CHALMERS

# Criteria Evaluations for Chromatographic Technique

## Comparison of High-Performance Liquid Chromatography and Supercritical Fluid Chromatography

A Bachelor's Thesis within Chemical Engineering

ANNA BYLUND AND EMMA ANDERSSON

DEPARTMENT OF CHEMISTRY AND CHEMICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, 2020
www.chalmers.se

DEPARTMENT OF CHEMISTRY AND
CHEMICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020
www.chalmers.se

BACHELOR'S THESIS 2020

# Criteria Evaluations for Chromatographic Technique

Comparison of High-Performance Liquid Chromatography and Supercritical Fluid Chromatography

ANNA BYLUND AND EMMA ANDERSSON


CHALMERS

Criteria Evaluations for Chromatographic Technique
Comparison of High-Performance Liquid Chromatography and Supercritical Fluid Chromatography

ANNA BYLUND AND EMMA ANDERSSON

© ANNA BYLUND AND EMMA ANDERSSON, 2020.

Department of Chemical and Chemistry Engineering
Division of Chemistry and Biochemistry
CHALMERS UNIVERISTY OF TECHNOLOGY
SE-412 96 Gothenburg
Sweden
Telephone +46 (0) 31-772 10 00

**Sammanfattning**

Inom läkemedelsindustrin används olika metoder för att analysera substanser för vidare utveckling till läkemedel. En metod som ofta har använts är vätskekromatografi men under senare år har superkritiskt vätskekromatografi presenterats på marknaden. Rapporten är baserad på en litteraturstudie, originalplanen var att utföra laborativa tester men på grund av COVID-19 har detta inte var möjligt. Med hjälp av vetenskapliga artiklar och redan gjorda experiment har data för relevanta substanser tagits fram för att utveckla en multivariabelanalys. Multivariabel analysen genererade fyra kluster som har analyserats för att hitta den fysikaliska-kemiska rymden, för att avgöra när vardera metoden är bäst att användas. Utifrån den fysikaliska-kemiska rymden har multivariabelanalysen använts för att analysera nya substanser och placera dessa i den fysikaliska-kemiska rymden som skapades av multivariabelanalysen. Målet är att finna skillnaderna mellan metoderna och avgöra när vardera metoden är att föredra. Utifrån vätskekromatografin har högpresterande vätskekromatografi och ultrahögpresterande vätskekromatografi utvecklats. Separationen sker i instrumentets kolonn på grund av skillnaden i exempelvis polaritet, det vill säga analytens förmåga att integrera med kolonnmaterialet, samt pKa och densitet. Utvecklingen av instrumentet har gett en ökning av upplösning, hastighet och känslighet. Jämfört med högpresterande vätskekromatografi finns det några parametrar som är viktiga och skiljer sig med ultrahögpresterande vätskekromatografi. Trycket är en av de viktigaste förbättringarna inom högpresterande vätskekromatografi, resultatet av utvecklingen har gett högre och smalare toppar vilket ger en lägre detektionsgräns och förbättrad upplösning mellan olika analyter. Superkritisk vätskekromatografi är i sin tur utvecklat från högpresterande vätskekromatografi. Största skillnaden är att superkritisk vätskekromatografi använder sig av en mobil fas som är högkomprimerad $CO_2$, vilket gör metoden både billigare och mer miljövänlig. Superkritisk vätskekromatografi används numera rutinmässigt inom utvecklingen av läkemedel men har inte alltid varit användbart på grund av exempelvis lägre känslighet och högre komplexitet. Slutsatsen utifrån multivariabelanalysen är att superkritisk vätskekromatografi bör användas för polära ämnen och reversed phase vätskekromatografi bör användas för opolära ämnen.

**Abstract**

The development of liquid chromatography has resulted in new technologies with higher performance. Today high performance liquid chromatography and ultra high-performance liquid chromatography are commonly used in pharmaceutical development. A new type of chromatography is supercritical fluid chromatography, which is a method that can be useful in pharmaceutical development. The main aim is to analyze the difference between the different techniques and to evaluate which method that is preferred for different compounds. This is done by a literature study, where experimental data are collected from the literature. Calculations and a multivariable analysis have been performed based on the data collected to evaluate a model that can answer which method that should be used. Based on the multivariable analysis, a conclusion could be drawn. A cluster analysis was done to define the physicochemical space and to find similarities between different compounds. The parameters have been analyzed to evaluate their influence on the clustering and to find out where in the space the substances can be placed. The multivariable analysis has then been used to analyze new substances and to place them based on the physicochemical space and to evaluate which method should be used to separate the components. The conclusion is that reversed phase liquid chromatography should be used to separate nonpolar substances and supercritical fluid chromatography is used for polar substances.

**Keywords:** LC; HPLC; SFC; UHPLC; Liquid chromatography; High performance liquid chromatography; Supercritical fluid chromatography; Ultra high-performance liquid chromatography; Pharmaceutical industry

**List of Abbreviations**

HPLC- High Performance Liquid Chromatography
SFC- Supercritical Fluid Chromatography
LC- Liquid Chromatography
UHPLC- Ultra high-performance liquid chromatography
UV- Ultraviolet
MS- Mass Spectrometry
S/N ratio- Signal-to-Noise ratio
RP-LC- Reversed Phase Liquid Chromatography
PCA- Principal Component Analysis
HCA- Hierarchical Clustering Analysis

**Acknowledgements**
We would like to thank AstraZeneca for the opportunity to do our bachelor thesis at your company. Many thanks to Jenny Ottosson and Magnus Klarqvist at AstraZeneca for assisting us with information and supervising during the project.

We would also like to thank our examiner Maria Abrahamsson at Chalmers University of Technology for feedback during the project. We would also like to thank Per Malmberg at Chalmers University of Technology for the supervising during the project.

# Table of Contents

# 1 Introduction

Chromatographic techniques have several advantages within different aspects of early drug development to find product properties and quality control [1][2]. Both high performance liquid chromatography (HPLC) and supercritical fluid chromatography (SFC) are analytical methods for analyzing and separating compounds. The main different between the two methods is that SFC is using a highly compressed gas as the mobile phase, $CO_2$. How can SFC be used more in the pharmaceutical industry?

Liquid chromatography (LC) is a commonly used analytical technique in analytical chemistry and is used for quality control and separation of different substances in liquids are performed using LC [1]. The sample is a liquid and the method use a mobile phase to transport the sample into a column into the chromatographic unit. There are mainly two different types of LC, HPLC and Ultra high-performance liquid chromatography (UHPLC).

The HPLC technique can both separate, identify and quantify the contents in a sample [1]. This makes it very useful for many types of samples that exist in liquid form. Through the development of columns, UHPLC has achieved a higher speed and higher sensitivity than the previous chromatographic techniques. It is a useful method but is not as widely used as HPLC today. Both HPLC and UHPLC is using the same principles for separations and can be used in many different projects for example in the early pharmaceutical development. Both HPLC and UHPLC can be used with ultraviolet (UV) detector and/or mass spectrometry (MS).

It is crucial for the pharmaceutical industry that effective analytical separation techniques are available, since all steps in the pharmaceutical process needs analysis [3][4]. The purpose of making a pharmaceutical analysis is to decide the quality of the pharmaceutical substances and the pharmaceuticals, this is done using different analytical methods. Different chromatographic techniques are used daily for this purpose and during the last year a big interest for SFC has been growing. The method has emerged by using already existing chromatographic instruments [5]. A new generation of advanced SFC instruments. SFC has faced challenges with low UV sensitivity, limited reliability and poor qualitative results. These limitations have been resolved with advanced technology. $CO_2$ alongside other factors is today seen to cause global warming, it could therefore be odd to see it as a green solvent in chemistry. $CO_2$ is used as the mobile phase in SFC and is seen as a green product because it is recycled and contributes low costs [2][3]. This is especially important in preparative scale when huge amount of mobile phase is used. SFC is today used for achiral separation and purification in the pharmaceutical industry.

## 1.1 Aim

The main aim of this thesis is to analyze the difference between LC and SFC and to evaluate which of the methods that is to prefer for different substances. The focus is to investigate how the

retention and signal- to- noise ratio (S/N ratio) can affect the result and what it is that determines if a substance get retention or not. The goal is to evaluate how SFC could be used more in the pharmaceutical industry. This will be done by a literature study, a cluster analysis and a multivariable analysis by analyzing already published experiments.

### 1.2 Constraints

The thesis will be based on a literature study because it is not possible to do laboratory work due to COVID-19. Focus will be to identify gaps for future potential project. It will be based on already performed experiments found by literature searches to analyze the physicochemical space, for example density, pKa and solubility in pure water, for different substances. When each method is to be used will be based on the physicochemical space.

While it is possible to apply LC and SFC in many different aspects, we have decided to only analyze reversed phase LC and achiral separation using SFC. The main argument to only analyze reversed phase LC (RP-LC) is because 75% of all LC that today is used is reversed phase [1].

### 1.3 Disposition

This thesis includes introduction, two theory sections, method, results and discussion, conclusions and future work. In addition, there are also references and appendix as the last two sections of the report.

The first section of this thesis is the introduction which consists of background, aim and constraints. The two theory sections that follows are relatively extensive because a literature study was done instead of practical lab work. In the first theoretical section, chapter 2 describing chromatography with subheadings on LC, HPLC and UHPLC and a comparison of the LC techniques. Chapter 3 continues in the same way, but this section is about SFC. After that, the method is described both from the literature study but also from the modeling as well as the software used. The end of the report consists of the result based on the aspects of theory, data and modeling. The discussion is based on the result but treats it in a larger perspective to reach a conclusion. As this thesis became theoretical and not practical as the idea was before COVID-19 erupted, the future work part will be about the practical tests that we could not do due to the circumstances. This will be discussed in chapter 7.

## 2 Chromatography

Chromatography is a method that is used to separate compounds of different kinds from each other [6]. The term was first mentioned by the Russian Mikhail Tswett in 1906, but the first analysis was done by James and Martin in 1952 [6]. The first chromatographic method was gas chromatography, but today several different techniques exist, such as, LC, HPLC, UHPLC and size exclusion chromatography. Many of the different chromatographic techniques are used in the pharmaceutical industry [2]. Today, there is much research on the new technology SFC and how it can be used more in the pharmaceutical industry.

### 2.1 Liquid chromatography

Today LC is one of the main methods in analytical chemistry. LC is used as a key-technology in the production and development of pharmaceuticals [7].

Column chromatography uses a column that has particles packed in it which is the stationary phase [1]. The analyte will separate due to its different capacity to interact with the stationary phase, due to the polarity. The analytes that are interacting more strongly with the surface will slow down and analytes less retained by the stationary phase will elute with higher speed i.e. shorter time. This results in a separation between the different compounds.

The stationary phase is a solid surface [7]. The stationary phase exists at the packing material, which is located at the surface in the column. The packing material is used to make the specific surface area per unit volume as large as possible, which leads to improved mass transfer and column efficiency. The packing material in the LC column is usually silica particles in the size of 5–10 μm. The separation mechanism is based on polarity. There are mainly two different methods that is connected to polarity, normal-phase and reversed-phase [1]. Normal-Phase uses a polar stationary phase and a non-polar mobile phase. Reversed-phase is the opposite of normal-phase, and therefore it has a polar mobile phase and a non-polar stationary phase. In Table 1 the different between normal phase and reversed phase is showed and which type of polarity that is used in the mobile phase and stationary phase.

| **Separation Mode** | **Stationary phase (particle)** | **Mobile phase (solvent)** |
|---|---|---|
| Normal phase | Polar | Non-polar |
| Reversed phase | Non-polar | Polar |

**Table 1:** The different between Normal phase and reversed phase, with respect to polarity in the mobile and stationary phase [1].

The mobile phase is a liquid that moves through the column [7]. It is characterized by its solubility, viscosity, how easy it mixes with other solvents and some other parameters. LC and the mobile phase are driven by a convective flow viscosity through the column, which is dependent of the

inner diameter of the column. The packing material in LC is usually native silica or alumina with a particle size of 30-40 μm.

The volume, length and inner diameter of the column will affect the time it takes to pass through the column [1]. This can affect the accuracy of the measurements. It is important to adapt this to the purpose. It can also affect the peaks in the chromatogram.

## 2.4 Development of the LC technique

The LC technique has been developed in several steps [5][8]. It started as LC but has later developed into HPLC and then UHPLC.

### 2.4.1 Comparison between LC and UHPLC

UHPLC is a new type of separation technique in LC and was introduced in 2004 [8]. It has achieved a dramatic increase in resolution, speed and sensitivity. This was possible through development in the instrumentation and column technology. The LC- technology has earlier used a smaller particle size to improve the separation. But with the small particles, the instruments have disadvantages to analyze and separate the particles and get the pressure in the correct limited range. UHPLC made this possible with improved technology that generates the possibility to use smaller particles in the columns together with higher pressure. It leads to minimize the band spreading and results in a higher performance and quality of the analytical data. The instrument and column design are very important, and both the intra-column and extra-column needs to operate at optimal conditions. UHPLC can resist higher back pressure than LC which results in higher speed, sensitivity and resolution [9]. In Table 3, the main operating conditions is showed fore LC and UHPLC.

| | **LC** | **UHPLC** |
|---|---|---|
| **Injected analyte volume (μL)** | 1-100 | 0.5-5 |
| **Flow rate (mL/min)** | 0.2-10 | 0.01-2 |
| **Analysis time (min)** | 10-60 | 1-10 |
| **Maximum pressure (psi)** | 6000 | 15000 |
| **Particle size (μm)** | 3.5-10 | ~1.7 |

**Table 3:** The main operating conditions for LC and UHPLC [9].

### 2.4.2 Comparison between HPLC and UHPLC

Both HPLC and UHPLC is an analytical method that separates different compounds in liquids [10]. The mobile phase is pumped into the system by a high-pressure pump which regulates the flow rate of the mobile phase. The sample is injected into the system and transported to the column by the constantly flowing mobile phase. The column contains packed material that is needed for the separation of the analytes, which is called the stationary phase. When the analytes come out of

the column they are detected by a detector, the most commonly used is UV and MS, and the result is obtained in a chromatogram. See Figure 1 for HPLC system.

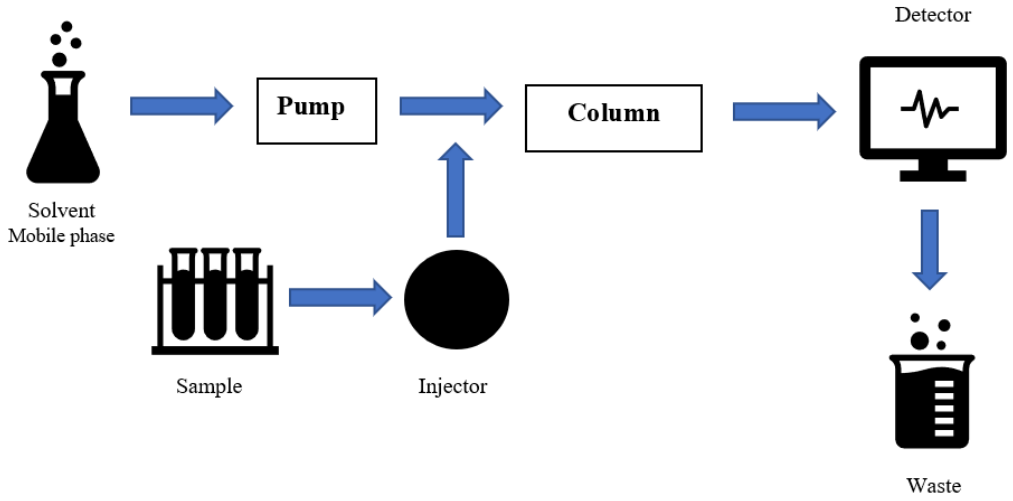



**Figure 1:** An HPLC system and its devices [1].

The development of UHPLC is a small modification of HPLC, both the techniques are based on LC [11][12]. UHPLC uses higher speed and smaller particles to reach a higher number of resolvable peaks to obtain higher effectivity and resolution, it generates a higher signal-to-noise ratio (S/N-ratio) due to the reduction in band spreading that appears [13][14]. UHPLC reduce the volume of the mobile phase, up to 80% compared to HPLC. The column can be connected in series while the fused-core particles require a lower back pressure. [11][13].

The most common mobile phase in reversed phase HPLC is a buffer aqueous solution or a mixture of a buffer aqueous solution and a polar solvent like methanol or acetonitrile [15]. The smaller particle size results in smaller volumes of the injected mobile phase and faster flow rate [16][17].

The stationary phase in HPLC usually consists of hydrocarbons of varying length [18]. Some of the most usual is C18, C8 and C4. The C18 column is preferred as it has good hydrophobic separation power and has a high coverage of the surface area. Packing material in the stationary phase used in UHPLC are often 1–1.5 μm silica-based nonporous particles [9].

In comparison with HPLC, UHPLC has a higher column temperature at 65°C [12]. The higher column temperature reduces the mobile phase viscosity, which results in an increased flow rate and diffusion coefficient. This phenomenon can appear without loss in effectivity or increase in the column back pressure. UHPLC also requires a shorter column with the dimension of 50 × 2.1 mm, and uses a small injection volume at 2 μL and a low flow rate at 0.5 mL/min.

In UHPLC, there are several aspects to consider in order to achieve the desired result. It can be both instruments and other parameters that can be optimized [8]. An important parameter in UHPLC is the particle size, small particles generates less band spreading. A wider band results in a lower and wider peak in the chromatogram which is unwanted and result in loss in sensitivity and resolution. The inner diameter and the length of a column affects the band spreading and the peak shape [10]. With a minimized band spreading, that results in higher effectivity and sensitivity.

## 3 Supercritical fluid chromatography

SFC is a separation technique that is based on HPLC, its equipment is almost identical to HPLC [2]. Using SFC, complex solutions can be separated, and the amount of the individual components can be identified. The analyte is injected to a high-pressure stream which transports the sample into a column. The biggest difference between HPLC and SFC is the mobile phase, SFC is using a highly compressed gas. When using a highly compressed gas, a back pressure needs to be used to ensure that the mobile phase retains its compressed state throughout the process. To ensure this, a back-pressure regulator is used.

### 3.1 Mobile phase

$CO_2$ is usually used as the mobile phase, it is preferable because it is cheap, relatively safe, accessible, the critical accessible critical value, a green product and miscible with highly polar solutions [2][3][5]. $CO_2$ is seen as a green product since it is recycled from other industries, and it contributes to low costs.

Pure $CO_2$ is not sufficiently polar for most of the compounds that is used in the pharmaceutical industry, therefore all analyzes is carried out with a mixture of $CO_2$ and a modifier under pressure [3][4]. It is necessary to add a modifier to improve the analyte solubility, elution and the analysis time. The modifier plays a major role as it changes the critical values for $CO_2$, but it can also create interactions with the analyte and increase the affinity for the mobile phase, resulting in elution.

Many strong acids and bases, and most amphoteric compounds does not elute or elute as distorted peaks when eluted with binary mixtures of $CO_2$ and a modifier [2][3][5]. If the addition of the modifier is not enough to enable elution of the analytes with good peak shape, it is important to add a third component in the mobile phase. A small amount of a strong polar component, dissolved in the modifier, is often added to improve the peak shape.

Density is the primary factor affected by change in mobile phase strength, other factors also affected by mobile phase strength are retention and selectivity [4]. The density depends on three parameters, mobile phase composition, system pressure and column temperature. If the pressure increases at constant temperature, the solution strength and the adsorption of the mobile phase composition on the surface of the stationary phase change. That results in a decrease of the retention factor. For pure $CO_2$, which is compressible, variation in density can be large and lead to important variations in retention.

### 3.2 Stationary phase

The stationary phase that is used in HPLC can also be used in SFC [3]. The choice of stationary phase has a big influence on how fast the analysis is done since some stationary phases retain the analytes more than other. The stationary phase can be changed by reducing the particle size which reduces analysis time [2]. Most analyzes with SFC use porous, high purity, silica or

bonded phases with porous silica. When pharmaceutical compounds including polar components, columns with polar packing material should be used to obtain desired retention. In pharmaceutical compounds having polar components, columns which select the polar are to obtain adequate retention. This category of columns includes, among others, silica, diol, aminopropyl and cyanopropyl.

### 3.3 The instrumentation

The system currently used for SFC analysis is developed from HPLC by a binary pump where one pump delivers $CO_2$ and the other delivers an organic modifier, usually an alcohol, such as methanol [5][19]. See Figure 2 for instrument setup.

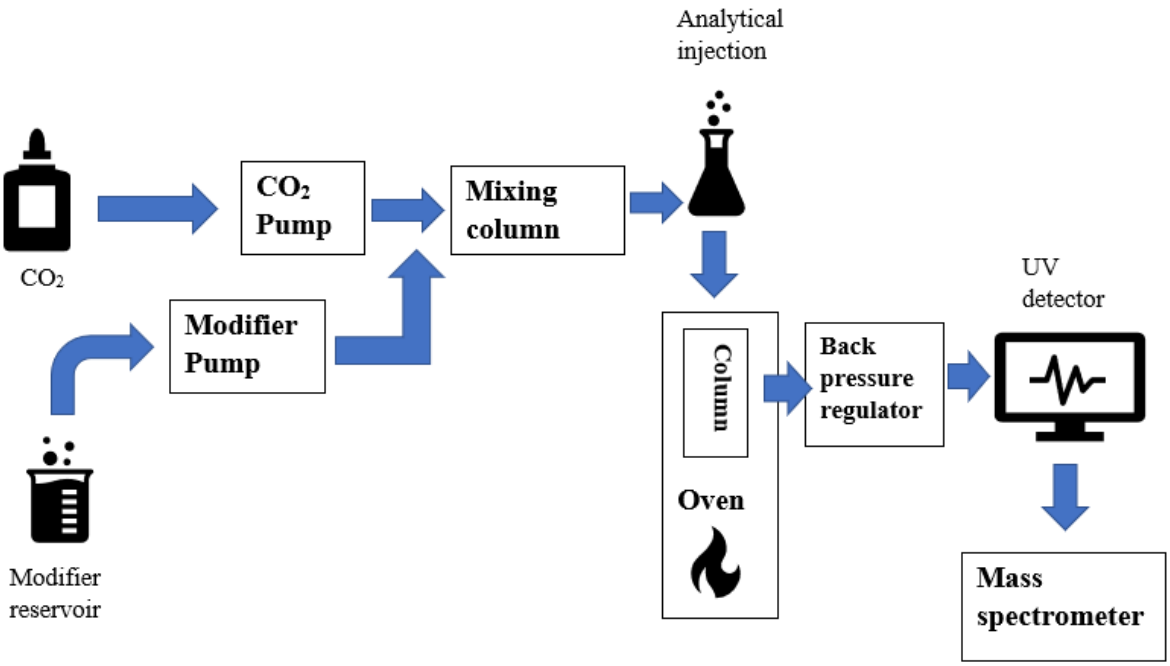



**Figure 2**: SFC instrument [19]

To control the pressure in the column, a dynamic back pressure regulator is used to maintain the desired pressure at the column outlet [5]. The dynamic back pressure regulator advantage is independent control of the outlet pressure without the flow rate and viscosity [4][5][20].

It has been difficult to use the UV detectors together with SFC because of its high detection limit. [3][21]. Improvements in the technique has today made SFC-UV more effective but the detection limits are still controlled by the S/N-ratio. Some of the reasons to the noise is fluid density and density variations in the UV cell due to pressure fluctuation that is a result of the back-pressure regulation in the pump system. The noise occurs due to the pressure fluctuations and this affects how well the UV detection can perform. Refractive index and density changes occur. This becomes a problem when specific profiling and quantification sensitivity is to be analyzed in very low

concentrations. The temperature can also affect the S/N-ratio, were temperatures over 40°C can generate a higher noise and therefore cooling after the column should be done [5].

# 4 Method

This thesis works involved literature study, followed by selection of compounds to evaluate the physicochemical space of substances for each method. The modelling will be based on the physicochemical properties that will be evaluated form the literature study.

## 4.1 Principal component analysis

Principal component analysis (PCA) is a commonly used multivariable analysis method. IT is commonly used for exploratory analysis of a dataset where its main purpose usually is dimensionality reduction. Another function of PCA is modelling since the method can be used to find relevant variables responsible for the highest variation in the dataset.

A score plot is used, it is a plane were different components are plotted in correlation to each other, seen in Figure 6. A score plot shows cluster of samples based on their inherent similarities. Substances that are close to each other in the score plot, has something similar and the substances that are far from each other has differences. The physicochemical parameters that are used, generates an axis for each component in different directions, this generates a space were the substances can be plotted. The knowledge of different parameters is called loadings, see Figure 7. The loading plot is therefore the different parameters in correlation to each other, which generates a plane. These components can be read from vectors that are called p[1] and p[2], which is the x-axis and y-axis. The distance to origin is an important part, the further distance, the stronger impact on the model the variable has. The variables can be both positive and negative, were negative means that it has an inverse correlation to the positive value and exist on the opposite side.

### 4.1.1 Hierarchical Clustering Analysis

A Hierarchical clustering analysis (HCA) was done, which means that large datasets are grouped due to relationships between parameters. The groups are called clusters. The clustering analysis makes natural cluster so there is greater similarity within the clusters and less similarity between different clusters. There are no predefined classes in a cluster analysis that is clustered, but they are formed after the analysis depending on what data is available. See the HCA in Figure 3.

## 4.2 Literature study

The literature study is based on data collection from several different sources. The main sources that has been used in the analysis come from Chalmers library and its search engines. In addition, Google Scholar have been used to find relevant scientific articles. The supervisors from AstraZeneca has also contributed relevant sources in the subject.

A careful selection has been made to use relevant sources in the report. The relevance of the sources is based on the boundaries of the thesis, to be able to address the aims a good way. It is also based on the credibility of the sources where it has been very important to have high quality

and credibility. The data is collected from scientific articles and books. Apart from relevance, no other filters have been used.

Due to COVID-19, data collection by laboratory tests has not been possible as part of this work. All data has been collected from scientific articles with data and substances. The choice of substances to be analyzed was based on which data was available in the scientific articles. The data that was considered to be relevant were data available for both SFC and RP-LC in order to compare the different methods. Data for caffeine, theobromine, theophylline and uracil were only available for SFC, the substances was considered to be relevant because a chromatogram was available, and it was easy to find physicochemical properties for the substances. These four substances are used to test the multivariable model, to in the future be able to use the model with new substances and to directly know which method should be used.

A data collection has been made to find theoretical values of the substances that has been analyzed with ACD Labs [22]. It has been made to be able to compare theoretical values and calculated values. The theoretical values that has been collected from the webpage DrugBank and has then been compared to the calculated values from ACD Labs. A comparison has been done to decide which values the modelling should be based on. See Table 4, 5 and 6 for calculated- and theoretical values. [22]

#### 4.2.1 Parameters

During the literature study, there are certain parameters that have been important to look for in order to obtain a good data base to do our analyzes. In the literature, chromatograms have been important for substances that have been analyzed with both SFC and RP-LC in order to compare data. Based on the chromatogram, retention time and peak width have been analyzed to see how the retention time is affected depending on which method has been used.

### 4.3 Physicochemical space

The physicochemical properties of selected compounds were predicted using available software. For some of the compounds data was found in the literature study and compared to calculated values. The structures used were analyzed through the BIOVIA Draw software and with the help of ACD Labs calculations were made. See results and parameters of the calculations in Appendix 3-12 and Table 4 and 5.

To evaluate the physicochemical space, different parameters are used to analyze different physicochemist properties of substances.

The pH value on the mobile phase has a huge effect on the separation and retention [10][23]. A change in pH affect the pKa value. The acid-base property, pKa value, is important in the development of pharmaceuticals because it affects the solubility, absorption, metabolism and distribution [23]. The pKa value is a function of physicochemical properties. In LC the pKa value

will be based on the relation between the capacity factors and the pH of the mobile phase. In LC the pKa value can only be determined in circumstances where suitable retention is obtained.

The Log S value is the solubility of a substance. Solubility describes how much of a certain substance that will be dissolved in a solvent [24]. Lipophilicity parameter, Log P, is a reference parameter and are commonly used in tests of a series of parameters in the pharmaceutical industry [25]. It uses as a standard reference parameter and is a logarithm partition coefficient.

### 4.3.1 Software

Different software has been used to calculate data based on structures that was found in the literature.

The BIOVIA Draw software were used to draw the different structures found in the literature study. The version that was used is BIOVIA Draw 16.1. See Appendix 2 for the structures, see Appendix 1 for specific details for BIOVIA Draw 16.1.

ACD Labs is a software used to calculate physicochemical properties. The molecule that was found in the literature was drawn in BIOVIA and was then used in ACD Labs 2018.2. The software can calculate several different parameters of the substance, see Appendix 13-16 for the different settings. These are the base for the result, see Appendix 3-12.

SIMCA 15.02 were used to cluster the data of the different parameters to be able to analyze the different substances and possible see trends between the substances. The clusters were used to find a physicochemical space and be a tool to determine what parameter values other substances should have, due to the physicochemical space. The software identifies critical parameters that is helpful to use when to determine parameter values of new substances that have not been analyzed with the method. The data used for the cluster, see Table 4.

## 5 Result and Discussion

The result is based on the multivariable modeling which is based on collected data and structures from the literature study. The structures and chromatograms that were found in the literature study has been analyzed and some of the parameters have been measured. The peak width for both RP-LC and SFC has been measured by us in the literature, so there has not been calculated values in the literature but there have been available chromatograms. See Appendix 17. The retention time for the substances that the multivariable analysis is based on is values found in the literature, both for SFC and RP-LC. See Table 4. The retention time and the peak width for the four other substances, uracil, caffeine, theobromine and theophylline, were not available in the literature so these are calculated by us based on the chromatograms. See Appendix 18.

### 5.1 Data

Collected data from the literature study and calculated data from ACD Labs.

#### 5.1.1 Data from ACD Labs

The data presented in Table 4 and 5 were calculated using ACD Labs. All parameters except retention time and peak width are calculated, for both RP-LC and SFC [26]. See Appendix 3-12. Retention time and peak width are measured by us from literature, see Appendix 17. They are calculated using chromatography from the literature.

| Parameters | Mometasone furoate | Compound 1 | Compound 3 | Compound 5 | Compound 7 | Compound 8 |
|---|---|---|---|---|---|---|
| **Molecular formula** | $C_{27}H_{30}Cl_2O_6$ | $C_{27}H_{28}Cl_2O_7$ | $C_{27}H_{30}Cl_2O_6$ | $C_{27}H_{29}ClO_4$ | $C_{27}H_{29}ClO_6$ | $C_{28}H_{33}ClO_6$ |
| **Formula weight (g/mol)** | 521,43 | 535,41 | 427,36 | 468,97 | 484,97 | 501,01 |
| **Composition** | 62,19% C<br>5,80% H<br>13,60% Cl<br>18,41% O | 60,57% C<br>5,27% H<br>13,24% Cl<br>20,92% O | 61,83% C<br>6,60% H<br>16,59% Cl<br>14,98% O | 69,15% C<br>6,23% H<br>7,56% Cl<br>17,06 % O | 66.87% C<br>6,03% H<br>7,31% Cl<br>19,79% O | 67,12% C<br>6,64% H<br>7,08% Cl<br>19,16% O |
| **Molar volume ($cm^3$)** | 379,20 | 376,00 | 316,50 | 362,90 | 367,10 | 384,00 |
| **Density ($g/cm^3$)** | 1,37 | 1,42 | 1,35 | 1,29 | 1,32 | 1,30 |
| **Polarizability ($cm^3$)** | $51,69*10^{-24}$ | $51,74*10^{-24}$ | $42,91*10^{-24}$ | $49,04*10^{-24}$ | $49,21*10^{-24}$ | $51,61*10^{-24}$ |

| | | | | | | |
|---|---|---|---|---|---|---|
| **pKa 1** | 13,02 | 12,84 | 13,46 | | | 13,48 |
| **pKa 2** | | | 12,80 | | | |
| **Log S** | -5,74 | -5,58 | -4,92 | -6,58 | -6,15 | -5,74 |
| **Log P** | 4,27 | 3,28 | 2,91 | 5,26 | 4,39 | 4,60 |
| **Solubility in pure water (mol/l)** | $1,80*10^{-6}$ | $2,60*10^{-6}$ | $1,20*10^{-5}$ | $2,60*10^{-7}$ | $7,20*10^{-7}$ | $1,80*10^{-6}$ |
| **Retention, SFC (min) [31]** | 9,08 | 7,73 | 9,55 | 5,10 | 5,91 | 4,84 |
| **Retention, RH-LC (min) [26]** | 24,03 | 18,69 | 11,39 | 32,85 | 22,50 | 27,49 |
| **Peak width, SFC (mm) [26]** | 14,00 | 11,50 | 11,50 | 7,00 | 11,00 | 6,00 |
| **Peak width, RP-LC (mm) [26]** | 19,50 | 16,00 | 12,00 | 22,00 | 16,50 | 16,00 |

**Table 4:** Theoretical values for the physicochemical space [36].

Calculated data, based on structures, with ACD Labs. All parameters except retention time and peak width are calculated with ACD Labs. There is no RP-LC data for these compounds, only SFC. Retention time and peak width are measured from literature [2], see Appendix 18.

| **Parameters** | **Caffeine** | **Uracil** | **Theobromine** | **Theopylline** |
|---|---|---|---|---|
| **Molecular formula** | $C_8H_{10}N_4O_2$ | $C_4H_4N_2O_2$ | $C_7H_8H_4O_2$ | $C_7H_8N_4O_2$ |
| **Formula weight (g/mol)** | 194,19 | 112,09 | 180,16 | 180,16 |
| **Composition** | 49,48% C,<br>5,19% H,<br>28,85% N,<br>16,48% O | 42,86% C,<br>3,60% H,<br>24,99% N,<br>28,55% O | 46,67% C,<br>4,48% H,<br>31,10% N,<br>17,76% O | 46,67% C,<br>4,48% H,<br>31,10% N,<br>17,76% O |
| **Molar volume ($cm^3$)** | 133,30 | 84,80 | 112,90 | 112,00 |

| | | | | |
|---|---|---|---|---|
| **Density ($g/cm^3$)** | 1,45 | 1,32 | 1,47 | 1,60 |
| **Polarizability ($cm^3$)** | $19,97*10^{-24}$ | $9,91*10^{-24}$ | $17,10*10^{-24}$ | $17,86*10^{-24}$ |
| **pKa 1** | 0,52 | 15,60 | 8,60 | 9,90 |
| **pKa 2** | -3,83 | 8,83 | 1,84 | 0,41 |
| **pKa 3** | | -4,19 | -4,09 | |
| **Log S** | 0,56 | -1,95 | -1,28 | -1,39 |
| **Log P** | -0,13 | -0,71 | -0,17 | -0,72 |
| **Solubility in pure water (mol/l)** | $2,70*10^{-2}$ | $1,10*10^{-2}$ | $5,20*10^{-2}$ | $4,00*10^{-2}$ |
| **Retention, SFC (min) [2]** | 1,00 | 1,20 | 0,90 | 1,30 |
| **Peak width, SFC (mm) [2]** | 3,0 | 3,50 | 3,50 | 2,50 |

**Table 5:** Theoretical values for the physicochemical space [2].

### 5.1.2 Theoretical data

Data based on the literature study, collected from the website DrugBank [22]. These values are theoretical values and not calculated with the same software as the values in Table 4 and 5.

| **Parameters** | **Mometasone furoate** | **Caffeine** | **Uracil** | **Theobromine** | **Theopylline** |
|---|---|---|---|---|---|
| **Molecular formula** | $C_{27}H_{30}Cl_2O_6$ | $C_8H_{10}N_4O_2$ | $C_4H_4N_2O_2$ | $C_7H_8H_4O_2$ | $C_7H_8N_4O_2$ |
| **Formula weight (g/mol)** | 521,43 | 194,19 | 112,09 | 180,16 | 180,16 |
| **Polarizability ($cm^3$)** | $52,31*10^{-24}$ | $18,95*10^{-24}$ | $9,37*10^{-24}$ | $16,85*10^{-24}$ | $16,86*10^{-24}$ |
| **pKa Acidic** | 13,84 | -0,92 | 9,77 | 9,28 | 7,82 |
| **pKa Basic** | -3,10 | | -5,50 | -0,91 | -0,78 |
| **Log S** | -4,70 | -1,20 | -0,63 | -1,30 | -0,90 |
| **Log P1** | 4,27 | -0,24 | -1,20 | -0,46 | -0,26 |
| **Log P2** | 5,06 | 0,55 | -0,86 | -0,77 | -0,77 |
| **Solubility in pure water (mol/l)** | 0,01 | 11,00 | 26,50 | 9,74 | 22,90 |

**Table 6:** Theoretical values for the physicochemical space [22].

The difference between the values calculated with ACD Labs, that is based on the substances, and theoretical values are large. See Table 4, 5 and 6. There are also many values that were calculated with ACD Labs that could not be found as theoretical values. The literature values were found in different sources, see Table 6. To do as reliable modeling as possible, the values calculated with ACD Labs were used to cover a larger width of parameters compared to the theoretical values.

### 5.2 Modeling

The result of the HCA categorized the substances into four different clusters. See the four clusters in Figure 3. The clusters are divided in four different steps, the first step divides compound 3 from all the other substances, called group 1 in Figure 3. This means that compound 3 has fewer parameters in common with the other compounds. In step 2, the substances are divided into two different clusters. Mometasone furoate and compound 1 are grouped into one cluster, called group 2. Compound 5 are in step 3 divided into a separate cluster, called group 3. Compound 7 and compound 8 are grouped to the last cluster, called group 4. See clusters in Figure 3.

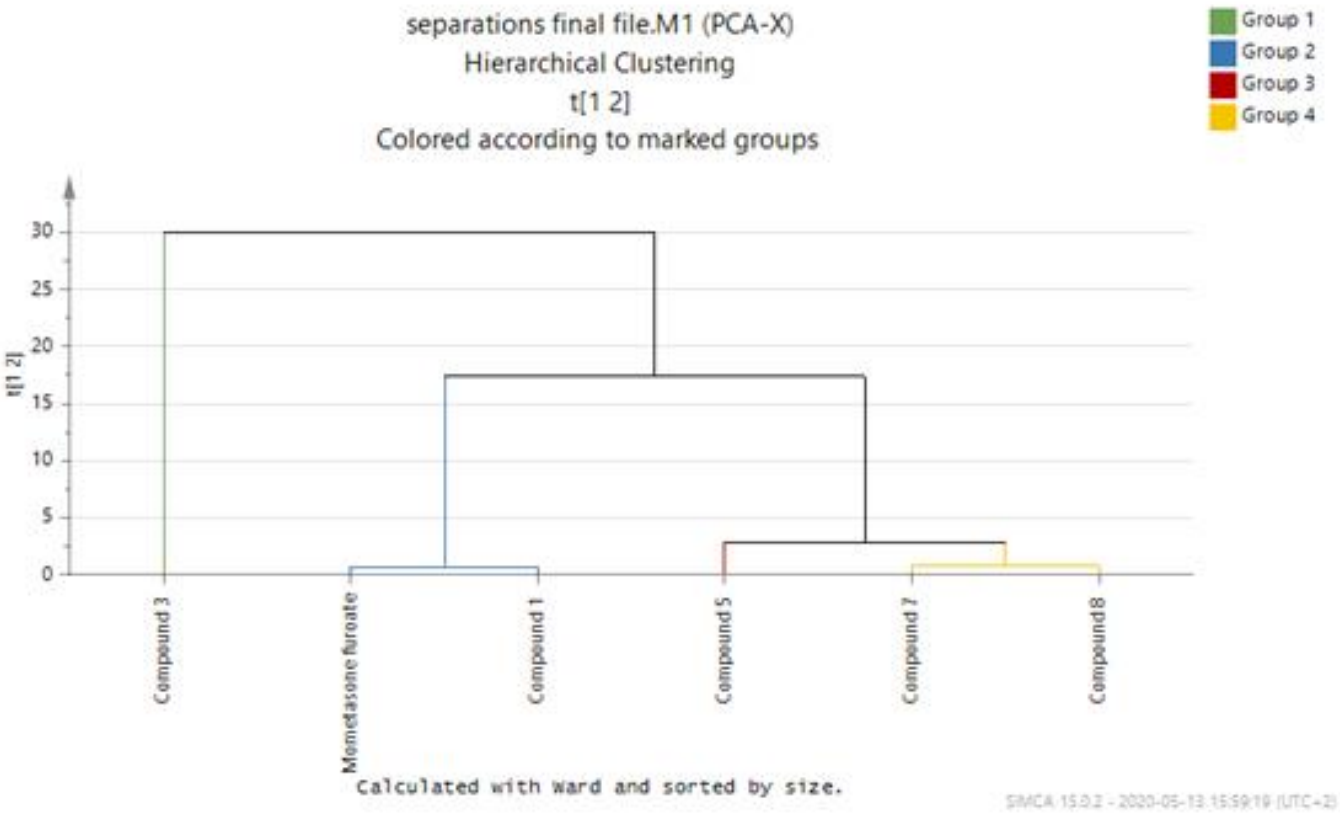



**Figure 3:** The four clusters that was found in the HCA.

What parameters that are considered as significant in this HCA, Figure 3, is individual for this analysis. Since there are only a few substances that are analyzed, it can be other clusters if more substances are clustered. Different parameters can be significant for different clusters.

Figure 4 shows the first step in the PCA which is component 1, which is the parameters that affects the compounds and how big the variation is within the parameters. Component 1 is the parameters that affects the clustering in x-axis, shown in Figure 8 as p[1]. The bars through the boxes are a 95% confidence interval.

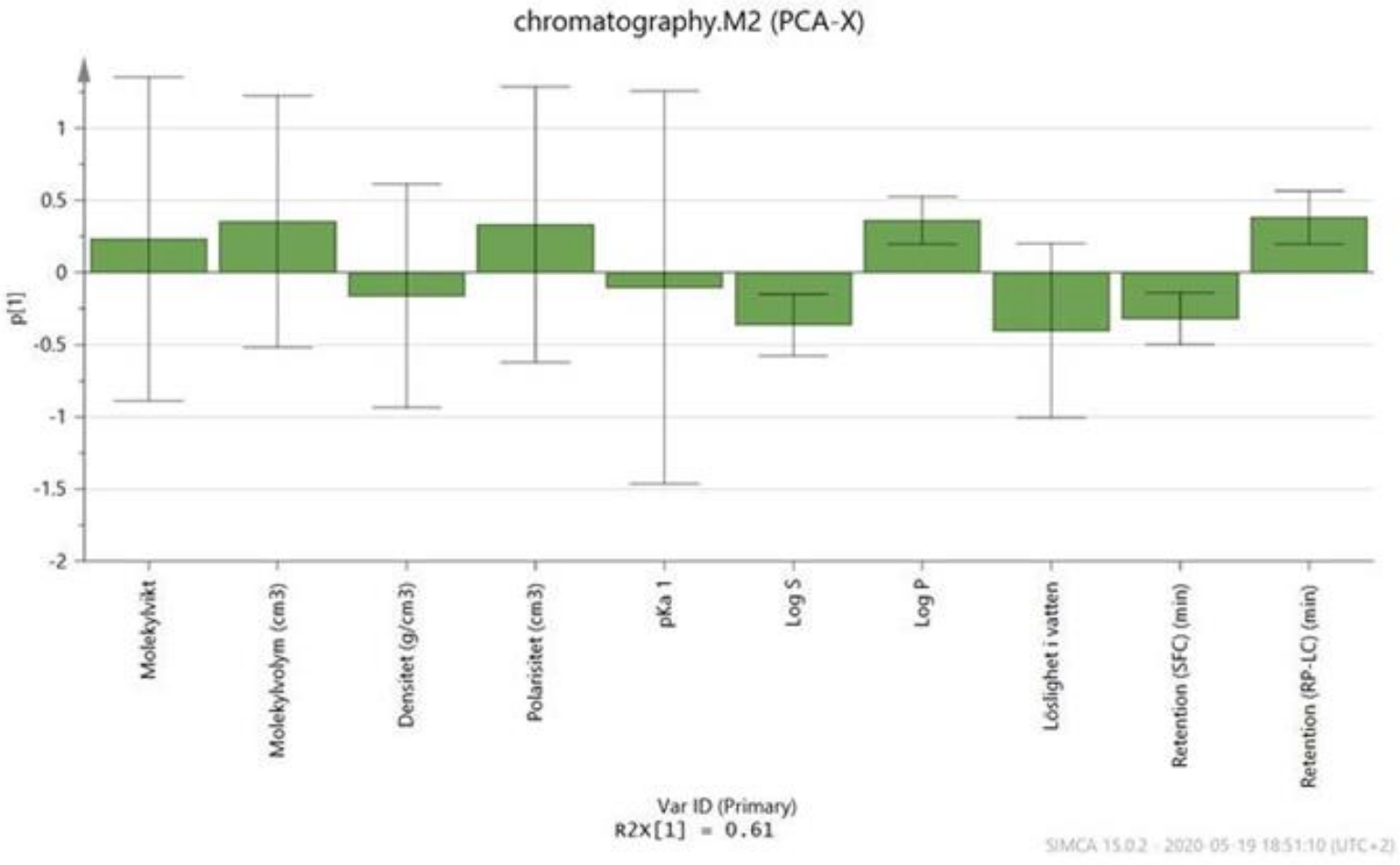



**Figure 4**: How the parameters affects the different compounds, Component 1.

The retention in SFC and RP-LC in Figure 4 can be read to be opposite to each other. A reason for this can be that the methods are used for different types of compounds. RP-LC uses a polar mobile phase and a nonpolar stationary phase, while SFC uses $CO_2$ which is a nonpolar mobile phase and a polar stationary phase.

Figure 5 shows component 2, which are the parameters that affect the clusters on the y-axis, shown in Figure 8 as p [2].

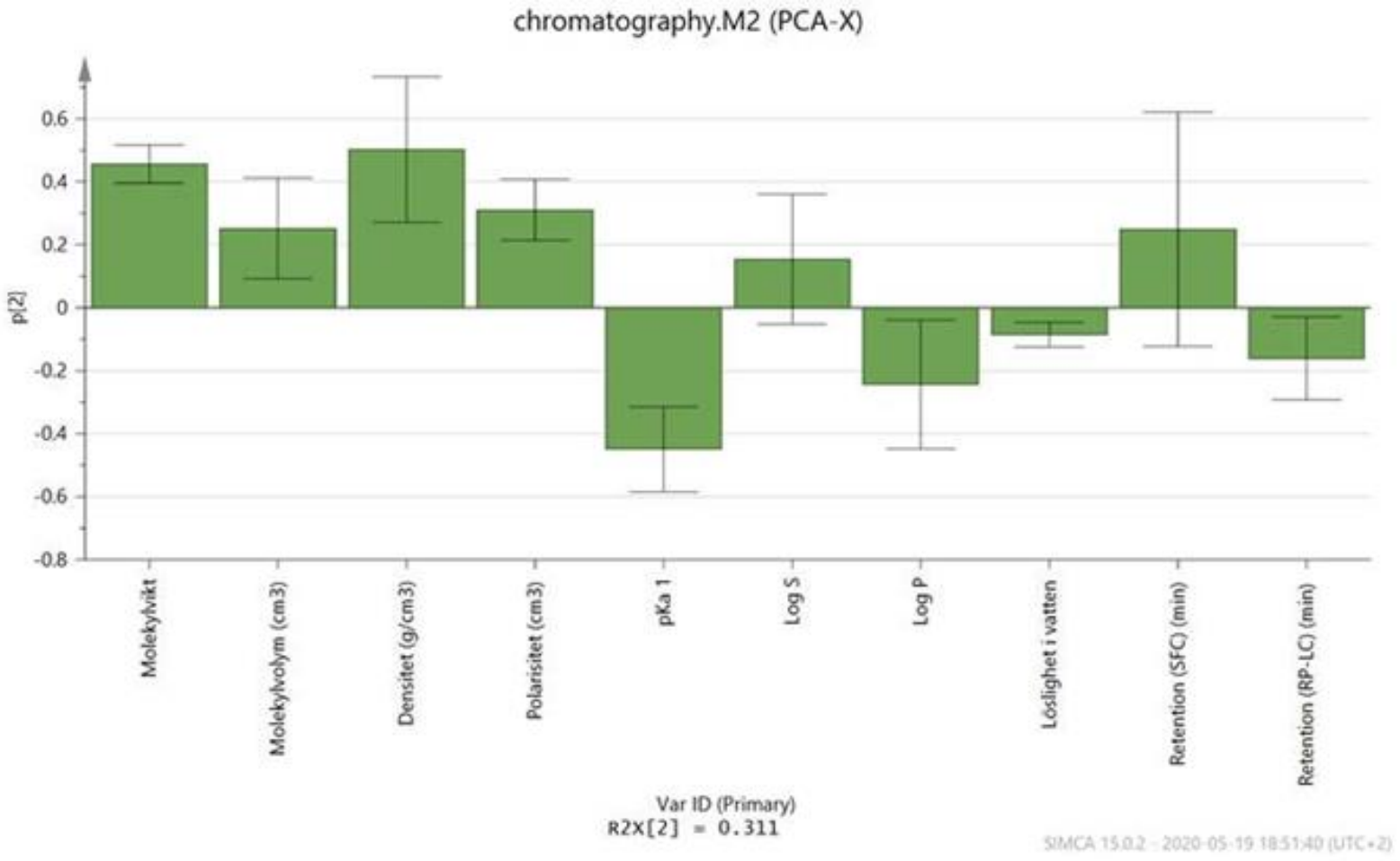



**Figure 5:** How the parameters affects the different compounds, Component 2.

Component 2 shown in Figure 5 represent the y-axis in Figure 8. The different parameters are shown, and larger response means that the parameter has greater impact on the compounds. These responses can be both positive like molecular weight and density or it can be negative as pKa 1 and Log P. It is the area of the response that is important and not if it is positive or negative. It can be seen that the biggest impact is represented by density and pKa 1. The density is positive in Figure 5 which means that if it is a significant parameter it appears in the positive direction on the y-axis in Figure 8. The pKa 1 value is negative in Figure 5 which means that it has negative impact and appears on the negative side on the y-axis in Figure 8. In Figure 7 it can be seen that density and pKa 1 values are reversed to each other, which means that compounds which is more impacted by one of these factors will come closer to that point. A compound which is highly affected by the density will appear high at the y-axis. In Figure 5 you can see that the attention time in SFC is greater than in RP-LC, it can be seen that they are each other's opposites. SFC retetion is positive

and RP-LC in negative. This means that they are affected by different parameters. The SFC retention response is higher than in RP-LC which means that it has a larger variation.

The parameter that has smallest affect in Figure 5 is the solubility in pure water. This means that this parameter has a small impact on the y-axis in Figure 8. On the other hand, it can be seen in Figure 4 that this parameter has a huge negative impact. This means that the solubility in pure water are represented at the x-axis in Figure 8 but has small impact on the y-axis. It can be seen in Figure 4 that the polarizability also has a huge impact. The polarizability and solubility in pure water are two physiochemical parameters that are highly connected to each other. Water is a polar compound, which means that other polar compounds has a higher solubility in pure water then nonpolar compounds. Log P is another parameter in Figure 4 that has a huge impact in the positive direction. Log P are calculated from the lipophilicity of the compound which also is correlated to the polarizability and solubility in pure water. Water is a lipophilic substance, and this makes other lipophilic substances easier to react with water. On the other hand, compounds that are fat hydrophilic has a low solubility in pure water. From Figure 4 it can therefore be seen that the x-axis in Figure 8 are highly affected by the solubility in pure water, which in turn are affected by the polarizability and lipophilicity.

Figure 6 shows a score plot and how the six different compounds are plotted in correlation to each other. The most striking feature in the score plot is the separation of compound 3 compared to the other compounds, which means that compound 3 is influenced by different parameters than the other compounds.

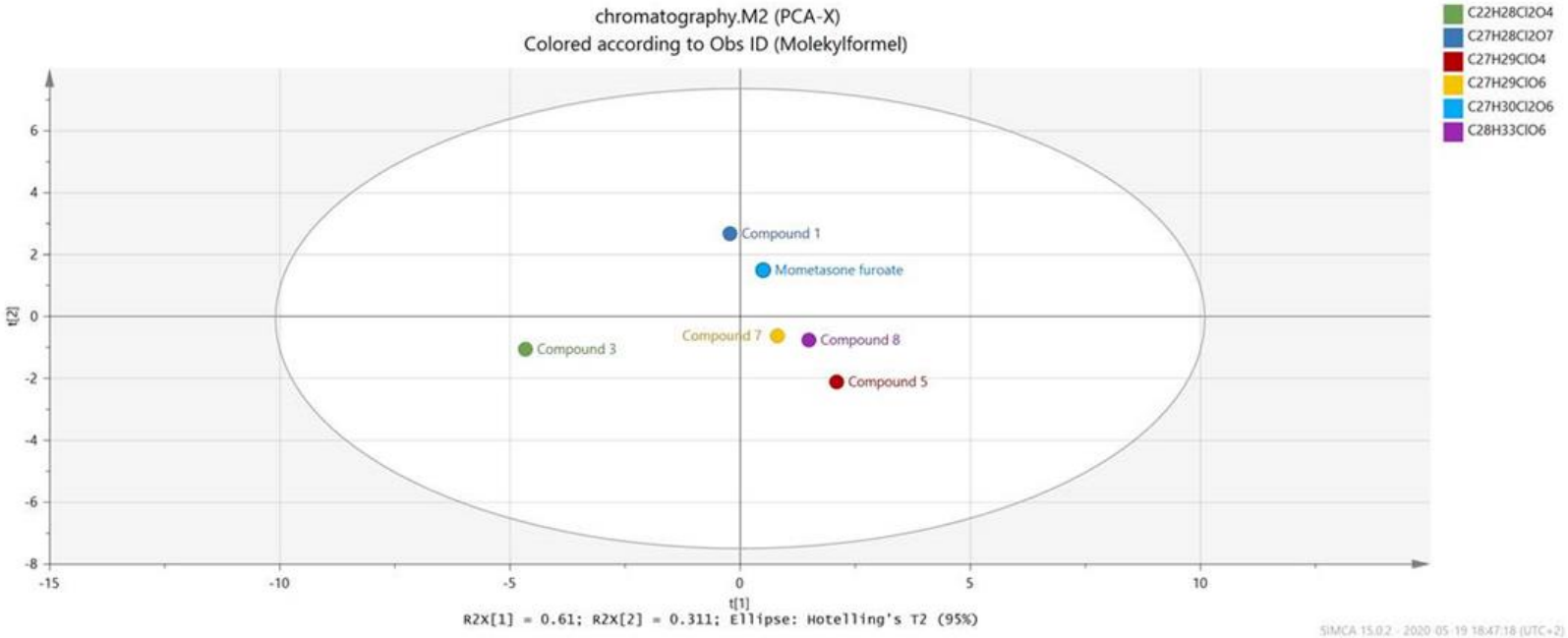


**Figure 6:** The six different compounds in a score plot.

Figure 7 shows a loading plot of the different parameters that influence the analysis, and how the parameters are related to each other. The axis p[1] is component 1, see Figure 4 and the axis p[2] is component 2, see Figure 5. Based on components 1 and 2, different compounds can be plotted in correlation to how much they are affected by the physicochemical parameters.

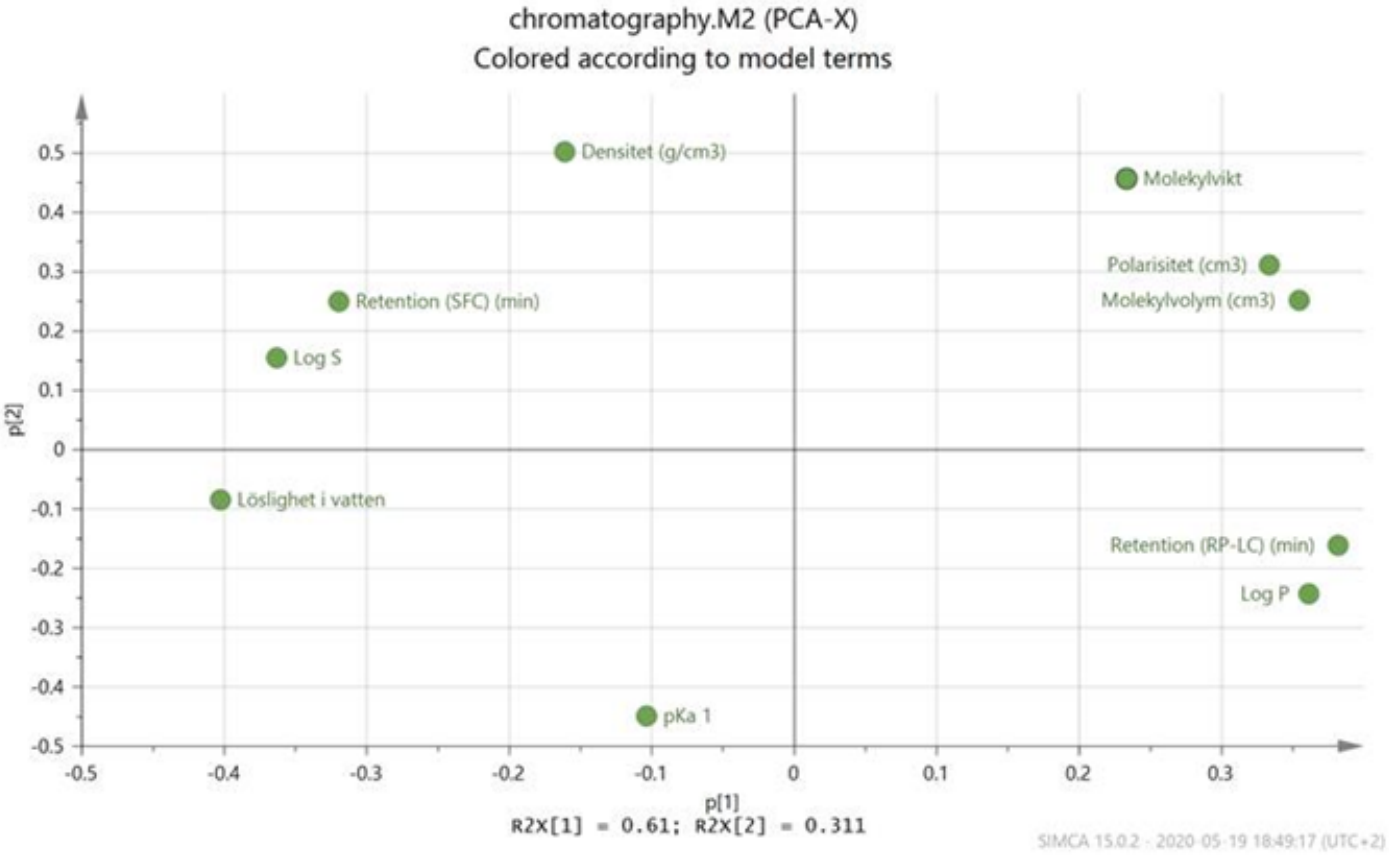



**Figure 7:** A loadings plot of the different parameters that affects the clusters.

Figure 8 shows a biplot and the different physicochemical parameters that was used to categorize the different clusters and why the substances was clustered the way it is. It shows how the substances distributes itself in the multidimensional space and the physicochemical space.

At the x-axis, a clear connection between water solubility and Log P can be read. An example that can be read from the biplot is that a high lipophilicity means that the substance exists on the right of the x-axis near the Log P, while a negative lipophilicity exists to the left and has a higher solubility in pure water. The other parameters work in the same way, the closer to the parameter point a compound is, the more the compound is affected by that parameter. It can be seen that compound 3 is the one closest to solubility in pure water which means that this compound is highly affected on that parameter. Since high Log P is found high up on the x-axis, lipophilic substances are affected accordingly. By looking at the retention times for RP-LC and SFC, it is seen that it has reverse retention order as the retention time is affected by different parameters. It can also be seen that pKa 1 spreads less in p[2], y-axis. Compound 5 is greatly affected by Log P since it is affected accordingly, but also the retention time in RP-LC.

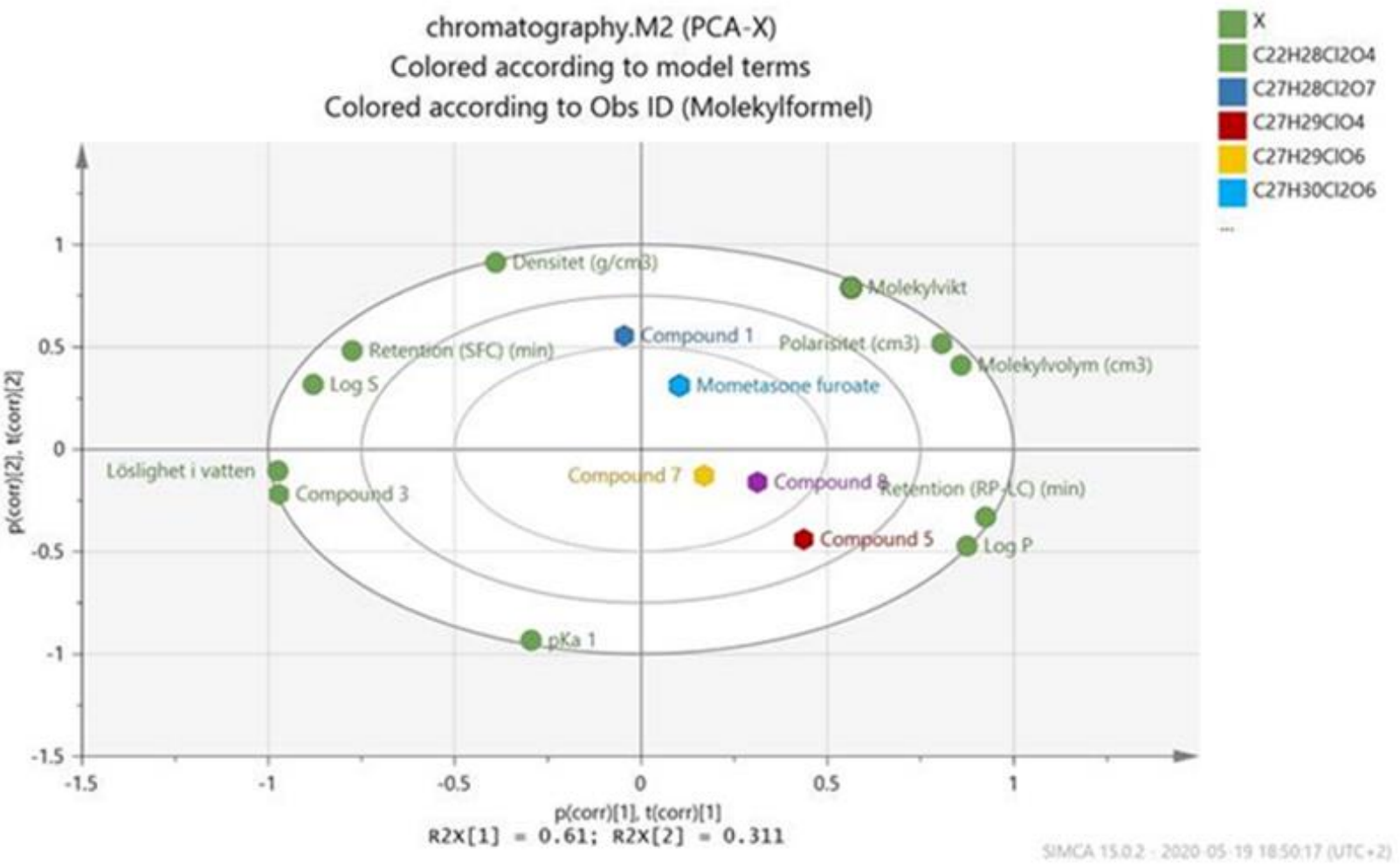



**Figure 8**: Physicochemical space and the relationships to the clusters.

In Figure 8, Compound 1 and Momentasone Furoate are drawn upwards on the y-axis. That means that some parameter is affecting both substances in that direction and that they have that parameter in common. Density and formula weight are high up on the y-axis and both parameters are quite similar between the substances, see Table 4. It is also seen that pKa 1 is the opposite parameter of density but Compound 1 and Momentasone Furoate are not drawn downwards on the y-axis. That means that density is a parameter that affects more than pKa 1, in cluster 2.

Compound 7 and 8 creates cluster 4, see Figure 3. In Figure 8, they are both in the same height at the y-axis. They have one parameter in common, the retention time in RP-LC, see Table 4. That is a parameter that affects cluster 4. Compound 5 differs from Compound 7 and 8 at the y-axis, see Figure 8. That difference might be the parameter that separates them into two different clusters, see Figure 3. All three of them has quite similar retention time in RP-LC, see Table 4, and therefore all of them are drawn to the right. One parameter that differs is the formula weight and that can affect the correlation between the three parameters in Figure 8.

### 5.3 Influence on the retention time

In Figure 8 Compound 3 is closest to the point solubility in pure water. This means that the parameter is an important part for this substance. This compound is the only substance that has two hydroxy groups, see Appendix 2, which increases the solubility in pure water. This can be one of the main parameters why Compound 3 are separated from the other parameters in the first step in Figure 3. In the study of hydroxy groups, Compound 1, see Appendix 2, and Mometasone furoate, see Appendix 2, has one hydroxy group which may be one of the main reasons for the separation in step 2, see Figure 3. The other three compounds, Compound 5, 7 and 8 did not include any hydroxy groups, see Appendix 2. This results in a similarity between the molecules and can contribute to relatively equal and low solubility in pure water.

The x-axis in Figure 8 is represented by the solubility in pure water, the polarizability and lipophilicity of the compounds. RP-LC and SFC uses the opposite compounds in their separation which means that that are correlated in the reversed place to each other on the x-axis in Figure 8. This is the reason why different retention time are reversed between RP-LC and SFC were a compound that travels faster in RP-LC will travel slowest in SFC. See Table 4 for the retention times and Appendix 17. This result in that compounds that are preferred in RP-LC will exist to the right at the x-axis which is the positive side, while compounds that are preferred in SFC will appear to the left on the negative side.

To provide longer retention and longer separation in RP-LC, substances that are more lipophilic should be used. This can be good as it gives a slow separation, but if it becomes too long a separation it will not be time efficient. The choice between SFC and RP-LC, depends on how long the retention time will be and how good the separation will be. If you choose SFC it gives a shorter

retention time, but it can also contribute to the separation going too fast and you do not get a clear result.

The model only includes six different substances, were only four of them has pKa values. This does not produce a result with the same credibility, as a model that includes more substances with pKa values. The pKa value can be a factor that can be improved in the model if more values were available. None of the compounds are close to pKa in Figure 8. This does not mean that pKa is a parameter that is not important.

The model will be a worse model compared to if more data and substances were available and a wider model could be done and analyzed. This is because only using six compounds, the physicochemical space will not be that wide. The more substances, the more data and the more credible the model becomes. When using a PCA model, there will always be a result no matter how much data that is available. Therefore, it is always important to take this into account that the more data, the more credible the model becomes.

### 5.4 Use of the model

Based on the model, four other compounds have been analyzed. The physicochemical parameters that the analysis is based on is calculated data from ACD Labs, see Table 4 and 5.

In Figure 9 the placement of the four compounds are placed, they are Caffeine, Theophylline, Theobromine and Uracil. The physicochemical space is the same as in Figure 8. According to the literature, SFC is a good method when nonpolar substances should be separated and according to the model the nonpolar substances will occur on the left side of the x-axis. The opposite is that polar substances will occur on the right side and should be separated with RP-LC/UHPLC.

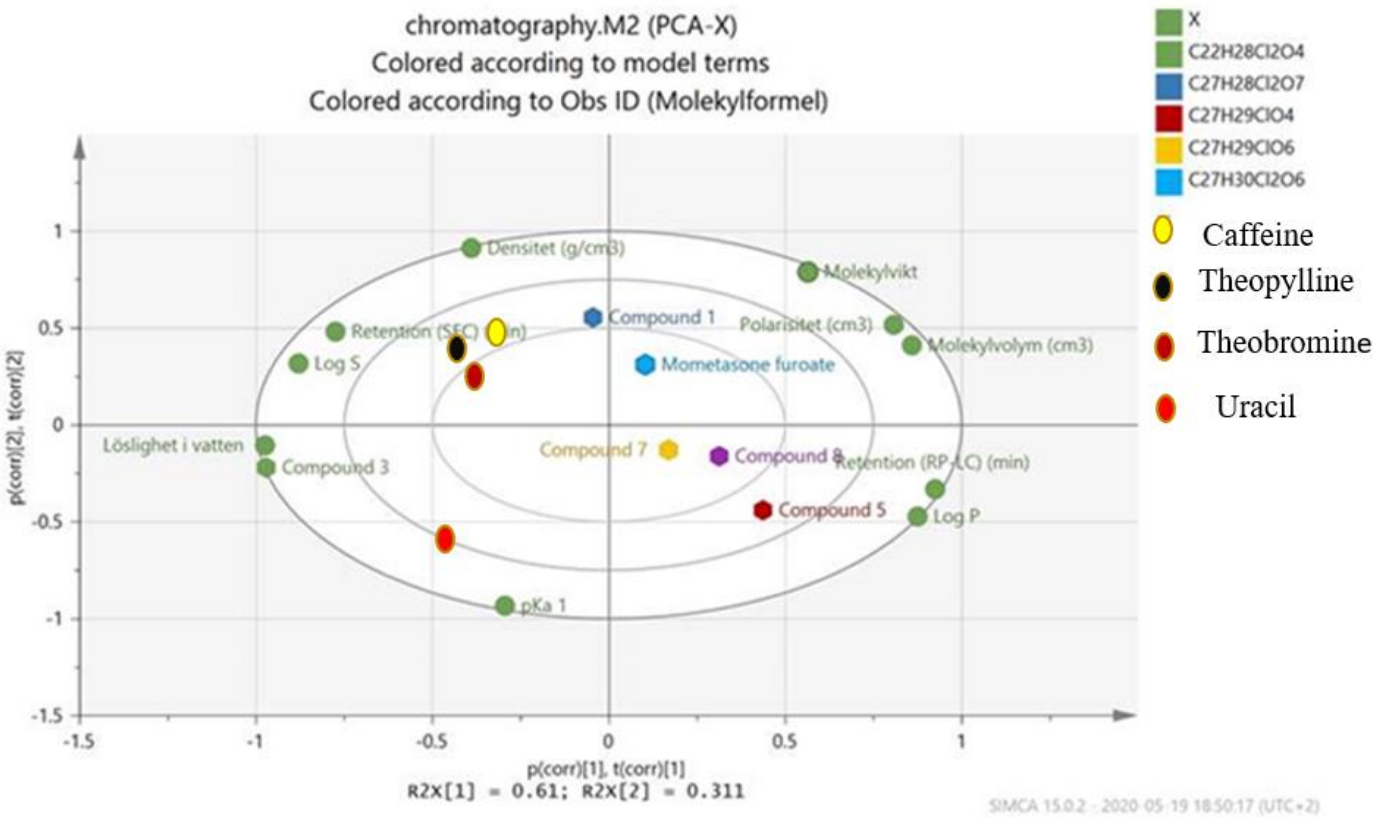



**Figure 9:** Placement of Caffeine, Theophylline, Theobromine and Uracil based on Figure 8.

### 5.4.1 Caffeine

Caffeine was compared both with the other compounds that already was plotted in Figure 8, but also from the values of the parameters in Table 4. On the x-axis, mainly the parameters solubility in pure water, Log P and Polarizability were studied, but also other parameters were used. It has a low and negative Log P value, low polarizability and low solubility in pure water. From these values, it is clear that caffeine will lie to the left, on the negative side of the x-axis. On the y-axis mainly density and pKa 1 values are studied, but also other parameters are represented. The pKa 1 value of caffeine is the lowest pKa 1 value of all compounds. The density is relatively low and seems to have less impact then the pKa 1 value. From these parameters, the caffeine suvstance will appear in upper part of the y-axis and has a positive value. The caffeine compound can be seen in Figure 9 as the yellow point.

### 5.4.2 Uracil

Uracil has a higher negative value on Log P and a lower polarizability then caffeine which means that those parameters are more to the left then in caffeine. But on the other hand, it has a lower solubility in pure water then caffeine which affects by moving the point to the right. As a result of these parameters Uracil will lie on the negative side of the x-axis, and a little bit more to the left than Caffeine. The density is low and similar to the other compounds. The pKa 1 value of Uracil is much higher than caffeine and highest of all compounds. This means that it will has a huge impact and appear on the negative side of the y-axis. Uracil can be found on the negative side of both the x and y-axis, and lowest of all compounds, se Figure 9.

### 5.4.3 Theobromine

Theobromine has a lower pKa 1 value than Uracil but a higher pKa 1 value than Caffeine. That means that the pKa 1 in Figure 9 will draw the theobromine in the negative direction at the y-axis. The density is high compared to the other substances, see Table 4 and 5, and that means that the density in Figure 9 will draw the theobromine in a positive direction at the y-axis. Log P and Solubility in pure water are each other's opposites, Theobromine has lower solubility in pure water than both Uracil and Caffeine. That means that it will not be drawn as much to the left on the x-axis, but the Log P value is negative and that means that Theobromine will be on the negative side at the x-axis since there is two parameters that draws the substance to the left. See the placement of Theobromine in Figure 9.

### 5.4.4 Theophylline

Theophylline has a lower pKa 1 value than Uracil but a higher pKa 1 value than Theobromine, and the density is the highest value of all the substances, see Table 4 and 5. That means that the density will affect Theophylline more than pKa 1, so Theophylline will be in a positive direction at the y-axis. The Log P value is the opposite to Solubility in pure water and Theophylline has a negative Log P value, which means that Theophylline will draw to the left on the x-axis. The solubility in pure water is lower than Theobromine and the Log P value is more negative for Theobromine compared to Theophylline. Therefore, it will be more to the left on the x-axis. See placement of Theophylline in Figure 9.

## 5.5 Error sources

The calculated data that is used, is only calculated by one software so the values are not confirmed by another software. It is important to be aware that these are theoretical values that are calculated on bases of mathematical functions and that this does not necessarily give the exact value in practice. When using different programs which is similar to ACD Labs, the result would probably not be exactly the same, depending on the programming and the mathematical that it is based on. Since the possibility of using several software not has been able, literature values have been used for comparision. Although all values have not been found in theory at the webpage DrugBank, it is clear that the values that were found differ significantly. This means that it is difficult to

determine which values that gives the most accurate result.. This can cause a large deviation depending on the data collection of the physicochemical properties.

Another error source is that only a few substances has been used in the modeling, which results in a very narrow and non-general physicochemical space. To get it more general, more substances with a big variation of physicochemical parameters should be analyzed. All substances analyzed in this thesis is quite similar to each other, which makes the modelling non-general.

One error source that is needed to take into account is that missing data is risky in small models. As seen in Table 4 and 5, the pKa values were not able to be calculated for all the substances. It was not able to be calculated due to some structures, this has a lot of impact since there is not so much data to base the model on. This is an error source because it can give an uncertain model, due to the missing data.

## 6 Conclusion

In RP-LC, a non-polar phase and a water-based or solution-based mobile phase are used. The mobile phase is polar and the stationary phase is non-polar. Therefore, if the substance appears on the positive side of component 1, it is more lipophilic and should be separated with RP-LC to get a longer retention time and longer separation. In SFC, a slightly polar stationary phase and compressed $CO_2$ as the mobile phase are used. If a substance appears on the negative side of component 1, it is more polar and should be separated with SFC to get a longer retention time and possibly increased separation. This is confirmed with both the model and the literature study. What determines if there occurs retention is whether to use the right mobile and stationary phase in relation to what is to be separated. Which, in summary, is the use of the right method.

SFC can be used more in the pharmaceutical industry if the S/N-ratio is improved by using both UV and MS as detectors. This is to increase the selectivity and to be able to analyze smaller compounds with lower concentrations.

## 7 Future work

This thesis was planned to be done as a practical thesis. Due to prevailing circumstances of COVID-19, it has been done as a theoretical thesis on distance. Next step is to use the modelling and physicochemical space and to do a laboratory test to conform the theoretical and calculated data. To confirm the calculated data, another software should calculate the same parameters to ensure the response in this thesis.

As seen in Figure 3 and 4, some of the parameters does not affect the modelling to such an extent. These parameters can be removed to optimize the model, to model only significant parameters.

**References**

[1] Beginners Guide to liquid Chromatography [Internet]. Netherlands: Waters; [cited 23 march 2020]. Available from: https://www.waters.com/waters/en_SE/HPLC---High-Performance-Liquid-Chromatography-Explained/nav.htm?cid=10048919&locale=en_SE

[2] Terry. A. Berger. Supercritical Fluid Chromatography: Primer. USA: Agilent Technologies; 2015

[3] Bertin. S, Lemasson. E, West. C. Journal of Separation Science [e-book]. Weinheim: Wiley-VCH Verlag GmbH & Co; 2015 [cited 23 March 2020] Available from: https://onlinelibrary-wiley-com.proxy.lib.chalmers.se/doi/epdf/10.1002/jssc.201501062

[4] Desfontaine. V, Francotte. E Guillarme. D, Nováková. L: Journal of Pharmaceutical and Biomedical Analysis [internet] ELSEVIER; 2015 [cited 9 April 2020]. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0731708515001752?via%3Dihub

[5] Lesellier. E, Wets. C: Journal of chromatography A [internet]. Orleans: ELSEVIER; 2014 [cited 23 March 2020]. Available from: https://www.journals.elsevier.com/journal-of-chromatography-a/

[6] Dhanarasu. S. Chromatography and its applications [e-book]. Rijeka: InTech; 2012 [cited 15 April 2020]. Available from: http://library.um.edu.mo/ebooks/b28050630.pdf

[7] Lamotte. S, Machtejevas. E, Unger. K.K. Column technology in liquid chromatography: Liquid Chromatography- Fundamentals and Instrumentation [e-book]. 2nd Edition. Amsterdam: ELSEVIER Inc; 2017 [cited 6 April 2020]. Chapter 3. Available from: [https://app.knovel.com/web/view/khtml/show.v/rcid:kpLCFIVE01/cid:kt011FJS61/viewerType:khtml//root_slug:3-column-technology-in-liquid-chromatography/url_slug:column-technology-in?kpromoter=federation&b-toc-cid=kpLCFIVE01&b-toc-root-slug=&b-toc-url-slug=column-technology-in&b-toc-title=Liquid%20Chromatography%20-%20Fundamentals%20and%20Instrumentation%2C%20Volume%201%20(2nd%20Edition)&page=2&view=collapsed&zoom=1

[8] Beginner's Guide to UPLC [Internet]. Netherlands: Waters; [cited 23 March 2020]. Available from: https://www.waters.com/waters/en_SE/UPLC---Ultra-Performance-Liquid-Chromatography-Beginner%27s-Guide/nav.htm?cid=134803622&locale=en_SE

[9] Armenta. S, De La Guardia. M. Downsizing the Methods. [Internet] ELSEVIER; [cited 14 April 2020] Available from: https://www.sciencedirect.com/science/article/pii/B9780444537096000070

[10] Kazakevich Y, Lobrutto R. HPLC for Pharmaceutical Scientists [e-book]. New Jersey: John Wiley & Sons Inc., 2007. [cited 8 April 2020] Available from: https://books.google.se/books?hl=sv&lr=&id=csAk-MpeTocC&oi=fnd&pg=PR5&dq=hplc+in+pharmaceutical+industry&ots=Dn83Um0IQn&sig=x6DkIMmbWWJmqSEmOK3w9WJFix8&redir_esc=y#v=onepage&q=hplc%20in%20pharmaceutical%20industry&f=false

[11] Ultra Performance Liquid Chromatography [Internet] ELSEVIER; [updated 2020; cited 14 April 2020] Available from: https://www.sciencedirect.com/topics/chemistry/ultra-performance-liquid-chromatography

[12] Farrokhi. V, Li. S, McShane. A. J, Nemati. R, Yao. X. An Overview of Quantitative Proteomic Approaches [Internet] ELSEVIER; [updated 2020; cited 15 April 2020] Available from: https://www.sciencedirect.com/science/article/pii/B9780444626516000052

[13] Ultra Performance Liquid Chromatography [Internet] ELSEVIER; [updated 2020; cited 14 April 2020] Available from: https://www.sciencedirect.com/topics/chemistry/ultra-performance-liquid-chromatography

[14] Rochfort. S.J, Trenery. V.C. Natural Products Research and Metabolomics [Internet] Werribee: ELSEVIER; [cited 19 April 2020] Available from: https://www.sciencedirect.com/science/article/pii/B9780080453828002112

[15] Harvey. D, High-Performanc Liquid chromatography [Internet]. Greencastle: DePauw University; 2019 [updated 5 June 2019, cited 8 april 2020]. Available from: https://chem.libretexts.org/Courses/Northeastern_University/12%3A_Chromatographic_and_Electrophoretic_Methods/12.5%3A_High-Performance_Liquid_Chromatography

[16] Chesnut. S.M, Salisbury. J.J. The role of UHPLC in pharmaceutical development [Internet]. Weinheim: Wiley; 2007. [cited 12 May 2020] Available from: https://onlinelibrary.wiley.com/doi/epdf/10.1002/jssc.200600505

[17] Cielecka-Piontek. J, Zalewski. P, Jelińska. A. UHPLC: The Greening Face of Liquid Chromatography [Internet]. Chromatographia; 2013. [cited 12 May 2020] Available from: https://link.springer.com/article/10.1007/s10337-013-2434-6

[18] A quick look at reversed phase HPLC [internet]. Phenomenex; 2020 [cited 25 April 2020]. Available from: https://www.phenomenex.com/hplc-column/reversed-phase-hplc-column

[19] Alexander, A.J. SFC Instrument Modification to Allow Greater Flexibility in Method Development by Generating Mixtures of Solvents and Modifiers On-Line for Mobile Phase B Optimization [Internet]. New Brunswick: Springer; 2012. [cited 29 April 2020] Available from: https://link-springer-com.proxy.lib.chalmers.se/content/pdf/10.1007/s10337-012-2291-8.pdf

[20] Collier. S.M, Hill. J.F, Tarafder. A. Addressing robust operation and reproducibility in SFC. I: Rossé.G. Supercritical Fluid Chromatography [e-book]. Philadelphia: Walter De Gruyter; 2019 [cited 7 May 2020]. Page 73-94. Available from: https://www-degruyter-com.proxy.lib.chalmers.se/view/title/523517

[21] Bertin. S, Lemasson. E, West. C. Journal of Separation Science [e-book]. Weinheim: Wiley-VCH Verlag GmbH & Co; 2015 [cited 23 March 2020] Available from: https://onlinelibrary-wiley-com.proxy.lib.chalmers.se/doi/epdf/10.1002/jssc.201501062

[22] DrugBank[Internet] DrugBank: 2020; [updated 2020; cited 29 April 2020] Available from: https://www.drugbank.ca/

[23] Babić. S, Horvat. A.J.M, Ka Mutavžić Pavlović. D, Kaštelan-Macan. M. Determination of $pK_a$ values of active pharmaceutical ingredients [Internet]. Zagreb: ELSEVIER; 2009. [cited 10 May 2020] Available from: https://www.sciencedirect.com/science/article/pii/S0165993607001884

[24] Skyner. R.E, McDonagh. J.L, Groom. C.R, Van Mourik. T, Mitchell. J.B.O. A review of methods for the calculation of solution free energies and the modelling of systems of solution [Internet]. Cambridge: Royal Society of Chemistry; 2015. [cited 11 May 2020] Available from: https://pubs.rsc.org/en/content/articlehtml/2015/cp/c5cp00288e

[25] Baczek. T, Haber. P, Kaliszan. R,Siluk. D, Valk. K. Lipophilicity and $pK_a$ estimates from gradient high-performance liquid chromathography [Internet]. ELSEVIER; 2009. [cited 12 May 2020] Available from: https://www.sciencedirect.com/science/article/pii/S0021967302001140

[26] Donovan. B, Liu. O, Wang. Z, Zhang. H. Development of an orthogonal method for mometasone furoate impurity analysis using supercritical fluid chromatography [Internet]. Summit: Elsevier; 2011 [cited 27 April] Available from: https://www-sciencedirect-com.proxy.lib.chalmers.se/science/article/pii/S0021967311002421?via%3Dihub

# 8 Appendix

## Appendix 1, BIOVIA Draw


**Figure A1.1:** BIOVIA Draw software, version 16.1

**Appendix 2, Substances drawn in BIOVIA Draw**

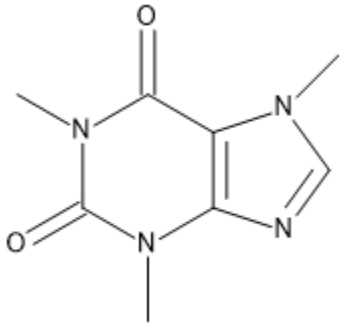



**Figure A2.1:** Caffeine

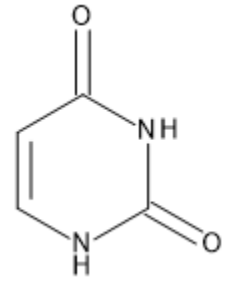



**Figure A2.2:** Uracil

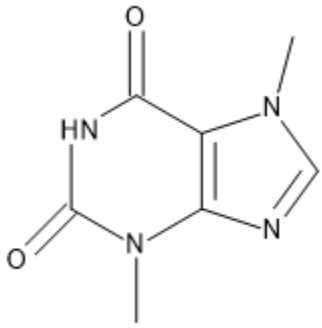



**Figure A2.3:** Theobromine

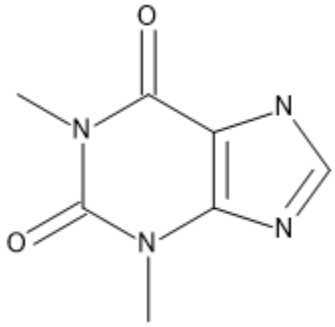



**Figure A2.4:** Theophylline

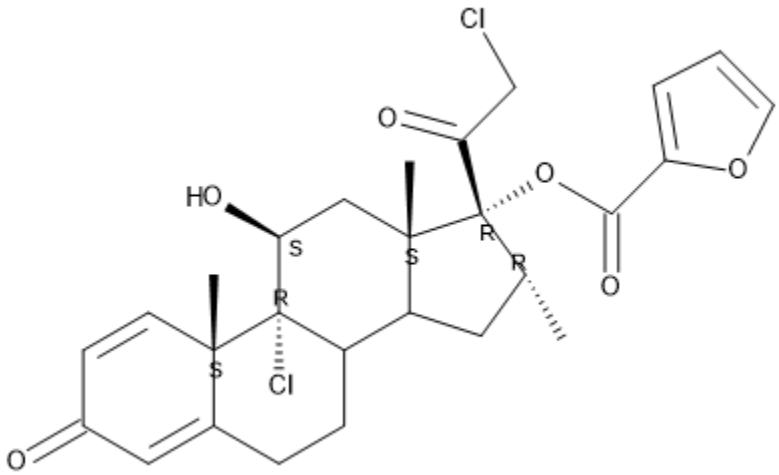



**Figure A2.5:** Mometasone Furoate

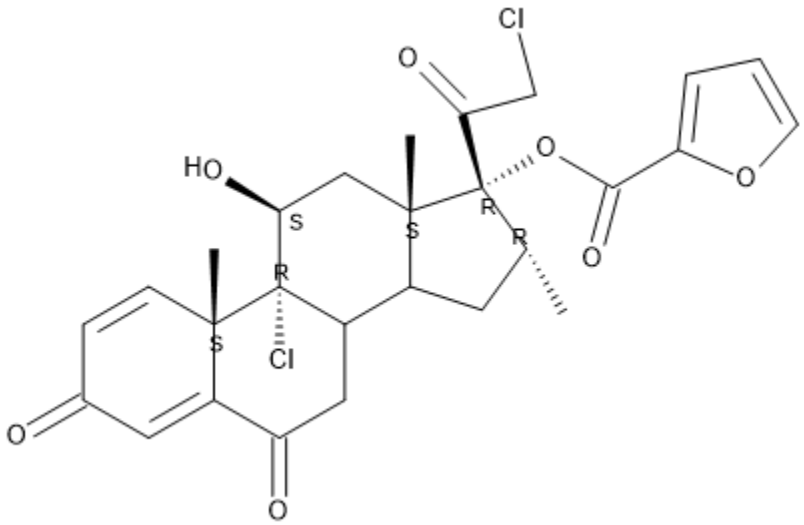



**Figure A2.6:** Compound 1

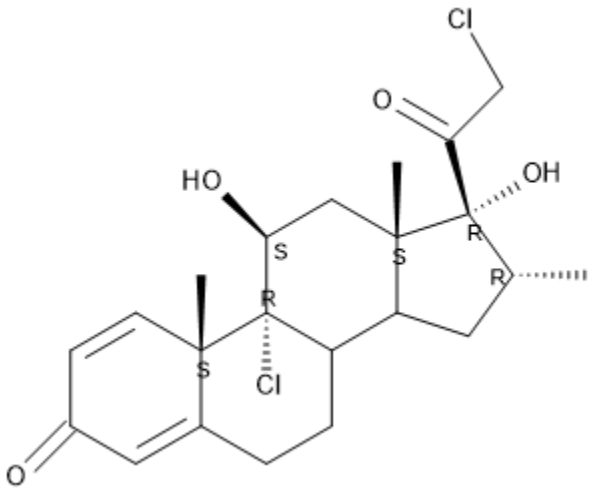



**Figure A2.7:** Compound 3

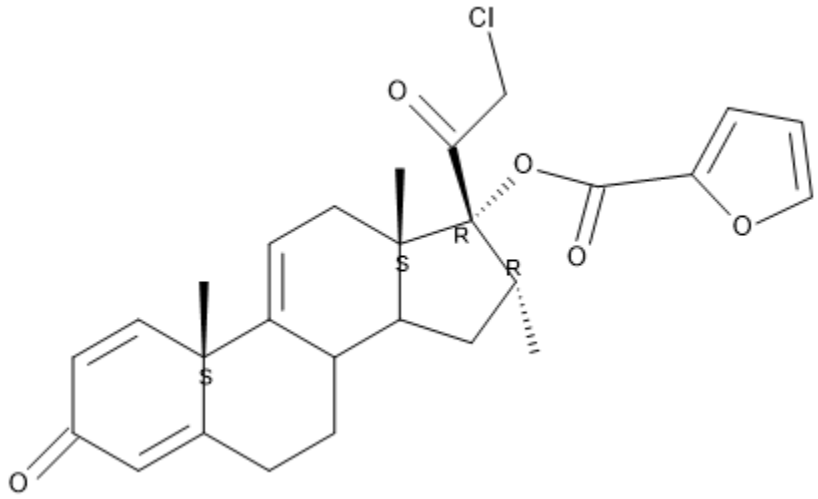



**Figure A2.8:** Compound 5

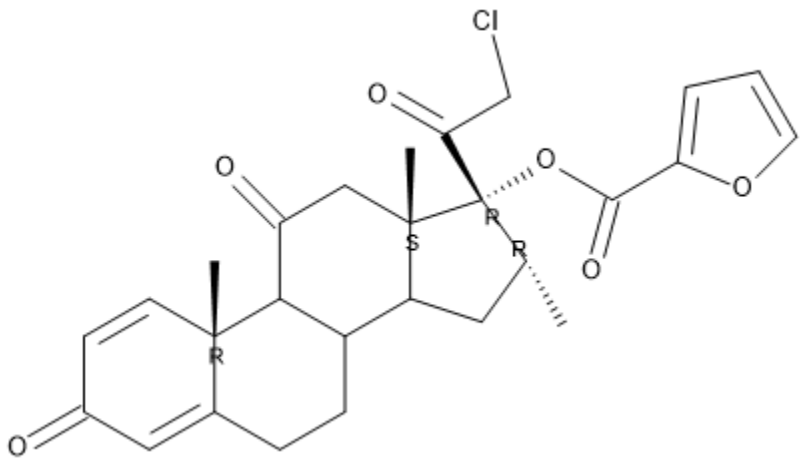



**Figure A2.9:** Compound 7

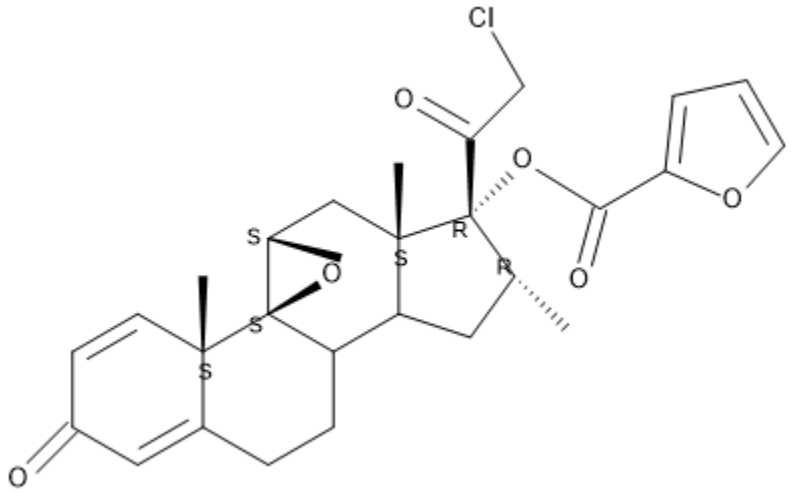
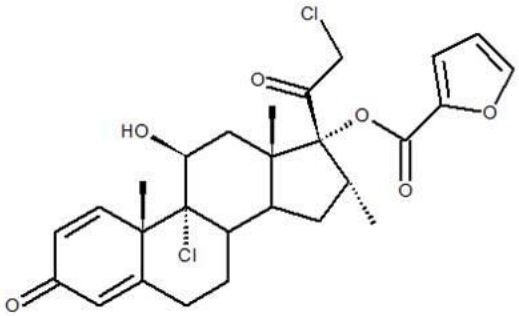



**Figure A.10:** Compound 8

**Appendix 3, Mometasone furoate**




| | |
|---|---|
| Molecular Formula: | $C_{27}H_{30}Cl_2O_6$ |
| Formula Weight: | 521.4295 |
| Composition: | C(62.19%) H(5.80%) Cl(13.60%) O(18.41%) |
| Molar Refractivity: | 130.40 ± 0.4 $cm^3$ |
| Molar Volume: | 379.2 ± 5.0 $cm^3$ |
| Parachor: | 1041.7 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.603 ± 0.03 |
| Surface Tension: | 56.9 ± 5.0 dyne/cm |
| Density: | 1.37 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 51.69 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 12 |
| Monoisotopic Mass: | 520.141944 Da |
| Nominal Mass: | 520 Da |
| Average Mass: | 521.4295 Da |
| M+: | 520.141395 Da |
| M-: | 520.142493 Da |
| [M+H]+: | 521.149221 Da |
| [M+H]-: | 521.150318 Da |
| [M-H]+: | 519.13357 Da |
| [M-H]-: | 519.134668 Da |

**Figure A3.1:** Calculations of physicochemical parameters

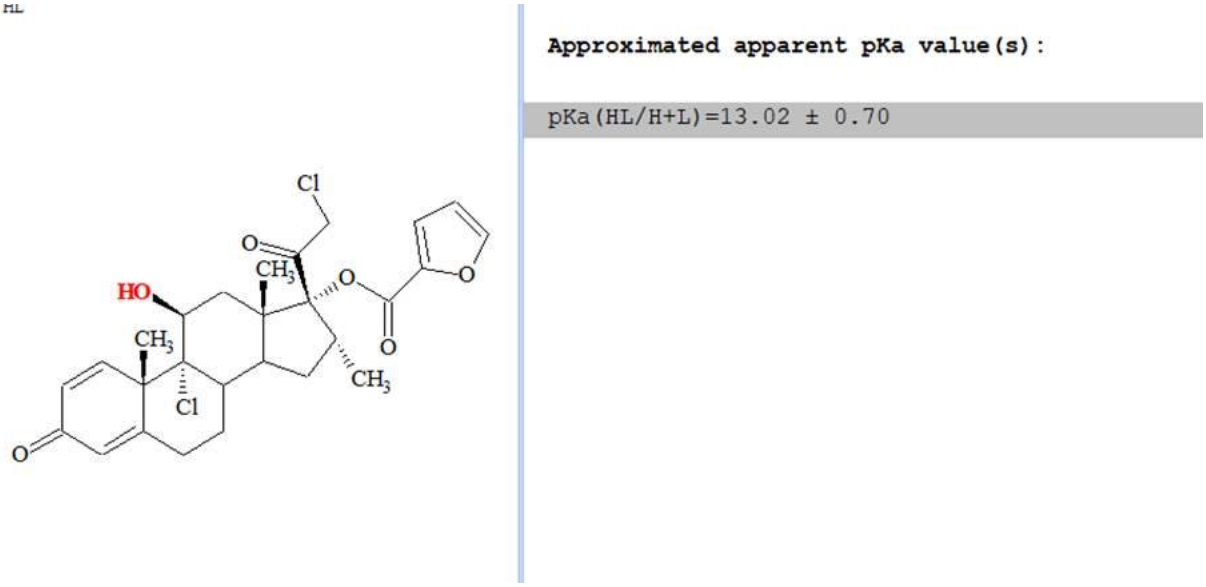
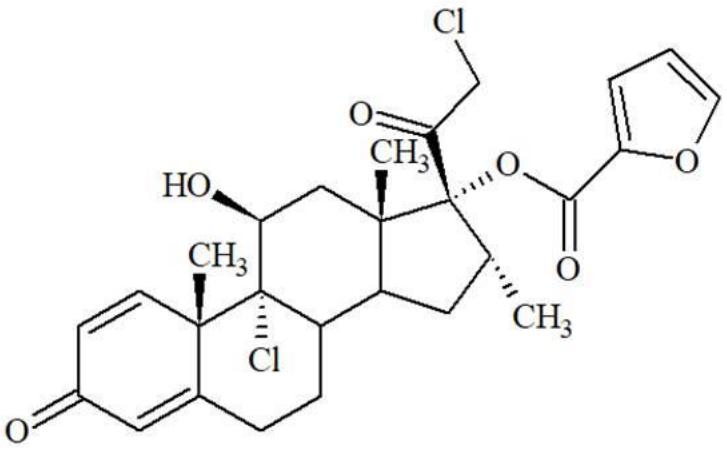



**Figure A3.2:** Calculated pKa value




**Calculated Solubility In Pure Water**

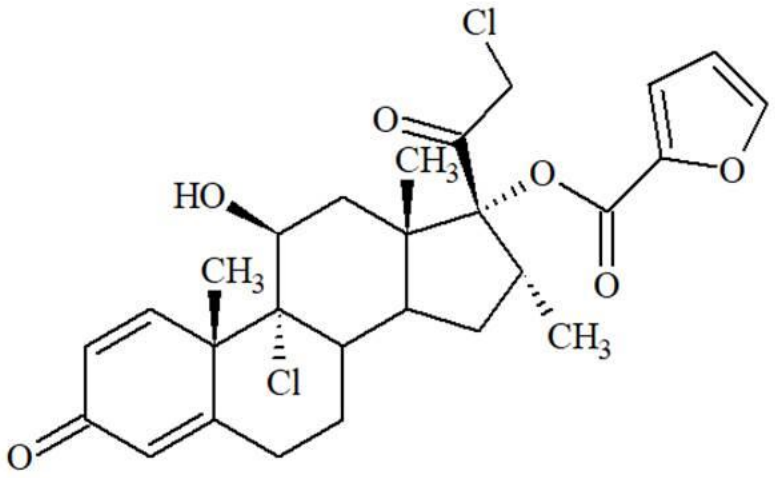
Solubility (mol/l) = 1.8e-6 at pH = 7.00

**Log(S) for neutral form: -5.74**

**Monoanions:**

Log(S) (Charge on O (or S) near aliphatic): -1.64 ± 1.00

**Figure A3.3:** Calculated solubility in pure water and Log S




c = 100.0%; c(H2O) = 0.0%; c(org) = 100.0%

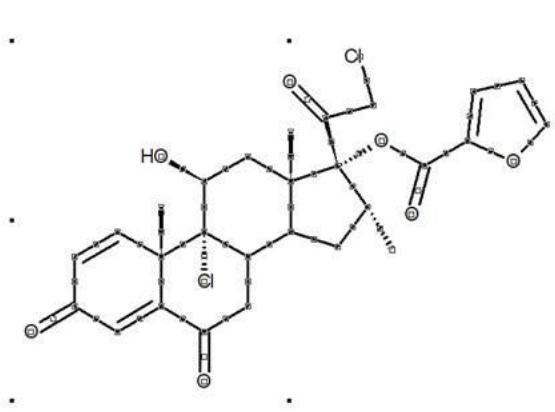
**LogP for neutral form: 4.27 ± 0.58**

**Monoanions:**

LogP (Charge on O (or S) near aliphatic): 0.17 ± 1.00

**Figure A3.4:** Calculated Log P

## Appendix 4, Compound 1



| | |
|---|---|
| Molecular Formula: | $C_{27}H_{28}Cl_2O_7$ |
| Formula Weight: | 535.41302 |
| Composition: | C(60.57%) H(5.27%) Cl(13.24%) O(20.92%) |
| Molar Refractivity: | 130.51 ± 0.4 $cm^3$ |
| Molar Volume: | 376.0 ± 5.0 $cm^3$ |
| Parachor: | 1046.8 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.610 ± 0.03 |
| Surface Tension: | 60.0 ± 5.0 dyne/cm |
| Density: | 1.42 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 51.74 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 13 |
| Monoisotopic Mass: | 534.121209 Da |
| Nominal Mass: | 534 Da |
| Average Mass: | 535.413 Da |
| M+: | 534.12066 Da |
| M-: | 534.121757 Da |
| [M+H]+: | 535.128485 Da |
| [M+H]-: | 535.129582 Da |
| [M-H]+: | 533.112835 Da |
| [M-H]-: | 533.113932 Da |

**Figure A4.1:** Calculations of physicochemical parameters

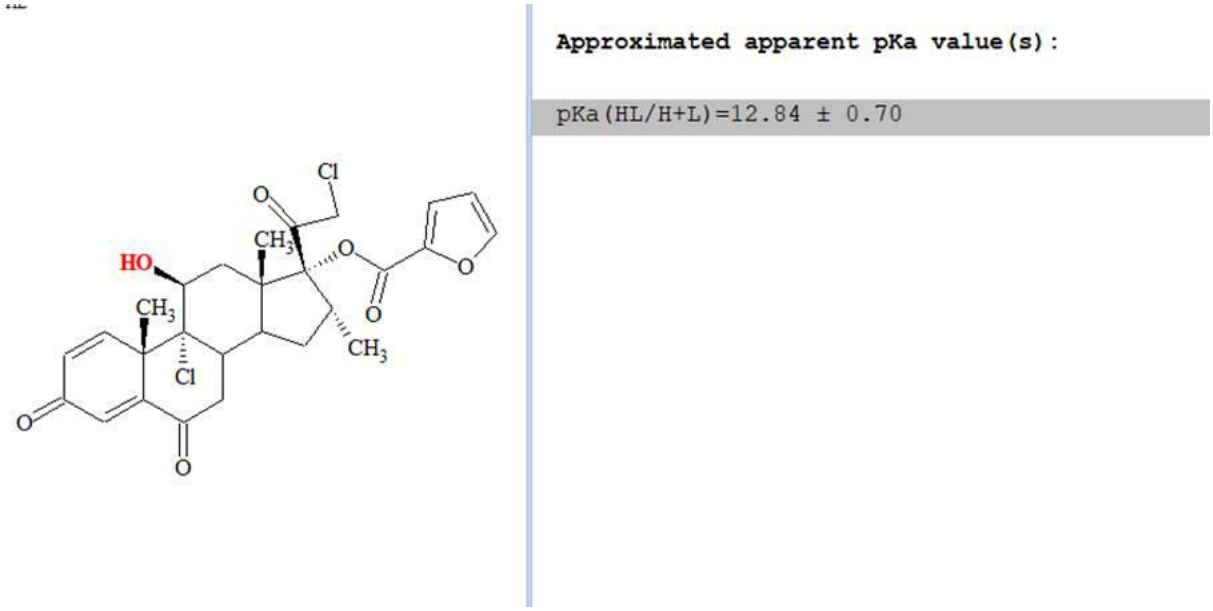
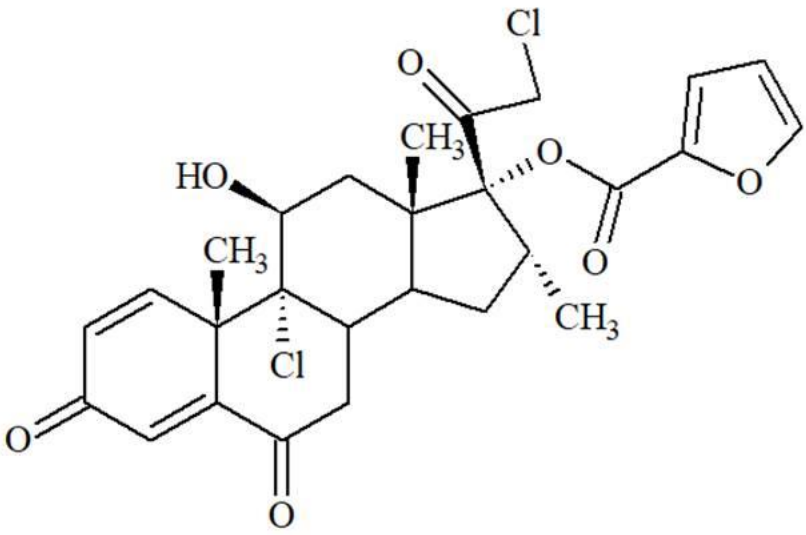



**Figure A4.2:** Calculated pKa value




**Calculated Solubility In Pure Water**

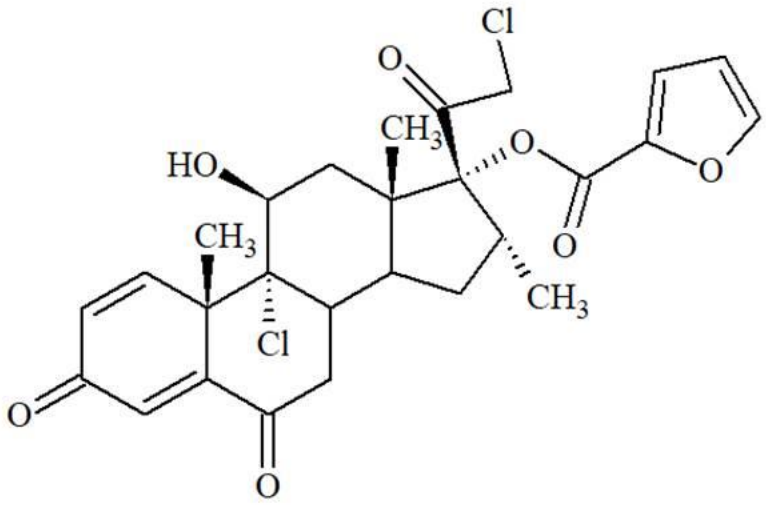
Solubility (mol/l) = 2.6e-6 at pH = 7.00

**Log(S) for neutral form: -5.58**

**Monoanions:**

Log(S) (Charge on O (or S) near aliphatic): -1.48 ± 1.00

**Figure A4.3:** Calculated solubility in pure water and Log S




c = 100.0%; c(H2O) = 0.1%; c(org) = 99.9%

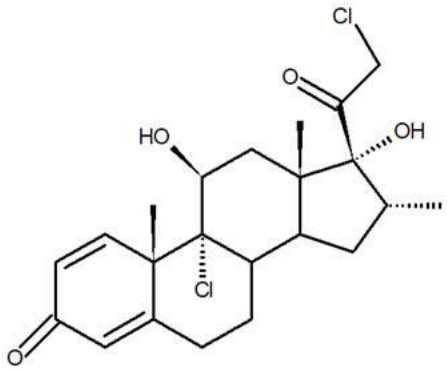
**LogP for neutral form: 3.28 ± 0.59**

**Monoanions:**

LogP (Charge on O (or S) near aliphatic): -0.82 ± 1.00

**Figure A4.4:** Calculated Log P

## Appendix 5, Compound 3




| | |
|---|---|
| Molecular Formula: | $C_{22}H_{28}Cl_2O_4$ |
| Formula Weight: | 427.36132 |
| Composition: | C(61.83%) H(6.60%) Cl(16.59%) O(14.98%) |
| Molar Refractivity: | 108.25 ± 0.4 $cm^3$ |
| Molar Volume: | 316.5 ± 5.0 $cm^3$ |
| Parachor: | 864.0 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.599 ± 0.03 |
| Surface Tension: | 55.5 ± 5.0 dyne/cm |
| Density: | 1.35 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 42.91 ± 0.5 $10^{-24} cm^3$ |
| RDBE: | 8 |
| Monoisotopic Mass: | 426.136465 Da |
| Nominal Mass: | 426 Da |
| Average Mass: | 427.3613 Da |
| M+: | 426.135916 Da |
| M-: | 426.137013 Da |
| [M+H]+: | 427.143741 Da |
| [M+H]-: | 427.144838 Da |
| [M-H]+: | 425.128091 Da |
| [M-H]-: | 425.129188 Da |

**Figure A5.1:** Calculations of physicochemical parameters.

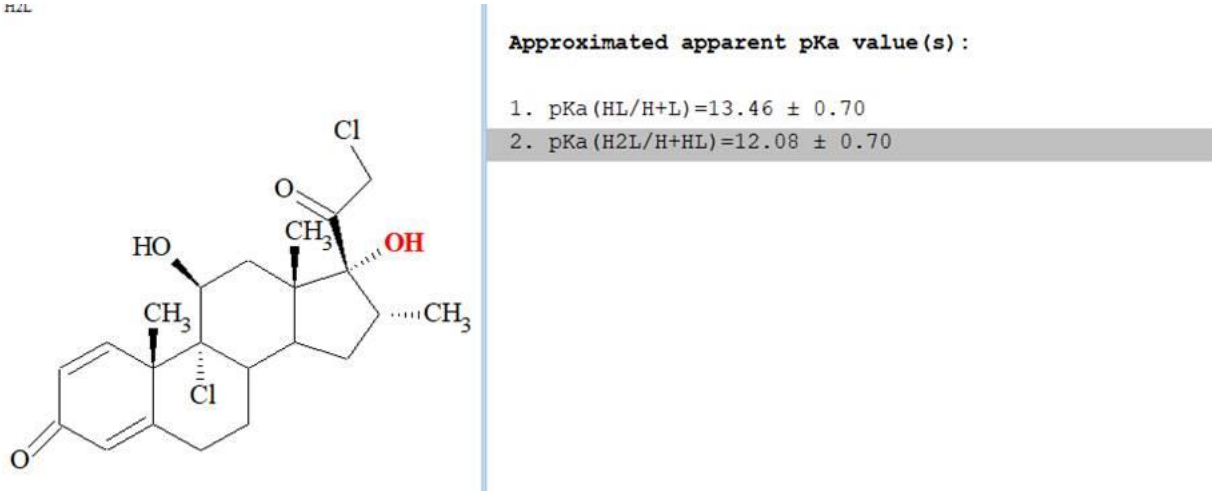
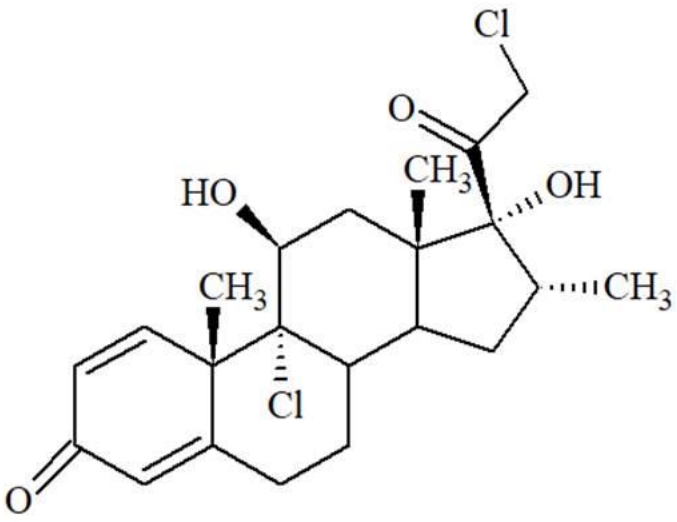



**Figure A5.2**: Calculated pKa value




**Calculated Solubility In Pure Water**

Solubility (mol/l) = 1.2e-5 at pH = 7.00

**Log(S) for neutral form: -4.92**

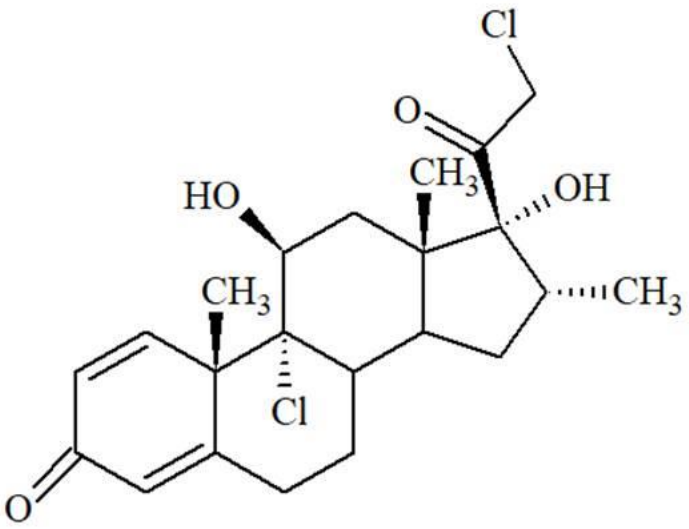
**Monoanions:**

Log(S) (Charge on O (or S) near aliphatic): -0.82 ± 1.00

**Other charged forms:**

Log(S) (summing charge: -2): 0.18 ± 1.50

**Figure A5.3:** Calculated solubility in pure water and Log S




c = 100.0%; c(H2O) = 0.1%; c(org) = 99.9%

**LogP for neutral form: 2.91 ± 0.55**

**Monoanions:**

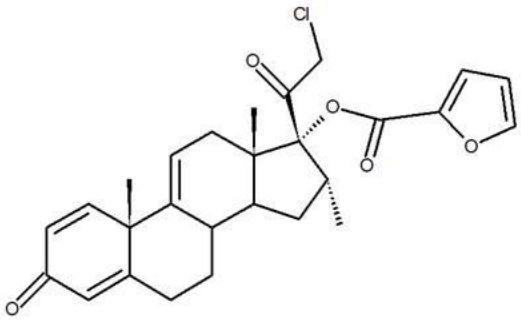
LogP (Charge on O (or S) near aliphatic): -1.19 ± 1.00

**Other charged forms:**

LogP (summing charge: -2): -2.19 ± 1.50

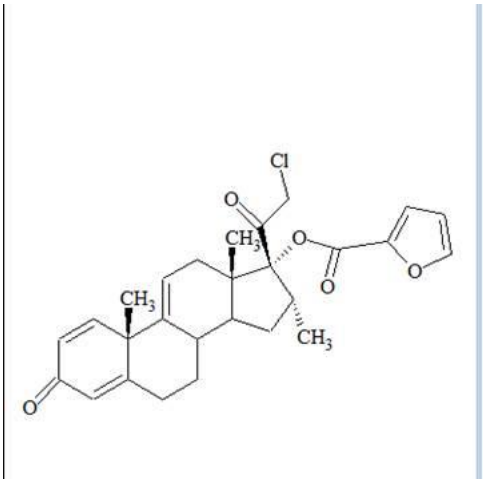
**Figure A5.4:** Calculated Log P

## Appendix 6, Compound 5



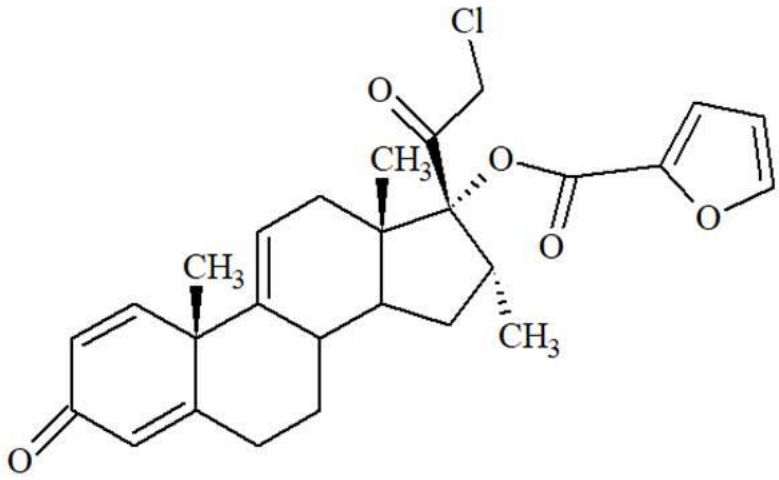

| | |
|---|---|
| Molecular Formula: | $C_{27}H_{29}ClO_5$ |
| Formula Weight: | 468.96916 |
| Composition: | C(69.15%) H(6.23%) Cl(7.56%) O(17.06%) |
| Molar Refractivity: | 123.72 ± 0.4 $cm^3$ |
| Molar Volume: | 362.9 ± 5.0 $cm^3$ |
| Parachor: | 975.4 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.597 ± 0.03 |
| Surface Tension: | 52.1 ± 5.0 dyne/cm |
| Density: | 1.29 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 49.04 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 13 |
| Monoisotopic Mass: | 468.170352 Da |
| Nominal Mass: | 468 Da |
| Average Mass: | 468.9692 Da |
| M+: | 468.169803 Da |
| M-: | 468.1709 Da |
| [M+H]+: | 469.177628 Da |
| [M+H]-: | 469.178725 Da |
| [M-H]+: | 467.161978 Da |
| [M-H]-: | 467.163075 Da |

**Figure A6.1**: Calculations of physicochemical parameters.




```
Approximated apparent pKa value(s) are not calculated
```

**Figure A6.2:** Calculated pKa value

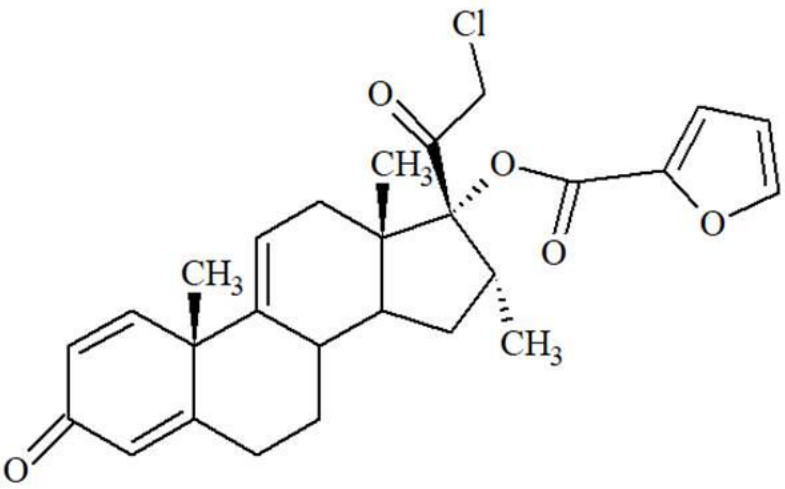



**Calculated Solubility In Pure Water**

Solubility (mol/l) = 2.6e-7 at pH = 7.00

**Log(S) for neutral form: -6.58**

**Figure A6.3:** Calculated solubility in pure water and Log S

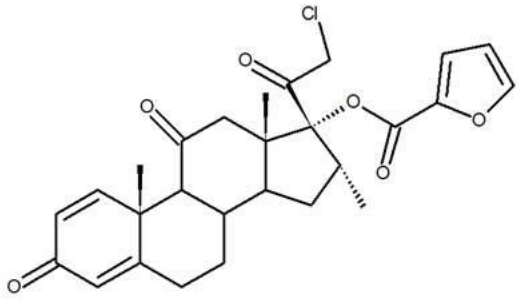



c = 100.0%; c(H2O) = 0.0%; c(org) = 100.0%

**LogP for neutral form: 5.26 ± 0.57**

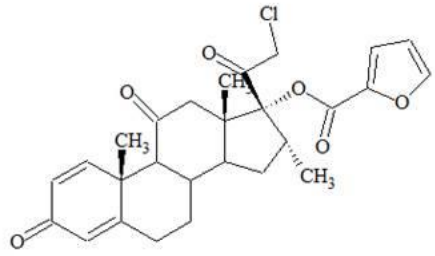
**Figure A6.4:** Calculated Log P

**Appendix 7, Compound 7**



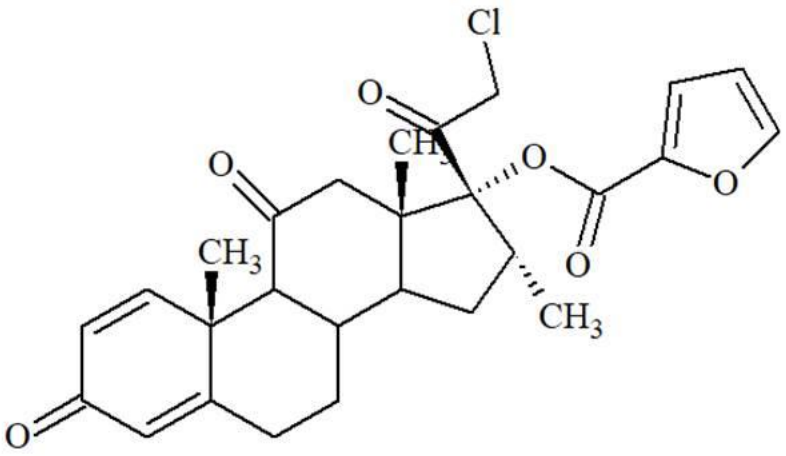

| | |
|---|---|
| Molecular Formula: | $C_{27}H_{29}ClO_6$ |
| Formula Weight: | 484.96856 |
| Composition: | C(66.87%) H(6.03%) Cl(7.31%) O(19.79%) |
| Molar Refractivity: | 124.13 ± 0.4 $cm^3$ |
| Molar Volume: | 367.1 ± 5.0 $cm^3$ |
| Parachor: | 994.6 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.591 ± 0.03 |
| Surface Tension: | 53.8 ± 5.0 dyne/cm |
| Density: | 1.32 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 49.21 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 13 |
| Monoisotopic Mass: | 484.165266 Da |
| Nominal Mass: | 484 Da |
| Average Mass: | 484.9686 Da |
| M+: | 484.164718 Da |
| M-: | 484.165815 Da |
| [M+H]+: | 485.172543 Da |
| [M+H]-: | 485.17364 Da |
| [M-H]+: | 483.156893 Da |
| [M-H]-: | 483.15799 Da |

**Figure A7.1:** Calculations of physicochemical parameters.




```
Approximated apparent pKa value(s) are not calculated
```

**Figure A7.2:** Calculated pKa value

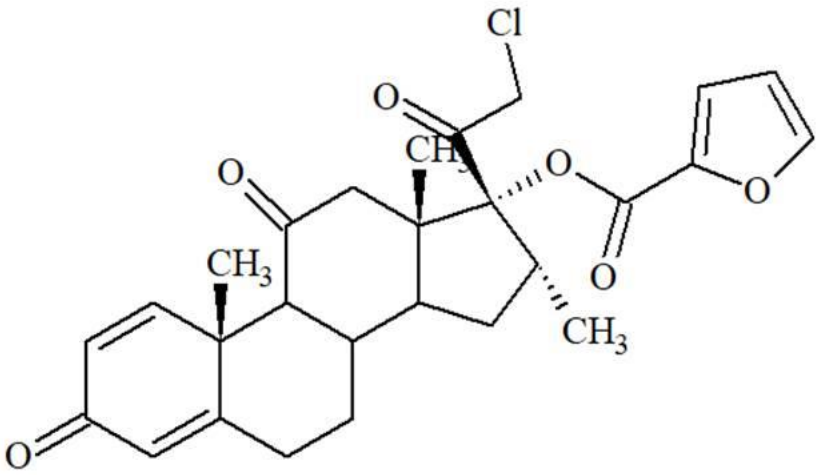



**Calculated Solubility In Pure Water**

Solubility (mol/l) = 7.2e-7 at pH = 7.00

**Log(S) for neutral form: -6.15**

**Figure A7.3:** Calculated solubility in pure water and Log S

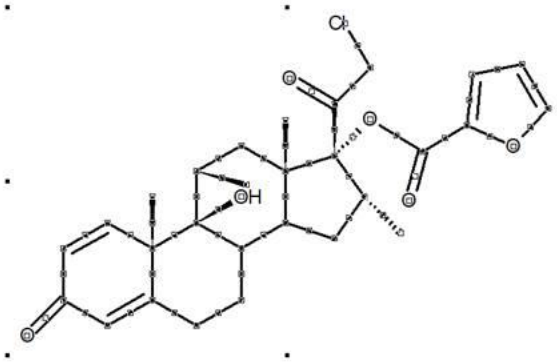



c = 100.0%; c(H2O) = 0.0%; c(org) = 100.0%

**LogP for neutral form: 4.39 ± 0.65**

**Figure A7.4:** Calculated Log P

## Appendix 8, Compound 8



| | |
|---|---|
| Molecular Formula: | $C_{28}H_{33}ClO_6$ |
| Formula Weight: | 501.01102 |
| Composition: | C(67.12%) H(6.64%) Cl(7.08%) O(19.16%) |
| Molar Refractivity: | 130.19 ± 0.4 $cm^3$ |
| Molar Volume: | 384.0 ± 5.0 $cm^3$ |
| Parachor: | 1042.8 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.593 ± 0.03 |
| Surface Tension: | 54.3 ± 5.0 dyne/cm |
| Density: | 1.30 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 51.61 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 12 |
| Monoisotopic Mass: | 500.196566 Da |
| Nominal Mass: | 500 Da |
| Average Mass: | 501.011 Da |
| M+: | 500.196018 Da |
| M-: | 500.197115 Da |
| [M+H]+: | 501.203843 Da |
| [M+H]-: | 501.20494 Da |
| [M-H]+: | 499.188193 Da |
| [M-H]-: | 499.18929 Da |

**Figure A8.1:** Calculations of physicochemical parameters.

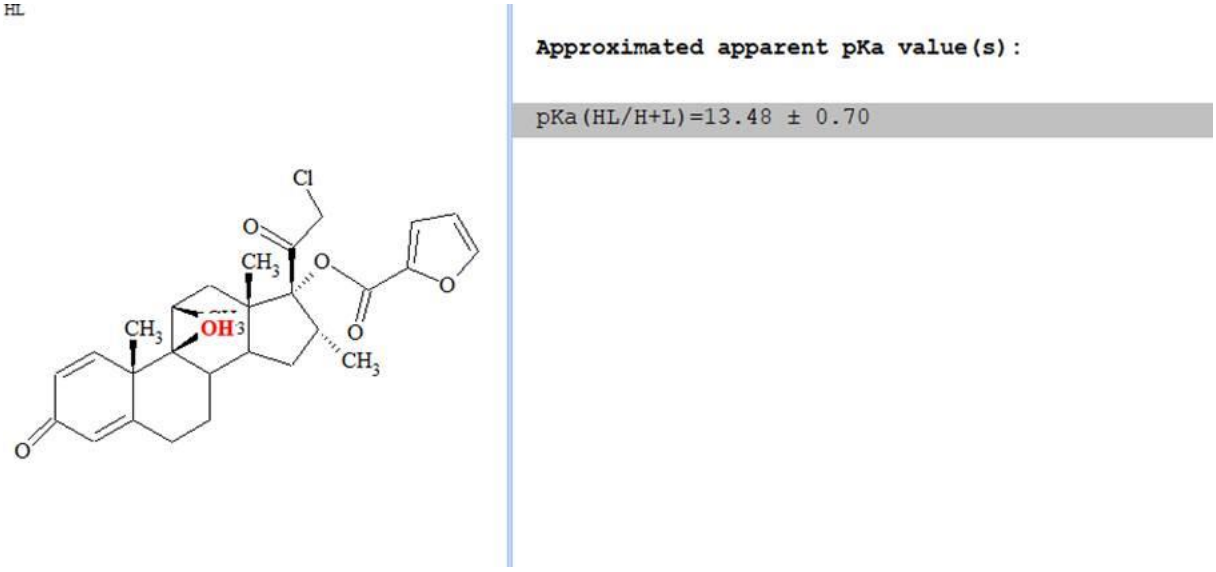
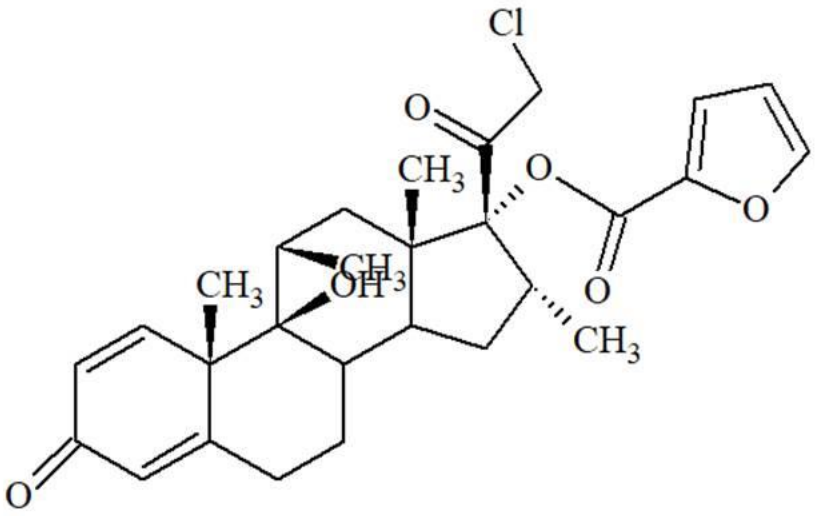



**Figure A8.2:** Calculated pKa value




**Calculated Solubility In Pure Water**

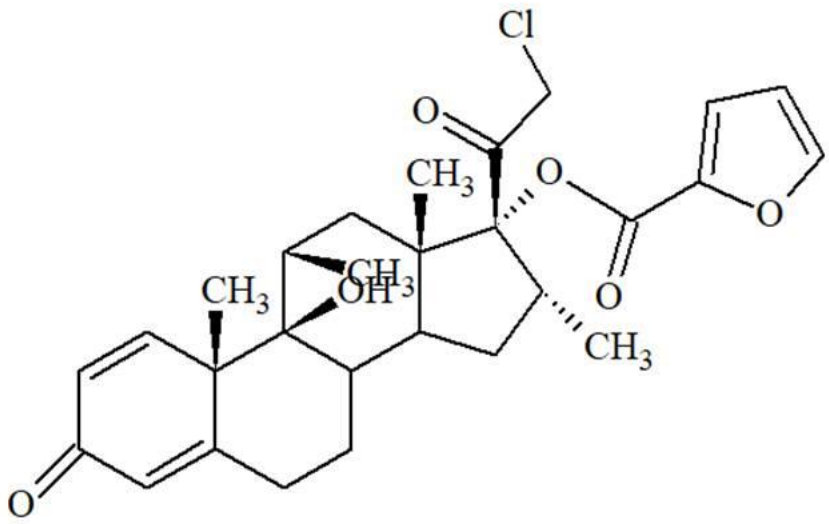
Solubility (mol/l) = 1.8e-6 at pH = 7.00

**Log(S) for neutral form: -5.74**

**Monoanions:**

Log(S) (Charge on O (or S) near aliphatic): -1.64 ± 1.00

**Figure A8.3:** Calculated solubility in pure water and Log S




c = 100.0%; c(H2O) = 0.0%; c(org) = 100.0%

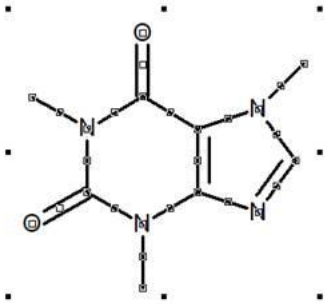
**LogP for neutral form: 4.60 ± 0.56**

**Monoanions:**

LogP (Charge on O (or S) near aliphatic): 0.50 ± 1.00

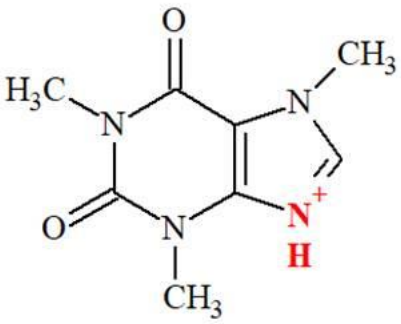
**Figure A8.4**: Calculated Log P

**Appendix 9, Coffeine**



| | |
|---|---|
| Molecular Formula: | $C_8H_{10}N_4O_2$ |
| Formula Weight: | 194.1906 |
| Composition: | C(49.48%) H(5.19%) N(28.85%) O(16.48%) |
| Molar Refractivity: | 50.38 ± 0.5 $cm^3$ |
| Molar Volume: | 133.3 ± 7.0 $cm^3$ |
| Parachor: | 364.5 ± 8.0 $cm^3$ |
| Index of Refraction: | 1.679 ± 0.05 |
| Surface Tension: | 55.7 ± 7.0 dyne/cm |
| Density: | 1.45 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 19.97 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 6 |
| Monoisotopic Mass: | 194.080376 Da |
| Nominal Mass: | 194 Da |
| Average Mass: | 194.1906 Da |
| M+: | 194.079827 Da |
| M-: | 194.080924 Da |
| [M+H]+: | 195.087652 Da |
| [M+H]-: | 195.088749 Da |
| [M-H]+: | 193.072002 Da |
| [M-H]-: | 193.073099 Da |

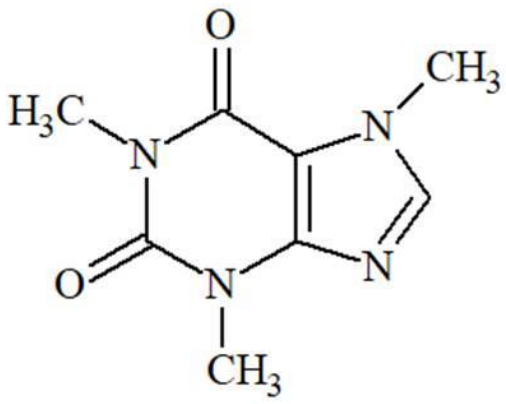
**Figure A9.1:** Calculations of physicochemical parameters.



**Approximated apparent pKa value(s):**

1. pKa(HL/H+L)=0.52 ± 0.70
2. pKa(H2L/H+HL)=-3.83 ± 0.20

**Figure A9.2:** Calculated pKa value




**Calculated Solubility In Pure Water**

Solubility (mol/l) = 0.27 at pH = 7.14

**Log(S) for neutral form: -0.56**

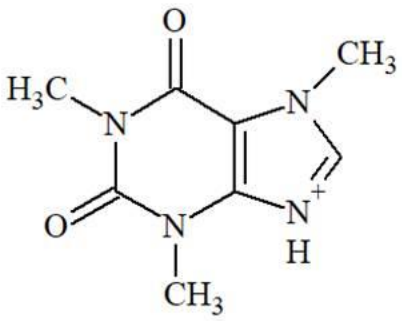
**Monocations:**

Log(S) (Charge in aromatic): 1.94 ± 1.00

**[1]** *Solubility, Mol/L:* 0.7922 ( 80 °C )
*Solubility, original:* 181.82 g/l water
Book *The Merck Index*, **1996.**
*Five point evaluation:*

**[2]** *Solubility, Mol/L:* 2.06 ( 100 °C )
*Solubility, original:* 666.67 g/l water
Book *The Merck Index*, **1996.**
*Five point evaluation:*

**Figure A9.3:** Calculated solubility in pure water and Log S




c = 65.8%; c(H2O) = 65.6%; c(org) = 0.2%

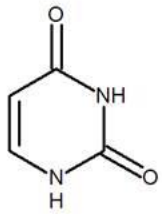
**LogP for neutral form: -0.13 ± 0.37**

**Monocations:**

LogP (Charge in aromatic): -2.63 ± 1.00

**Figure A9.4:** Calculated Log P

**Appendix 10, Uracil**



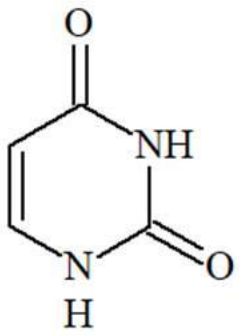
| | |
|---|---|
| Molecular Formula: | $C_4H_4N_2O_2$ |
| Formula Weight: | 112.08676 |
| Composition: | C(42.86%) H(3.60%) N(24.99%) O(28.55%) |
| Molar Refractivity: | 25.00 ± 0.3 $cm^3$ |
| Molar Volume: | 84.8 ± 3.0 $cm^3$ |
| Parachor: | 215.0 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.501 ± 0.02 |
| Surface Tension: | 41.3 ± 3.0 dyne/cm |
| Density: | 1.321 ± 0.06 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 9.91 ± 0.5 $10^{-24} cm^3$ |
| RDBE: | 4 |
| Monoisotopic Mass: | 112.027277 Da |
| Nominal Mass: | 112 Da |
| Average Mass: | 112.0868 Da |
| M+: | 112.026729 Da |
| M-: | 112.027826 Da |
| [M+H]+: | 113.034554 Da |
| [M+H]-: | 113.035651 Da |
| [M-H]+: | 111.018904 Da |
| [M-H]-: | 111.020001 Da |

**Figure A10.1:** Calculations of physicochemical parameters.

**Approximated apparent pKa value(s):**

1. pKa(HL/H+L)=15.60 ± 0.20
2. pKa(H2L/H+HL)=8.83 ± 0.10
3. pKa(H3L/H+H2L)=-4.19 ± 0.20

**Figure A10.2**: Calculated pKa value




**Calculated Solubility In Pure Water**

Solubility (mol/l) = 0.011 at pH = 5.13

**Log(S) for neutral form: -1.95**

**Monoanions:**

Log(S) (Charge on N): 0.05 ± 1.00

**Other charged forms:**

Log(S) (summing charge: -2): 1.05 ± 1.50

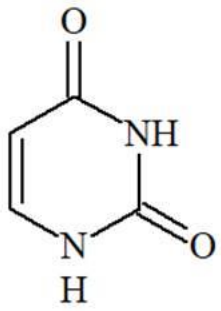
**[1]** *Solubility, Mol/L:* 0.0319 ( 25 °C )

*Solubility, original:* 3.58 g/l water

Book *The Merck Index*, **1996.**

*Five point evaluation:*

**Figure A10.3:** Calculated solubility in pure water and Log S




c = 100.0%; c(H2O) = 83.6%; c(org) = 16.4%

**LogP for neutral form: -0.71 ± 0.29**

**Monoanions:**

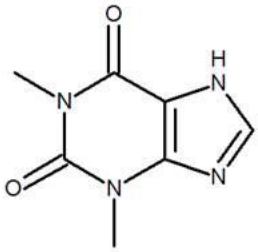
LogP (Charge on N): -2.71 ± 1.00

**Other charged forms:**

LogP (summing charge: -2): -3.71 ± 1.50

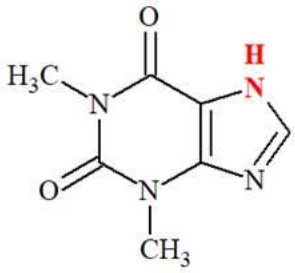
**Figure A10.4:** Calculated Log P

**Appendix 11, Theophylline**



| | |
|---|---|
| Molecular Formula: | $C_7H_8N_4O_2$ |
| Formula Weight: | 180.16402 |
| Composition: | C(46.67%) H(4.48%) N(31.10%) O(17.76%) |
| Molar Refractivity: | 43.14 ± 0.3 $cm^3$ |
| Molar Volume: | 122.9 ± 3.0 $cm^3$ |
| Parachor: | 352.4 ± 6.0 $cm^3$ |
| Index of Refraction: | 1.619 ± 0.02 |
| Surface Tension: | 67.6 ± 3.0 dyne/cm |
| Density: | 1.465 ± 0.06 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 17.10 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 6 |
| Monoisotopic Mass: | 180.064726 Da |
| Nominal Mass: | 180 Da |
| Average Mass: | 180.164 Da |
| M+: | 180.064177 Da |
| M-: | 180.065274 Da |
| [M+H]+: | 181.072002 Da |
| [M+H]-: | 181.073099 Da |
| [M-H]+: | 179.056352 Da |
| [M-H]-: | 179.057449 Da |

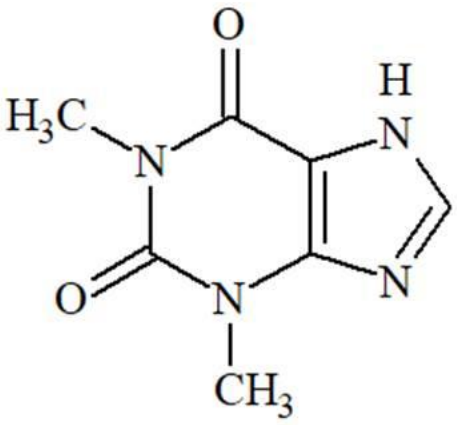
**Figure A11.1:** Calculations of physicochemical parameters.



**Approximated apparent pKa value(s):**

1. pKa(HL/H+L)=8.60 ± 0.50
2. pKa(H2L/H+HL)=1.84 ± 0.70
3. pKa(H3L/H+H2L)=-4.04 ± 0.20

**Figure A11.2:** Calculated pKa value

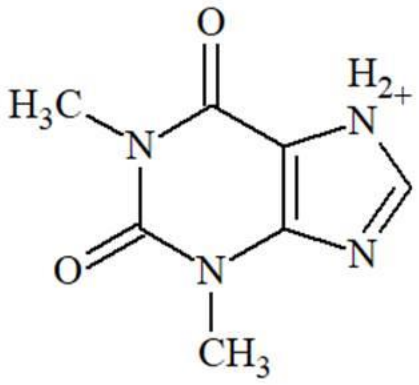



**Calculated Solubility In Pure Water**

Solubility (mol/l) = 0.052 at pH = 5.33

**Log(S) for neutral form: -1.28**

**Figure A11.3:** Calculated solubility in pure water and Log S




c = 71.1%; c(H2O) = 71.0%; c(org) = 0.2%

**LogP for neutral form: -0.17 ± 0.31**

**Monoanions:**

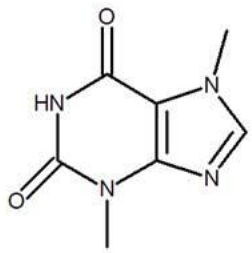
LogP (Charge on N): -2.17 ± 1.00

**Monocations:**

LogP (Charge in aromatic): -2.67 ± 1.00

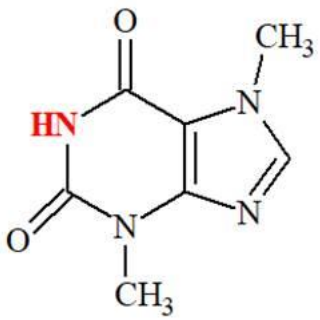
**Figure A11.4:** Calculated Log P

**Appendix 12, Theobromine**




| | |
|---|---|
| Molecular Formula: | $C_7H_8N_4O_2$ |
| Formula Weight: | 180.16402 |
| Composition: | C(46.67%) H(4.48%) N(31.10%) O(17.76%) |
| Molar Refractivity: | 45.05 ± 0.5 $cm^3$ |
| Molar Volume: | 112.0 ± 7.0 $cm^3$ |
| Parachor: | 319.0 ± 8.0 $cm^3$ |
| Index of Refraction: | 1.737 ± 0.05 |
| Surface Tension: | 65.7 ± 7.0 dyne/cm |
| Density: | 1.60 ± 0.1 $g/cm^3$ |
| Dielectric Constant: | Not available |
| Polarizability: | 17.86 ± 0.5 $10^{-24}cm^3$ |
| RDBE: | 6 |
| Monoisotopic Mass: | 180.064726 Da |
| Nominal Mass: | 180 Da |
| Average Mass: | 180.164 Da |
| M+: | 180.064177 Da |
| M-: | 180.065274 Da |
| [M+H]+: | 181.072002 Da |
| [M+H]-: | 181.073099 Da |
| [M-H]+: | 179.056352 Da |
| [M-H]-: | 179.057449 Da |

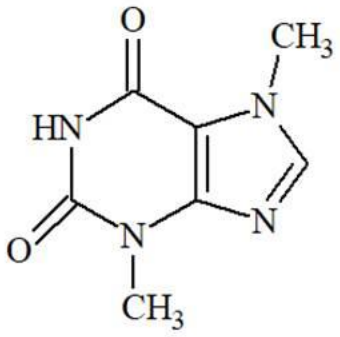
**Figure A12.1:** Calculations of physicochemical parameters.




**Approximated apparent pKa value(s):**

1. pKa(HL/H+L)=9.90 ± 0.50
2. pKa(H2L/H+HL)=0.41 ± 0.70

**Figure A12.2:** Calculated pKa value

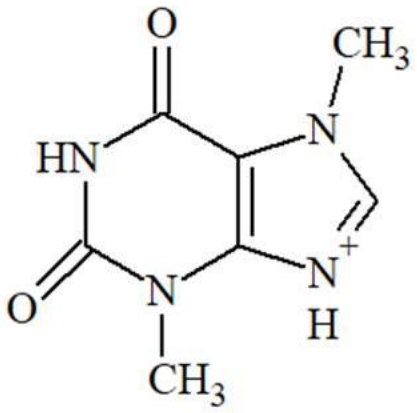



**Calculated Solubility In Pure Water**

Solubility (mol/l) = 0.04 at pH = 5.67

**Log(S) for neutral form: -1.39**

**Figure A12.3:** Calculated solubility in pure water and Log S




c = 68.5%; c(H2O) = 68.5%; c(org) = 0.0%

**LogP for neutral form: -0.72 ± 0.55**

**Monoanions:**

LogP (Charge on N): -2.72 ± 1.00

**Monocations:**

LogP (Charge in aromatic): -3.22 ± 1.00

**Figure A12.4:** Calculated Log P

**Appendix 13, ACD Labs Log D**


**Figure A13.1:** ACD Labs, calculate Log D.

**Appendix 14, ACD Labs pKa**

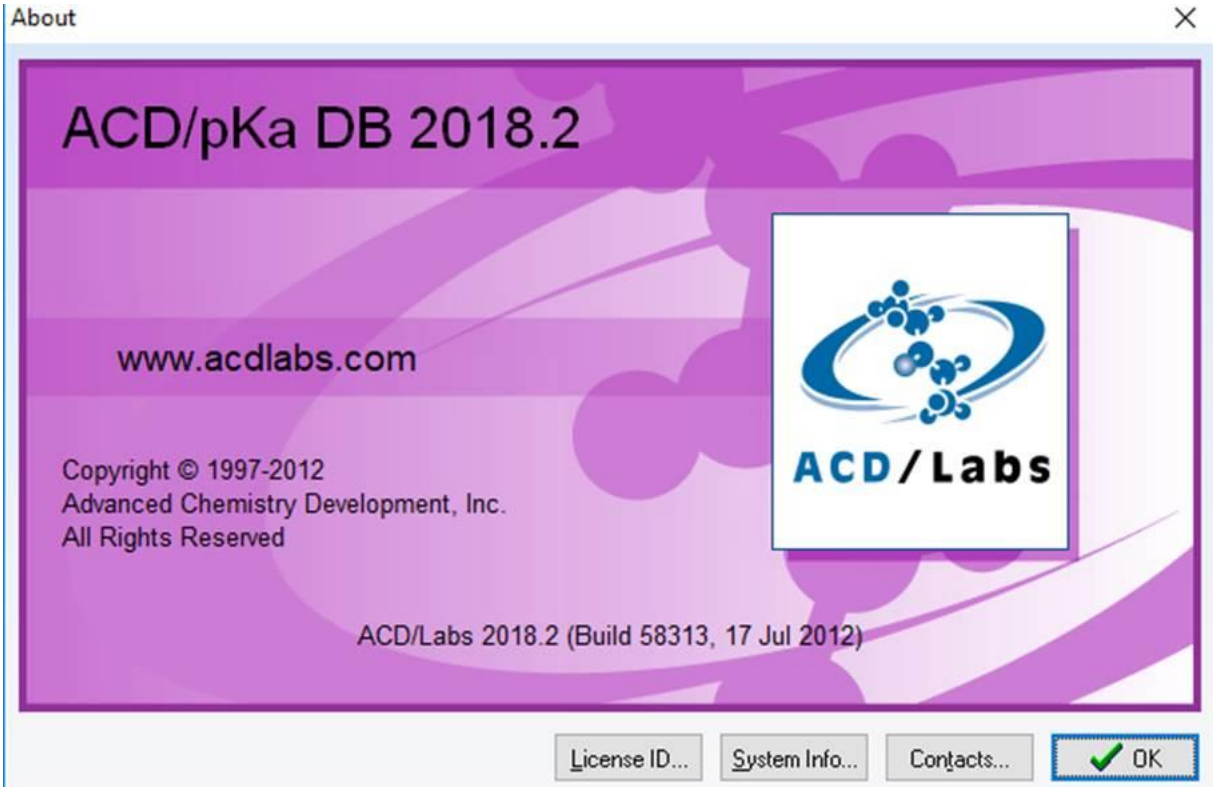



**Figure A14.1:** ACD Labs, calculate pKa

**Appendix 15, ACD Labs Solubility**


**Figure A15.1:** ACD Labs, calculate the solibility

**Appendix 16, ACD Labs ChemSketch**


**Figure A16.1:** ACD Labs, ChemSketch

**Appendix 17, Retention time and peak width SFC and RP-LC**

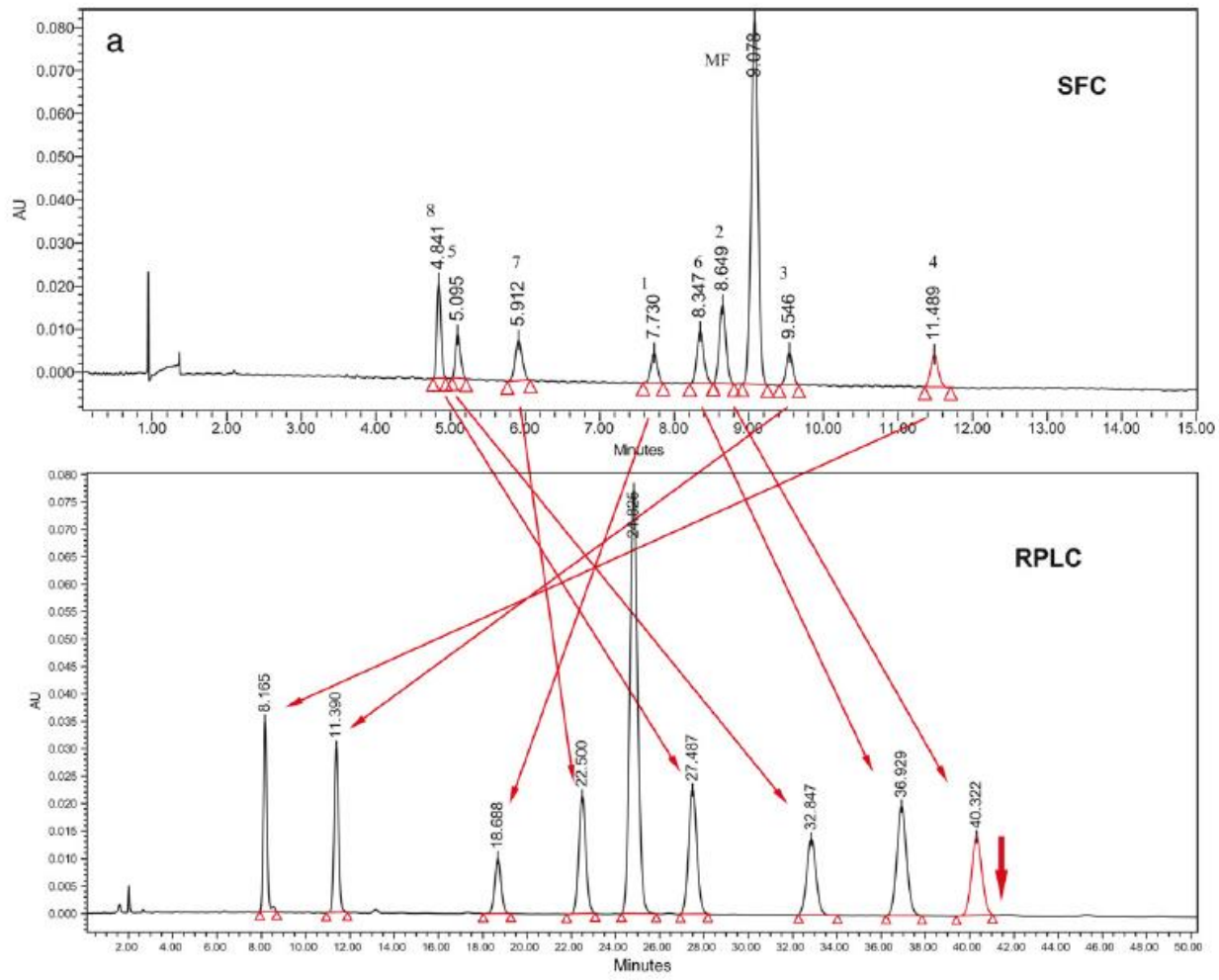



**Figure A17.1:** Retention time and peak width, SFC and RPLC

**Appendix 18, Retention time and peak width SFC**

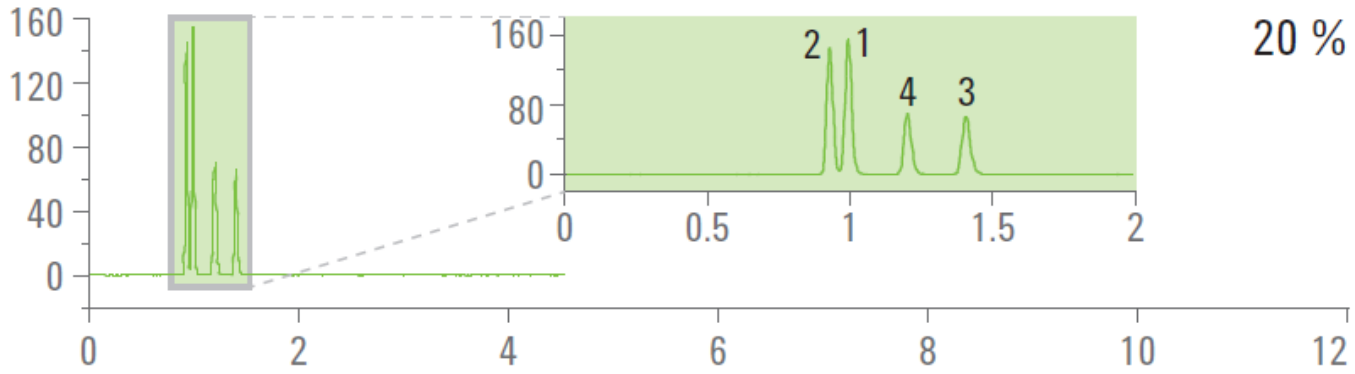



**Figure A18.1:** 1= Caffeine, 2= Theophylline, 3= Theobromine, 4= Uracil