]		9	0.2874 (±0.0134) (21)[b]	-0.0155 (±0.0015) (-11)[b]	0.9905	0.0159
6[a]		n/a	n/a	n/a	n/a	n/a
7		9	0.7360 (±0.0066) (111)[b]	-0.0299 (±0.0004) (-67)[b]	0.999	0.0125
8		9	0.9591 (±0.0169) (57)[b]	-0.0290 (±0.0007) (-40)[b]	0.9953	0.0276
9		10	0.8799 (±0.0124) (71)[b]	-0.0284 (±0.0006) (-49)[b]	0.9972	0.0202
10		9	1.3137 (±0.0221) (59)[b]	-0.0394 (0.0009) (-43)[b]	0.9967	0.0279
11		9	1.4838 (±0.0209) (71)[b]	-0.0363 (±0.0007) (-49)[b]	0.998	0.0249

Regression analysis of log k against φ. n: number of data points; log k_θ and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 10 (continued). Results of polycratic studies of test compounds using EDMA as an additive in the modifier.

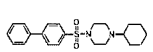
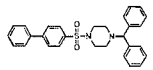
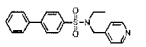
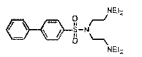
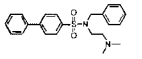
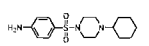
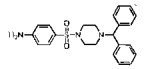
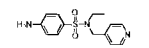
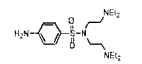
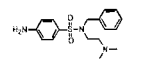
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
12		10	1.5012 (±0.0252) (60)[b]	-0.0352 (±0.0009) (-41)[b]	0.9966	0.0334
13		8	0.6256 (±0.0110) (57)[b]	-0.0297 (±0.0008) (-39)[b]	0.9955	0.0194
14		9	0.9052 (±0.0146) (62)[b]	-0.0295 (±0.0007) (-42)[b]	0.9963	0.0234
15		7	0.7703 (±0.00500) (154)[b]	-0.0385 (±0.0003) (-118)[b]	0.9996	0.0045
16		10	0.6851 (±0.0112) (61)[b]	-0.0361 (±0.0009) (-40)[b]	0.9966	0.0171
17		10	0.6205 (±0.0148) (42)[b]	-0.0298 (±0.0010) (-29)[b]	0.9926	0.0196
18[a]		n/a	n/a	n/a	n/a	n/a
19		9	0.5600 (±0.0187) (30)[b]	-0.0238 (±0.0010) (-23)[b]	0.9831	0.0231
20[a]		n/a	n/a	n/a	n/a	n/a
21[a]		n/a	n/a	n/a	n/a	n/a
22[a]		n/a	n/a	n/a	n/a	n/a

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were uot taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 10 (continued). Results of polycratic studies of test compounds using EDMA as an additive in the modifier.

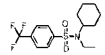
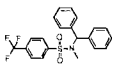
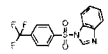
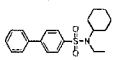
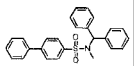
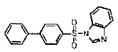
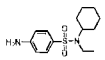
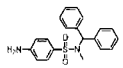
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
23		8	0.8732 (±0.0112) (78)[b]	-0.0332 (±0.0006) (-59)[b]	0.9976	0.0158
24		10	1.1720 (±0.0121) (97)[b]	-0.0338 (±0.0005) (-63)[b]	0.9984	0.0212
25		9	0.9909 (±0.01350) (73)[b]	-0.0384 (±0.0007) (-57)[b]	0.9969	0.0191
26		10	2.0957 (±0.0343) (61)[b]	-0.0521 (±0.0343)(-54)[b]	0.9973	0.0161
27		10	0.9112 (±0.0111) (82)[b]	-0.0362 (±0.0006) (-63)[b]	0.9976	0.0148
28		9	1.3975 (±0.0218) (64)[b]	-0.0437 (±0.0009) (-47)[b]	0.9964	0.0313
29		10	1.4422 (±0.0216) (67)[b]	-0.0338 (±0.0007) (-50)[b]	0.9968	0.0265
30		7	1.4016 (±0.0199) (71)[b]	-0.0448 (±0.0009) (-52)[b]	0.9983	0.0198
31[a]		8	1.5460 (±0.027) (57)[b]	-0.0516 (±0.0012) (-42)[b]	0.9963	0.0344
32		8	1.4499 (±0.0232) (63)[b]	-0.0462 (±0.0011) (-44)[b]	0.9972	0.0296

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 11. Results of polycratic studies of test compounds using NH_4OAc as an additive in the modifier.

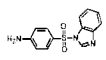
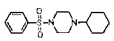
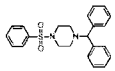
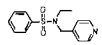
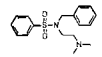
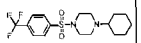
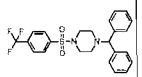
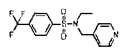
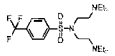
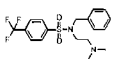
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
1[a]		9	0.2906 (±0.0118) (25)[b]	-0.0138 (±0.0011) (-12)[b]	0.9905	0.0106
2		9	0.7270 (±0.0081) (89)[b]	-0.0259 (±0.0005) (-57)[b]	0.9978	0.0177
3		9	0.5656 (±0.0104) (55)[b]	-0.0220 (±0.0007) (-31)[b]	0.9957	0.0198
4[a]		n/a	n/a	n/a	n/a	n/a
5[a]		10	0.2527 (±0.0116) (22)[b]	-0.0144 (±0.0011) (-13)[b]	0.976	0.0182
6[a]		n/a	n/a	n/a	n/a	n/a
7		10	0.6555 (±0.0073) (90)[b]	-0.0248 (±0.0005) (-54)[b]	0.9979	0.0154
8		10	0.9466 (±0.0113) (84)[b]	-0.0284 (±0.0005) (-56)[b]	0.9977	0.0199
9		10	0.8341 (±0.0093) (90)[b]	-0.0269 (±0.0005) (-58)[b]	0.9979	0.0182
10		10	1.3198 (±0.0200) (66)[b]	-0.0406 (±0.0009) (-46)[b]	0.9961	0.035
11		10	1.4800 (±0.0230) (64)[b]	-0.0366 (±0.0008) (-43)[b]	0.997	0.0332

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 11 (continued). Results of polycratic studies of test compounds using NH_4OAc as an additive in the modifier.

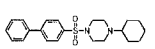
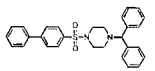
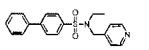
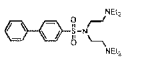
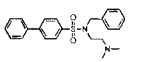
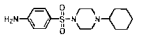
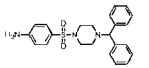
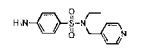
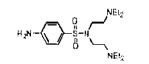
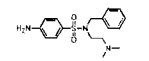
Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
12		10	1.4638 (±0.0230) (64)[b]	-0.0355 (±0.0008) (-46)[b]	0.9967	0.0305
13		9	0.6523 (±0.0087) (75)[b]	-0.0302 (±0.0006) (-53)[b]	0.9969	0.0162
14		10	0.8828 (±0.0140) (63)[b]	-0.0278 (±0.0006) (-45)[b]	0.9958	0.0231
15		9	0.7703 (±0.0050) (154)[b]	-0.0385 (±0.0003) (-118)[b]	0.9996	0.0045
16[a]		7	0.8441 (0.0172) (50)[b]	-0.0375 (±0.0012) (-30)[b]	0.9985	0.0142
17		10	0.6586 (±0.0198) (33)[b]	-0.0290 (±0.0013) (-22)[b]	0.9908	0.0262
18[a]		n/a	n/a	n/a	n/a	n/a
19		10	0.5223 (±0.0165) (32)[b]	-0.0211 (±0.0009) (-25)[b]	0.9832	0.0219
20[a]		n/a	n/a	n/a	n/a	n/a
21[a]		n/a	n/a	n/a	n/a	n/a
22[a]		n/a	n/a	n/a	n/a	n/a

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

Table 11 (continued). Results of polycratic studies of test compounds using NH_4OAc as an additive in the modifier.

Compounds	Structures	n	log k_0 (±Std. errors) (t, P)[b]	S (±Std. errors) (t, P)[b]	r^2	Std. errors of estimates
23		9	0.8856 (±0.0155) (57)[b]	-0.0329 (±0.0008) (-42)[b]	0.9955	0.0219
24		10	1.1623 (±0.0176) (66)[b]	-0.0326 (±0.0008) (-42)[b]	0.9968	0.0309
25		9	0.9735 (±0.0157) (62)[b]	-0.0364 (±0.0008) (-46)[b]	0.996	0.0223
26[a]		10	2.1188 (±0.1009) (21)[b]	-0.0490 (±0.0028) (-17)[b]	0.9879	0.0473
27		10	0.9317 (±0.0190) (49)[b]	-0.0339 (±0.0010) (-34)[b]	0.9948	0.0253
28		10	1.3816 (±0.0324) (43)[b]	-0.0405 (±0.0013) (-31)[b]	0.9906	0.0513
29		10	1.4671 (±0.0275) (53)[b]	-0.0335 (±0.0009) (-39)[b]	0.9955	0.0337
30		8	1.4301 (±0.0289) (50)[b]	-0.0435 (±0.0012) (-37)[b]	0.9954	0.0332
31[a]		8	1.5730 (±0.0970) (16)[b]	-0.0453 (±0.0040) (-11)[b]	0.9651	0.1114
32		10	1.4436 (±0.0380) (38)[b]	-0.0407 (±0.0015) (-27)[b]	0.9893	0.0603

Regression analysis of log k against φ. n: number of data points; log k_0 and S: regression coefficients of equation (4.1); t: values of the t-test of significance; P: significance level of each term of the equations; r^2: squared correlation coefficient.

[a] Compounds were not taken into account in the multiple regression against molecular descriptors.

[b] Where not given, significance P was less than 1×10^{-4}.

4.3. Correlation of retention characteristics with molecular descriptors.

4.3.1. Multiple regression analysis.

Log k_0 and φ_0 were regressed against the three calculated molecular descriptors obtained with Spartan'02 (dipole moment, μ, electron excess charge of the most negatively charged atom, δ_{min}, and molecular surface area, A) to derive model QSRR equations of the type:

$$X = a\mu + bA + c\delta_{\min} + d \qquad (4.3)$$

where X is the studied retention characteristics and a, b, c and d are regression coefficients characteristic of the system.

Multiple regression analysis equations were derived using Microsoft Excel 2003 (Microsoft, Redmond, USA) and the statistical validity of the results was assessed by calculation of multiple correlation coefficients (R), standard errors of estimate (s), significance levels of each term of the whole equations (P) and values of the F-test of significance (F).

4.3.2. Results.

Multiple regression analysis was used to evaluate how measured retention characteristics correlate with molecular descriptors of the test analytes and to derive the coefficients of equation (4.3). One equation describing the retention was obtained for each modifier.

When no additive was used, log k_0 and φ_0 were seen to correlate with the descriptors (multiple correlation coefficients R of 0.90 and 0.83, respectively; significance level $P < 4\times10^{-3}$) while S showed no correlation with calculated μ, A and δ_{min} ($R = 0.51$, $P = 0.321$). With the addition of EDMA in the modifier, correlation was observed between log k_0, S and φ_0 and the descriptors ($R = 0.92$, 0.85 and 0.82, respectively; $P < 2\times10^{-4}$). The best correlations of log k_0 and φ_0 with

the descriptors were obtained with the ammonium acetate additive (R = 0.94 and 0.87 respectively; $P < 4 \times 10^{-5}$). In the presence of NH_4OAc, S also showed correlation with the descriptors as compared with the results obtained in the absence of additive (R = 0.82, $P = 7 \times 10^{-4}$). Values of the coefficients and statistical parameters are shown in Table 12.

Table 12. Results of multiple regression analysis with pure MeOH, MeOH + 0.1 % EDMA and MeOH + 0.6 mM NH_4OAc as modifier.

Modifier	n	a	b	c	d	*R*	*s*	*P*	*F*
MeOH	16	Regression of log k_0 against μ, A, δ_{min} ($\log k_0 = a\mu + bA + c\delta_{min} + d$)							
		0.108	0.002	-1.26	-1.31	0.898	0.157	3.0×10^{-4}	15.2
		Regression of S against μ, A, δ_{min} ($S = a\mu + bA + c\delta_{min} + d$)							
		-0.0004	9.5×10^{-7}	0.011	-0.021	0.513	0.004	0.321	1.31
		Regression of φ_0 against μ, A, δ_{min} ($\varphi_0 = a\mu + bA + c\delta_{min} + d$)							
		2.81	0.056	-28.7	-29.4	0.829	5.32	0.004	8.06
MeOH + 0.1 % EDMA	23	Regression of log k_0 against μ, A, δ_{min} ($\log k_0 = a\mu + bA + c\delta_{min} + d$)							
		0.029	0.001	-1.68	-1.06	0.923	0.136	4.6×10^{-8}	36.2
		Regression of S against μ, A, δ_{min} ($S = a\mu + bA + c\delta_{min} + d$)							
		0.003	-2.2×10^{-5}	0.032	-0.013	0.856	0.004	1.1×10^{-5}	17.4
		Regression of φ_0 against μ, A, δ_{min} ($\varphi_0 = a\mu + bA + c\delta_{min} + d$)							
		3.06	0.025	-22.8	-15.7	0.815	4.31	9.7×10^{-5}	12.5
MeOH + 0.6 mM NH_4OAc	21	Regression of log k_0 against μ, A, δ_{min} ($\log k_0 = a\mu + bA + c\delta_{min} + d$)							
		0.014	0.001	-1.64	-0.909	0.946	0.113	1.5×10^{-8}	48.5
		Regression of S against μ, A, δ_{min} ($S = a\mu + bA + c\delta_{min} + d$)							
		0.001	3.9×10^{-6}	0.025	-0.019	0.805	0.004	4.0×10^{-4}	10.4
		Regression of φ_0 against μ, A, δ_{min} ($\varphi_0 = a\mu + bA + c\delta_{min} + d$)							
		1.82	0.045	-26.5	-17.8	0.876	3.38	1.2×10^{-5}	18.7

Measured retention characteristics against molecular descriptors μ, A and δ_{min}. n: number of studied compounds; a, b, c and d: regression coefficients of equation (4.3); R: multiple correlation coefficient; s: standard error of estimate; P: significance level; and F: F-test of significance.

In order to better visualise the results, predicted values of the retention characteristics were calculated, using the derived equation and plotted against experimental values; the resultant graphs are shown in Figure 27.

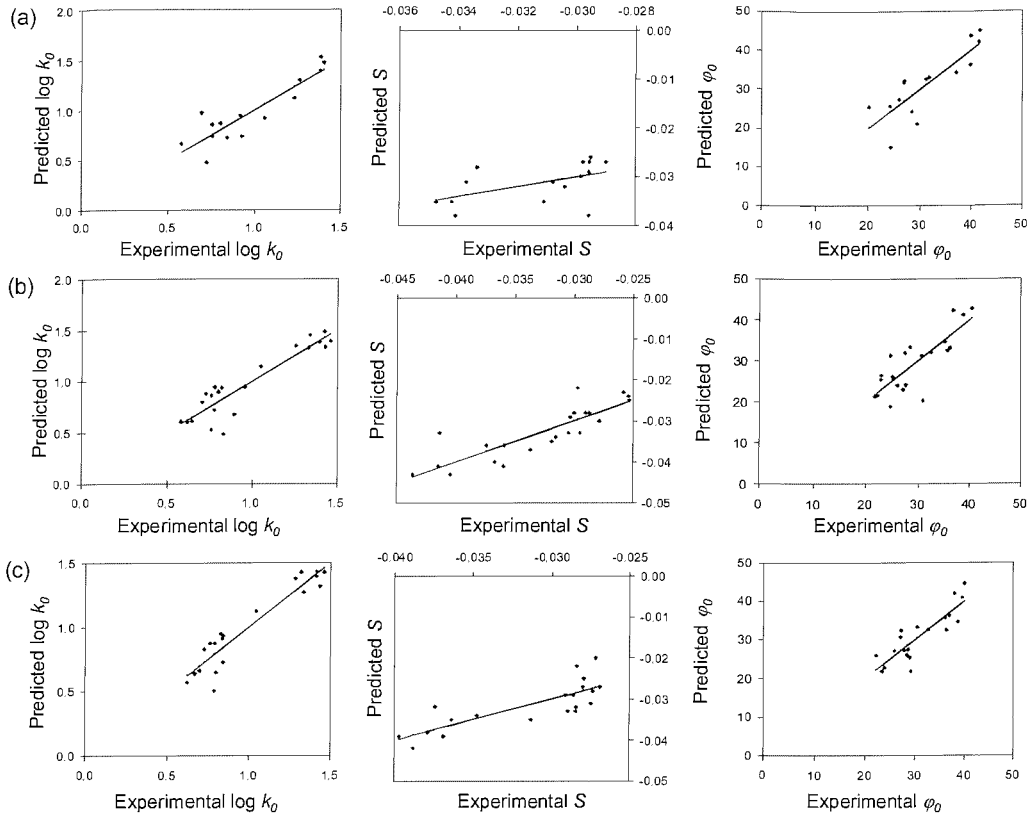


Figure 27. Plots of experimental values of log k_0, S and φ_0 *vs.* their predicted values. (a) Modifier: pure MeOH; (b) Modifier: 0.1 % EDMA in MeOH; (c) Modifier: 0.6 mM NH_4OAc in MeOH

In Figure 27a and 27c, a possible clustering of data points could be interpreted, especially in the case of S and to a lesser extent for log k_0. This could be seen as a sign that the correlation is not genuine but rather due to chance. However, when a larger data set was studied (*i.e.* with EDMA as additive, Figure 27b), data points are distributed more evenly along the best-fit line, suggesting a genuine correlation. In the case of φ_0 no clustering was observed.

Because φ_0 is the easiest parameter to interpret, it is of interest to compare predicted and experimental values on a bar chart (Figure 28). For most of the compounds, the predicted φ_0 differs from the experimental value by 5 % modifier or less.

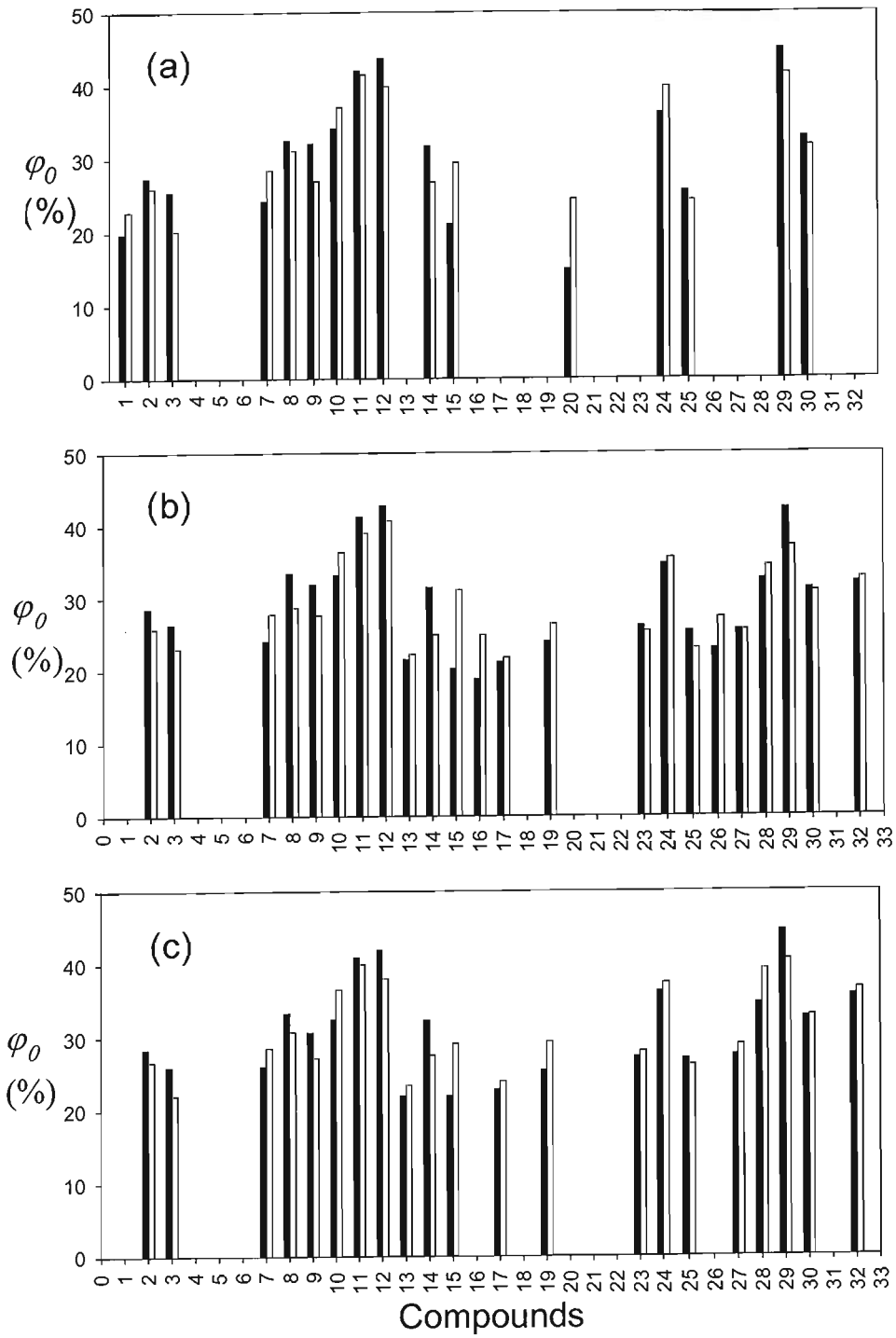


Figure 28. Bar charts comparison of experimental (black) and predicted (white) values of φ_0. (a) Modifier: pure MeOH; (b) Modifier: 0.1 % EDMA in MeOH; (c) Modifier: 0.6 mM NH_4OAc in MeOH.

4.3.3. Extension to gradient elution.

This study has been undertaken under isocratic conditions. Isocratic methods are often used in pharmaceutical analysis in diverse stages of the drug development.[94] However, real samples presenting a diversity of chemicals structures may necessitate a gradient elution, hence the importance of being able to predict gradient retention times. Although true retention factors cannot strictly be measured under gradient conditions, "apparent" retention factors *k'* can be measured and used to quantify retention.

Apparent retention factors *k'* of the test analytes were measured using a 5 to 50 % modifier gradient elution. Other conditions were kept identical to polycratic study conditions: 100 bar, 35°C, 4 mL min^{-1}, 4 µL injection volume.

Subsequently, experimental log *k'* were regressed against experimental log k_0. It appeared that log *k'* is well correlated with log k_0 ($r > 0.92$, $P < 3.8 \times 10^{-8}$, $F > 79$, Table 13). As a consequence, a model predicting log k_0 would be suitable for the prediction of gradient elution log *k'*, since the latter can be deduced from the former. Prediction of log *k'* can be achieved by using the same model as the one described herein.

Table 13. Regression of gradient elution log *k'* against polycratic log k_0, with pure MeOH, MeOH + 0.1 % EDMA and MeOH + 0.6 mM NH_4OAc as modifier.

Modifier	a	b	*r*	*s*	*P*	*F*
MeOH	$\log k' = a \log k_0 + b$					
	0.572	0.167	0.922	0.085	3.8×10^{-8}	79.5
MeOH + 0.1 % EDMA	$\log k' = a \log k_0 + b$					
	0.453	0.306	0.955	0.048	5.5×10^{-12}	206
MeOH + 0.6 mM NH_4OAc	$\log k' = a \log k_0 + b$					
	0.507	0.245	0.991	0.023	5.5×10^{-17}	939

a and b: regression coefficients in log *k'* = a log k_0 + b; *r*: correlation coefficient; *s*: standard error of estimate; *P*: significance level; and *F*: *F*-test of significance.

4.4. Conclusion.

The polycratic study of a library of thirty-two sulfonamides was undertaken using a 2-EP stationary phase with three different modifiers: pure MeOH, MeOH

with 0.1 % EDMA and MeOH with 0.6 mM NH_4OAc. The aim was to assess the linearity of the relationship between the logarithm of retention factor, log k, and the proportion of modifier, φ, in the mobile phase. Over the range $0 < \log k < 1$, and providing that φ was kept above 10 %, this relationship was found to be linear for the studied compounds.

The linearity of $\log k = f(\varphi)$ permitted the computation of intercepts $\log k_0$ and slope S of the curves and the calculation of φ_0. These quantities are characteristic of the retention of each analyte. They could subsequently be correlated, by means of multiple regression analysis, with calculated molecular descriptors introduced by Kaliszan: namely μ, A and δ_{min}. Predicted values calculated using the derived equations are in good agreement with experimental values. Whilst calculation of the descriptors using more sophisticated computation techniques may give slightly better results, the techniques used in this study have the advantage of being readily achieved using a simple PC, whereas a more complicated algorithm would be more costly in terms of hardware and time.

The extension of the prediction capability of the model to gradient elution retention times was highlighted by showing that polycratic $\log k_0$ is correlated with gradient elution $\log k'$. Were the model to be validated, this means that it would allow for the prediction of both isocratic and gradient retention times.

The scope of the model is somewhat restricted by the fact that some of the initial compounds had to be withdrawn from the study due to peak tailing or splitting. It has been shown that the use of ammonium acetate as an additive improved the peak shapes of tailing compounds. The effect of higher concentration of this additive on the retention of test analytes was investigated, as described in the next chapter.

Chapter 5.

The Effect of Increasing Concentration of Ammonium Acetate as an Additive in SFC

5.1. Introduction.

As seen in the previous chapter, the use of a modifier in the mobile phase in SFC is not always sufficient to afford elution of very polar or basic compounds, *e.g.* amines. Most often, a third component (the additive) is added to the mobile phase. It is often an acidic or basic compound typically added in a small amount (<1 % v/v) in order to increase efficiency of separation and obtain symmetrical, well-defined peaks.[17] Additives can be chosen according to the nature of the analytes of interest. Elution of carboxylic acids, for instance, will be improved by addition of trifluoroacetic acid (TFA) or citric acid, whilst the elution of bases will be improved by using aliphatic amines, *e.g.* isopropylamine (IPA), diethylamine (DEA), ethyl-dimethyl-amine (EDMA) or triethylamine (TEA).[18] The use of volatile ammonium salts, such as ammonium acetate, has been recently studied by Pinkston and co-workers.[95-97] The main advantage of ammonium salts additives over basic and acidic additives are their better compatibility with mass spectrometric detection. While acidic and basic additives commonly lead to ion suppression, no such phenomenon is observed with ammonium salts.

The mechanism by which these additives influence retention is not definitively understood. Two mechanisms seem to be involved: (i) the deactivation of active silanol groups on the stationary phase and (ii) an ion-pair interaction between the analyte and the additive.

The effects on retention and peak shape of increasing concentrations of ammonium acetate in the mobile phase have been studied on three stationary

phases: bare silica (Si), 2-ethyl-pyridine (2-EP) and endcapped 2-ethyl-pyridine (2-EP-EC). The aim was to investigate the mechanism through which ammonium acetate (NH_4OAc) modifies retention of the analytes. The hypothesis of an ion-pairing mechanism has been tested by evaluating the effect of the absence of the additive in the mobile phase compared with its presence in the sample solvent.

For the purpose of the study, four sulfonamides were chosen from the SOTLIB library: sulfonamides **16**, **26**, **31** and **32**. In addition to these four compounds, three widely used drug molecules have been included in the study: (*S*)-(+)-naproxen, (*2S,3S*)-(+)-*cis*-diltiazem hydrochloride and (±)-atenolol. Structures of the probe analytes are shown in Figure 29.

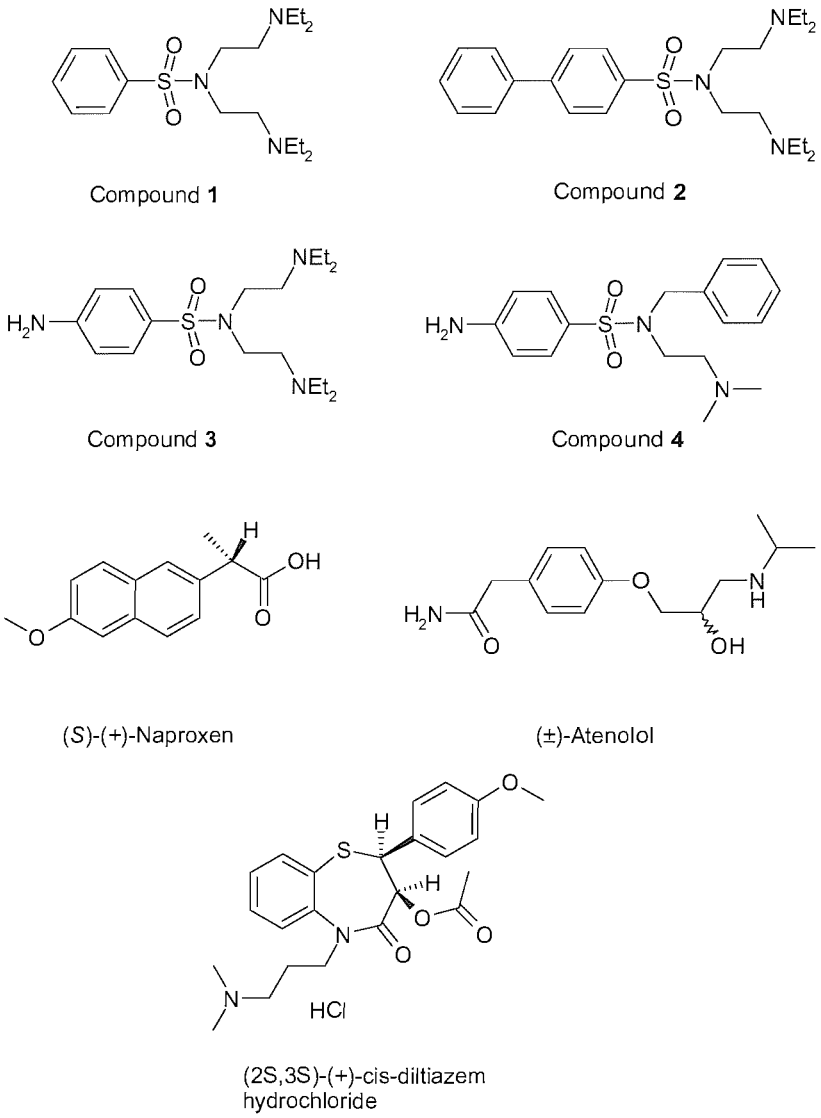


Figure 29. Structures of probe analytes.

Analytes were chosen to exhibit distinctive behaviour on the 2-EP stationary phase, *e.g.* compound **31** and atenolol exhibited badly tailing peak shapes in the absence of additive, making them almost invisible on the UV trace at the used concentrations, while naproxen showed a symmetric peak shape in the presence or absence of NH_4OAc. Incidentally, naproxen is also the only one of the test compounds to feature an acidic functionality. Other compounds showed intermediate behaviours, *i.e.* slight to pronounced tailing in the absence of additive. Examples of these peak shapes are shown in Figure 30.

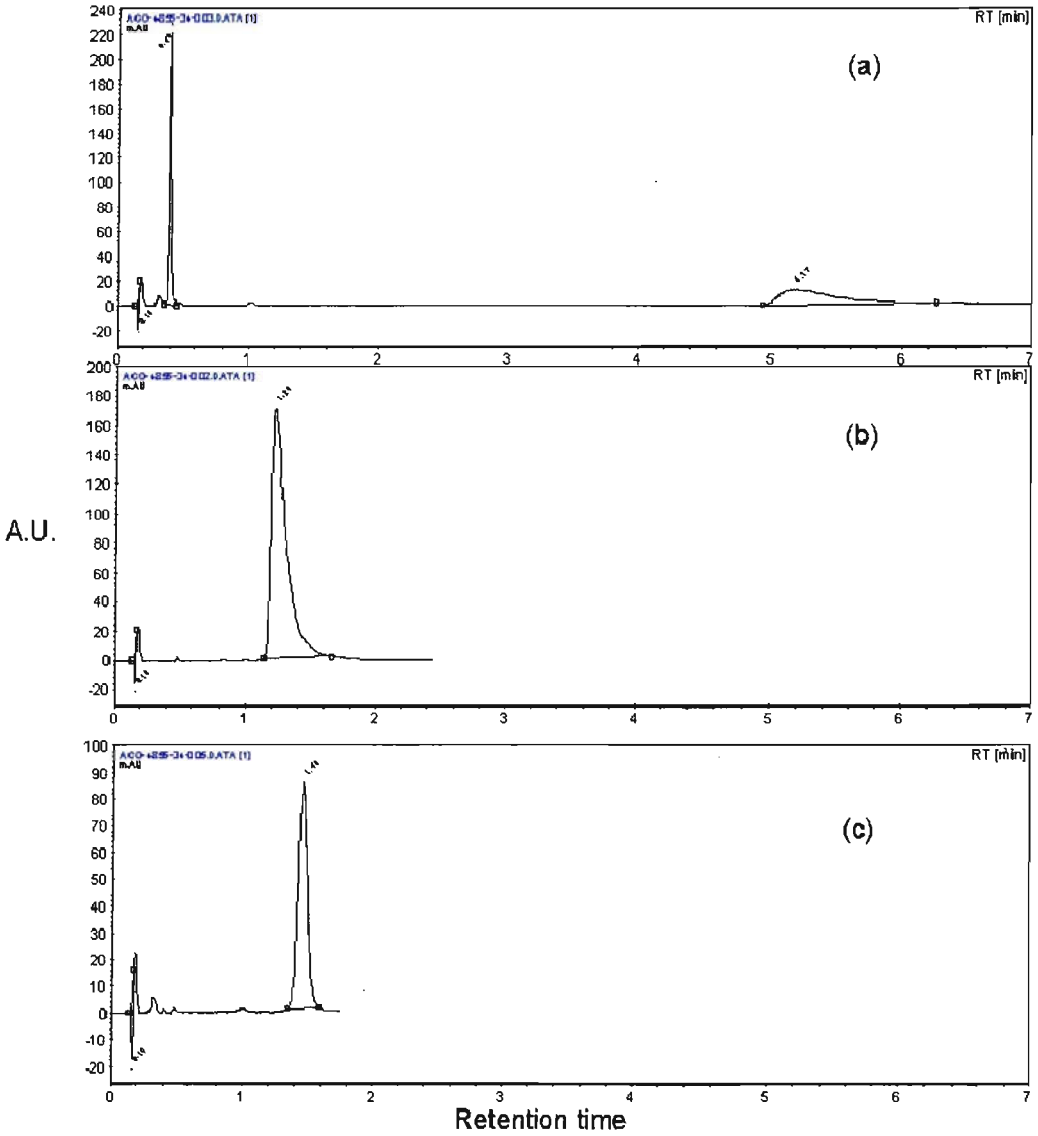


Figure 30. UV traces of three analytes on 2-EP column, 100 bar, 35°C, 4 mL min^{-1}, 10 % MeOH in CO_2, no additive; (a) compound **31**, $As._{USP}$ = 3.18; (b) compound **26**, $As._{USP}$ = 2.50; (c) naproxen, $As._{USP}$ = 1.06.

The effect of additive concentration on the peak shape of the analytes has been evaluated by measuring the peak asymmetry according to the United States Pharmacopoeia ($As._{USP}$):

$$As._{USP} = \frac{w_{5\%}}{2 \times f} \qquad (5.1)$$

where $w_{5\%}$ is the peak width at 5 % of peak height and f is the first half width at 5 % of peak height.

5.2. Results.

5.2.1. Additive in the modifier.

The first series of analyses were undertaken by adding different concentrations of NH_4OAc into the modifier. Retention factors, k, and asymmetry, $As._{USP}$, were measured at nine different concentrations on 2-EP and 2-EP-EC stationary phases and at six different concentrations on the Si column. Results are shown in Table 14, where k and $As._{USP}$ values are mean values of at least three replicate experiments and standard deviations are given in brackets. The observed standard deviations for k and $As._{USP}$ are high in several instances *e.g.* compound **31** with no additive using 2-EP column or diltiazem with no additive using Si column. For reason of clarity, error bars are not shown in Figure 32 to 34 but are present in Figure 35. This variability could be seen as an indication of unreliability of the system and/or of the experimental results. However, there is evidence that it is not the case. Firstly, analytes that are not affected by the presence of the additive (like naproxen) showed good repeatability with low standard deviations. Secondly, significant variations are consistently observed in cases where k and $As._{USP}$ are high. The variability of retention factor for highly retained analytes is accounted for by their sensitivity to the presence of residual additive in the column. In fact, these compounds exhibited higher retention factors in the first series of experiments

(when the stationary phase had not yet been subjected to the presence of the additive) than in the subsequent replicate experiments. This suggests that the retention obtained in replicate experiments was lowered by the presence of residual additive inside the chromatographic column. The existence of such memory effect has been demonstrated on the Si stationary phase by Pinkston and co-workers.[96] Concerning the variability of $As._{USP}$, the same argument can be made in addition to the fact that strict definition of peak boundaries is made difficult for badly tailing peaks, resulting in an inherent uncertainty in the calculation of $As._{USP}$.

The information encoded in asymmetry values is therefore qualitative rather than quantitative: values comprised between 0.9 and 1.1 are to be viewed as indicators of satisfactory peak symmetry whereas values outside that range indicate asymmetric peaks.

Table 14. Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the modifier of the mobile phase.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Compound 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
2-EP	0.0	4	7.73 (2.03)[b]	2.24 (0.26)	11.50 (2.61)	3.93 (1.20)	43.54 (16.95)	4.38 (1.15)	22.89 (3.95)	2.50 (1.12)	7.80 (0.09)	1.05 (0.02)	4.55 (0.17)	1.37 (0.24)	25.94 (1.68)	1.21 (n/a)[c]
	0.3	4	6.13 (1.51)	2.40 (0.87)	10.75 (3.19)	2.72 (0.87)	45.50 (12.73)	3.39 (0.43)	20.86 (3.77)	1.15 (0.12)	7.88 (0.23)	1.13 (0.04)	4.04 (0.48)	1.34 (0.19)	25.19 (1.94)	flat[d] (n/a)[c]
	0.6	4	4.84 (0.80)	3.28 (1.67)	7.94 (0.63)	4.91 (2.40)	44.48 (15.60)	2.97 (0.22)	18.81 (2.54)	1.07 (0.09)	7.81 (0.13)	1.15 (0.05)	3.79 (0.03)	1.51 (0.21)	23.33 (n/a)[c]	flat[d] (n/a)[c]
	1.0	4	4.04 (0.36)	3.23 (1.99)	6.69 (0.21)	3.98 (1.86)	36.56 (7.37)	3.88 (1.86)	18.09 (2.63)	1.05 (0.04)	7.84 (0.06)	1.14 (0.06)	3.84 (0.37)	1.33 (0.21)	20.81 (0.97)	1.79 (n/a)[c]
	2.0	4	3.48 (0.39)	2.24 (0.90)	5.78 (0.73)	3.10 (1.51)	26.88 (3.97)	3.03 (1.38)	17.48 (2.65)	1.00 (0.03)	7.88 (0.09)	1.15 (0.07)	3.75 (0.35)	1.28 (0.18)	20.11 (0.82)	1.58 (0.39)
	5.0	4	3.08 (0.50)	1.37 (0.36)	5.08 (0.73)	1.67 (0.49)	20.58 (4.14)	3.31 (0.70)	16.72 (2.50)	0.96 (0.02)	7.92 (0.06)	1.12 (0.06)	3.75 (0.35)	1.12 (0.10)	18.50 (1.80)	1.39 (0.16)
	10	4	3.05 (0.37)	1.17 (0.19)	4.72 (0.64)	1.27 (0.22)	18.80 (3.85)	2.03 (0.60)	16.59 (2.23)	0.95 (0.01)	7.94 (0.11)	1.11 (0.05)	3.80 (0.33)	1.00 (0.03)	17.19 (2.06)	1.17 (0.26)
	15	4	3.14 (0.66)	1.04 (0.11)	4.94 (0.99)	1.11 (0.13)	19.47 (5.55)	1.44 (0.23)	16.97 (2.54)	0.96 (0.03)	8.03 (0.23)	1.07 (0.05)	3.91 (0.46)	0.93 (0.05)	16.27 (2.60)	1.12 (0.14)
	30	4	3.09 (0.62)	0.89 (0.12)	4.86 (0.97)	0.96 (0.07)	18.60 (5.04)	1.21 (0.07)	16.61 (2.34)	0.94 (0.01)	8.14 (0.43)	1.03 (0.04)	3.91 (0.47)	0.89 (0.03)	16.63 (2.52)	0.96 (0.06)

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulted in a very low peak height preventing calculation of asymmetry.

Table 14 (continued). Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the modifier of the mobile phase.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Compound 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
2-EP-EC	**0.0**	3	2.25 (0.19)	4.40 (1.64)	3.75 (0.47)	2.72 (1.16)	13.69 (2.90)	2.61 (0.87)	11.25 (0.84)	2.23 (0.68)	5.75 (0.17)	1.12 (0.04)	2.73 (0.18)	1.24 (0.12)	11.31 (1.25)	1.36 (n/a)[c]
	0.3	3	2.04 (0.14)	2.16 (0.38)	3.48 (0.16)	5.13 (1.36)	11.60 (1.28)	3.16 (0.69)	10.69 (0.66)	1.81 (0.09)	5.79 (0.20)	1.13 (0.03)	2.60 (0.13)	1.22 (0.06)	9.77 (0.89)	2.99 (0.10)
	0.6	3	1.96 (0.13)	1.59 (0.22)	3.21 (0.22)	3.58 (1.71)	10.92 (1.37)	2.72 (0.48)	10.44 (0.77)	1.52 (0.05)	5.81 (0.17)	1.13 (0.02)	2.56 (0.17)	1.15 (0.06)	9.61 (0.80)	2.66 (0.61)
	1.0	3	1.92 (0.10)	1.26 (0.16)	3.06 (0.31)	2.65 (1.10)	10.33 (1.40)	2.20 (0.12)	10.38 (0.78)	1.25 (0.02)	6.01 (0.32)	1.16 (0.01)	2.54 (0.13)	1.13 (0.09)	9.56 (0.82)	1.91 (0.40)
	2.0	3	1.85 (0.13)	1.04 (0.11)	3.00 (0.31)	0.96 (0.03)	9.98 (1.32)	1.68 (0.17)	10.29 (0.76)	1.08 (0.04)	5.88 (0.17)	1.17 (0.04)	2.54 (0.13)	1.21 (0.13)	9.52 (0.95)	1.29 (0.14)
	5.0	3	1.83 (0.16)	0.82 (0.06)	3.01 (0.41)	0.83 (0.01)	9.65 (1.16)	1.16 (0.08)	10.19 (0.74)	1.00 (0.05)	5.94 (0.17)	1.13 (0.02)	2.52 (0.16)	1.03 (0.05)	9.29 (0.87)	1.24 (0.51)
	10	3	1.81 (0.13)	0.80 (0.05)	2.88 (0.25)	0.84 (0.03)	9.46 (1.08)	0.99 (0.08)	10.08 (0.73)	0.95 (0.03)	5.98 (0.16)	1.14 (0.02)	2.50 (0.12)	1.05 (0.05)	9.19 (0.92)	0.93 (0.05)
	15	3	1.81 (0.13)	0.78 (0.02)	2.88 (0.25)	0.80 (0.04)	9.35 (1.07)	0.94 (0.07)	10.04 (0.72)	0.94 (0.06)	6.00 (0.19)	1.15 (0.00)	2.55 (0.06)	1.04 (0.06)	9.15 (0.92)	0.87 (0.03)
	30	3	1.81 (0.13)	0.79 (1.64)	2.92 (0.30)	0.81 (0.01)	9.25 (1.01)	0.87 (0.05)	9.96 (0.69)	0.87 (0.05)	5.96 (0.19)	1.09 (0.02)	2.50 (0.12)	0.99 (0.05)	9.47 (0.61)	0.79 (0.04)

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulted in a very low peak height preventing calculation of asymmetry.

Table 14 (continued). Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the modifier of the mobile phase.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Componnd 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
Si	**0.0**	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.08 (0.72)	2.05 (0.53)	139.93 (50.86)	flat[d] (n/a)[c]	no peak	no peak
	1.0	3	109.11 (n/a)[c]	flat[d] (n/a)[c]	122.64 (20.94)	6.68 (1.96)	no peak	no peak	78.21 (19.33)	3.00 (1.61)	3.04 (0.33)	1.89 (0.56)	21.48 (4.64)	4.82 (1.43)	216.83 (n/a)[c]	flat[d] (n/a)[c]
	5.0	3	55.40 (23.65)	4.78 (1.65)	61.56 (31.27)	7.77 (2.21)	no peak	no peak	42.80 (10.82)	2.16 (0.39)	3.17 (0.20)	1.96 (0.15)	13.81 (2.96)	2.25 (0.42)	110.44 (24.64)	flat[d] (n/a)[c]
	15.0	3	32.28 (12.30)	5.32 (1.18)	38.91 (15.00)	4.90 (0.94)	199.56 (n/a)[c]	flat[d] (n/a)[c]	35.58 (8.67)	1.43 (0.12)	3.39 (0.11)	2.03 (0.09)	12.17 (2.48)	1.48 (0.14)	86.41 (20.28)	2.06 (0.62)
	30.0	3	40.56 (32.60)	4.90 (1.49)	43.50 (29.49)	3.19 (1.12)	177.03 (30.76)	flat[d] (n/a)[c]	46.48 (30.61)	1.14 (0.18)	4.88 (2.15)	1.76 (0.66)	12.72 (4.83)	1.15 (0.20)	73.43 (29.58)	1.90 (0.78)
	60.0	3	19.44 (4.19)	3.19 (0.10)	21.43 (5.38)	2.48 (0.71)	126.24 (21.99)	3.23 (n/a)[c]	24.37 (2.10)	1.13 (0.06)	3.98 (0.22)	1.52 (0.11)	8.56 (0.50)	1.17 (0.03)	43.72 (3.65)	1.85 (0.17)

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulted in a very low peak height preventing calculation of asymmetry.

Initially, the behaviour of the analytes in the absence of an additive was evaluated. Using the 2-EP column, all seven analytes were eluted without additive, although compound **31**, compound **32** and atenolol exhibited late elution ($k > 20$) and peak tailing in such conditions. The use of the endcapped 2-EP (2-EP-EC) considerably reduces the retention, with k decreasing below 15. However, peak splitting was observed for that stationary phase. This was not the case on the non-endcapped 2-EP column. As far as the Si column is concerned, only naproxen was successfully eluted in the absence of the additive. Diltiazem did elute from the column but as a flat peak with a k value of *ca.* 140. An example of flat peak is shown in Figure 31 .

Addition of a small amount (1 mM) of the additive was sufficient to significantly reduce retention times and improve peak shapes of tailing compounds on all three stationary phase. Indeed, the addition of 1 mM NH_4OAc in the modifier caused all the compounds, except compound **31**, to elute from the Si column, although elution was late and with strongly tailing peak shape. Conversely, the additive had virtually no effect at all on the retention and the peak shape of the non-tailing compounds, as shown in Figure 31.

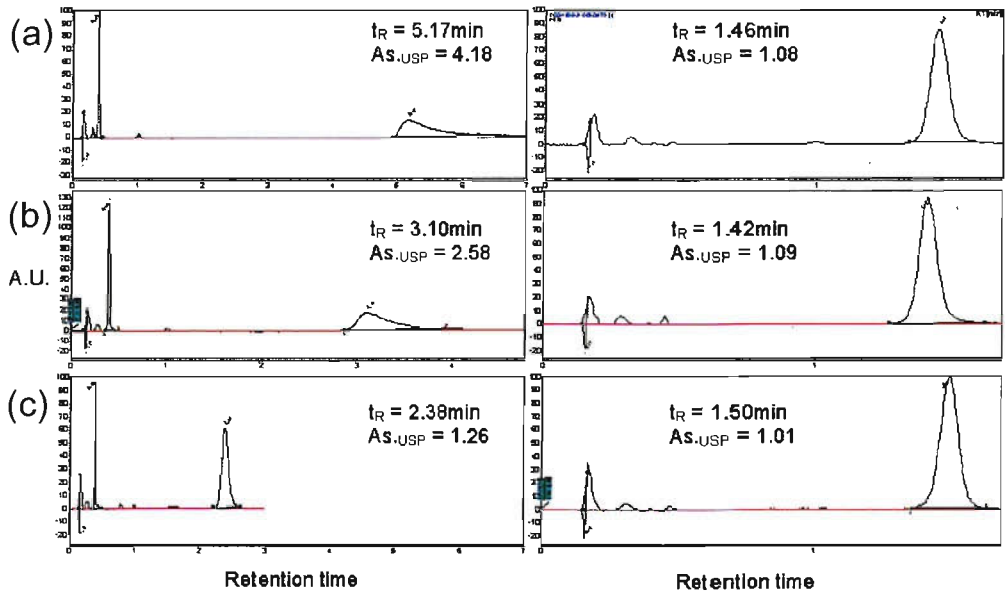


Figure 31. Evolution of retention time and peak shape with increasing amounts of NH_4OAc in the modifier, 2-EP column, 10 % v/v modifier in CO_2, 100 bar, 35°C, 4 mL min^{-1}. Left: compound **31**; right: naproxen. Modifier: (a) 0.3 mM NH_4OAc in MeOH; (b) 5 mM NH_4OAc in MeOH; (c) 30 mM NH_4OAc in MeOH.

Increasing the concentration of the additive gradually decreased both retention and asymmetry towards discrete minimum values. Figure 32 shows that

the capacity factors of each compound reach their minima when the additive is present in the modifier at or above a concentration which is dependent on the nature of the stationary phase. For the 2-EP column, minimum retention is reached at NH_4OAc concentration above 10 mM, whereas on the 2-EP-EC column a concentration of 5 mM NH_4OAc proved sufficient. On the Si column, increasing concentration up to 60 mM NH_4OAc in the modifier decreased retention factors gradually. Addition of higher additive concentration in the modifier may achieve shorter retention on the Si column.

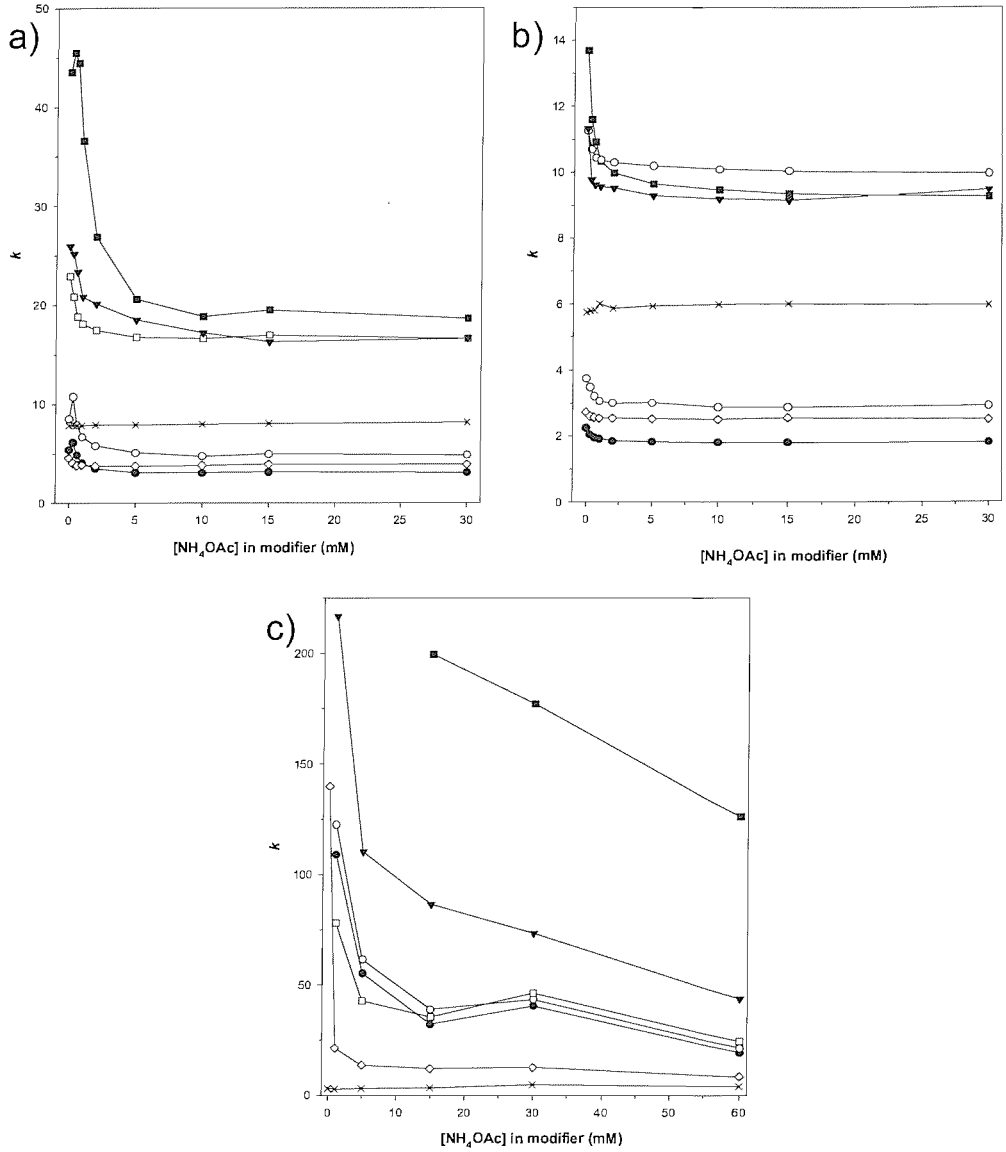


Figure 32. Effect of the addition of increasing concentration of ammonium acetate in the modifier on capacity factors. Conditions: 100 bar, 35°C, 4 mL min^{-1} of 10 % v/v MeOH+NH_4OAc in CO_2. ●: compound **16**, ○: compound **26**, ■: compound **31**, □: compound **32**, ×: naproxen, ◊: diltiazem, ▼: atenolol. a) 2-EP column, b) 2-EP-EC column, c) Si column.

Figure 33 illustrates the effect of additive concentration on peak asymmetry. As observed with retention, asymmetry decreased with increasing concentration of additive. Minimum values of asymmetry were reached at 15 mM NH_4OAc for 2-EP column and 10 mM for 2-EP-EC, *i.e.* at slightly higher concentrations than those at which minimum retention values were reached. Again, on Si stationary phase, no stabilization could be achieved up to 60 mM NH_4OAc in the modifier. However, concerning asymmetry, the aim is not only to reach a minimum value but also to achieve symmetry of the peak. On 2-EP stationary phase, peak shapes of all compounds evolved towards symmetry (*i.e.* $As._{USP} = 1$). This is not the case on 2-EP-EC column for which only naproxen and diltiazem exhibited perfect symmetry; whilst asymmetries of the other compounds converged towards values lower than 1, *i.e.* fronting peak shapes. On the Si column, peak symmetry was achieved for compound **32** and diltiazem; other analytes exhibiting gradually improved peak shape without reaching symmetry.

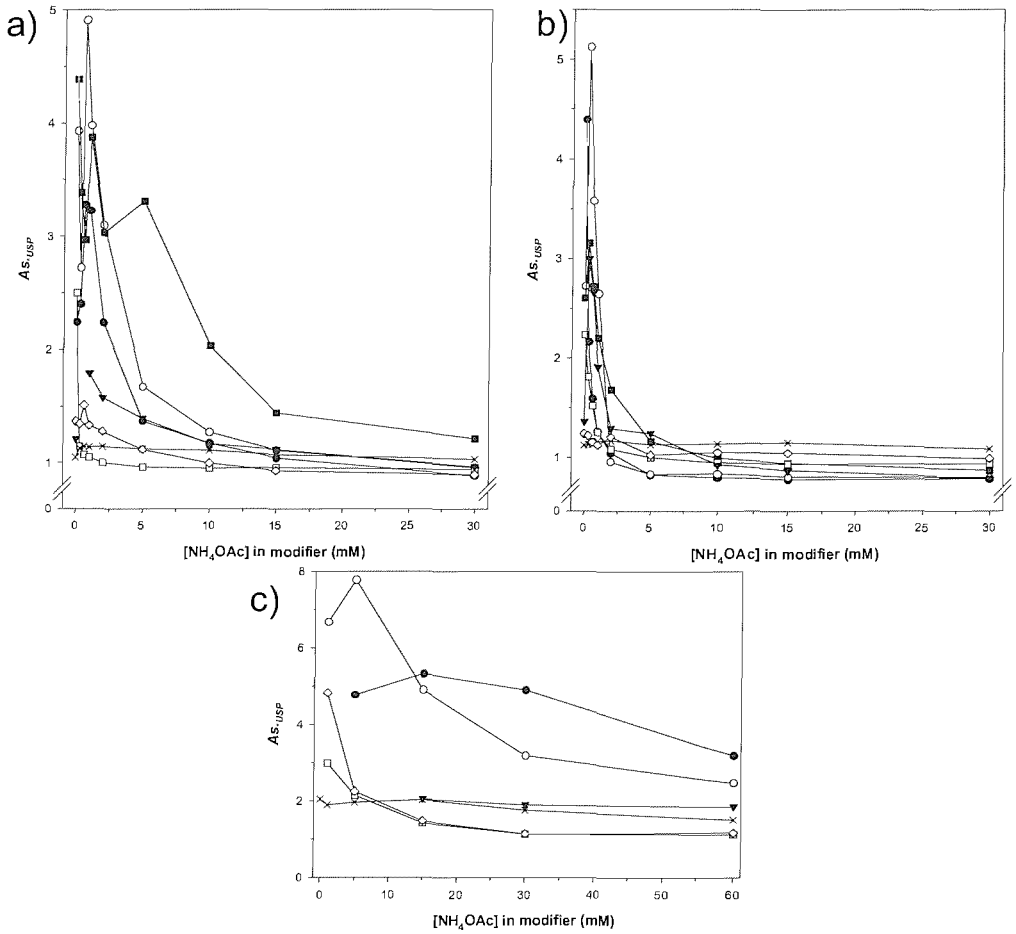


Figure 33. Effect of the addition of increasing concentration of ammonium acetate in the modifier on asymmetry. Conditions: 100 bar, 35°C, 4 mL min^{-1} of 10 % v/v MeOH + NH_4OAc in CO_2. ●: compound **16**, ○: compound **26**, ■: compound **31**, □: compound **32**, ×: naproxen, ◊: diltiazem, ▼: atenolol. a) 2-EP column, b) 2-EP-EC column, c) Si column.

5.2.2. Additive in sample.

A second series of analyses was undertaken adding NH_4OAc at different concentrations directly to the sample solvent. k and $As._{USP}$ were measured at five different concentrations on all three stationary phases. Results are shown in Table 15, where k and $As._{USP}$ values are means of at least three replicate experiments. Standard deviations are given in brackets and previous comments apply (*vide supra*). When added to the sample solvent, the additive proved to have different effects with regards to the stationary phase. On the 2-EP column, reduced retention times were observed, and, furthermore, retention factors reached a minimum value at a concentration of only 5 mM, to compare with the 10 mM needed when the additive was present in the modifier, although shorter retention was achieved in the latter case. However, the presence of NH_4OAc in the sample did not improve peak shapes in the same fashion as seen when added in the modifier. The addition of NH_4OAc in the sample did not influence the elution of the analytes from the 2-EP-EC phase. Improved peak symmetry was observed using the latter for sample additive concentrations above 10 mM.

Table 15. Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the sample solvent.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Compound 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
2-EP	**0.0**	3	5.39 (2.03)[b]	2.24 (0.26)	8.50 (2.61)	3.93 (1.20)	43.54 (16.95)	4.38 (1.15)	23.50 (4.60)	2.48 (1.37)	7.83 (0.07)	1.04 (0.02)	4.60 (0.16)	1.25 (0.08)	25.94 (1.68)	1.21 (n/a)[c]
	2.0	3	4.50 (0.83)	2.03 (0.29)	7.27 (1.24)	2.71 (0.08)	35.69 (7.92)	3.52 (1.12)	21.13 (1.62)	1.64 (0.08)	7.81 (0.00)	1.06 (0.04)	4.67 (0.31)	1.12 (0.06)	24.35 (2.82)	1.54 (0.35)
	5.0	3	3.63 (0.47)	1.88 (0.47)	5.88 (0.80)	2.01 (0.56)	27.46 (4.53)	4.29 (0.91)	19.19 (1.50)	1.44 (0.05)	7.83 (0.04)	1.09 (0.02)	4.35 (0.22)	1.12 (0.04)	21.65 (1.88)	1.92 (0.72)
	10	3	3.56 (0.47)	2.05 (0.45)	5.73 (0.73)	2.07 (0.18)	26.29 (4.45)	4.28 (0.56)	18.98 (1.47)	1.33 (0.03)	7.85 (0.04)	1.10 (0.02)	4.33 (0.22)	1.10 (0.03)	21.33 (1.76)	1.98 (0.21)
	15	3	3.50 (0.47)	2.02 (0.46)	5.63 (0.70)	1.86 (0.21)	25.44 (4.28)	4.13 (1.10)	18.79 (1.41)	1.31 (0.02)	7.83 (0.04)	1.11 (0.01)	4.29 (0.19)	1.08 (0.04)	20.83 (1.76)	2.05 (0.31)

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulting very low peak height prevented calculation of asymmetry.

Table 15 (continued). Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the sample solvent.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Compound 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
2-EP-EC	**0.0**	3	1.98 (0.24)	3.05 (0.21)	3.19 (0.44)	3.24 (1.57)	11.65 (1.32)	2.68 (0.71)	10.81 (0.27)	1.62 (0.10)	5.75 (0.17)	1.12 (0.03)	2.63 (0.06)	1.11 (0.00)	9.90 (1.43)	1.36 (n/a)[c]
	2.0	3	1.92 (0.13)	1.70 (0.60)	3.10 (0.30)	2.60 (0.16)	11.81 (1.23)	2.76 (1.09)	10.88 (0.62)	1.48 (0.11)	5.85 (0.04)	1.16 (0.03)	2.65 (0.10)	1.07 (0.06)	10.15 (1.02)	1.80 (0.64)
	5.0	3	2.00 (0.27)	1.25 (0.47)	3.21 (0.37)	1.21 (0.06)	11.31 (1.17)	2.23 (0.55)	10.63 (0.25)	1.49 (0.07)	5.79 (0.20)	1.14 (0.04)	2.60 (0.04)	1.05 (0.05)	10.06 (0.98)	1.73 (0.61)
	10	3	2.04 (0.24)	1.09 (0.29)	3.17 (0.34)	1.11 (0.04)	11.06 (1.16)	2.17 (0.44)	10.52 (0.25)	1.39 (0.08)	5.77 (0.18)	1.15 (0.08)	2.56 (0.06)	1.08 (0.09)	10.13 (0.80)	1.60 (0.49)
	15	3	1.98 (0.24)	0.91 (0.02)	3.10 (0.30)	1.23 (0.35)	10.77 (0.95)	1.49 (0.16)	10.40 (0.18)	1.30 (0.08)	5.79 (0.20)	1.15 (0.08)	2.54 (0.04)	1.07 (0.02)	9.90 (0.61)	1.37 (0.08)

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulting very low peak height prevented calculation of asymmetry.

Table 15 (continued). Retention factor and asymmetry values of the test compounds measured on three stationary phases for different concentrations of ammonium acetate in the sample solvent.

Columns	[NH_4OAc] in modifier (mM)	n[a]	Compound 16		Compound 26		Compound 31		Compound 32		Naproxen		Diltiazem		Atenolol	
			k	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*	*k*	*As.USP*
Si	**0.0**	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.08 (0.72)	2.05 (0.53)	139.93 (50.83)	flat[d] (n/a)	no peak	no peak
	2.0	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.60 (0.13)	2.63 (0.39)	97.97 (4.20)	flat[d] (n/a)	no peak	no peak
	5.0	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.56 (0.35)	2.69 (0.29)	43.00 (15.24)	2.29 (0.53)	no peak	no peak
	10	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.54 (0.32)	2.58 (0.25)	32.80 (13.48)	4.80 (3.29)	no peak	no peak
	15	3	no peak	no peak	no peak	no peak	no peak	no peak	no peak	no peak	3.63 (0.46)	2.37 (0.04)	16.39 (3.32)	13.73 (6.33)	no peak	no peak

[a] number of replicate experiments.
[b] values in brackets indicate standard deviation.
[c] unavailability of standard deviation is due to the fact that only one experiment allowed for the measurement of the given value.
[d] "flat" means that the tailing peak shape resulting very low peak height prevented calculation of asymmetry.

For instance, peak splitting was reduced to the presence of a shoulder or removed altogether. Nevertheless, peak symmetry was not improved as significantly as with the additive in the modifier. Adding additive in the sample had no effect on the retention of naproxen on Si column, neither did it allow for the elution of the other compounds to the exclusion of diltiazem. For this latter analyte, *k* decreased from *ca.* 140 in the absence of modifier to less than 20 in the presence of 60 mM NH_4OAc in the sample solvent. These results are illustrated in Figure 34.

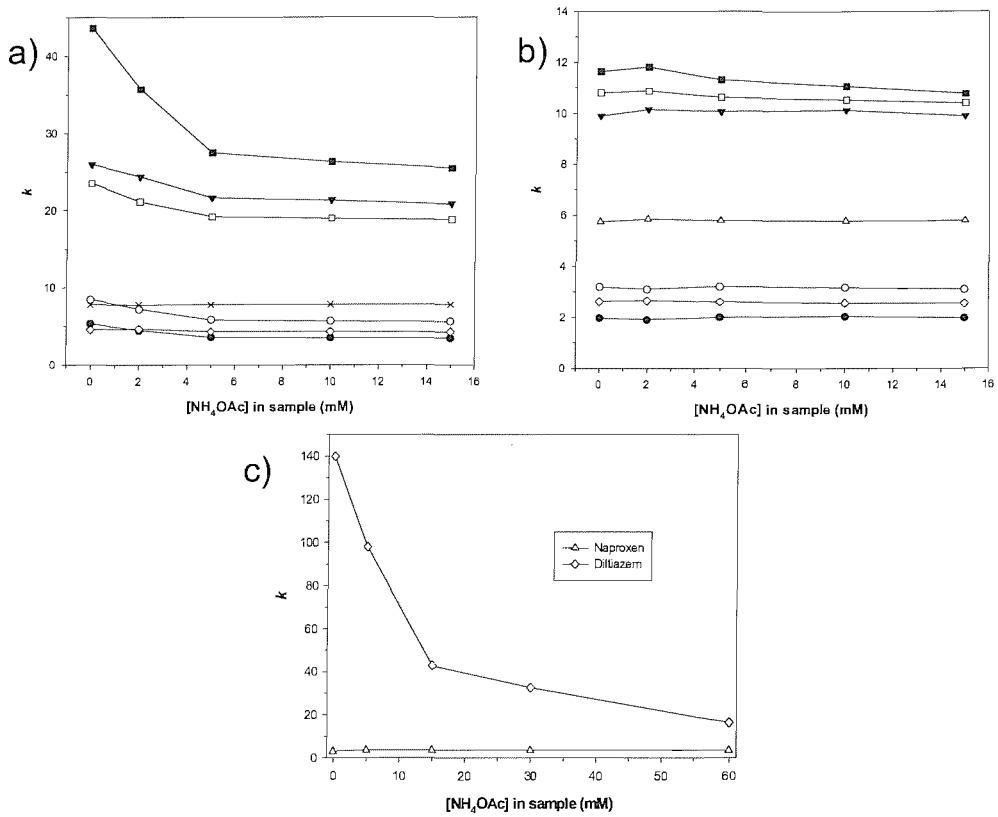


Figure 34. Effect of the addition of increasing concentration of ammonium acetate in the sample solvent. Conditions: 100 bar, 35°C, 4 mL min^{-1} of 10 % v/v MeOH in CO2. ●: compound **16**, ○: compound **26**, ■: compound **31**, □: compound **32**, ×: naproxen, ◊: diltiazem, ▼: atenolol. a) 2-EP column, b) 2-EP-EC column, c) Si column.

5.2.3. Summary.

In summary, using the 2-EP stationary phase, the addition of ammonium acetate in the modifier reduced retention and improved peak shape at

concentration of 15 mM and above. When added in the sample solvent, the additive reduced retention of test analytes at lower concentration (5 mM), although to a smaller extent and with limited or no effect on peak symmetry.

With the 2-EP-EC, compounds were less retained than on the non-endcapped 2-EP phase, but peak splitting was observed in the absence of additive. The presence of NH_4OAc in the modifier reduced retention and improved peak shape at a concentration of 5 mM and above. Nevertheless, NH_4OAc in the sample affected neither retention nor peak shape of the test analytes.

A comparison of the effect of the additive in modifier and sample for both the endcapped and non-endcapped 2-EP stationary phases is shown in Figure 35.

With a Si column, the addition of ammonium acetate in the modifier afforded the elution of all compounds compared with the absence of additive when only naproxen was eluted. Although retention remained high for most compounds and symmetry was not always achieved, the additive nevertheless produced significant improvement of chromatographic behaviour. On the other hand, when added in the sample solvent, and even at high concentration (60 mM), the additive did not permit the elution of the test compounds, with the exception of diltiazem.

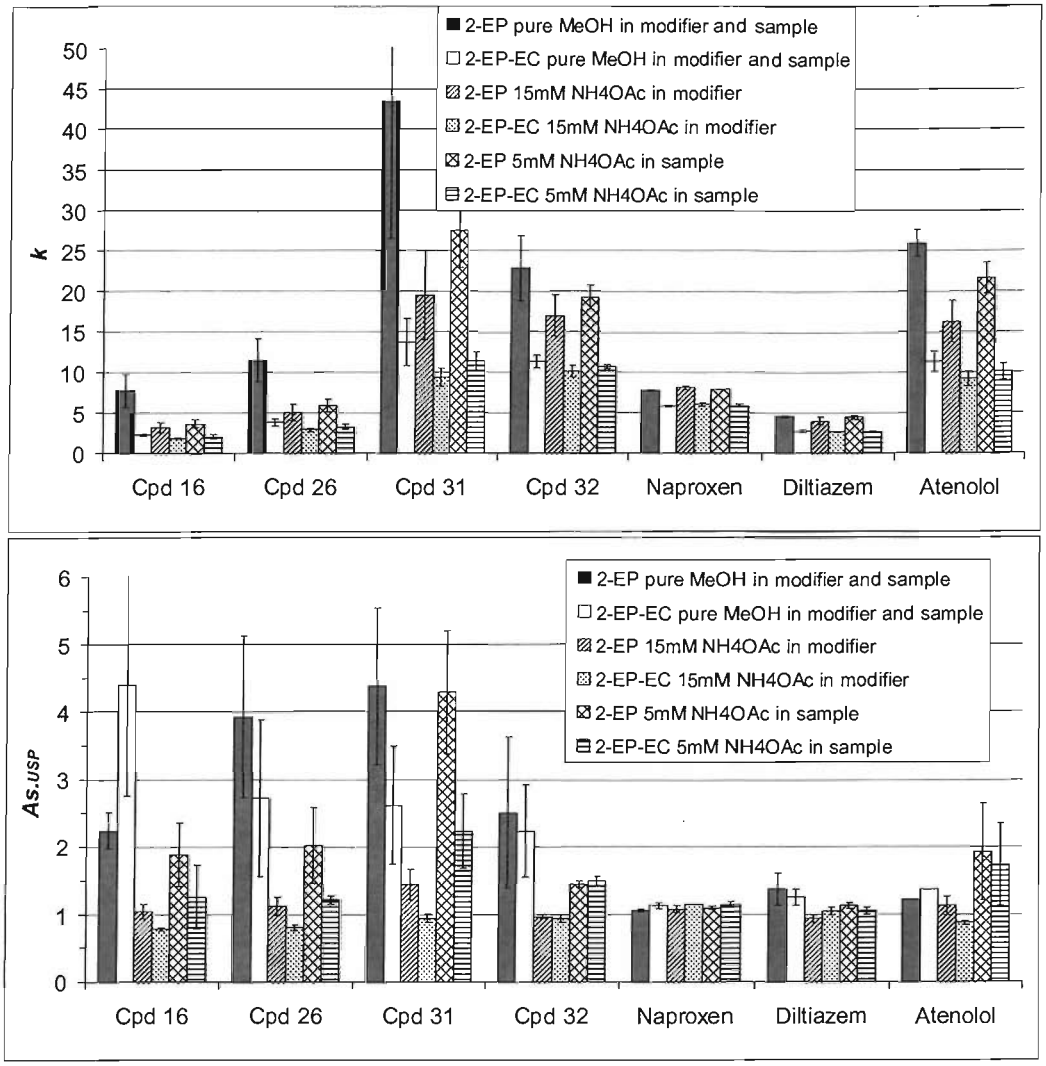


Figure 35. Comparison of the effect of the addition of additive in the modifier and in the sample for both 2-EP and 2-EP-EC stationary phases. a) effect on retention factors; b) effect on asymmetry.

5.3. Discussion.

Literature suggests that the additive affects retention by covering free silanols.[96] This is consistent with the fact that less additive is required to achieve minimum retention on 2-EP-EC than on 2-EP, for silanols are present in fewer number on the endcapped phase. Moreover, it has been proposed that several retention mechanisms could be involved in the elution of the analytes on 2-EP stationary phases.[4] Firstly, the 2-EP stationary phase could be partially deactivated by hydrogen bonding between free silanols and the nitrogen atom of the pyridine

groups. Secondly, the bonded aromatic ring would prevent interaction of analytes with free silanols. Thirdly, the nitrogen atom of the pyridine groups would be partially protonated due to the acidity of CO_2/MeOH mobile phase and, therefore, a positively charged analyte peak would exhibit short retention time and fronting shape.

Diltiazem was the only compound to be unequivocally protonated before elution (purchased as the hydrochloric salt). It was less retained on 2-EP-EC than on 2-EP and very strongly retained on Si. The latter fact is explained by the strong interaction between the silanols of the fully hydroxylated silica phase and the positive charge born by the solute. On the 2-EP phase, the retention is probably due to interaction of the solute with the pyridine rings and the free silanols. In agreement with the retention mechanisms proposed above,[4] it is hypothesized that, on 2-EP-EC, there being fewer free silanols, there is less hydrogen bonding opportunities between the free silanols and the nitrogen atoms of the pyridyl group. The nitrogen atoms would therefore be more available for protonation by the acidic mobile phase and thus cause shorter retention of positively charged analytes (Figure 36). This would be consistent with the fact that several analytes exhibited split peaks on the 2-EP-EC column. The splitting could indeed be explained by the co-existence of two distinct retention mechanisms for these analytes on the endcapped column: (i) interaction with free silanols and (ii) interaction with positively charged nitrogen atoms. Adding ammonium acetate should deactivate the silanols and therefore favour the second mechanism, thus leading to the observed disappearance of splitting (only one mechanism left) and fronting peak shapes (repulsion between positively charged analytes and nitrogen atoms of the pyridyl group).

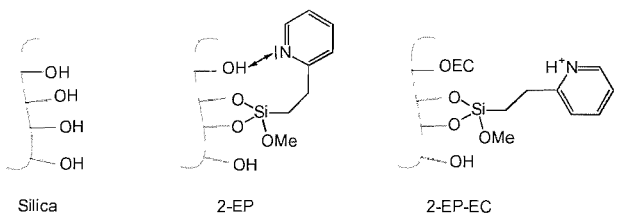


Figure 36. Left: interactions between free silanols and nitrogen of the pyridine group on a 2-EP stationary phase. Right: on the endcapped 2-EP these interactions do not occur, therefore allowing a bigger proportion of nitrogen to be protonated by the acidic mobile phase. (NB: EC stands for endcapping; the exact nature of the endcapping agent is unknown due to proprietary concerns)

The retention of naproxen is not affected by the additive on any of the stationary phase and it exhibited the shortest retention on the Si phase. Moreover, naproxen was the only acidic compound of the test set and is thought to be in its uncharged form due to the acidity of the mobile phase. The retention mechanism for this compound is thus believed to be mainly due to π-π interactions. However, the fact that it was less retained on 2-EP-EC than on 2-EP and slightly retained on bare silica suggests that the free silanols must be at least partially involved in the retention process.

Addition of NH_4OAc in the sample solvent led to diverse effects with regard to the nature of the stationary phase. With the 2-EP column, retention was reduced but to a smaller extent than when added in the modifier and the effect on symmetry was rather limited. With 2-EP-EC, addition of NH_4OAc to the sample did not have any significant effect on retention. Only for Diltiazem, when using the Si phase, was retention affected. The effect of the additive in the sample is thought to alter retention by the formation of ion pairs between analytes and ammonium acetate. Because addition of NH_4OAc to the modifier improves retention and peak shape is better than its addition to the sample, the additive is thought to act predominantly by deactivating the free silanols of the stationary phase.

Concluding remarks

The suitability of SFC for pharmaceutical applications has been discussed in the literature. In the present work, a library of thirty-two drug-like sulfonamides was designed and synthesised with the aim of studying the retention behaviour of pharmaceutical compounds in SFC. The technique was confirmed as being suitable for the analysis of the compounds of interest. Affording fast (less than 3 minutes) and reliable elution with symmetric peak shapes of the analytes, provided that the appropriate experimental conditions were used.

With regard to the stationary phase, 2-ethyl-pyridyl (2-EP), cyanopropyl (CN), diol and bare silica (Si) were studied in preliminary investigations. The 2-EP column proved to be the most useful, affording the elution of all analytes, whilst the most basic compounds failed to elute from CN, diol and Si columns under the conditions used. Nevertheless, when using pure methanol as a modifier, even with the 2-EP stationary phase, some of the analytes exhibited tailing peak shapes, which prompted the use of additives in the modifier.

A common way to improve peak shape and shorten retention time in SFC is to use an additive in the modifier of the mobile phase. One advantage of SFC is the ease of scaling up from analytical to preparative scale and the fact that purified fraction are collected in modifier (usually methanol) and, therefore, are readily evaporated to yield the compound of interest. Thus, the use of an additive can be seen as deleterious to the easy recovery of purified compounds. For that reason, preference is given to volatile additives used in as small amounts as possible. As a consequence, ethyl-dimethyl-amine (EDMA) and ammonium acetate (NH_4OAc) were added in the modifier in low concentrations (0.1 % v/v EDMA and 0.6 mM NH_4OAc in MeOH). Already at these low concentrations, the presence of the additives in the modifier afforded an improvement in the peak shapes of tailing compounds (as well as it afforded the elution from diol or Si stationary phases, for instance, of compounds that did not elute with pure methanol as modifier), allowing for the analysis of most of the SOTLIB library compounds. Nevertheless, the additive concentration study described in Chapter 5 showed that using ammonium acetate at concentrations as high as 10 or 15 mM in the modifier may

be of interest for the elution of particularly difficult compounds, whilst not affecting the retention of analytes presenting symmetric peak shapes. Whether or not the use of such concentration of additive would be deleterious to the recovery of pure compounds in the case of a preparative method is worth investigation but could not be determined during this project. The study of increasing concentrations of basic additive, like EDMA, would also be interesting to assess whether or not results similar to those obtained with ammonium acetate would be observed.

Using polycratic studies, it was shown that the logarithm of the retention factor, log k, varies linearly with the modifier proportion, φ, in the mobile phase, provided that log k is kept within a given range ($0 < \log k < 1$) and that φ is kept above 10 %. This linearity allowed for the measurement of retention characteristics $\log k_0$, S and φ_0 of the SOTLIB compounds. These characteristics were subsequently shown to correlate with the calculated molecular descriptors of the analytes (μ, A and δ_{min}) that had been used by Kaliszan and co-workers for the description of retention of drug molecules. Although Kaliszan's model had been used to describe HPLC retention of drug molecules, using C_{18} stationary phases, this work suggests that this model is applicable to the description of retention in SFC using 2-EP stationary phase and CO_2-MeOH mobile phases. Kaliszan's descriptors present the advantage of being meaningful in terms of retention mechanism. However, one must consider the validity of such calculations. The model used herein to calculate molecular descriptors are widely used in molecular modelling but they are indeed rather simple. Are the calculated numeric values genuine or are they only broad approximations? Further, how does that influence the validity of the results? When dealing with such questions, one must clearly define the aim of the project. In the present case, the aim was to determine whether retention could be correlated with easily calculated descriptors and whether a simple model could help predicting retention of unknown analytes from their molecular structures. Ideally, the calculations should be undertaken as quickly as possible, with limited knowledge of computational chemistry and limited computational facilities (for instance, during this project the computer used was not powerful enough to calculate the descriptors with the most advanced computational models offered by Spartan'02). For that matter, whether or not the

calculated descriptors correspond to a physical reality is not relevant. The only relevant question is: do the calculated descriptors allow for reliable prediction of retention whatever the compound structure? Of course, in a more fundamental setting, where the aim of the calculations would be in-depth understanding of retention mechanisms, approximate descriptors would not do, and advanced (most likely tailor-made) computational models should be used, although this would probably be more the scope of a computational chemistry project.

This study also remains limited in terms of diversity of the test set. Firstly, some of the initial compounds had to be withdrawn due to peak splitting or tailing. Secondly, all test compounds are structurally related. The results obtained in Chapter 5, with regards to the improvement in chromatographic behaviour of problematic analytes when using increasing concentrations of ammonium acetate as an additive, are encouraging for the extension of the model to withdrawn compounds. Nevertheless, the study would have to be extended in two ways to ensure that Kaliszan's model can be applied universally to SFC. Firstly, the test set should be extended to acidic analytes and diverse chemical structures to see whether or not a unified equation is still valid for all structures. Secondly, other stationary phase should be studied to give insight into the differences in retention mechanisms. The descriptors used herein will probably be suitable for description of retention using stationary phases with similar retention mechanism to 2-EP (like CN), leading to similar equations with slight differences in regression coefficients. On the contrary, these descriptors might not be of any use for the prediction of retention using other stationary phases involving distinct retention mechanisms (*e.g.* diol, Si) for which identification of other relevant descriptors would be needed.

To conclude, the model exposed herein describes the retention of sulfonamides in SFC using a 2-EP stationary phase and CO_2-MeOH mobile phases. It highlights the fact that dipole–dipole and dipole-induced dipole interactions, polar and hydrogen-bonding type interactions and dispersive interactions are of importance for the retention of these analytes in SFC using these experimental conditions. Whether or not this model generalises to any type of analyte has not yet been determined. Nevertheless it proves that simple molecular calculations can be used to predict retention in SFC for drug-like compounds as has previously been shown in HPLC. Armed with this knowledge the Pharmaceutical industry could re-assess their use of organic solvent–hungry

HPLC methods and protocols, and consider using more ecologically friendly SFC approaches.

Experimental

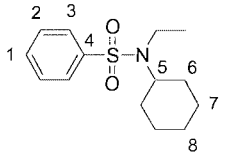
I. Synthesis.

Reactions were monitored by thin-layer chromatography on silica gel (pre-coated PF_{254} Merck plates); the spots were examined with UV light and when necessary visualized with ninhydrin. Preparative thin-layer chromatography was performed using Merck silica 60 PF_{254} on glass plates. NMR spectra were recorded on a Bruker DRX 300 MHz spectrometer. Melting points were determined on a Gallenkamp apparatus and are uncorrected.

*General procedure for the synthesis of sulfonamides **1, 2, 4, 5, 7** and **8**.*

Triethylamine (3.3 mmol, 0.33 g, 1.1 eq.) was added to a solution of the appropriate amine (3.0 mmol, 1.0 eq.) in CH_2Cl_2 (15 mL) and the mixture was stirred at room temperature for five minutes. Then the appropriate sulfonyl chloride (3.0 mmol, 1.0 eq.) was added. The mixture was stirred at room temperature for 7 to 12 hours depending on the reactants, washed with aqueous HCl (1M) (2x15 mL) and brine (15 mL). The aqueous layers were combined and extracted with CH_2Cl_2 (15 mL). The organic layers were combined, dried over $MgSO_4$ and concentrated *in vacuo*. A small amount of the crude was either purified by preparative TLC using CH_2Cl_2 as an eluant or recrystallized, depending on its purity to afford sulfonamides **1**, **2**, **4**, **5**, **7** and 8.

*N-Cyclohexyl-N-ethyl-benzenesulfonamide (**1**)*



N-Ethylcyclohexylamine and benzenesulfonyl chloride were reacted according to the general procedure to afford *N*-cyclohexyl-*N*-ethyl-benzenesulfonamide (**1**) (59 % yield) as a white crystalline solid, m.p. 61-63°C (EtOH).

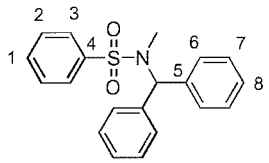
LR-MS (ESI+): m/z 268.2 $[M+H]^+$, 290.2 $[M+Na]^+$

HR-MS (ESI+): m/z 290.1186 $[M+H]^+$ (calculated: 290.1185, error: 0.28 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 7.85 (dt, 2H, H_3, J = 6.9 and 1.8 Hz), 7.66-7.53 (m, 3H, H_1 and H_2), 3.62 (tt, 1H, H_5, J = 11.5 and 3.7 Hz), 3.27 (q, 2H, $C\underline{H}_2CH_3$, J = 7.0 Hz), 1.77-1.73 (m, 2H), 1.63-1.29 (m, 8H, 7 aliphatic H and H_8), 1.23 (t, 3H, $CH_2C\underline{H}_3$, J = 7.0 Hz), 1.11 (tt, 1H, H_8, J = 12.6 and 3.5 Hz).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 143.1 (C_4), 133.6 (C_1), 130.3 (C_2), 127.9 (C_3), 59.5 (C_5), 39.6 ($\underline{C}H_2CH_3$), 32.8 (C_6), 27.2 (C_7), 26.4 (C_8), 18.2 ($CH_2\underline{C}H_3$).

N-Benzhydryl-N-methyl-benzenesulfonamide (**2**)[98]



N-(Diphenylmethyl)methyl-amine and benzenesulfonyl chloride were reacted according to the general procedure to afford *N*-benzhydryl-*N*-methyl-benzenesulfonamide (**2**) (54 % yield) as a white crystalline solid, m.p. 72-73°C (hexane). Literature data were in agreement.[98]

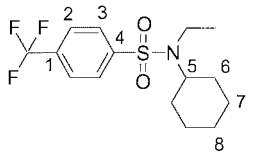
LR-MS (ESI+): m/z 360.1 $[M+Na]^+$, 697.3 $[2M+Na]^+$

HR-MS (ESI+): m/z 360.1035 $[M+H]^+$ (calculate : 360.1028, error : 1.89 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 7.77 (dt, 2H, H_3, J = 6.9 and 1.6 Hz), 7.61 (tt, 1H, H_1, J = 1.5 and 7.3 Hz), 7.50 (tt, 2H, H_2, J = 1.4 and 7.5 Hz), 7.29-7.25 (m, 6H, H_6 or H_7 and H_8), 7.08-7.04 (m, 4H, H_6 or H_7), 6.42 (s, 1H, CH), 2.70 (s, 3H, CH_3).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 141.2 (C_4), 139.8 (C_5), 133.8 (C_1), 130.2 (C_2 or C_3 or C_8), 129.9 (C_6 or C_7), 129.4 (C_6 or C_7), 128.8 (C_2 or C_3 or C_8), 128.3 (C_2 or C_3 or C_8), 65.7 (CH), 32.0 (CH_3).

N-Cyclohexyl-N-ethyl-4-trifluoromethyl-benzenesulfonamide (4)



N-ethylcyclohexylamine and 4-trifluoromethylbenzenesulfonyl chloride were reacted according to the general procedure to afford *N*-cyclohexyl-*N*-ethyl-4-trifluoromethyl-benzenesulfonamide (**4**) (73 % yield) as a white crystalline solid, m.p. 106-108°C (cyclohexane).

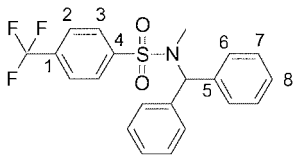
LR-MS (ESI+): m/z 358.1 $[M+Na]^+$

HR-MS (ESI+): m/z 336.1245, 20 % $[M+H]^+$; 358.1059, 60 % $[M+Na]^+$; 693.2277, 100 % $[2M+Na]^+$ (calculated: 358.1059, error: -0.05 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 8.05 (d, 2H, H_3, J = 8.2 Hz), 7.90 (d, 2H, H_2, J = 8.2 Hz), 3.67 (tt, 1H, H_5, J = 3.9 and 11.4 Hz), 3.30 (q, 2H, $C\underline{H}_2CH_3$, J = 6.9 Hz), 1.80-1.75 (m, 2H), 1.65-1.29 (m, 7H), 1.25 (t, 3H, $CH_2C\underline{H}_3$, J = 6.9 Hz), 1.12 (tt, 1H, H_8, J = 3.6 and 12.6 Hz).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 147.0 (C_4), 128.6 (C_2 and C_3), 127.4 (CF_3), 59.8 (C_5), 39.8 ($\underline{C}H_2CH_3$), 32.9 (C_6), 27.2 (C_7), 26.3 (C_8), 18.1 ($\underline{C}H_3$).

N-Benzhydryl-N-methyl-4-trifluoromethyl-benzenesulfonamide (5)



N-(Diphenylmethyl)methyl-amine and 4-trifluoromethylbenzenesulfonyl chloride were reacted according to the general procedure to afford *N*-benzhydryl-*N*-methyl-4-trifluoromethyl-benzenesulfonamide (**5**) (70 % yield) as a white crystalline solid, m.p. 86-89°C (cyclohexane).

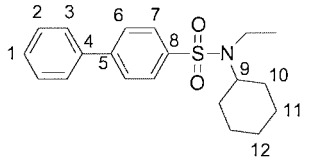
LR-MS (ESI+): m/z 428.4 $[M+Na]^+$, 833.7 $[2M+Na]^+$

HR-MS (ESI+): m/z 167.0856, 35 % $[Ph_2CH]^+$; 428.0905, 60 % $[M+Na]^+$; 833.1956, 100 % $[2M+Na]^+$ (calculated: 428.0902, error: 0.48 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 7.92 (d, 2H, H_3, J = 8.4 Hz), 7.77 (d, 2H, H_2, J = 8.4 Hz), 7.29-7.26 (m, 6H, H_8 and H_6 or H_7), 7.10-7.05 (m, 4H, H_6 or H_7), 6.43 (s, 1H, CH), 2.78 (s, 3H, CH_3).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 145.1 (C_4), 139.7 (C_1 and C_5), 130.9, 130.8, 130.1 (C_6 or C_7), 129.8 (C_6 or C_7), 129.0, 127.5 (CF_3), 66.3 (CH), 32.5 (CH_3).

Biphenyl-4-sulfonic acid cyclohexyl-ethyl-amide (7)



N-ethylcyclohexylamine and biphenyl-4-sulfonyl chloride were reacted according to the general procedure to afford biphenyl-4-sulfonic acid cyclohexyl-ethyl-amide (**7**) (53 % yield) as a white crystalline solid, m.p. 119-120°C (cyclohexane).

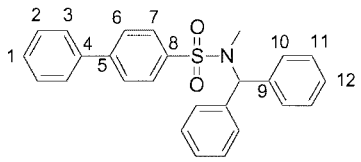
LR-MS (ESI+): m/z 344.4 $[M+H]^+$, 366.4 $[M+Na]^+$

HR-MS (ESI+): m/z 344.1687, 40 % $[M+H]^+$; 366.1506, 95 % $[M+Na]^+$; 709.3137, 100 % $[2M+Na]^+$ (calculated: 366.1498, error: 2 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 7.88 (s, 4H, aromatic H), 7.73 (d, 2H, aromatic H, J = 7.5 Hz), 7.55-7.40 (m, 3H, H_1 and 2 H aromatic), 3.65 (tt, 1H, H_9, J = 11.1 and 4.0 Hz), 3.25 (q, 2H, $C\underline{H}_2CH_3$, J = 6.9 Hz), 1.70 (d, 2H, J = 12.0 Hz), 1.58-1.22 (m, 9H), 1.20 (t, 3H, $CH_2C\underline{H}_3$, J = 7.0 Hz), 1.10-1.02 (m, 1H, one H_{12}).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 143.8 (quaternary C), 140.1 (quaternary C), 138.3 (quaternary C), 129.1, 128.5 (C_1), 127.4, 127.1, 127.0, 57.5 (C_9), 38.0 ($C\underline{H}_2CH_3$), 31.0 (C_{10}), 25.5 (C_{11}), 24.7 (C_{12}), 17.6 ($CH_2C\underline{H}_3$).

Biphenyl-4-sulfonic acid benzhydryl-methyl-amide (8)



N-(Diphenylmethyl)methyl-amine and biphenyl-4-sulfonyl chloride were reacted according to the general procedure to afford biphenyl-4-sulfonic acid benzhydryl-methyl amide (**8**) (30 % yield) as a white crystalline solid, m.p. 97-99°C (diethyl ether).

LR-MS (ESI+): m/z 436.4 $[M+Na]^+$

HR-MS (ESI+): m/z 436.1342, 100 % $[M+Na]^+$; 849.2799 $[2M+Na]^+$ (calculated: 436.1341, error: 0.15 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.70 (dt, 2H, biphenyl aromatic H, J = 8.7 and 1.8 Hz), 7.55-7.48 (m, 4H, biphenyl aromatic H), 7.43-7.33 (m, 3H, H_1 and biphenyl aromatic H), 7.20-7.15 (m, 6H, H_{10} or H_{11} and H_{12}), 7.05-7.00 (m, 4H, H_{10} or H_{11}), 6.43 (s, 1H, CH), 2.66 (s, 3H, CH_3).

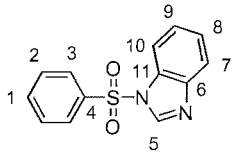
^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 144.2 (quaternary C), 138.4 (quaternary C), 137.5 (quaternary C), 137.3 (C_9), 128.0 (C_{12} or C aromatic biphenyl), 127.8 (C_{10} or C_{11}), 127.3 (2 peaks, C_1 and C_{10} or C_{11}), 126.7 (C_{12} or C aromatic biphenyl), 126.6 (C_{12} or C aromatic biphenyl), 126.4 (C_{12} or C aromatic biphenyl), 126.3 (C_{12} or C aromatic biphenyl).

*General procedure for the synthesis of sulfonamides **3**, **6** and **9**.*

Triethylamine (3.3 mmol, 0.33g, 1.1 eq.) was added to a solution of benzimidazole (3 mmol, 0.35g, 1 eq.) in THF (15 mL) and the mixture was stirred at room temperature for five minutes. The appropriate sulfonyl chloride (3 mmol, 1.0 eq.) was added, resulting in a white of triethyl-ammonium chloride being formed. The mixture was stirred at room temperature overnight and filtered. The filtrate was concentrated *in vacuo*, the residue diluted with ethyl acetate (15 mL) and washed with aqueous HCl (1M) (2x15 mL) and brine (1x15 mL). The aqueous layers were combined and washed with ethyl acetate (1x15 mL). The organic layers were combined, dried over $MgSO_4$ and concentrated *in vacuo*. A small amount of the

crude material was either purified by preparative TLC using CH_2Cl_2 as an eluant or recrystallized, depending on its purity to afford neutral sulfonamides **7**-**9**.

*1-Benzenesulfonyl-1H-benzoimidazole (**3**)*[99]



Benzenesulfonyl chloride and benzimidazole were reacted according to general procedure to afford 1-benzenesulfonyl-1H-benzoimidazole (**3**) (36 % yield) as a white crystalline solid, m.p. 98-99°C (petroleum ether, literature: 95°C in petroleum ether[99]).

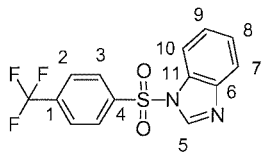
LR-MS (ESI+): m/z 259.1 $[M+H]^+$

HR-MS (ESI+): m/z 259.0537 $[M+H]^+$ (calculate : 259.0536, error : 0.48 ppm), 539.0839 $[2M+Na]^+$

^{1}H-NMR (300 MHz, MeOD, δ ppm): 8.68 (s, 1H, H_5), 8.11 (dt, 2H, H_3, J = 7.2 and 1.2 Hz), 7.91 (dt, 1H, H_1, J = 7.2 and 1.2 Hz), 7.71-7.66 (m, 2H, H_7 and H_{10}), 7.59 (tt, 2H, H_2, J = 7.2 and 1.5 Hz), 7.41 (qd, 1H, H_8 or H_9, J = 7.5 and 1.2 Hz), 7.38 (qd, 1H, H_8 or H_9, J = 7.5 and 1.2 Hz).[vii]

^{13}C-NMR (75 MHz, MeOD, δ ppm): 144.7 (C_4), 143.4 (C_5), 138.7 (C_6 or C_{11}), 136.3 (C_1), 131.9 (C_6 or C_{11}), 131.1 (C_2), 128.5 (C_3), 127.0 (C_8 or C_9), 126.2 (C_8 or C_9), 121.4 (C_7 or C_{10}), 113.8 (C_7 or C_{10}).

*1-(4-Trifluoromethyl-benzenesulfonyl)-1H-benzoimidazole (**6**)*



4-Trifluoromethylbenzenesulfonyl chloride and benzimidazole were reacted according to general procedure to afford 1-(4-trifluoromethyl-benzenesulfonyl)-

[vii] NB: both qd are actually ddd with peak overlapping.

1H-benzoimidazole (**6**) (52 % yield) as a white crystalline solid, m.p. 100-102°C (cyclohexane).

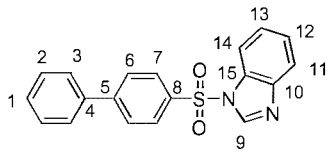
LR-MS (ESI+): m/z 327.2 $[M+H]^+$

HR-MS (ESI+): m/z 327.0409, 100 % $[M+H]^+$; 409.1632, 20 % (?); 675.0597, 20 % $[2M+Na]^+$ (calculated: 237.04095, error: 0.12 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 8.73 (s, 1H, H_5), 8.33 (d, 2H, H_2, J = 8.9 Hz), 7.95-7.91 (m, 3H, H_3 and H_7 or H_{10}), 7.72 (d, 1H, H_7 or H_{10}, J = 7.5 Hz), 7.45 (qd, 1H, H_8 or H_9, J = 7.2 and 1.2 Hz), 7.42 (qd, 1H, H_8 or H_9, J = 7.5 and 1.5).[viii]

^{13}C-NMR (75 MHz, MeOD, δ ppm): 144.8 (quaternary C), 143.4 (C_5), 142.2 (quaternary C), 137.0 (quaternary C), 131.7 (quaternary C), 129.5 (C_2 and C_3), 128.2 (CF_3), 127.2 (C_8 or C_9), 126.5 (C_8 or C_9), 121.7 (C_7 or C_{10}), 113.8 (C_7 or C_{10}).

*1-(Biphenyl-4-sulfonyl)-1H-benzoimidazole (**9**)*



Biphenyl-4-sulfonyl chloride and benzimidazole were reacted according to general procedure to afford 1-(biphenyl-4-sulfonyl)-1H-benzoimidazole (**9**) (47 % yield) as a white crystalline solid, m.p. 146-149°C (methanol).

LR-MS (ESI+): m/z 335.3 $[M+H]^+$

HR-MS (ESI+): m/z 335.0849, 100 % $[M+H]^+$; 691.1453, 33 % $[2M+Na]^+$ (calculated: 335.0849, error: 0.06 ppm)

^{1}H-NMR (300 MHz, MeOD, δ ppm): 8.91 (s, 1H, H_9), 8.24 (dt, 2H, J = 8.7 and 1.9 Hz), 7.96-7.91 (m, 3H), 7.78 (d, 1H, H_1), 7.72-7.67 (m, 2H), 7.52-7.37 (m, 5H).

^{13}C-NMR (75 MHz, MeOD, δ ppm): 146.8 (quaternary C), 143.5 (quaternary C), 142.4 (C_9), 137.7 (quaternary C), 135.1 (quaternary C), 130.1 (quaternary C), 129.1 (aromatic biphenyl), 129.0 (C_1 or aromatic benzimidazole), 128.3 (aromatic biphenyl), 127.9 (aromatic biphenyl), 127.3 (aromatic biphenyl), 125.7 (C_1 or

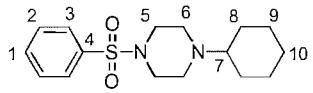
[viii] NB: both qd are actually ddd with peak overlapping.

aromatic benzimidazole), 124.9 (C_1 or aromatic benzimidazole), 120.7 (C_1 or aromatic benzimidazole), 112.4 (C_1 or aromatic benzimidazole).

General procedure for the synthesis of sulfonamides ***13-27.***

Triethylamine (3.3 mmol, 0.33g, 1.1 eq.) was added to a solution of the appropriate amine (3 mmol, 1 eq.) in CH_2Cl_2 (15 mL) and the mixture was stirred at room temperature for five minutes. Then the appropriate sulfonyl chloride (3 mmol, 1 eq.) was added. The mixture was stirred at room temperature for 4 to 12 hours depending on the reactants and the solvent was evaporated *in vacuo*. A small amount of the crude was purified by preparative TLC using CH_2Cl_2 (compounds **14**, **18**, **19**, **21** and **24**), CH_2Cl_2 94/6 (compound **23**) or CH_2Cl_2/MeOH 9/1 (compounds **13**, **15-17**, **20**, **22** and **25-27**) as eluant and recrystallized where possible to afford basic sulfonamides **13-27**.

1-Benzenesulfonyl-4-cyclohexyl-piperazine (**13**)[100]



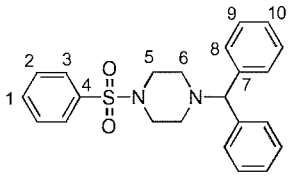
Benzenesulfonyl chloride and 1-cyclohexyl-piperazine were reacted according to general procedure to afford 1-benzenesulfonyl-4-cyclohexyl-piperazine (**13**) (64 % yield) as a white crystalline solid after recrystallization from methanol.

HR-MS (ESI+): m/z 309.1624, 100 % $[M+H]^+$ (calculated: 309.1631, error: -2.19 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.75 (dt, 2H, H_3, J = 6.6 and 1.5 Hz), 7.61-7.48 (m, 3H, H_1 and H_2), 3.02 (t, 4H, H_5 or H_6, J = 4.8 Hz), 2.63 (t, 4H, H_5 or H_6, J = 4.8 Hz), 2.22 (tt, 1H, H_7, J = 10.5 and 2.4 Hz), 1.79-1.74 (m, 4H, H_8), 1.27-1.03 (m, 6H, H_9 and H_{10}).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 134.4 (C_4), 131.7 (C_1), 128.0 (C_2 or C_3), 126.9 (C_2 or C_3), 62.3 (C_7), 47.1 (C_5 or C_6), 45.5 (C_5 or C_6), 27.9 (C_8 or C_9), 25.2 (C_{10}), 24.7 (C_8 or C_9).

1-Benzenesulfonyl-4-benzhydryl-piperazine (**14**)



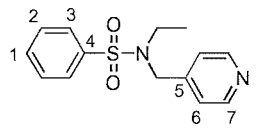
Benzenesulfonyl chloride and 1-(diphenylmethyl)piperazine were reacted according to general procedure to afford 1-benzenesulfonyl-4-benzhydryl-piperazine (**14**) (69 % yield) as a white crystalline solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 393.1640, 100 % $[M+H]^+$; 415.1457, 84 % $[M+Na]^+$; 807.3102, 68 % $[2M+Na]^+$ (calculated: 393.1631, error: 2.1 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.70 (dt, 2H, H_3, J = 6.9 and 1.5 Hz), 7.56-7.45 (m, 3H, H_1 and H_2), 7.26-7.23 (m, 4H, H_8 or H_9), 7.19-7.07 (m, 6H, H_{10} and H_8 or H_9), 4.14 (s, 1H, C<u>H</u>), 2.96 (t, 4H, H_5 or H_6, J = 4.8 Hz), 2.39 (t, 4H, H_5 or H_6, J = 4.9 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 141.0 (C_7), 134.6 (C_4), 131.8 (C_1), 128.0 (C_2 or C_3), 127.6 (C_8 or C_9), 126.8 (C_2 or C_3), 126.7 (C_8 or C_9), 126.2 (C_{10}), 74.6 (<u>C</u>H), 49.9 (C_5 or C_6), 45.2 (C_5 or C_6).

N-Ethyl-N-pyridin-2-ylmethyl-benzenesulfonamide (**15**)



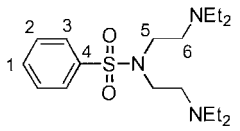
Benzenesulfonyl chloride and 4-(ethylaminomethyl)pyridine were reacted according to general procedure to afford *N*-ethyl-*N*-pyridin-2-ylmethyl-benzenesulfonamide (**15**) (49 yield) as a light brown oil.

HR-MS (ESI+): m/z 277.0998, 100 % $[M+H]^+$; 299.0824, 15 % $[M+Na]^+$ (calculated: 277.1005, error: -2.44 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 8.55 (d, 2H, H_7, J = 5.7 Hz), 7.84 (dt, 2H, H_2 or H_3, J = 6.9 and 1.5 Hz), 7.61-7.50 (m, 3H, H_1 and H_2 or H_3), 7.25 (d, 2H, H_6, J = 5.7 Hz), 4.35 (s, 2H, CH_2), 3.24 (q, 2H, $C\underline{H}_2CH_3$, J = 7.2 Hz), 0.95 (t, 3H, $CH_2C\underline{H}_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 150.1 (C_7), 146.2 (C_5), 139.9 (C_4), 132.7 (C_1), 129.2 (C_2 or C_3), 127.1 (C_2 or C_3), 122.6 (C_6), 50.2 (CH_2), 43.3 ($\underline{C}H_2CH_3$), 13.4 ($CH_2\underline{C}H_3$).

N,N-bis-(2-diethylamino-ethyl)-benzenesulfonamide (**16**)



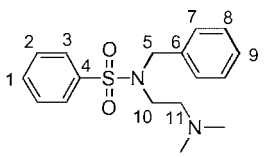
Benzenesulfonyl chloride and *N,N,N',N'*-tetraehtyldiethylenetriamine were reacted according to general procedure to afford *N,N*-bis-(2-diethylamino-ethyl)-benzenesulfonamide (**16**) (65 % yield) as a light yellow oil.

HR-MS (ESI+): m/z 356.2358, 100 % $[M+H]^+$ (calculated: 356.2366, error: -2.18 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.82 (dt, 2H, H_2 or H_3, J = 6.9 and 1.4 Hz), 7.60-7.47 (m, 3H, H_1 and H_2 or H_3), 3.28 (t, 4H, H_5 or H_6, J = 7.5 hz), 2.74 (t, 4H, H_5 or H_6, J = 7.5 Hz), 2.60 (q, 8H, $C\underline{H}_2CH_3$, J = 7.2 Hz), 1.05 (t, 12H, $CH_2C\underline{H}_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 139.4 (C_4), 132.6 (C_1), 129.2 (C_2 or C_3), 127.1 (C_2 or C_3), 52.3 (C_6), 47.4 (C_5), 47.3 ($\underline{C}H_2CH_3$), 11.2 ($CH_2\underline{C}H_3$).

N-Benzyl-N-(2-dimethylamino-ethyl)-benzenesulfonamide (**17**)



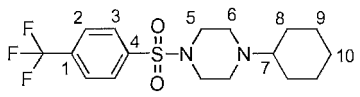
Benzenesulfonyl chloride and *N'*-benzyl-*N,N*-dimethylethylenediamine were reacted according to general procedure to afford *N*-benzyl-*N*-(2-dimethylamino-ethyl)-benzenesulfonamide (**17**) (70 % yield) as a light yellow oil.

HR-MS (ESI+): m/z 319.1477, 100 % $[M+H]^+$ (calculated: 319.1475, error: 0.7 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.85 (dt, 2H, H_2 or H_3, J = 6.9 and 1.5 Hz), 7.59-7.45 (m, 3H, H_1 and H_2 or H_3), 7.28-7.22 (m, 5H, H_7, H_8 and H_9), 4.35 (s, 2H, H_5), 3.16 (t, 2H, H_{10} or H_{11}, J = 7.5 hz), 2.19 (t, 2H, H_{10} or H_{11}, J = 7.5 Hz), 2.01 (s, 6H, CH_3).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 139.1 (C_4), 135.3 (C_6), 131.5 (C_1), 128.1 (aromatic C), 127.6 (aromatic C), 127.4 (aromatic C), 126.9 (C_9), 126.2 (aromatic C), 56.9 (C_{11}), 51.7 (C_5), 44.7 (C_{10}), 44.4 (CH_3).

1-Cyclohexyl-4-(4-trifluoromethyl-benzenesulfonyl)-piperazine (**18**)



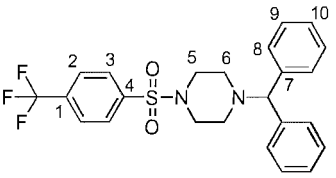
4-Trifluromethylbenzenesulfonyl chloride and 1-cyclohexyl-piperazine were reacted according to general procedure to afford 1-cyclohexyl-4-(4-trifluoromethyl-benzenesulfonyl)-piperazine (**18**) (63 % yield) as a white solid after recrystallization from cyclohexane.

HR-MS (ESI+): m/z 377.1500, 100 % $[M+H]^+$ (calculated: 377.1505, error: -1.26 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.88 (d, 2H, aromatic H, J = 8.1 Hz), 7.79 (d, 2H, aromatic H, J = 8.1 Hz), 3.03 (t, 4H, H_5 or H_6, J = 4.8 Hz), 2.64 (t, 4H, H_5 or H_6, J = 4.9 Hz), 2.24 (t, 1H, H_7, J = 9.9 Hz), 1.78-1.75 (m, 4H, H_8), 1.23-1.07 (m, 6H, H_9 and H_{10}).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 139.2 (C_4), 134.3 (C_1), 128.3 (C_2 and C_3), 126.1 (CF_3), 63.4 (C_7), 48.1 (C_5 or C_6), 46.6 (C_5 or C_6), 28.9 (C_8), 26.1 (C_{10}), 25.7 (C_9).

1-Benzhydryl-4-(4-trifluoromethyl-benzenesulfonyl)-piperazine (**19**)



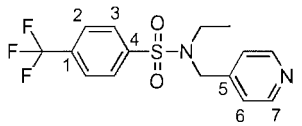
4-Trifluromethylbenzenesulfonyl chloride and 1-(diphenylmethyl)piperazine were reacted according to general procedure to afford 1-benzhydryl-4-(4-trifluoromethyl-benzenesulfonyl)-piperazine (**19**) (69 % yield) as a white solid after recrystallization from cyclohexane.

HR-MS (ESI+): m/z 167.0857, 76 % $[Ph_2CH]^+$; 461.1500, 100 % $[M+H]^+$ (calculated: 461.1505, error: 1.13 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.82 (d, 2H, H_2 or H_3, J = 8.4 Hz), 7.75 (d, 2H, H_2 or H_3, J = 8.4 Hz), 7.27-7.23 (m, 4H, H_8 or H_9), 7.19-7.08 (m, 6H, H_{10} and H_8 or H_9), 4.16 (s, 1H, C<u>H</u>), 2.98 (t, 4H, H_5 or H_6, J = 4.6 Hz), 2.40 (t, 4H, H_5 or H_6, J = 4.8 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 141.8 (C_7), 139.5 (C_4), 134.3 (C_1), 128.8 (C_8 or C_9), 128.3 (C_2 or C_3), 127.7 (C_8 or C_9), 127.2 (C_2 or C_3), 126.3 (C_{10}), 126.2 (<u>C</u>F_3), 75.6 (<u>C</u>H), 50.9 (C_5 or C_6), 46.3 (C_5 or C_6).

N-Ethyl-N-pyridin-4-ylmethyl-4-trifluoromethyl-benzenesulfonamide (**20**)



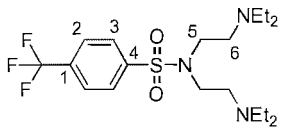
4-Trifluromethylbenzenesulfonyl chloride and 4-(ethylaminomethyl)pyridine were reacted according to general procedure to afford *N*-ethyl-*N*-pyridin-4-ylmethyl-4-trifluoromethyl-benzenesulfonamide (**20**) (40 % yield) as a light yellow oil.

HR-MS (ESI+): m/z 345.0877, 100 % $[M+H]^+$ (calculated: 345.0879, error: -0.51 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 8.57 (bs, 2H, H_7), 7.95 (d, 2H, H_2 or H_3, J = 8.4 Hz), 7.79 (d, 2H, H_2 or H_3, J = 8.4 Hz), 7.27 (bs, 2H, H_6), 4.36 (s, 2H, CH_2), 3.25 (q, 2H, C<u>H</u>$_2$CH$_3$, J = 7.2 Hz), 0.96 (t, 3H, CH$_2$C<u>H</u>$_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 149.0 (C_7), 144.9 (C_5), 142.5 (C_4), 133.3 (C_1), 126.6 (C_2 and C_3), 125.4 (CF_3), 121.7 (C_6), 49.2 (CH_2), 42.4 ($\underline{C}H_2CH_3$), 12.4 ($CH_2\underline{C}H_3$).

N,N-Bis-(2-diethylamino-ethyl)-4-trifluoromethyl-benzenesulfonamide (**21**)



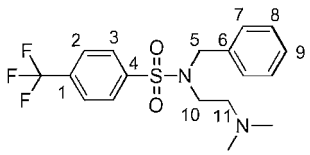
4-Triflurомethylbenzenesulfonyl chloride and *N,N,N',N'*-tetraehtyldiethylenetriamine were reacted according to general procedure to afford *N,N*-bis-(2-diethylamino-ethyl)-4-trifluoromethyl-benzenesulfonamide (**21**) (79 % yield) as a yellow oil.

HR-MS (ESI+): m/z 424.2243, 100 % $[M+H]^+$ (calculated: 424.2240, error: 0.76 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.98 (d, 2H, H_2, J = 8.1 Hz), 7.76 (d, 2H, H_3, J = 8.4 Hz), 3.25 (t, 4H, H_5 or H_6, J = 7.5 Hz), 2.62 (t, 4H, H_5 or H_6, J = 7.5 Hz), 2.51 (q, 8H, $C\underline{H}_2CH_3$, J = 7.1 Hz), 0.99 (t, 12H, $CH_2C\underline{H}_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 143.7 (C_4), 127.6 (C_2 and C_3), 126.1 (CF_3), 52.4 (C_6), 47.6 (C_5), 47.5 ($\underline{C}H_2CH_3$), 11.8 ($CH_2\underline{C}H_3$).

N-Benzyl-N-(2-dimethylamino-ethyl)-4-trifluoromethyl-benzenesulfonamide (**22**)



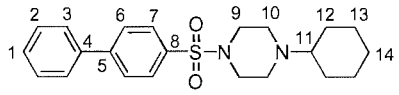
4-Triflurомethylbenzenesulfonyl chloride and *N'*-benzyl-*N,N*-dimethylethylene-diamine were reacted according to general procedure to afford *N*-benzyl-*N*-(2-dimethylamino-ethyl)-4-trifluoromethyl-benzenesulfonamide (**22**) (58 % yield) as a white solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 387.1346, 100 % $[M+H]^+$ (calculated: 387.1349, error: 0.7 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.92 (d, 2H, H_2 or H_3, J = 8.1 Hz), 7.70 (d, 2H, H_2 or H_3, J = 8.1 Hz), 7.28-7.18 (m, 5H, H_7, H_8 and H_9), 4.35 (s, 2H, H_5), 3.16 (t, 2H, H_{10}, J = 7.2 Hz), 2.17 (t, 2H, H_{11}, J = 7.2 Hz), 1.98 (s, 6H, 2 CH_3).

^{13}C-NMR (75 MHz, $CDCl_3D$, δ ppm): 143.9 (C_4), 135.7 (C_1), 128.7 (aromatic C), 128.4 (2 aromatic C), 128.1 (C_9), 127.7 (aromatic C), 126.1 (CF_3), 57.7 (C_{11}), 52.5 (C_5), 45.6 (C_{10}), 45.4 (2 CH_3).

1-(Biphenyl-4-sulfonyl)-4-cyclohexyl-piperazine (**23**)



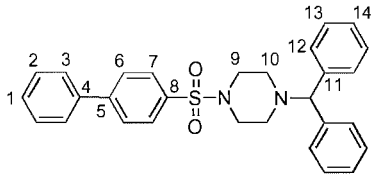
Biphenyl-4-benzenesulfonyl chloride and 1-cyclohexyl-piperazine were reacted according to general procedure to afford 1-(biphenyl-4-sulfonyl)-4-cyclohexyl-piperazine (**23**) (81 % yield) as a white solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 385.1942, 58 % $[M+H]^+$; 408.1714, 100 % (?); 407.1761, 36 % $[M+Na]^+$; 769.3778, 22 % $[2M+H]^+$ (calculated: 385.1944, error: 0.64 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.85 (dt, 2H, aromatic H, J = 8.4 and 1.8 Hz), 7.70 (dt, 2H, aromatic H, J = 8.4 and 1.8 Hz), 7.58 (dt, 2H, aromatic H, J = 6.9 and 1.8 Hz), 7.52-7.38 (m, 3H, H_1 and aromatic H), 3.06 (t, 4H, H_9, J = 4.5 Hz), 2.66 (t, 4H, H_{10}, J = 4.8 Hz), 2.27-2.20 (m, 1H, H_{11}), 1.80-1.75 (m, 2H, H_{12}), 1.27-1.03 (m, 3H, H_{13} and H_{14}).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 145.8 (quaternary aromatic C), 139.4 (quaternary aromatic C), 133.8 (quaternary aromatic C), 129.0 (aromatic C), 128.5 (C_1), 128.4 (aromatic C), 127.6 (aromatic C), 127.3 (aromatic C), 63.3 (C_{11}), 48.1 (C_9 or C_{10}), 46.6 (C_9 or C_{10}), 29.9 (C_{12}), 26.2 (C_{14}), 25.8 (C_{13}).

1-Benzhydryl-4-(biphenyl-4-sulfonyl)-piperazine (**24**)



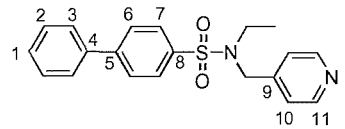
Biphenyl-4-benzenesulfonyl chloride and 1-(diphenylmethyl)piperazine were reacted according to general procedure to afford 1-benzhydryl-4-(biphenyl-4-sulfonyl)-piperazine (**24**) (75 % yield) as a white solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 469.1939, 100 % $[M+H]^+$ (calculated: 469.1944, error: 1.14 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.74 (d, 2H, biphenyl aromatic H, J = 8.1 Hz), 7.67 (d, 2H, biphenyl aromatic H, J = 8.4 Hz), 7.56 (d, 2H, biphenyl aromatic H, J = 7.5 Hz), 7.45-7.33 (m, 3H, H_1 and biphenyl aromatic H), 7.25 (d, 4H, H_{12} or H_{13}, J = 7.2 Hz), 7.18-7.05 (m, H_{12} or H_{13} and H_{14}), 4.51 (s, 1H, C<u>H</u>), 3.00 (bs, 4H, H_9 or H_{10}), 2.41 (t, 4H, H_9 or H_{10}, J = 4.5 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 145.7 (biphenyl quaternary aromatic C), 142.0 (C_{11}), 139.2 (biphenyl quaternary aromatic C), 134.2 (biphenyl quaternary aromatic C), 129.0 (biphenyl aromatic C), 128.6 (C_{12} or C_{13}), 128.5 (C_1), 128.4 (biphenyl aromatic C), 127.7 (C_{12} or C_{13}), 127.6 (biphenyl aromatic C), 127.3 (biphenyl aromatic C), 127.2 (C_{14}), 75.7 (<u>C</u>H), 51.0 (C_9 or C_{10}), 46.3 (C_9 or C_{10}).

Biphenyl-4-sulfonic acid ethyl-pyridin-4-ylmethyl-amide (**25**)



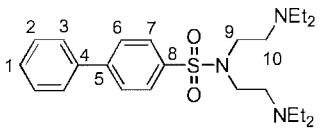
Biphenyl-4-benzenesulfonyl chloride and 4-(ethylaminomethyl)pyridine were reacted according to general procedure to afford biphenyl-4-sulfonic acid ethyl-pyridin-4-ylmethyl-amide (**25**) (48 % yield) as a pale pink solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 353.1315, 100 % $[M+H]^+$; 375.1151, 16 % $[M+Na]^+$; 705.2665, 30 % $[2M+H]^+$ (calculated: 353.1318, error: 0.9 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 8.49 (d, 2H, H_{11}, J = 6.0 Hz), 7.83 (dt, 2H, biphenyl aromatic H, J = 8.1 and 1.5 Hz), 7.66 (d, 2H, biphenyl aromatic H, J = 8.4 and 1.8 Hz), 7.54 (dt, 2H, aromatic H, J = 6.9 and 1.5 Hz), 7.44-7.34 (m, 3H, H_1 and biphenyl aromatic H), 7.20 (d, 2H, H_{10}, J = 5.7 Hz), 4.31 (s, 2H, CH_2), 3.20 (q, 2H, $C\underline{H}_2CH_3$, J = 7.2 Hz), 0.92 (t, 3H, $CH_2C\underline{H}_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 150.1 (C_{11}), 146.2 (C_9), 145.7 (quaternary aromatic C), 139.2 (quaternary aromatic C), 138.3 (quaternary aromatic C), 129.1 (aromatic C), 128.6 (C_1), 127.84 (aromatic C), 127.6 (aromatic C), 127.3 (aromatic C), 122.7(C_{10}), 50.3 (CH_2), 43.4 ($\underline{C}H_2CH_3$), 13.5 ($CH_2\underline{C}H_3$).

Biphenyl-4-sulfonic acid bis-(2-diethylamino-ethyl)-amide (**26**)



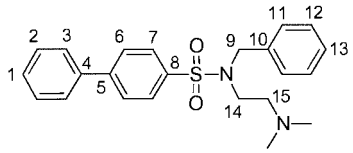
Biphenyl-4-benzenesulfonyl chloride and 4-(ethylaminomethyl)pyridine were reacted according to general procedure to afford biphenyl-4-sulfonic acid bis-(2-diethylamino-ethyl)-amide (**26**) (80 % yield) as a yellow oil.

HR-MS (ESI+): m/z 432.2681, 100 % $[M+H]^+$ (calculated: 432.2679, error: 0.52 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.89 (d, 2H, H_7, J = 8.4 Hz), 7.69 (d, 2H, H_6, J = 8.4 Hz), 7.59 (dt, 2H, H_3, J = 6.6 and 1.8 Hz), 7.50-7.37 (m, 3H, H_1 and H_2), 3.29 (t, 4H, H_9 or H_{10}, J = 7.5 Hz), 2.70 (t, 4H, H_9 or H_{10}, J = 7.6 Hz), 2.57 (q, 8H, $C\underline{H}_2CH_3$, J = 7.2 Hz), 1.03 (t, 12H, $CH_2C\underline{H}_3$, J = 7.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 145.4 (quaternary aromatic C), 139.4 (quaternary aromatic C), 138.2 (quaternary aromatic C), 129.0 (aromatic C), 128.4 (C_1), 127.7 (2 aromatic C), 127.3 (aromatic C), 52.4 (C_{10}), 47.6 (C_9), 47.4 ($\underline{C}H_2CH_3$), 11.6 ($CH_2\underline{C}H_3$).

Biphenyl-4-sulfonic acid benzyl-(2-dimethylamino-ethyl)-amide (**27**)



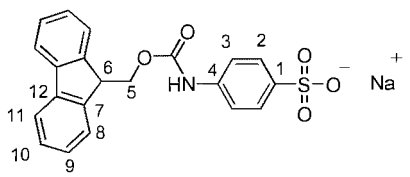
Biphenyl-4-benzenesulfonyl chloride and 4-(ethylaminomethyl)pyridine were reacted according to general procedure to afford biphenyl-4-sulfonic acid benzyl-(2-dimethylamino-ethyl)-amide (**27**) (55 % yield) as a white solid after recrystallization from ethanol.

HR-MS (ESI+): m/z 301.1412, 20 % (?); 395.1783, 100 % $[M+H]^+$ (calculated: 395.1788, error: 1.2 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.94 (d, 2H, H_7, J = 8.4 Hz), 7.73 (d, 2H, H_6, J = 8.4 Hz), 7.65-7.60 (m, 2H, H_2 or H_3), 7.52-7.42 (m, 3H, H_1 and H_2 or H_3), 7.33-7.30 (m, 5H, aromatic H), 4.42 (s, 2H, H_9), 3.23 (t, 2H, H_{14}, J = 7.5 Hz), 2.26 (t, 2H, H_{15}, J = 7.4 Hz), 2.06 (s, 6H, 2 CH_3).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 145.4 (quaternary aromatic C), 139.4 (quaternary aromatic C), 138.6 (quaternary aromatic C), 136.3 (quaternary aromatic C), 129.1 (aromatic C), 128.6 (aromatic C), 128.5 (aromatic C), 127.9 (C_{13}), 127.7 (aromatic C), 127.6 (aromatic C), 127.3 (aromatic C), 57.9 (C_9), 52.7 (C_{15}), 45.8 (C_{14}), 45.5 (2 CH_3).

Synthesis of sodium 4-(9H-fluoren-9-ylmethoxycarbonylamino)-benzenesulfonate



Sulfanilic acid (20 mmol, 3.46 g, 1.2 eq.) was diluted in saturated aqueous $NaHCO_3$ (20 mL), 9-Fluorenylmethyl chloroformate (16.6 mmol, 3.46 g, 1 eq.) was added and the mixture stirred overnight at room temperature. The precipitated yellowish solid was filtered off, washed with diethylether (Et_2O) and dried under reduced

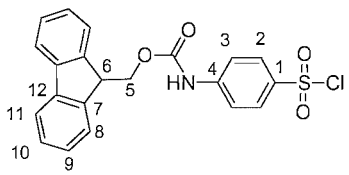
pressure to afford sodium 4-(9H-fluoren-9-ylmethoxycarbonylamino)-benzenesulfonate (60 % yield) as a white solid.

HR-MS (ESI+): m/z 394.07546, 100 % (calculated: 394.07546, error: 0.01ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 9.76 (s, 1H, NH), 7.90 (d, 2H, H_2 or H_3, J = 7.2 Hz), 7.75 (d, 2H, H_2 or H_3, J = 7.5 Hz), 7.50 (d, 2H, aromatic H, J = 8.4 Hz), 7.45-7.32 (m, 6H, aromatic H), 4.48 (d, 2H, H_5, J = 6.3 Hz), 4.29 (t, 1H, H_6, J = 6.6 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 153.4 (C=O), 143.7 (C_7), 142.5 (C_4), 140.8 (C_{12}), 139.1 (C_1), 127.6 (aromatic C), 127.1 (aromatic C), 126.6 (aromatic C), 126.2 (aromatic C), 120.1 (aromatic C), 112.2 (C_3), 65.3 (C_5), 46.6 (C_6).

(4-Chlorosulfonyl-phenyl)-carbamic acid 9H-fluoren-9-ylmethyl ester

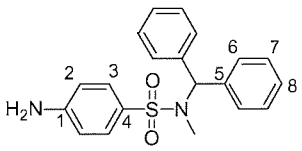


Sodium 4-(9H-fluoren-9-ylmethoxycarbonylamino)-benzenesulfonate was (7.2 mmol, 3.0 g, 1.0 eq.) was diluted in dry DMF/toluene 1:10 (110 mL), thionyl chloride (28.8 mmol, 3.43 g, 4 eq.) was added. Reaction mixture stirred overnight at room temperature, poured into water (80 mL) and then neutralized with saturated $NaHCO_3$. Organic layer was washed with water (2x80 mL) and brine (1x80 mL) dried over $MgSO_4$ and solvent was removed under reduced pressure to yield a yellowish solid. The solid was washed with Et_2O and hexane and filtered off to afford (4-chlorosulfonyl-phenyl)-carbamic acid 9H-fluoren-9-ylmethyl ester (77 %) as white solid.

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.94 (d, 2H, H_2 or H_3, J = 9.0 Hz), 7.79 (d, 2H, H_8 or H_{11}, J = 7.5 Hz), 7.60 (d, 2H, H_8 or H_{11}, J = 7.5 Hz), 7.55 (d, 2H, H_2 or H_3, J = 8.7 Hz), 7.43 (t, 2H, H_9 or H_{10}, J = 7.3 Hz), 7.31 (td, 2H, H_9 or H_{10}, J = 7.5 and 1.2 Hz), 7.02 (bs, 1H, NH), 4.64 (d, 2H, H_5, J = 6.3 Hz), 4.27 (t, 1H, H_6, J = 6.2 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 152.2 (C=O), 144.1 (C_1), 143.3 (C_7 or C_{12}), 141.4 (C_7 or C_{12}), 138.1 (C_4), 128.7 (aromatic C), 128.0 (aromatic C), 127.2 (aromatic C), 124.7 (aromatic C), 120.1 (C_{11}), 118.1 (C_3), 67.3 (C_5), 46.9 (C_6).

4-Amino-N-benzhydryl-N-methyl-benzene sulfonamide ***(11)***



Triethylamine (1.32 mmol, 0.13 g, 1.1 eq.) was added to a solution of *N*-(diphenylmethyl)methylamine (1.32 mmol, 0.26 g, 1.1 eq.) in dry CH_2Cl_2 (20 mL) and the mixture was stirred at room temperature for five minutes. Then (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (1.2 mmol, 0.50 g, 1.0 eq.) was added. The mixture was stirred at room temperature for 30 hours before addition of piperidine (1.20 mmol, 0.10 g, 1.0 eq.). Reaction mixture was stirred 30 min at room temperature, the washed with HCl 1N (2x20 mL) and brine (20 mL). Organic layer was dried over $MgSO_4$ and solvent removed *in vacuo*. When diluting a small amount of crude in CH_2Cl_2 for purification, it appeared that part of it was not soluble. All crude was then diluted in CH_2Cl_2 and precipitate filtered off. Evaporation of filtrate *in vacuo* led to 350 mg of crude as a yellowish solid. Purification of the crude by prepTLC using a mixture CH_2Cl_2:MeOH 9:1 as eluant afford 4-amino-*N*-benzhydryl-*N*-methyl-benzene sulfonamide (**11**) (37 % yield) as a white solid.

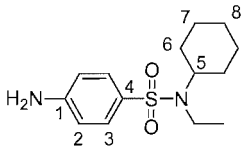
LR-MS: ESI+ m/z 375, 100 % $[M+Na]^+$; ESI- m/z 351, 100 % $[M-H]^-$

HR-MS (ESI+): m/z 375.1147, 52 % $[M+Na]^+$; 453.1288, 78 % $[M+101]^+$; 727.2438, 100 % $[2M+Na]^+$ (calculated: 375.1137, error: 2.62 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.43 (d, 2H, H_3, J = 8.7 Hz), 7.22-7.16 (m, 6H, aromatic H), 7.05-7.01 (m, 4H, aromatic H), 6.51 (dt, 2H, H_2, J = 2.0 Hz), 6.37 (s, 1H, CH), 4.00 (bs, 2H, NH_2), 2.57 (s, 3H, CH_3).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 150.3 (C_1), 138.7 (C_5), 129.3 (C_3), 128.8 (C_6 or C_7), 128.2 (C_6 or C_7), 127.4 (C_8), 113.6 (C_2), 64.0 (CH), 31.0 (CH_3).

4-Amino-N-cyclohexyl-N-ethyl benzenesulfonamide (**10**)



N-Ethylcyclohexylamine (6.0 mmol, 3.0 eq., 0.76g) was diluted in dry CH_2Cl_2 (40 mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (2.0 mmol, 1.0 eq., 0.83 g) was added. Reaction mixture was stirred at room temperature for 3 days, filtered and solvent was removed *in vacuo* to yield a yellowish solid. A small amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 9:1 as eluant to yield 4-Amino-*N*-cyclohexyl-*N*-ethyl benzenesulfonamide (**10**) (82 % yield) as a white solid.

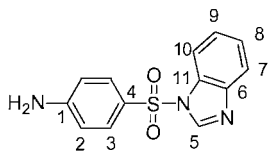
LR-MS: ESI+ m/z 305, 100 % $[M+Na]^+$; ESI- m/z 281, 100 % $[M-H]^-$

HR-MS (ESI+): m/z 305.1298, 19 % $[M+Na]^+$; 383.1441, 64 % $[M+101]^+$; 587.2708, 100 % $[2M+Na]^+$ (calculated: 305.1294, error: 1.25 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.58 (d, 2H, H_3, J = 8.7 Hz), 6.64 (d, 2H, H_2, J = 8.7 Hz), 4.08 (bs, 2H, NH_2), 3.58 (tt, 1H, H_5, J = 11.4 and 3.6 Hz), 3.18 (q, 2H, $C\underline{H}_2CH_3$, J = 7.0 Hz), 1.75-1.69 (m, 2H), 1.65-1.55 (m, 3H), 1.39-1.23 (m, 4H, 3H + one H_8), 1.21 (t, 3H, $CH_2C\underline{H}_3$, J = 7.0 Hz), 1.02 (tt, 1H, one H_8, J = 12.6 and 3.5 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 149.9 (C_1), 130.4 (C_4), 128.8 (C_3), 114.0 (C_2), 57.6 (C_5), 38.1 ($\underline{C}H_2CH_3$), 31.7 (C_6), 26.1 (C_7), 25.4 (C_8), 17.7 ($CH_2\underline{C}H_3$).

4-(Benzimidazole-1-sulfonyl)-phenylamine (**12**)



Benzimidazole (6.0 mmol, 3.0 eq., 0.71 g) was diluted in dry THF (20 mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (2.0 mmol, 1.0 eq., 0.83 g) was added. Reaction mixture was stirred at room temperature for 24 hours, filtered and solvent was removed *in vacuo* to yield yellowish oil. A small

amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 9:1 as eluant to yield 4-(benzimidazole-1-sulfonyl)-phenylamine (**12**) (62 % yield) as a white solid.

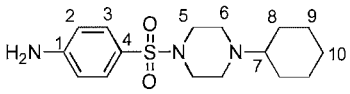
LR-MS: ESI+ m/z 274 $[M+H]^+$; 296 $[M+Na]^+$;'ESI- m/z 281, 100 % $[M-H]^-$

HR-MS (ESI+): m/z 274.0648, 100 % $[M+H]^+$; 328.0733, 19 % $[M+MeOH+NH_4]^+$; 374.0612, 52 % $[M+101]^+$; 569.1057, 37 % $[2M+Na]^+$; 842.1808, 16 % $[3M+Na]^+$ (calculated: 274.0645, error: 1.18 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 8.36 (s, 1H, H_5), 7.85-7.81 (m, 1H, aromatic H), 7.77-7.72 (m, 3H, H_3 and aromatic H), 7.40-7.26 (m, 2H, aromatic H), 6.60 (dt, 2H, H_2, J = 8.7 and 2.4 Hz), 4.33 (bs, 2H, NH_2).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 152.33 (C_1), 144.0 (C_6 or C_{11}), 141.3 (C_5), 130.8 (C_6 or C_{11}), 129.7 (C_3), 125.6 (C_8 or C_9), 124.6 (C_4), 124.4 (C_8 or C_9), 120.8 (C_7), 114.1 (C_2), 112.5 (C_{10}).

*4-(4-Cyclohexyl-piperazine-1-sulfonyl)-phenylamine (**28**)*



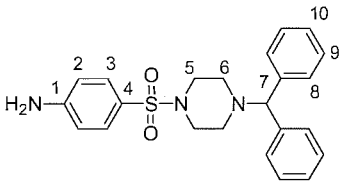
1-Cyclohexyl-piperazine (6.0 mmol, 3 eq., 1.01 g) was diluted in dry CH_2Cl_2 and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (2.0 mmol, 1.0 eq., 0.83 g) was added. Reaction mixture was stirred overnight at room temperature, filtered and solvent was removed *in vacuo* to yield yellowish solid. A small amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 95:5 as eluant to yield 4-(4-cyclohexyl-piperazine-1-sulfonyl)-phenylamine (**28**) (50 % yield) as pale yellow oil.

LR-MS: ESI+ m/z 324 $[M+H]^+$; ESI- 322 $[M-H]^-$

HR-MS (ESI+): m/z 324.1739 $[M+H]^+$ (calculated: 324.1740, error: 0.42 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.43 (d, 2H, H_3, J = 8.7 Hz), 6.58 (d, 2H, H_2, J = 8.7 Hz), 4.06 (bs, 2H, NH_2), 2.90 (t, 4H, H_5 or H_6, J = 4.7 Hz), 2.54 (t, 4H, H_5 or H_6, J = 4.9 Hz), 2.18-2.11 (m, 4H, H_7), 1.25-0.95 (m, 10H, H_8, H_9 and H_{10}).

*4-(4-Benzhydryl-piperazine-1-sulfonyl)-phenylamine (**29**)*



1-(Diphenylmethyl)piperazine (6.0 mmol, 3.0 eq., 1.51 g) was diluted in dry CH_2Cl_2 (20 mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (2.0 mmol, 1.0 eq., 0.83 g) was added. Reaction mixture was stirred for 2 days at room temperature. Piperidine (2.0 mmol, 1.0 eq., 0.17 g) was added and the reaction mixture was stirred at room temperature overnight, filtered and solvent was removed *in vacuo* to yield yellowish pasty solid. A small amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 9:1 as eluant to yield 4-(4-benzhydryl-piperazine-1-sulfonyl)-phenylamine (**29**) (80 % yield) as a white solid.

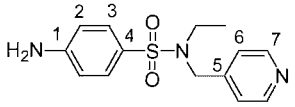
LR-MS: ESI+ m/z 408 $[M+H]^+$; ESI- 407 $[M-H]^-$

HR-MS (ESI+): m/z 408.1739, 100 % $[M+H]^+$; 815.3425, 25 % $[2M+H]^+$ (calculated: 408.1740, error: -0.19 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.45 (d, 2H, H_3, J = 8.4 Hz), 7.26 (dd, 4H, H_8, J = 8.7 and 1.2 Hz), 7.16 (dd, 4H, H_9, J = 7.2 and 0.6 Hz), 7.08 (tt, 2H, H_{10}, J = 7.1 and 1.7 Hz), 6.63 (d, 2H, H_2, J = 8.7 Hz), 4.14 (s, 1H, CH), 4.08 (bs, 2H, NH_2), 2.91 (t, 4H, H_5, J = 4.6 Hz), 2.38 (t, 4H, H_6, J = 4.8 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 150.6 (C_1), 142.1 (C_7), 130.0 (C_3 or C_{10}), 128.6 (C_8 or C_9), 127.7 (C_8 or C_9), 127.1 (C_3 or C_{10}), 123.6 (C_4), 114.0 (C_2), 75.7 (CH), 50.9 (C_6), 46.3 (C_5).

*4-Amino-N-ethyl-N-pyridin-4-yl-benzenesulfonamide (**30**)*



4-(Ethylaminomethyl)pyridine (6.0 mmol, 3.0 eq., 0.82 g) was diluted in dry CH_2Cl_2 (20 mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (2.0 mmol, 1.0 eq., 0.83 g) was added. Reaction mixture was stirred overnight at

room temperature. Reaction mixture was filtered and solvent was removed *in vacuo* to yield yellowish solid. A small amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 94:6 as eluant to yield 4-Amino-*N*-ethyl-*N*-pyridin-4-yl-benzenesulfonamide (**30**) (70 % yield) as a white solid.

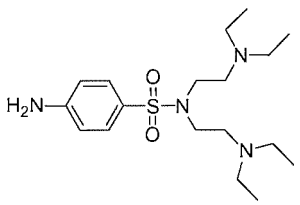
LR-MS: ESI+ m/z 292 $[M+H]^+$; ESI- 290 $[M-H]^-$

HR-MS (ESI+): m/z 292.1117, 100 % $[M+H]^+$; 314.0937, 25 % $[M+Na]^+$; 583.2159, 67 % $[2M+H]^+$; 605.1931, 29 % $[2M+Na]^+$

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 8.54 (bs, 2H, H_7), 7.60 (d, 2H, H_3, J = 6.6 Hz), 7.26 (d, 2H, H_6, J = 5.4 Hz), 6.68 (d, 2H, H_2, J = 6.9 Hz), 4.29 (s, 2H, CH_2), 3.18 (q, 2H, $C\underline{H}_2CH_3$, J = 7.1 Hz), 0.95 (t, 3H, $CH_2C\underline{H}_3$, J = 6.1 Hz).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 149.7 (C_1 or C_5), 148.9 (C_7), 145.8 (C_1 or C_5), 128.2 C_3 or C_6), 126.8 (C_4), 121.7 (C_3 or C_6), 113.1 (C_2), 49.2 ($\underline{C}H_2\Phi$), 42.2 ($\underline{C}H_2CH_3$), 12.4 ($CH_2\underline{C}H_3$).

Amino-N,N-bis-(2-diethylamino-ethyl)-benzenesulfonamide (***31***)

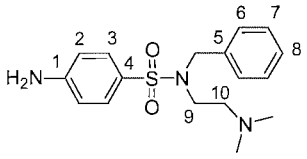


N,N,N',N'-tetraehtyldiethylenetriamine (3.6 mmol, 3.0 eq., 0.78g) was diluted in dry CH_2Cl_2 (20mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (1.2 mmol, 1.0 eq., 0.50g) was added. Reaction mixture was stirred overnight at room temperature. Piperidine (1.2 mmol, 1.0 eq., 0.10g) was added and the reaction mixture was stirred at room temperature for one hour. Reaction mixture was filtered and solvent was removed *in vacuo* to yield yellowish solid. The compound could not be successfully purified.

LR-MS: ESI+ m/z 371 $[M+H]^+$; ESI- 369 $[M-H]^-$

HR-MS (ESI+): m/z 371.2478 $[M+H]^+$ (calculated: 371.2475, error: -0.62 ppm)

*4-Amino-N-benzyl-N-(2-dimethylamino-ethyl)-benzenesulfonamide (**32**)*



N'-Benzyl-*N*,*N*-dimethylethylene-diamine (4.2 mmol, 3.0 eq., 0.75 g) was diluted in dry CH_2Cl_2 (20 mL) and (4-chlorosulfonyl-phenyl)-carbamic acid *9H*-fluoren-9-ylmethyl ester (1.4 mmol, 1.0 eq., 0.58 g) was added. Reaction mixture was stirred overnight at room temperature. Piperidine (1.4 mmol, 1.0 eq., 0.12 g) was added and the reaction mixture was stirred at room temperature overnight. Reaction mixture was filtered and solvent was removed *in vacuo* to yield yellowish pasty solid. A small amount of the crude was purified by prepTLC using a mixture CH_2Cl_2:MeOH 94:6 as eluant to yield 4-Amino-*N*-benzyl-*N*-(2-dimethylamino-ethyl)-benzenesulfonamide (**32**) (75 % yield) as a white solid.

LR-MS: ESI+ m/z 334 $[M+H]^+$; ESI- 332 $[M-H]^-$

HR-MS (ESI+): m/z 334.1578, 100 % $[M+H]^+$; 689.2900, 5 % $[2M+Na]^+$ (calculated: 334.1584, error: -1.72 ppm)

^{1}H-NMR (300 MHz, $CDCl_3$, δ ppm): 7.56 (d, 2H, H_3, J = 8.7 Hz), 7.25-7.18 (m, 5H, aromatic H), 6.62 (d, 2H, H_2, J = 8.7 Hz), 4.23 (s, 2H, CH_2), 4.13 (bs, 2H, NH_2), 3.07 (t, 2H, H_9 or H_{10}, J = 7.5 Hz), 2.18 (t, 2H, H_9 or H_{10}, J = 7.6 Hz), 1.99 (s, 6H, 2 CH_3).

^{13}C-NMR (75 MHz, $CDCl_3$, δ ppm): 161.5 (quaternary aromatic C), 149.6 (quaternary aromatic C), 135.6 (quaternary aromatic C), 128.3 (aromatic C), 127.5 (aromatic C), 127.4 (aromatic C), 126.7 (C_8), 113.1 (C_2), 56.9 (C_{10}), 51.8 (CH_2), 44.7 (C_9), 44.4 (2 CH_3).

II. Descriptors calculations.

II.1 Spartan '02.

The molecular descriptors dipole moment, μ, electron excess charge of the most negatively charged atom, δ_{min}, and molecular surface area, A, were calculated using Spartan'02 (Wavefunction, Inc., Irvine, USA). The geometry of the molecules was first optimized to a minimum using an empirical molecular mechanics based on Merck Molecular Force Field (MMFF). The surface area, A, calculated by the software corresponds to the van der Waals surface of the molecule. μ and δ_{min} were calculated for single point energy at ground state, using a Hartree-Fock method with a 3-21G* basis set, starting from MMFF conformer and AM1 geometry. The settings were as follows:

single point geometry at ground state
Hartree-Fock 3-21G(*)
start from MMFF conformer and AM1 geometry
subject to symmetry
total charge: neutral
multiplicity: singlet

II.2 Hyperchem 7.52.

Molecular surface area and solvent-accessible surface areas were also calculated using the modelling package Hyperchem 7.52 (Hyper-Cube, Waterlooo, Canada). The geometry of the molecules was first optimized to a minimum using an empirical molecular mechanics based on MM+ Force Field. This step was followed by semi-empirical calculations based on the AM1 model. The surface area, A, calculated by the software corresponds to the van der Waals surface of the molecule. The solvent-accessible surface areas were calculated for three different

solvents: water, MeOH and CO_2, with the following probe radii: 1.4 Å, 1.7 Å and 1.8 Å, respectively.[86,89] The settings were as follows:

total charge = 0
spin multiplicity = 0
convergence limit = 1.10^{-8}
iteration limit = 50
spin paiting = RHF
state = lowest

III. Chromatography.

III.1. Preliminary study.

III.1.1. Chemicals.

The preliminary study included thirty-two sulfonamides (compounds **1** to **32**), the synthesis of which is described in the synthesis section. HPLC grade MeOH was purchased from Fisher Scientific (Loughborough, UK) and SFC grade CO_2 from BOC Gases (Guildford, UK). Trietylamine (TEA) and diethylamine (DEA) were purchased from Sigma-Aldrich (Gillingham, UK) and used without further purification.

III.1.2. Stationary phases.

Experiments were undertaken on bare 2-ethyl-pyridine (2-EP), Cyano, Diol and Silica (Si) stationary phases. Column dimensions were 250 mm length by 4.6 mm I.D., 6 µm particle size, 60 Å pore size. All columns were donated by Princeton Chromatography (Cranbury, NJ, USA).

III.1.3. Instrumentation.

SFC analyses were undertaken on a Berger Minigram system (Mettler-Toledo Autochem, Newark, DE, USA) equipped with a Knauer k-2501 variable wavelength UV detector (Knauer, Berlin, Germany); both were donated for the project by Mettler-Toledo AutoChem.

The SFC system was hyphenated to a Platform LCZ mass spectrometer (Waters/Micromass, Manchester, UK) fitted with an electrospray ionisation source (ESI).

In order to guarantee good ionisation of the analytes and prevent crystallisation of residuals in tubing, a make-up flow was pumped into the system by a HP 1050 HPLC system (Agilent, South Queensferry, UK). Both MS and HPLC systems were donated by GlaxoSmithKline (Harlow, UK).

The SFC, MS and HPLC systems were assembled as follows: a T-piece positioned immediately after the UV detector splits the flow into two. One part goes to waste while the rest is directed towards the mass spectrometer. Immediately before the MS-source inlet, a second tee-piece allows for the mixing of the chromatographic flow with the make-up flow from the HPLC pump. Figure 37 is a schematic of the SFC-MS system.

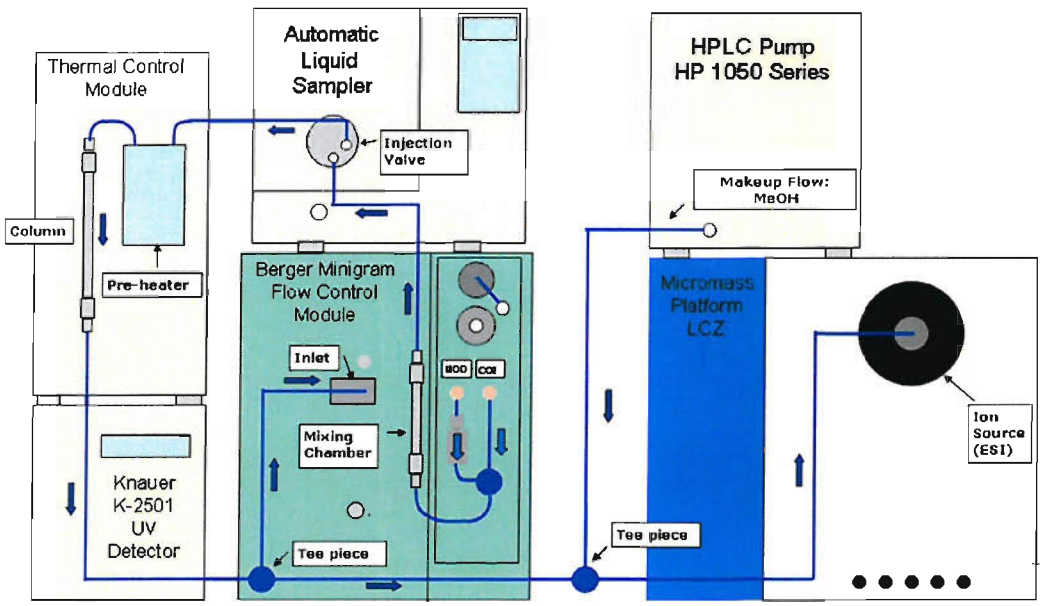


Figure 37. Schematic of the SFC-MS system.

III.1.4. Experimental conditions.

Analytes were dissolved in MeOH at a concentration of *ca.* 0.50 mg mL^{-1}. Chromatographic analyses were undertaken using isocratic elution, outlet pressure was set at 100 bar, oven temperature was 35°C, flow rate was 4 mL min^{-1} with a 4 μL injection volume.

The UV data were recorded at 254 nm for all analytes. Chromatograms were reprocessed using ProNTo software. Retention times (t_R) were measured at the peak apex.

Both positive and negative ion mass spectra were recorded and reprocessed using MassLynx. MS conditions were as follows: capillary voltage 3.5 kV, cone voltage, 20 V, ion energy 0.8 V, multiplier 550 V, analyzer vacuum 2.8×10^{-4} mPa, dissolvation gas flow 500 L h^{-1}, mass range m/z 100–800. The make-up flow consisted of pure MeOH at a rate of 0.1 mL min^{-1}.

III.2. Polycratic study.

III.2.1. Chemicals.

Sulfonamides were synthesized and purified in-house, the synthesis being described in the synthesis section. HPLC grade MeOH was purchased from Fisher Scientific (Loughborough, UK) and SFC grade CO_2 from BOC Gases (Guildford, UK). Ammonium acetate (NH_4OAc) and ethyl-dimethyl-amine (EDMA) were purchased from Sigma-Aldrich (Gillingham, UK) and used without further purification.

III.2.2. Stationary Phase.

Experiments were undertaken using a 2-ethylpyridine (2-EP) stationary phase, 50 mm length by 4.6 mm I.D., 5 μm particle size, 60 Å pore size. All columns were donated by Princeton Chromatography (Cranbury, NJ, USA).

III.2.3. Instrumentation.

See § III.1.3.

III.2.4. Experimental conditions.

Chromatographic analyses were undertaken using isocratic elution at different modifier concentrations in the mobile phase, outlet pressure was set at 100 bar, oven temperature was 35°C, flow rate was 4 mL min^{-1} with a 4 μL injection volume. The UV data were recorded at 254 nm and the chromatograms were reprocessed using ProNTo software.

The modifier consisted of pure methanol, a solution of 0.1 % volume/volume (v/v) of EDMA in MeOH or a solution of 0.6 mM NH_4OAc in MeOH. Each time the modifier was changed, the modifier line was purged with the new solvent and the system equilibrated by injecting 4 μL of pure methanol and pumping the new mobile phase through the column for 10 min. All analytes were dissolved at a concentration of *ca.* 0.25 mg mL^{-1} in MeOH.

The same conditions were applied for each gradient elution experiment. The modifier gradient was as follows: 5 % modifier held for 0.5 min, 5 % to 50 % modifier in 2.5 min (15 % min^{-1}), 50 % to 5 % in 0.5 min (90 % min^{-1}), 5 % held for 0.5 min, total run time 4 min.

Retention times (t_R) were measured at the peak apex. Dead times (t_0) were measured as the retention times if the first negative peak due to unretained solvent. Capacity factors k of analytes were calculated using the following formula:

$$k = \frac{t_R - t_0}{t_0}$$

III.2.5. Retention characteristics.

Three replicate polycratic studies of each analyte were undertaken to ensure reproducibility of the data. The mean values of log k for each percentage φ were calculated using SigmaPlot for windows version 10.0 (Systat Software, Inc., Richmond, USA). The linearity of the log k = f(φ) relationship for each analyte was assessed by linear regression analysis: regression equations were derived giving values of log k_0 and S. The statistical validity of the results was assessed by calculation of squared correlation coefficients (r^2), standard errors of estimates, significance levels of each term of the equations (P), and values of the t-test of significance (t). φ_0 values were subsequently determined.

III.2.6. Multiple regression analysis.

Log k_0 and φ_0 were regressed against the three calculated molecular descriptors obtained with Spartan'02 to derive model QSRR equations to be used for retention predictions. Multiple regression analysis equations were derived using Microsoft Excel 2003 (Microsoft, Redmond, USA) and the statistical validity of the results was assessed by calculation of multiple correlation coefficients (R), standard errors of estimate (s), significance levels of each term of the whole equations (P) and values of the F-test of significance (F).

III.3. Additive study.

III.3.1. Chemicals.

Sulfonamides (compounds **16**, **26**, **31** and **32**) were synthesized in-house according to protocols described in the synthesis section. (*S*)-(+)-Naproxen and (*2S,3S*)-(+)-*cis*-diltiazem hydrochloride were purchased from Fluka (Buchs, Switzerland); (±)-Atenolol and ammonium acetate (NH_4OAc) were purchased from Sigma-Aldrich (Gillingham, UK). Purchased compounds were used without further purification. HPLC grade MeOH was purchased from Fisher Scientific (Loughborough, UK) and SFC grade CO_2 from BOC Gases (Guildford, UK).

III.3.2. Stationary Phases.

Experiments were undertaken on bare silica (Si), 2-ethylpyridine (2-EP) and endcapped 2-ethyl-pyridine (2-EP-EC) stationary phases. Column dimensions were 50 mm length by 4.6 mm I.D., 5 μm particle size, 60 Å pore size. All columns were donated by Princeton Chromatography (Cranbury, NJ, USA).

III.3.3. Instrumentation.

See § III.1.3.

III.3.4. Experimental conditions.

All analytes were dissolved at a concentration of *ca.* 0.25 mg mL^{-1}. Sulfonamides were directly diluted in the sample solvent, while Naproxen, Diltiazem and Atenolol were first diluted in DMSO at a concentration of *ca.* 10 mg mL^{-1} and then diluted to *ca.* 0.25 mg mL^{-1} with the sample solvent. The sample solvent was either pure MeOH or a solution of NH_4OAc in MeOH, as described in the text.

Chromatographic analyses were undertaken using isocratic elution at 10 % v/v modifier in CO_2; the outlet pressure was set at 100 bar, oven temperature 35°C, flow rate 4 mL min^{-1}, using a 4 μL injection. The modifier consisted of either pure methanol or a solution of ammonium acetate in methanol. Each time the modifier was changed, the modifier line was purged with the new modifier and then the system was equilibrated by injecting 4 μL of pure methanol and pumping the new mobile phase through the column for 10 min.

The UV data were recorded at 254 nm for all analytes, with the exception of Atenolol for which signal was recorded at 220 nm. Chromatograms were reprocessed using ProNTo software.

Both positive and negative ion mass spectra were recorded and reprocessed using MassLynx. MS conditions were as follows: capillary voltage 3.5 kV, cone voltage, 20 V, ion energy 0.8 V, multiplier 550 V, analyzer vacuum 2.8×10^{-4} mPa, dissolvation gas flow 500 L h^{-1}, mass range m/z 100–800. The make-up flow consisted of pure MeOH at a rate of 0.1 mL min^{-1}.

Retention times (t_R) were measured at peak apex. Dead times (t_0) were measured as the retention times if the first negative peak due to unretained solvent (in the conditions used, t_0 was of *ca.* 0.16 min). Capacity factors k of analytes were calculated using the following formula:

$$k = \frac{t_R - t_0}{t_0}$$

The effect of additive concentration on the peak shape of the analytes has been evaluated by measuring the peak asymmetry (As.) according to the United States Pharmacopoeia (USP):

$$As._{USP} = \frac{w_{5\%}}{2 \times f}$$

where $w_{5\%}$ is the peak width at 5 % of peak height and f is the first half width at 5 % of peak height. The measurement was carried out using the functionality of ProNTo software.

References

1. Korfmacher, W. A., Principles and applications of LC-MS in new drug discovery. *Drug Discovery Today* **2005,** 10, (20), 1357-1367.

2. Lim, C. K.; Lord, G., Current developments in LC-MS for pharmaceutical analysis. *Biological & Pharmaceutical Bulletin* **2002,** 25, (5), 547-557.

3. Welch, C. J.; Leonard, W. R.; DaSilva, J. O.; Biba, M.; Albaneze-Walker, J.; Henderson, D. W.; Laing, B.; Mathre, D. J., Preparative chiral SFC as a green technology for rapid access to enantiopurity in pharmaceutical process research. *LC GC Europe* **2005,** 18, (5), 264-272.

4. Zheng, J.; Pinkston, J. D.; Zoutendam, P. H.; Taylor, L. T., Feasibility of supercritical fluid chromatography/mass spectrometry of polypeptides with up to 40-mers. *Analytical Chemistry* **2006,** 78, (5), 1535-1545.

5. Pinkston, J. D.; Wen, D.; Morand, K. L.; Tirey, D. A.; Stanton, D., Comparison of LC/MS and SFC/MS for Screening of a Large and Diverse Library of Pharmaceutically relevant compounds. *Analytical Chemistry* **2006,** 78, 7467-7472.

6. Zhao, Y.; Sandra, P.; Woo, G.; Thomas, S.; Gahm, K.; Semin, D., Packed Column Supercritical FLuid Chromatography-Mass Spectrometry for Drug Discovery Applications. *LC GC Europe* **2005,** 17, (4), 224-238.

7. Neue, U. D., *HPLC Columns. Theory, Technology and Pratice.* First ed.; Wiley-VCH, Inc.: 1997.

8. Héberger, K., Quantitative structure–(chromatographic) retention relationships. *Journal of Chromatography A* **2007,** 1158, (1-2), 273-305.

9. Klesper, E.; Corwin, A.; Turner, D. A., High Pressure Gas Chromatography above Critical Temperature. *Journal of Organic Chemistry* **1962,** 27, (2), 700-701.

10. Harris, C. M., The SFC comeback - Pharmaceuticals give supercritical fluid chromatography a fighting chance. *Analytical Chemistry* **2002,** 74, (3), 87A-91A.

11. Harrsch, P. B.; Bente, P. F.; Berger, T. A. Berger Instruments Business Briefing Report. www.bergersfc.com

12. Garzotti, M.; Hamdan, M., Supercritical fluid chromatography coupled to electrospray mass spectrometry: a powerful tool for the analysis of chiral mixtures.

Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences **2002,** 770, (1-2), 53-61.

13. Gere, D. R., Supercritical Fluid Chromatography. *Science* **1983,** 222, (4621), 253-259.

14. Yaku, K.; Morishita, F., Separation of drugs by packed-column supercritical fluid chromatography. *Journal of Biochemical and Biophysical Methods* **2000,** 43, (1-3), 59-76.

15. Pyo, D., New device for controlling the amount of methanol in carbon dioxide mobile phase for supercritical fluid chromatography. *Microchemical Journal* **2001,** 68, (2-3), 183-188.

16. Smith, R. M., Supercritical fluids in separation science - the dreams, the reality and the future. *Journal of Chromatography A* **1999,** 856, (1-2), 83-115.

17. Phinney, K. W., SFC of drug enantiomers. *Analytical Chemistry* **2000,** 72, (5), 204A-211A.

18. Berger, T. A., Separation of polar solutes by packed column supercritical fluid chromatography. *Journal of Chromatography A* **1997,** 785, (1-2), 3-33.

19. Berger, T. A.; Fogleman, K.; Staats, T.; Bente, P.; Crocket, I.; Farrell, W.; Osonubi, M., The development of a semi-preparatory scale supercritical-fluid chromatograph for high-throughput purification of 'combi-chem' libraries. *Journal of Biochemical and Biophysical Methods* **2000,** 43, (1-3), 87-111.

20. Berger, T. A.; Wilson, W. H., High-speed screening of combinatorial libraries by gradient packed-column supercritical fluid chromatography. *Journal of Biochemical and Biophysical Methods* **2000,** 43, (1-3), 77-85.

21. Chester, T. L.; Pinkston, J. D., Supercritical fluid and unified chromatography. *Analytical Chemistry* **2002,** 74, (12), 2801-2811.

22. He, P.; Yang, Y., Studies on the long-term thermal stability of stationary phases in subcritical water chromatography. *Journal of Chromatography A* **2003,** 989, (1), 55-63.

23. Yarita, T.; Nakajima, R.; Shibukawa, M., Superheated water chromatography of phenols using poly(styrene-divinylbenzene) packings as a stationary phase. *Analytical Sciences* **2003,** 19, (2), 269-272.

24. Gyllenhaal, O.; Karlsson, A., Enantiomeric separations of amino alcohols by packed-column SFC on Hypercarb with L-(+)-tartaric acid as chiral selector. *Journal of Biochemical and Biophysical Methods* **2002,** 54, (1-3), 169-185.

25. del Nozal, M. J.; Toribio, L.; Bernal, J. L.; Nieto, E. M.; Jimenez, J. J., Separation of albendazole sulfoxide enantiomers by chiral supercritical-fluid chromatography. *Journal of Biochemical and Biophysical Methods* **2002,** 54, (1-3), 339-345.

26. Anderson, M. E.; Aslan, D.; Clarke, A.; Roeraade, J.; Hagman, G., Evaluation of generic chiral liquid chromatography screens for pharmaceutical analysis. *Journal of Chromatography A* **2003,** 1005, (1-2), 83-101.

27. Deschamps, F. S.; Gaudin, K.; Lesellier, E.; Tchapla, A.; Ferrier, D.; Baillet, A.; Chaminade, P., Response enhancement for the evaporative light scattering detection for the analysis of lipid classes and molecular species. *Chromatographia* **2001,** 54, (9-10), 607-611.

28. Wenclawiak, B.; Otterbach, A., Carbon-based quantitation of pyrethrins by supercritical-fluid chromatography. *Journal of Biochemical and Biophysical Methods* **2000,** 43, (1-3), 197-207.

29. Brunelli, C.; Górecki, T.; Zhao, Y.; Sandra, P., Corona-Charged Aerosol Detection in Supercritical Fluid Chromatography for Pharmaceutical Analysis. *Analytical Chemistry* **2007,** 79, 2472-2482.

30. Randall, L. G.; Wahrhaftig, A. L., Direct Coupling of a Dense (Supercritical) Gas-Chromatograph to a Mass-Spectrometer Using a Supersonic Molecular-Beam Interface. *Review of Scientific Instruments* **1981,** 52, (9), 1283-1295.

31. Albert, K., Supercritical fluid chromatography proton nuclear magnetic resonance spectroscopy coupling. *Journal of Chromatography A* **1997,** 785, (1-2), 65-83.

32. Horton, D. A., The combinatorial synthesis of bicyclic privileged structures or privileged substructures. *Chemical Reviews* **2003,** 103, (3), 893-930.

33. Bannwarth, W.; Hinzen, B., *Combinatorial Chemistry. From Theory to Application.* 2nd ed.; Wiley-VCH: Weinheim, 2006; Vol. 26.

34. Deprez-Poulain, R.; Deprez, B., Facts, figures and trends in lead generation. *Current Topics in Medicinal Chemistry* **2004,** 4, (6), 569-580.

35. Schreiber, S. L., Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* **2000,** 287, (5460), 1964-1969.

36. Gaudilliere, B., Chapter 28. To market, to market - 2000. *Annual Reports in Medicinal Chemistry, Vol 36* **2001,** 36, 293-318.

37. Jorgensen, W. L., The many roles of computation in drug discovery. *Science* **2004,** 303, (5665), 1813-1818.

38. Rupasinghe, C. N.; Spaller, M. R., The interplay between structure-based design and combinatorial chemistry. *Current Opinion in Chemical Biology* **2006,** 10, (3), 188-193.

39. Rose, S.; Stevens, A., Computational design strategies for combinatorial libraries. *Current Opinion in Chemical Biology* **2003,** 7, (3), 331-339.

40. Walters, W. P.; Murcko, M. A., Prediction of 'drug-likeness'. *Advanced Drug Delivery Reviews* **2002,** 54, (3), 255-271.

41. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **1997,** 23, (1-3), 3-25.

42. Raevsky, O. A.; Schaper, K. J.; Seydel, J. K., H-bond contribution to octanol-water partition coefficients of polar compounds. *Quantitative Structure-Activity Relationships* **1995,** 14, (5), 433-436.

43. Abraham, M. H., Hydrogen-Bonding .31. Construction of a Scale of Solute Effective or Summation Hydrogen-Bond Basicity. *Journal of Physical Organic Chemistry* **1993,** 6, (12), 660-684.

44. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J., A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of Combinatorial Chemistry* **1999,** 1, (1), 55-68.

45. Evans, B. E.; Rittle, K. E.; Bock, M. G.; Dipardo, R. M.; Freidinger, R. M.; Whitter, W. L.; Lundell, G. F.; Veber, D. F.; Anderson, P. S.; Chang, R. S. L.; Lotti, V. J.; Cerino, D. J.; Chen, T. B.; Kling, P. J.; Kunkel, K. A.; Springer, J. P.; Hirshfield, J., Methods for Drug Discovery - Development of Potent, Selective, Orally Effective Cholecystokinin Antagonists. *Journal of Medicinal Chemistry* **1988,** 31, (12), 2235-2246.

46. Bondensgaard, K.; Ankersen, M.; Thogersen, H.; Hansen, B. S.; Wulff, B. S.; Bywater, R. P., Recognition of privileged structures by G-protein coupled receptors. *Journal of Medicinal Chemistry* **2004,** 47, (4), 888-899.

47. Hajduk, P. J.; Bures, M.; Praestgaard, J.; Fesik, S. W., Privileged molecules for protein binding identified from NMR- based screening. *Journal of Medicinal Chemistry* **2000,** 43, (18), 3443-3447.

48. Bemis, G. W.; Murcko, M. A., The properties of known drugs .1. Molecular frameworks. *Journal of Medicinal Chemistry* **1996,** 39, (15), 2887-2893.

49. Bemis, G. W., Properties of known drugs. 2. Side chains. *Journal of Medicinal Chemistry* **1999,** 42, (25), 5095-5099.

50. Domagk, G., A new class of disinfectants. *Deutsche Medizinische Wochenschrift* **1935,** 61, 829-832.

51. Tréfouël, J.; Nitti, F.; Bovet, D., Action of p-aminophenylsulfamide in experimental streptococcus infections of mice and rabbits. *Comptes Rendus des Seances de la Societe de Biologie et de Ses Filiales* **1935,** 120, 756-758.

52. Dost, K.; Jones, D. C.; Davidson, G., Determination of sulfonamides by packed column supercritical fluid chromatography with atmospheric pressure chemical ionisation mass spectrometric detection. *Analyst* **2000,** 125, (7), 1243-1247.

53. Chazalette, C.; Riviere-Baudet, M.; Supuran, C. T.; Scozzafava, A., Carbonic anhydrase inhibitors: Allylsulfonamide, styrene sulfonamide, N-allyl sulfonamides and some of their Si, Ge, and B derivatives. *Journal of Enzyme Inhibition* **2001,** 16, (6), 475-489.

54. Drews, J., Drug discovery: A historical perspective. *Science* **2000,** 287, (5460), 1960-1964.

55. Johnson, D. C.; Widlanski, T. S., Cerium(III) chloride-mediated reactions of sulfonamide dianions. *Journal of Organic Chemistry* **2003,** 68, (13), 5300-5309.

56. Enders, D.; Wallert, S.; Runsink, J., Asymmetric synthesis of b-amino cyclohexyl sulfonates, b-sultams and g-sultones. *Synthesis* **2003**, (12), 1856-1868.

57. Hanessian, S.; Sailes, H.; Therrien, E., Synthesis of functionally diverse bicyclic sulfonamides as constrained proline analogues and application to the design of potential thrombin inhibitors. *Tetrahedron* **2003,** 59, (35), 7047-7056.

58. Moree, W. J., Synthesis of Peptidosulfinamides and Peptidosulfonamides - Peptidomimetics Containing the Sulfinamide or Sulfonamide Transition-State Isostere. *Journal of Organic Chemistry* **1995,** 60, (16), 5157-5169.

59. Leung, D., Protease inhibitors: Current status and future prospects. *Journal of Medicinal Chemistry* **2000,** 43, (3), 305-341.

60. Levin, J. I.; Du, M. T.; DiJoseph, J. F.; Killar, L. M.; Sung, A.; Walter, T.; Sharr, M. A.; Roth, C. E.; Moy, F. J.; Powers, R.; Jin, G. X.; Cowling, R.; Skotnicki, J. S., The discovery of anthranilic acid-based MMP inhibitors. Part 1: SAR of the 3-position. *Bioorganic & Medicinal Chemistry Letters* **2001,** 11, (2), 235-238.

61. Inoue, J.; Cui, Y. S.; Sakai, O.; Nakamura, Y.; Kogiso, H.; Kador, P. F., Synthesis and aldose reductase inhibitory activities of novel N-nitromethylsulfonanilide derivatives. *Bioorganic & Medicinal Chemistry* **2000,** 8, (8), 2167-2173.

62. Ryckebusch, A.; Deprez-Poulain, R.; Debreu-Fontaine, M. A.; Vandaele, R.; Mouray, E.; Grellier, P.; Sergheraert, C., Parallel synthesis and anti-malarial activity of a sulfonamide library. *Bioorganic & Medicinal Chemistry Letters* **2002,** 12, (18), 2595-2598.

63. Binisti, C.; Assogba, L.; Touboul, E.; Mounier, C.; Huet, J.; Ombetta, J. E.; Dong, C. Z.; Redeuilh, C.; Heymans, F.; Godfroid, J. J., Structure-activity relationships in platelet-activating factor (PAF). 11-From PAF-antagonism to phospholipase A(2) inhibition: syntheses and structure-activity relationships in 1-arylsulfamido-2-alkylpiperazines. *European Journal of Medicinal Chemistry* **2001,** 36, (10), 809-828.

64. Booth, G., Aniline-2,4,6-Trisulphonylchloride and Derived Sulfonamides. *Synthetic Communications* **1983,** 13, (8), 659-661.

65. Uehling, D. E.; Donaldson, K. H.; Deaton, D. N.; Hyman, C. E.; Sugg, E. E.; Barrett, D. G.; Hughes, R. G.; Reitter, B.; Adkison, K. K.; Lancaster, M. E.; Lee, F.; Hart, R.; Paulik, M. A.; Sherman, B. W.; True, T.; Cowan, C., Synthesis and evaluation of potent and selective beta(3) adrenergic receptor agonists containing acylsulfonamide, sulfonylsulfonamide, and sulfonylurea carboxylic acid isosteres. *Journal of Medicinal Chemistry* **2002,** 45, (3), 567-583.

66. Greene, T. W., *Protective Groups in Organic Synthesis*. 3rd ed.; John Wiley & Sons Inc.: 1999.

67. Chino, M.; Wakao, M.; Ellman, J. A., Efficient method to prepare hydroxyethylamine-based aspartyl protease inhibitors with diverse P-1 side chains. *Tetrahedron* **2002,** 58, (32), 6305-6310.

68. Jameson, G. W.; Elmore, D. T., Affinity chromatography of bovine trypsin. *Biochemical Journal* **1974,** 141, 555-565.

69. Kaliszan, R., *Structure and Retention in Chromatography. A Chemometric Approach.* Harwood Academic Publishers: Amsterdam, 1997.

70. Martire, D. E.; Boehm, R. E., Unified Molecular Theory of Chromatography and Its Application to Supercritical Fluid Mobile Phases .1. Fluid Liquid (Absorption) Chromatography. *Journal of Physical Chemistry* **1987,** 91, (9), 2433-2446.

71. Martire, D. E., Unified Theory of Adsorption Chromatography - Gas, Liquid and Supercritical Fluid Mobile Phases. *Journal of Chromatography* **1988,** 452, 17-30.

72. Schoenmakers, P. J., Thermodynamic Model for Supercritical Fluid Chromatography. *Journal of Chromatography* **1984,** 315, 1-18.

73. Bartle, K. D.; Clifford, A. A.; Kithinji, J. P.; Shilstone, G. F., Studies of the Temperature-Dependence of Retention in Supercritical Fluid Chromatography. *Journal of the Chemical Society-Faraday Transactions I* **1988,** 84, 4487-4493.

74. West, C.; Lesellier, E., Characterization of stationary phases in subcritical fluid chromatography by the solvation parameter model I. Alkylsiloxane-bonded stationary phases. *Journal of Chromatography A* **2006,** 1110, (1-2), 181-190.

75. West, C.; Lesellier, E., Characterization of stationary phases in subcritical fluid chromatography by the solvation parameter model II. Comparisons tools. *Journal of Chromatography A* **2006,** 1110, (1-2), 191-199.

76. West, C.; Lesellier, E., Characterization of stationary phases in subcritical fluid chromatography with the solvation parameter model III. Polar stationary phases. *Journal of Chromatography A* **2006,** 1110, (1-2), 200-213.

77. West, C.; Lesellier, E., Characterization of stationary phases in subcritical fluid chromatography with the solvation parameter model IV. Aromatic stationary phases. *Journal of Chromatography A* **2006,** 1115, (1-2), 233-245.

78. Alvarez, G. A.; Baumann, W., Dielectric interactions and the prediction of retention times of pesticides in supercritical fluid chromatography with CO2. *Zeitschrift Fur Naturforschung Section a - A Journal of Physical Sciences* **2005,** 60, (1-2), 61-69.

79. Fatemi, M. H.; Baher, E., Prediction of retention factors in supercritical fluid chromatography using artificial neural network. *Journal of Analytical Chemistry* **2005,** 60, (9), 860-865.

80. Sadek, P. C.; Carr, P. W.; Doherty, R. M.; Kamlet, M. J.; Taft, R. W.; Abraham, M. H., Study of Retention Processes in Reversed-Phase High-Performance Liquid-Chromatography by the Use of the Solvatochromic Comparison Method. *Analytical Chemistry* **1985,** 57, (14), 2971-2978.

81. Abraham, M. H., Scales of Solute Hydrogen-Bonding - Their Construction and Application to Physicochemical and Biochemical Processes. *Chemical Society Reviews* **1993,** 22, (2), 73-83.

82. Baczek, T.; Kaliszan, R., Combination of linear solvent strength model and quantitative structure-retention relationships as a comprehensive procedure of approximate prediction of retention in gradient liquid chromatography. *Journal of Chromatography A* **2002,** 962, (1-2), 41-55.

83. Baczek, T.; Kaliszan, R., Predictive approaches to gradient retention based on analyte structural descriptors from calculation chemistry. *Journal of Chromatography A* **2003,** 987, (1-2), 29-37.

84. Kaliszan, R.; van Straten, M. A.; Markuszewski, M.; Cramers, C. A.; Claessens, H. A., Molecular mechanism of retention in reversed-phase high-performance liquid chromatography and classification of modern stationary phases by using quantitative structure-retention relationships. *Journal of Chromatography A* **1999,** 855, (2), 455-486.

85. Kaliszan, R., Retention data from affinity high-performance liquid chromatography in view of chemometrics. *Journal of Chromatography B* **1998,** 715, (1), 229-244.

86. Lee, B.; Richards, F. M., Interpretation of Protein Structures - Estimation of Static Accessibility. *Journal of Molecular Biology* **1971,** 55, (3), 379-400.

87. Bondi, A., van der Waals Volumes and Radii. *Journal of Physical Chemistry* **1964,** 68, (3), 441-451.

88. Bondi, A., *Physical Properties of Molecular Crystals, Liquids, and Glasses.* John Wiley & Sons, Inc.: New York, 1968.

89. Kovacs, H.; Mark, A. E.; Johansson, J.; Vangunsteren, W. F., The Effect of Environment on the Stability of an Integral Membrane Helix - Molecular-Dynamics Simulations of Surfactant Protein-C in Chloroform, Methanol and Water. *Journal of Molecular Biology* **1995,** 247, (4), 808-822.

90. Hehre, W. J., *A Guide to Molecular Mechanics and Quantum Chemical Calculations.* Wavefunction: Irvine, 2001.

91. Snyder, L. R., *Principles of Adsorption Chromatography*. First ed.; Marcel Dekker: New York, 1968.

92. Soczewinski, E., Mechanistic molecular model of liquid-solid chromatography - Retention-eluent composition relationships. *Journal of Chromatography A* **2002,** 965, (1-2), 109-116.

93. Schoenmakers, P. J.; Billiet, H. A. H.; De Galan, L., Influence of organic modifiers on the retention behaviour in reverse-phase liquid chromatography and its consequences for gradient elution. *Journal of Chromatography* **1979,** 185, 179-195.

94. Tiller, P. R.; Romanyshyn, L. A.; Neue, U. D., Fast LC/MS in the analysis of small molecules. *Analytical and Bioanalytical Chemistry* **2003,** 377, (5), 788-802.

95. Pinkston, J. D.; Stanton, D. T.; Wen, D., Elution and preliminary structure-retention modeling of polar and ionic substances in fluid chromatography using volatile ammonium salts as mobile hase additives. *Journal of Separation Science* **2004,** 27, (1-2), 115-123.

96. Zheng, J.; Glass, T.; Taylor, L. T.; Pinkston, J. D., Study of the elution mechanism of sodium aryl sulfonates on bare silica and a cyano bonded phase with methanol-modified carbon dioxide containing an ionic additive. *Journal of Chromatography A* **2005,** 1090, (1-2), 155-164.

97. Zheng, J.; Taylor, L. T.; Pinkston, J. D., Elution of cationic species with/without ion pair reagents from polar stationary phases via SFC. *Chromatographia* **2006,** 63, (5-6), 267-276.

98. Pelletier, G.; Powell, D. A., Copper-catalyzed amidation of allylic and benzylic C-H bonds. *Organic Letters* **2006,** 8, (26), 6031-6034.

99. Sawlewicz; Jasinska, *Roczniki Chemii* **1961,** 35, 165-168.

100. Iiyama, K.; Inaba, N. Heat-sensitive two-color adhesive recording label. 86-3610588, 19860327., 1986.