UPPSALA
UNIVERSITET

UPTEC X 19036
Examensarbete 30 hp
Augusti 2019

Chromatography of pharmaceutical peptides - contrasting SFC and HPLC

Joakim Bagge

UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

Chromatography of pharmaceutical peptides - contrasting SFC and HPLC

Joakim Bagge

This work is a comparison of a well-established and a novel, "green" and efficient technique to separate peptides of pharmaceutical interest. An attempt is made to derive the chromatographic retention behaviour in these techniques to a number of property descriptors derived from the linear sequence of amino acids. A set of therapeutic peptides were carefully chosen to be experimentally evaluated using in silico-based descriptor calculations. A principle component analysis was performed to assess the distribution of calculated descriptors for including peptides with variable properties. A diluent optimization study was also included to find the optimal diluent for peptides with minimal diluent effects and peak splitting phenomena. The results showed that the solvents tert-butanol and methanol performed best between 20-30 and 50 volumetric percent water as additive in SFC and HPLC, respectively. These diluents were then used for the peptides within the set to evaluate the retention and selectivity in HPLC and SFC. SFC performed well in terms of resolving power. In particular, SFC was able to separate Leuprolide and Triptorelin while HPLC was not. A comparison was also made in between the two stationary phases CN and XT, where a global selectivity was shown to be higher for CN.

This work does also assess a novel method for determining solubility of analytes in supercritical fluid. The method was evaluated using the pharmaceutical compounds caffeine and aspirin and then used to determine solubility of Leu-Enkephalin in 20% (v/v%) methanol. The solubility of caffeine was determined to be 0.45 mg ml-1 in pure SF-CO2 under 140 bar pressure and 3.9 mg ml-1 for aspirin in 2.4% methanol. Both values correlated well with measurements from four acknowledged papers within this field. Leu-Enkephalin was found to have a solubility of 1.90 mg ml-1 using a solvent corresponding to the initial phase condition of the gradient used for peptide analysis in SFC. Further experimental work is required before the method can be implemented as a useful tool in preparative chromatography, however the results presented here show the compatibility of assessing biomolecules in both pure SF-CO2 and mixed with modifier. The possibility to determine solubility with additional modifier infers an important step of including and evaluating these compounds creating a solid support to subsequent large scale separation.

Handledare: Linda Thunberg, Martin Enmark
Ämnesgranskare: Ulf Göransson
Examinator: Jan Andersson
ISSN: 1401-2138, UPTEC X19 036

Peptidläkemedelsanalys med miljövänligare teknik

I en värld med en ökande befolkning finns behov av ett tillräckligt utbud av medicin och hållbar läkemedelsutveckling för att kunna möta dagens och morgondagens sjukdomar. För att kunna framställa nya effektiva och säkra läkemedel sätts krav på innovativa, miljömässigt hållbara och tidseffektiva metoder. Idag utgör kromatografi som teknik en grundpelare för separation och rening av aktiva läkemedelssubstanser innan produkt når patient. Läkemedel delas generellt sätt upp i småmolekylära läkemedel vilka utgör majoriteten av alla läkemedel samt biologiska som intar en alltmer ökande andel av marknaden. Bland biologiska läkemedel har forskning länge bedrivits kring terapeutiska peptider vilka har visat på bindningsspecificitet och låg toxicitet som båda är nödvändiga egenskaper.

Superkritisk vätskekromatografi, förkortat SFC, fick sitt genombrott som teknik under åttiotalet men har under de senaste tio åren genomgått en rad förbättringar. Dessa har genererat instrumentation med ökad robusthet samt medfört kompabilitet med populära analystekniker som masspektrometri. Prefixet superkritisk syftar till användningen av en vätska som opererar under höga tryck och hög temperatur. Användningen av en så-kallad superkritisk mobil fas har visat sig ha kromatografiskt fördelaktiga egenskaper såsom låg viskositet och hög diffusivitet. Egenskaperna medför en 3-10 gånger effektivare separation vid jämförelse med konventionell vätskekromatografi. SFC tillämpar också koldioxid (CO_2) som mobilfas vilket gör tekniken till en miljövänlig teknik. Separation av terapeutiska peptider utförs i första hand via vätskekromatografi där mångårig etablerad forskning har optimerat processen till att kunna separera peptider på aminosyra-nivå. För SFC har separation av peptider ännu inte slagit igenom främst på grund av att de är svårlösliga i superkritisk CO_2. Dessutom är det svårt att hitta ett bra lösningsmedel för peptider som vid injektion genererar smala kromatografiska toppar, vilket är önskvärt. Ett sätt att öka lösligheten har varit att blanda CO_2 med ett organiskt lösningsmedel, som exempelvis metanol eller etanol. För vätskekromatografi löses oftast föreningar i samma lösningsmedel som mobilfasen utgör, vilket har visat minimera breddning av toppar. Att tillämpa detta på SFC är inte möjligt då mobilfasen är i ett slutet system med ett tryck över atmosfärstryck och med hög temperatur. Att lösa en förening under dessa förhållanden är alltså en stor utmaning.

Detta arbete har utförts hos Early Chemical Development på AstraZeneca i Mölndal och avdelningen för medicinalkemi på Uppsala Universitet, med inriktning mot att kontrastera den konventionella vätskekromatografin mot SFC för separation av peptider. Detta har gjorts för att utvärdera hur teknikerna presterar samt vilka potentiella fördelar dessa två tekniker har. I arbetet ingick dessutom en utvärdering av en ny föreslagen metod för att mäta löslighet i en superkritisk vätska.

Genom att förstå processerna i en kromatografisk separation kan metoden och de omgivande parametrarna optimeras för att på så sätt generera en säker och effektiv separation av aktiv

farmaceutisk substans (API) från biprodukter. Med detta följer ekonomiska fördelar för såväl läkemedelsbranschen som för patientgrupper. Detta kan ur globalt perspektiv gynna en ökad, laglig försäljning av medicinska preparat till lägre priser vilket i sin tur gör läkemedel tillgängligt för fler patienter, samtidigt som trycket för handel av illegalt framställda läkemedel kan minska.

Table of Contents

Abbreviations

α	Selectivity factor
ACN	Acetonitrile
AgII	Angiotensin II
AMP	Antimicrobial peptides
APCI	Atmospheric Pressure Chemical Ionization
BPR	Back Pressure Regulator
CO_2	Carbon dioxide
ESI	Electrospray Ionization
FDA	Food and Drug Administration (USA)
HDMS	High Definition Mass Spectrometer
HPLC	High Performance Liquid Chromatography
k	Retention factor
LC-MS	Liquid Chromatography - Mass Spectrometry
Leu-Enk	Leu-Enkephalin
MeOH	Methanol
modlAMP	molecular design laboratory's Antimicrobial Peptides
MS	Mass Spectrometry
PCA	Principal Component Analysis
PDA	Photodiode-array
QTOF	Quadrupole Time of Flight
RP-HPLC	Reversed-Phase High Performance Liquid Chromatography
t_R	Retention time
SFC	Supercritical Fluid Chromatography
SF-CO_2	Supercritical Fluid carbon dioxide
SiO_2	Silica-dioxide
t_0	hold-up time
t-BuOH	tert-butanol
TFA	Trifluoroacetic acid
UV-vis	Ultraviolet-visible spectrum

1 Project goal

This study examines the use and possible benefits of using SFC for separating pharmaceutical peptides instead of conventional HPLC. The use of SFC has in recent years expanded in the pharmaceutical industry where shorter purification time and the use of less expensive solvents are main attractions. Experiments were conducted on both SFC and HPLC where eight peptides that were injected in analytical and overloaded amounts and with methods that had been optimized for each technique. A number of diluents were investigated to assess peak shape improvements where a selection of common organic modifiers were used mixed with varying volume fraction of water. A novel approach for solubility determination in supercritical fluid was also assessed. The new method was evaluated using the two small molecule caffeine and aspirin before the method was applied to the peptides Leu-Enkephalin and Angiotensin II.

2 Background

2.1 Pharmaceutical peptides

The interest of using biological compounds as pharmaceuticals has increased the past decade. Among these "biologics" that comprise of nucleic acids, monoclonal antibodies are peptides considered to be a promising alternative as therapeutics. Peptides have natively important functions in the human body where they act as hormones, growth factors, neuro-transmitters and anti-infectives among others (Fosgerau & Hoffmann, 2015). These peptides are compelling to use as pharmaceuticals due to their high binding specificity and efficiency derived from their native structure but also due to quaternary complex formations when binding in to target ligand. Toxicity is generally low for eukaryotes due to their specific targeting and subsequent degradation into simple amino acids (Badiani, 2012). One example of a commercially available and therapeutically used peptide is Angiotensin II (AgII) which works as a vasoconstrictor, i.e. it induces contraction of blood vessels by binding to G-protein-coupled angiotensin II receptor type 1 and 2. The effect is increased blood pressure and the release of the sodium-regulating hormone aldosterone. The drug is aimed towards patients suffering from vasodilatory shock, a condition where blood pressure drops to critical levels causing failure of oxygen transfer to organs (Wong, 2016, Baker & Levien, 2018).

The production of these therapeutic peptides was historically performed through extraction and isolation of the sought peptide from animal tissue. This was initially how isolation of the first therapeutic peptide, insulin, was done. Over time, techniques have evolved which today permits the synthesis of peptides through coupling of amino acids. This has enabled the *de novo* design of several analogues of a lead peptide compound that are not naturally occurring. This is useful in peptide drug development where the lead peptide has interesting pharmaceutical activity but is easily degraded when entering living tissue (Lau & Dunn, 2018).

The progress and successfulness to commercialize biological drugs has been presented in recent literature, showing that a record of 17 new biological compounds comprising of enzymes and peptides were approved by FDA in 2018. Approximately 25% of all approved pharmaceuticals have between the years 2014-2018 corresponded to biological drugs (de la Torre & Albericio, 2019). As of 2017, 68 peptides have been approved in Europe, USA and Japan and 155 were currently in active clinical development (Lau & Dunn, 2018).

Scientific research is today conducted, apart from searching for new therapeutic peptides, to solve the antibiotic resistance through antimicrobial active peptides. Antimicrobial peptides (AMPs) are peptides with activity against bacteria and fungi by showing specific binding to general antibiotic targets such as microbial membrane and DNA. (Nguyen et al, 2011). Many conventional antibiotics to date target peptides in microbes however this strategy could potentially develop resistance too easily through mutation. The use of AMPs would presumably reduce the risk of evolving resistance (Peschel & Sahl, 2006).

The need for effective and simple tools for characterizing peptides and compounds has led to the development of a vast number of pieces of computer prediction software. For scientists, the need for application of different programming languages to develop open source code has led to a large offer of scripts that are dispersed and difficult to combine without proper programming skills (Müller et al., 2017). A python software package was in 2017 released by Müller et al. where descriptor calculation, activity prediction and analysis tools have been integrated into one package, containing nine modules, to facilitate characterization of peptides with particular focus on AMPs. Despite its main purpose, the package can be used to characterize any type of linear peptide regardless of what functionality it possesses. Through basic python programming skills, the user can easily import any of the modules and via any peptide linear sequence obtain predictive data about its characteristics.

2.2 Chromatography as separation technique

2.2.1 High Performance Liquid Chromatography

The elution of a compound will occur when the molecular interactions between solute and solvent are stronger than the interactions with the stationary phase. Since interactions can result from e.g. hydrophilic, hydrophobic and ion interactions, the elution strength is much dependent on the physico-chemical properties of the solute. Therefore, chromatography has been divided into several different techniques which is built on the specific property you wish to separate accordingly. Both normal-phase (NP-), reversed-phase (RP-) and ion-exchange (IEX-) liquid chromatography are applied techniques for peptide separation which also can be used orthogonally in series to enhance resolution power. For normal phase LC, the separation is carried out as polar molecules are retained by a polar stationary phase, and where a nonpolar mobile phase like hexane or heptane is used to decrease retention. For RPLC, the separation is according to hydrophobicity where the stationary phase is based on silica particles coated with alkyl carbon chains with lengths differing from 4 to 18 carbons. The mobile phase is usually organic such as ACN and MeOH and is mixed with a polar solvent e.g. water to adjust retention.

Ion-exchange chromatography is used to separate according to charge through ionic interactions. In the case of peptides, protonation of side groups may result in different total charge of the biomolecule, thus it is of importance to apply a proper pH in the mobile phase.

The line-up of mobile phase usually involves two phases, A and B that have orthogonal properties. For separations in RPLC a solvent A usually consists of water and solvent B an organic modifier that is miscible in A. A highly hydrophobic compound will have a strong affinity to the stationary phase and thus be retained for a longer period than a hydrophilic compound. The change in mobile phase composition through mixing of two solvents can influence the retention behavior of a compound. In this case, a larger volumetric ratio of ACN or MeOH that possess a high elution strength for hydrophobic compounds weaken the affinity to the column and therefore elute compounds faster. In the case of a separation problem where a sample contains many different compounds of similar size, and physico-chemical properties, the composition of the mobile phase has a more important role to optimize the separation at hand. This can be carried out through either isocratic or gradient elution mode, where the composition is either constant throughout the process or changed with time, respectively. Isocratic elution mode is very straight forward but is usually not sufficient for separating samples with several species that have similar intrinsic properties. A gradient elution can in practice take on any form throughout a chromatographic process depending on how the species are retained. The gradient can be designed to either linearly or non-linearly change the composition of solvent A and B through commissioning a constant or variable slope of change to the volumetric percentage of flow.

2.2.2 Supercritical Fluid Chromatography

From the roots of liquid chromatography, Supercritical Fluid Chromatography (SFC) was rapidly developed in the beginning of the 1980s. SFC is an alternative separation technique to conventional LC where compressed and heated carbon dioxide (CO_2) is used as main eluent. The state requires that both the temperature and pressure is high enough to make the transition from gas or liquid, to which the CO_2 enters the supercritical region (Figure 1). The chromatographic instrument is technically very much similar to a LC instrument with the exception of following components: a CO_2 inlet, a back-pressure regulator (BPR) to control and maintain wanted pressure in system and a separator. SF-CO_2 is nonpolar and has a very low elution strength, which is why an additional organic modifier usually is mixed. The organic modifier usually corresponds to one of many of the commonly used eluents such as methanol (MeOH) or acetonitrile (ACN). Depending on present conditions for the analysis and mobile phase composition, as binary or ternary mixture can result in a drop from supercritical to subcritical state thus affecting the chemical equilibrium with the analyte. After the liquid has transported analyte through the column, the flow is usually split between detector and waste. At waste, CO_2 is usually separated from other solvents. The separation is made possible through cyclonic separation where flow is directed through a cylinder with airstreams directed in a helical motion along the inside. Efficient cyclones are available today where the amount of recycled CO_2 usually exceeds 90%. At the department of R&D at AstraZeneca, a SuperSep 600

from NovaSep (Pompey, France) was reported to be used where ~90% of the CO_2 could be recycled, using 10% organic modifier in mobile phase (Lindskog et al., 2014). It is known that a larger amount of organic modifier in the mobile phase results in reduction of CO_2 acquisition.

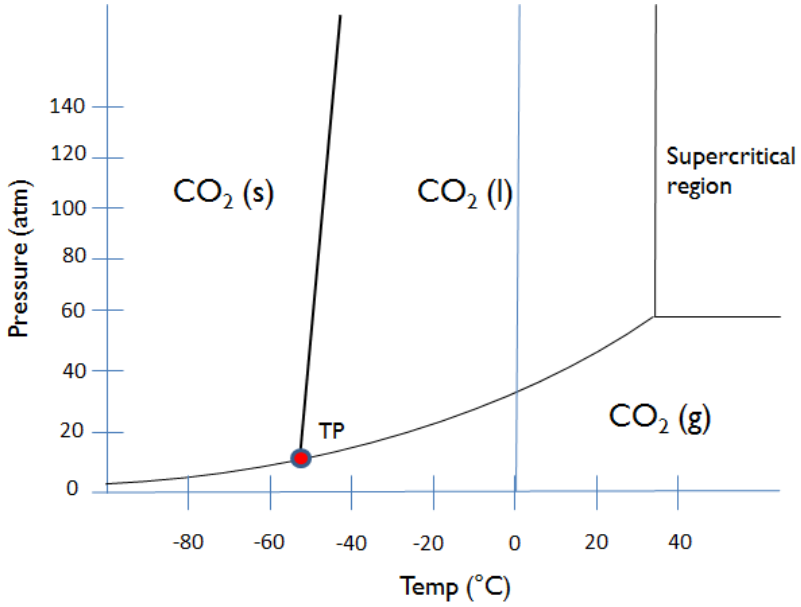

Figure 1. A phase diagram for carbon dioxide (CO_2). The triple point (TP) where the gas is in the three states solid, liquid and gas is denoted and the transitions between phases is depicted as lines. A supercritical liquid requires sufficiently high pressure and temperature to reach the supercritical region, depicted in the right-uppermost part of the diagram.

SFC has shown great use for separation of chiral compounds in both analytical and preparative scale, which is why the technique has become increasingly popular in the pharmaceutical industry. Approximately 40% of all pharmaceuticals are chiral, where a quarter of these are pure enantiomers (Webster, 2011). In terms of peptides, there are several reports suggesting that SFC can be used for rapid peptide separation however the applicability for preparative scale is yet to be assessed (Zheng et al. 2006, Tognarelli et al. 2010, Shao et al. 2016).

2.2.3 Detection with on-line hyphenated techniques

A chromatographic separation is usually followed by detection of the eluting compounds which normally is made either by absorbance in the UV-vis spectra or by hyphenated techniques such as mass spectrometry. Mass spectrometry, or MS, is a useful technique where incoming analytes are ionized by an ionization source in a vacuum chamber. The ionization induces charge to the analyte, and after passing a field with induced electric potential it will be separated according to the mass by charge ratio, m/z. There are numerous strategies that have been applied for ionizing compounds, such as electrospray ionization (ESI), atmospheric pressure photo ionization (APPI) and atmospheric pressure photo ionization (APCI). The mostly used ionization sources for SFC is ESI and APCI. APCI has in particular been preferred in combination with SFC (Chen, 2009). The ionization process in this technique involves the generation of H_3O^+ ions through discharge on the incoming fluid, induced by a voltage source. The H_3O^+ molecules protonate the incoming sample, which has in a preceding step been evaporated to an aerosol. The ESI technique is considered as a soft technique since the process

keeps analytes intact without fragmenting the incoming sample. Instead, the ionization is driven by applying potential over a liquid, creating an elliptic shaped droplet with a charged surface. The solvent within the droplet is while in motion evaporated, leading to an increase of charge density at the surface until the surface envelopes the molecule enabling separation by mass. The ionization process is well understood and since it conserves analytes present, it has become widely adopted for analyses of peptides and proteins, (Wilm et al. 2011, Zhang et al. 2010).

The detection of masses occurs in the subsequent mass analyzer which separates the ionized molecules according to their charge and mass. Two types of mass analyzer that are often found are quadrupole and time of flight (TOF) mass analyzers. A quadrupole mass analyzer consists of four metal rods that induces two sinusoidal electric fields with a 90 degrees phase shift. The shift results in a electric field that oscillates in the shape of a circle which causes charged compounds to move in a circular motion. The diameter of the motion correspond to the mass by charge ratio of the compound, in relationship to the strength of the electric field given in voltage per meter. In contrast, a TOF mass analyzer separates ions given their kinetic energies without the presence of an electric or magnetic field. The technique is based on the fundamental concept that kinetic energy equals the mass of the molecule in motion times its velocity squared, divided by two. This quantity is the same as the charged state of the ion multiplied by the voltage from the ion source. With knowledge that the time of flight is the length of the flight tube in the analyzer divided by the velocity of the particle, the time of flight can be expressed as being proportional to the square-root of its mass-to-charge-ratio. There is a limitation for larger molecules, as the increase in mass tends to result in longer time of flight. This issue has been solved through the incorporation of a reflector that is composed of several ring electrodes. These induces a strong voltage which decelerates the particle and causes it to reverse its motion. Fast and slow molecules that has the same mass-to-charge-ratio will move with different lengths into the flight tube, but as the reflector causes them to reverse they will reach the detector simultaneously.

Chen (2009) reported that coupling of on-line MS is preferable when using SFC in contrast to HPLC. A large innate pressure drop occurs within the system where the pressure set by the BPR is reduced to atmospheric pressure that is required for ESI and APCI ionization. This drop assists the nebulization of mobile phase to form aerosols which has led to speculations that MS is more suitable to use together with SFC. It is however known that decompression of CO_2 is an endotherm reaction, it is therefore necessary to increase temperature as heat is required as energy source to proceed the nebulization process.

2.2.4 Contrasting HPLC and SFC – as Technique and its Applications

The supercritical phase of carbon dioxide (SF-CO_2) has shown interesting physico-chemical properties such as low viscosity and high diffusivity which is considered optimal for chromatographic procedures. These properties implicate low back pressure on the chromatographic system, therefore enabling higher flow rates that would not be possible when using normal liquids. It has been shown that separation efficiency can be increased 3-10-fold with SFC in comparison to conventional LC (Webster, 2011). CO_2 is generally due to its low

toxicity used in food processing industry, which make it well-suited for pharmaceutical purposes (Cifti, 2012). It is also when compared to commonly used solvents a much cheaper alternative, where, to date, 20 litres of gaseous CO_2 costs 546 SEK while 2.5 litres of HPLC-graded MeOH costs 653 SEK[1].

The use of cyclones for recycling CO_2 implicates that much of the mobile phase, depending on the volumetric composition of the phase, can be re-used resulting in lower consumption of chemicals which implicates economical savings. Another beneficial aspect of using CO_2 is that in many processes, the gas is a wanted end product from combustion processes and not the least from refineries. The recycling of CO_2 can be considered to have an environmental positive effect due to the profit industry might do collecting it instead of releasing it out into the atmosphere.

The extended use of SFC for preparative scale separations at AstraZeneca R&D department was reported in 2014 by Lindskog et al. The department used at the time SFC for 50% of all achiral separations in milligram to gram scale and 85% of achiral samples in gram to kilogram scale. The main reason for increased use of SFC was based on the improved robustness of SFC instruments and the incorporation of MS. NPLC, which in this case was the alternative technique, is usually not used together with MS as ionization of alkane-based solvents can produce hazardous rest-products. The report concludes that operating costs were lower due to decreased solvent consumption and that the cost of column stationary phase was reduced since much smaller dimensions of columns give the same throughput as larger columns. Smaller particle sizes are tolerated in SFC and can be used with higher flowrate due to the lower viscosity of supercritical fluids. A result of this is shorter cycle time and increased purification productivity for same column dimensions.

There are very few scientific papers published to date that deals with contrasting HPLC and SFC. In an article by Vera et al. (2015), the selectivity was investigated and compared between the two techniques using linear polynuclear aromatic hydrocarbons using a number of different phenyl-type stationary phases. For separation of peptides, LC and in particular RPLC has always been considered an indispensable method (Gedela & Medicherla 2007, Isidro-Llobet et al. 2019). The many advantages SFC have as previously described has led to a swelling interest of adopting this technique for separating peptides in a rapid manner. In 2006, a separation of oligopeptides up to 40 amino acids and with varying physico-chemical properties was made using SFC. The separation was made using a mobile phase of MeOH and SF-CO_2 with TFA as additive. A separation was possible when using an ethylpyridine stationary phase, suggesting that SFC holds a promising future for separation of peptides (Zheng et al., 2006). In 2010, a separation of peptides in range of 238.2 to 1046.2 Da was performed on SFC within 12 minutes. Among the peptides analyzed were AgII and Leu-Enkephalin (Leu-Enk) where AgII was the largest in the set. The study showed nearly a five-fold improvement in efficiency where a 50 minutes long method was required to fully separate peptides (Tognarelli et al, 2010). A

[1] Prices are obtained from AGA and Merck/Sigma Aldrich, respectively.

separation of cyclic peptides with SFC has also been reported where five cyclosporin analogues comprising 11 amino acids each were evaluated. This was in particular interesting since separation with HPLC has been reported difficult and thus required multiple separation procedures. The study included screening of three columns with different column chemistries, where bare silica gave best results. An evaluation of mobile phase modifier was also done which concluded that MeOH was the optimal modifier. The elution order of the cyclosporin analogs were correlated to their calculated logP values apart from the analogue CsH which was eluted first. The report concluded that the structural dissimilarity of CsH compared to the other analogues could explain the different retention behavior of this particular cyclic peptide. When the experiment was repeated for HPLC, the CsH was retained longer than CsA, a homologue with equal logP value (Shao et al., 2016).

Enmark et al. (2018) reported a fundamental study where robustness of SFC was investigated for separation of peptides. In this report, a mixture of gramicidin analogues A,B,C and D were separated using both isocratic and gradient mode. A mobile phase was used containing MeOH as organic modifier with volumetric fraction of water in the range from 0 to 8.7%. In order to take column pressure drop into consideration, the density of the mobile phase was experimentally and accurately determined using a Corolis mass flow meter. The most important factors for retention were the volumetric fraction of modifier in mobile phase and the fraction of water with modifier. Here, a fraction of 8.7% water showed optimal separation of analogues compared with lower fractions or no water as additive. The robustness was found highest for gradient mode, where the increase of gradient linear slope from 1% to 13% min^{-1} showed improved robustness with decreasing relative error in retention.

Even if the quantity of articles relating to peptide separation on SFC is mentionable, there was however at time of this study no literature found available presenting the separation of protein digests. This leaves a gap of knowledge as SFC, mainly due to its advantages in efficiency, could be of interest for large scale peptide separation which would have applications for fields such as in proteomics.

2.3 Need for solubility and the purpose of diluents

Regardless of the technique at hand is LC or SFC, both require a liquidous mobile phase that will elute analytes. The eluent is represented by a - in ideal cases - miscible solution that is composed of one or several solvents. Each solvent can be characterized by polarity, density, dielectric constant among others. In the case of chromatography, it is also interesting to characterize a solvent due to its elution strength, which represents how strongly it elutes a compound. This is determined by the resulting interactions between the adsorbent, solvent and sample. The property can be regarded as an alternative to adsorption strength which represents how strongly a compound has adsorbed to a stationary phase. Both these extrinsic properties are given in an eluotropic series and are dependent on each other. The elution strength of a solvent can therefore highly vary depending on which stationary phase that is used. Equation 1 can be used to describe both the adsorption strength between a solute molecule to a phase or

the force of interaction with a solvent. When the forces between solute and phase are stronger than the forces between the individual components, the solute will either be dissolved or retained in a chromatographic system.

$$K_{AB} > K_{AA} \qquad K_{AB} > K_{BB} \tag{1}$$

Since both LC and SFC require a liquidous mobile phase, it is most convenient to inject the sample dissolved in a liquid as well. This injection solvent is commonly referred to as the diluent. The diluent can but is not required to have the same chemical composition as the mobile phase. In contrary, it can adopt any suitable composition for which the compound is soluble in. In the case of SFC, supercritical state can as previously stated only be reached when the pressure has superseded the critical point, which in turn must meet the requirement that the system is closed. To avoid this problem the sample is generally injected in the same diluent as it would in conventional LC.

There are a number of characteristics that are used to describe an ideal dilute:

- It should dissolve the analyte
- It should dissolve eventual impurities present
- It should not degrade analyte
- It should not affect detection of analyte
- It should contribute to improving peak shape

The first and second definition can appear as obvious as precipitation of analyte will prevent it from being injected and thus analysed. The impurities are of course required to be soluble if separation from the main product is to take place. In the case of peptides, an impurity usually corresponds to a similar peptide sequence derived from the synthesis with one or several aberrating amino acids. These differences could result in other hydrophobic and hydrophilic ratios which implies that impurities are not necessarily soluble in one diluent that dissolves the main product. The third definition relates to the potential overlap of absorbance spectra. Many organic solvents such as acetone and DMSO have good dissolving power, however the overlapping detection region makes it difficult to identify the fraction analyte of the detected peak. The final criteria refers to the diluent's effect on peak shape. To further explain the process when a diluent is injected into a mobile phase, the diluent can be regarded as a plug of solvent that is distinctly separated from the surrounding. When the plug reaches the column inlet, the diluent molecules will be dispersed and start to dissolve with the mobile phase along the surface of the plug. As it continues its path towards the end of the column, the plug will diminish in size as diluent is constantly mixed with mobile phase until the core is completely dissolved. It has been shown that if the solvent strength of the diluent is much higher than that of mobile phase, peaks becomes broadened and could show splitting behaviour (Enmark et al., 2015). The comparison of elution strength between mobile phase solvent and the chosen diluent solvent could therefore be of interest to predict peak shape. A number of solvents evaluated in

this work and their related polarity and elution strength for normal phase and reversed phase are given in Table 1.

Table 1. Values are taken from Solvents and solvent effects in Organic Chemistry (Reichardt 2003).

Solvent	Polarity	Elution strength, ε° (Al_2O_3)	Elution strength, ε° (TLC RP8 plate)
Water	1.00	Higher (>>1)	0
MeOH	0.762	0.95	1.67
2-PrOH	0.546	0.82	1.46
ACN	0.460	0.65	1.35
tert-BuOH	0.389	-	-

2.3.1 Methods for measuring solubility in supercritical fluids

Any chromatographic separation sets the requirement that the sample is soluble in the mobile phase. Available solubility techniques to date are categorized into dynamic and static solubility methods. The main difference between the two types of methods are the assumption upon reaching solubility equilibrium. Since SF-CO_2 have specific demands regarding environmental pressure and temperature, solubility measurement procedures are hard to perform without affecting the supercritical conditions. A popular theme has been to design this method around an occurring SFC system. There have been several experimental setups proposed under this premise that are published for measuring small molecule solubility. In a study by Gahm et al. (2011) they evaluated the applicability of a new method for measuring solubility of the xanthine derivatives caffeine and theophylline in four organic solvents MeOH, EtOH, IPOH and ACN. The instrumental setup consisted of ten solubility steel chambers with a 10-channel selection valve coupled to a HPLC system with UV detector which allowed for rapid solubility screening for several different conditions. The method was reported to be able to measure solubility of a single compound in SCF but also for a mixture with different thermal and baric conditions. In a study by Li et al. (2016) another method was reported for measuring solubility of caffeine solely in SF-CO_2 using an SFC instrument with or without post-column derivatization. The method was built on a coupling scheme using two Rheodyne valves that linked an equilibrating cell or stirred supercritical vessel (abbreviated SSV) together with an SFC instrument. Their study concluded that the method was a rapid way to integrate an already existing SFC instrument with minimal components to enable solubility measurements with organic compounds. In this work, a novel static solubility method is presented that only requires a simple recoupling of a single SFC system. The method was evaluated using caffeine and aspirin as reference compounds and thereafter applied to determine solubility of the two peptides Leu-Enk and AgII.

2.3.2 Calibration curve

To derive amount of a sample with unknown concentration, a calibration curve or standard curve is usually created. This was performed in this work to determine solubility of reference compounds and later peptides measured with the solubility method described above. A calibration curve represents a regression function built from experimental measurements of a given compound of known amounts and concentration (see Figure 2 below). The reason why one curve must be applied for each specific compound is that the signal strength of a detector varies dependently on the analyte to be detected. In the case of chromatography, one or more standard solutions can be prepared with varying concentration. The samples are then injected in a series of varying injection volumes from where different peak areas are obtained. These quantities are then for each measured sample plotted in a graph as a function of the corresponding amount of sample. A regression line is fitted for this plot in two dimensions where the function is based on minimizing the sum of squares of distance for all residuals, and the intercept is conveniently placed in origin as zero signal should correspond to a concentration of zero.

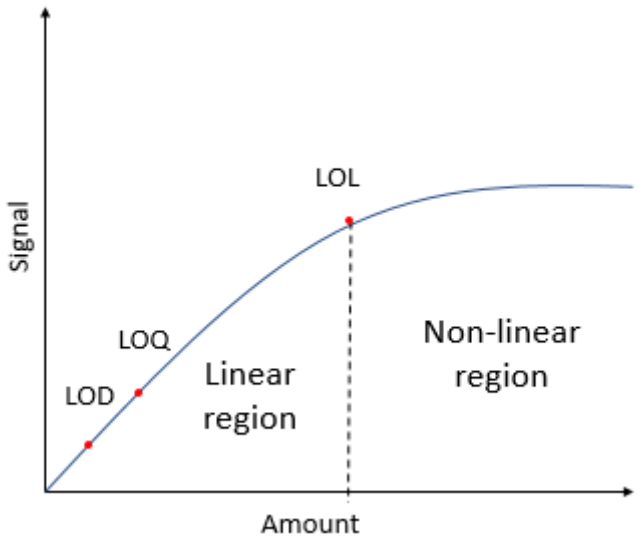

Figure 2. The theoretical shape and corresponding terms for a calibration curve. Explanation of abbreviations: LOD – Limit of Detection, LOQ – Limit of Quantification and LOL – Limit of Linearity.

To this date, peptides are generally separated in RP-HPLC and there is a large scientific support for this technique. The feasibility to convert into SFC would primarily mean an increase in efficiency with shorter analysis time and a lower consumption of organic solvents followed by lower costs. Analyses on a collection of well-chosen peptides with real-world pharmaceutical applications and with variable properties on both techniques could potentially give a first answer to if and in that case when SFC could be an applicable technique.

3 Theory

3.1 Chromatographic theory

A chromatographic procedure is graphically represented by an elution diagram or chromatogram, seen in figure Figure 3 below. The diagram presents the chromatographic behaviour of a compound in shape of a peak and is eluted as a function of time. In ideal cases, all eluted molecules of the same species would be eluted simultaneously at the same time, which would be represented as an infinitesimally thin peak in a chromatogram. In practice though, the peaks are analytically gaussian distributed around a maximum peak apex with equal width on both sides of the centre of the peak. In chromatographic theory, there are a number of chromatographic parameters that are necessary to quantify the performance of the process. The time elapsed from point of injection to the maximum height of a peak is defined as the peak's retention time. This entity is based on the complex relations of molecular interactions between the solute, mobile and stationary phase where affinity to the two phases will either retain or elute the compound. If the affinity between solute and stationary phase is stronger than that of solute and mobile phase, the solute will be retained in the column until the interactions with mobile phase will surpass and thus elute the compound. A way to normalize the retention time for several individual peaks is to calculate the relative retention for each, defined as retention factor. The retention factor k is the retention time for a peak in relation to the void or dead volume of a column which corresponds to the volume occupied only by the eluent and is defined in Equation 2. The void volume can also be quantified in terms of time and is then referred to as the hold-up time. Since the mobile phase should not have affinity to column stationary phase, the void is always eluted first.

$$k = \frac{t_R - t_0}{t_0} \tag{2}$$

Here t_R is the peak retention time and t_0 equals the hold-up time. In the case of a separation problem, the main approach is to have dissimilar retention time for each species. The separation of two peaks can be evaluated by the selectivity which corresponds to the quotient of the difference of their retention time and the hold-up time, given in Equation 3.

$$\alpha = \frac{t_{R,2} - t_0}{t_{R,1} - t_0} = \frac{k_2}{k_1} \tag{3}$$

Selectivity can take values from 1 where peaks are co-eluted to positive infinity. The larger the selectivity α is, the better the separation between the two peak apices.

A chromatographic peak is apart from how long it is retained in the column also evaluated by its shape. It is common to measure the width *w* of a peak measured as the distance between the first eluted part, the front to the end of the peak, termed tail. There are several well established

standard procedures to evaluate peak width such as peak width at half height or 5% of peak height and peak width at 2 standard deviations for equally distributed peaks. The peak width at 5% of peak height is in Figure 3 annotated for the two therapeutic peptides Desmopressin and Leu-Enk.

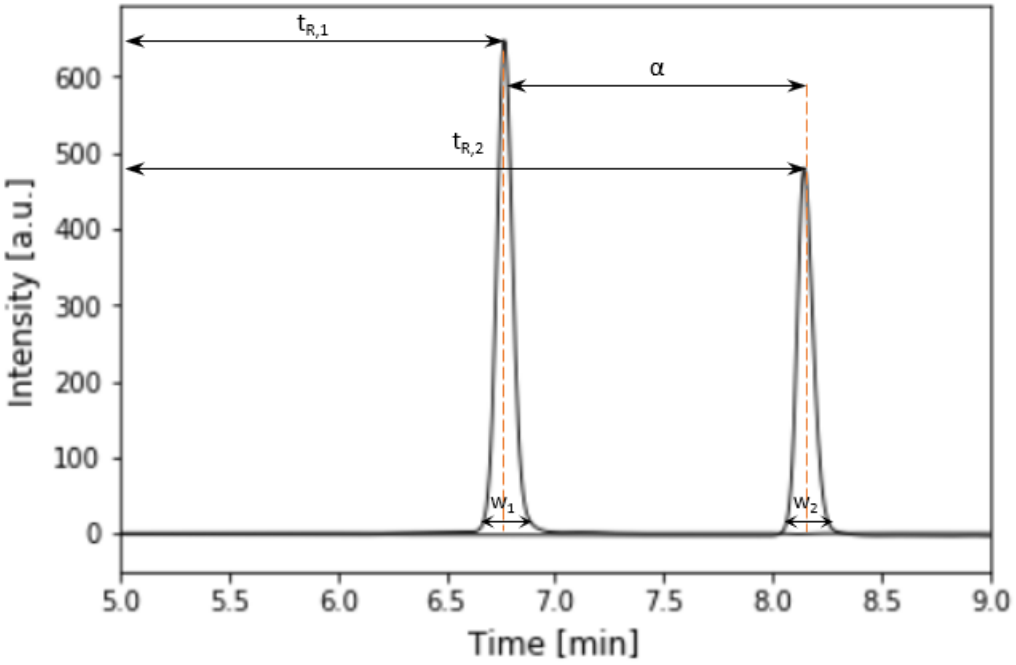

Figure 3. A chromatogram showing the separation of Desmopressin and Leu-Enk in HPLC. The peaks are narrow and has a gaussian disturbed shape suggesting the peptides were injected in analytical amounts. In this case, 2 uL of 1 ml min^{-1} concentration of each peptide were injected into the system. Retention time and peak width for each peak and selectivity for these two is given. Note that the selectivity factor is not the absolute distance between two peaks but is defined as the relation between their retention factors (see equation 3).

3.2 Charge descriptor calculation from amino acid sequence

Calculation of peptide charge in modlAMP is taken from Bjellqvist et al. (1982). This algorithm divides charged amino acids together with C- and N-terminus into groups of positive and negative pK-values depending on the state of charge. The sets are defined as followed:

$$positive\ pK = \{N-terminus = 9.38,\quad K = 10.67,\quad R = 12.10,\quad H = 6.04\}$$

$$negative\ pK = \{C-terminus = 15,\quad D = 3.71,\quad E = 4.15,\quad C = 8.14,\quad Y = 10.10\}$$

Here, amino acids are given in one-letter code. The concentration of the buffering impact for each amino acid is seen in Equation 4 below and is a modification of Hendelson-Hasselbalch equation,

$$C_{r,pos} = 10^{(pK-pH)},\ C_{r,neg} = 10^{(pH-pK)} \qquad (4)$$

where pH is determined by the actual condition. The charge is then calculated as the difference between the sums of the relative concentration *Cr / 1+ Cr* for each amino acid times the number of occurrences for that particular amino acid in the sequence (Equation 5).

$$charge = \sum aa_i \cdot \frac{10^{(pK_i - pH)}}{1 + 10^{(pK_i - pH)}} - \sum aa_i \cdot \frac{10^{(pH - pK_i)}}{1 + 10^{(pH - pK_i)}} \quad (5)$$

For cyclic peptides, this implicates that the charge of C-terminal and N-terminal is neutral due to the closure of the linear sequence through binding.

There is an increasing interest in proteomics to numerically predict retention time as function of a given peptide, given certain input conditions such as mobile phase and temperature. Proteomics and peptide analysis usually include liquid chromatography with on-line mass spectrometry where data such as absorbance and mass spectra generate information of compounds that are present. These data can sometimes result in obtained false-positives that can only be identified by additional data, e.g. by numerical predictions. In a study by Gilar et al. (2010) they presented a retention predicting algorithm that given a protein or peptide sequence calculates the cumulative contribution of each individual amino acid AA_i. This is given in Equation 6 below.

$$RT = (1 - a \cdot \ln L)\left(\sum b_i \cdot AA_i + b_0\right) \quad (6)$$

Here, L equals the length of the peptide sequence, bi is a coefficient related to the actual amino acid, a is an optimized coefficient and b_0 is the intercept of the model that was created in this work. The model from which this algorithm was built upon was constructed using retention data experimentally derived from 165 peptides. The algorithm does not take secondary structure into consideration as it is only related to the total amino acid composition. The experimental retention time and the predicted retention values using this model showed a highly linear correlation with a coefficient of determination equal to 0.9639.

4 Method and Materials

The following section renders the approach for a PCA analysis on the chosen peptide set and the experimental procedures for analysis of peptides and measurements with the newly proposed solubility method.

4.1 Descriptor calculation of peptides and PCA analysis

A script was written in Python language, implementing the open-source available package modlAMP (Müller et al., 2017). The script was designed to store name and sequence of the peptides included in the analytical set (Table 2) and to import a list of known fragment sequences from BSA digest[2]. The peptide sequences were obtained from AstraZeneca in three-letter code, therefore a custom script was written to convert this into one-letter code. This was carried out since modlAMP modules only were able to handle amino acid sequence in this format. Numerous descriptor calculators from sequence data were included such as average charge, hydrophobicity ratio and molecular weight. Equation 6 taken from article by Gilar (2010) was implemented in code and used to predict retention time for all peptides within the test set. The data was collectively depicted in 2-dimensional plots with hydrophobicity ratio set as x-axis to graphically examine the discrepancies or similarities between peptides (Figure 4).

Table 2. Selected peptides that were evaluated through their physico-chemical properties and that were to be included in the chromatographic study. The whole set can be found in Appendix D.

Peptide	Cyclic/ Linear	Sequence	Charge (pH 2)	Hydrophobicity ratio	MW
Angiotensin II	L	DRVYIHPF	2.566	0.375	1046.18
Carbetocin	C	YIQNCPLG	-7.32E-07	0.375	907.05
Desmopressin	C	YFQNCPRG	0.999	0.25	984.09
Gramicidin-Val-A	L	VGALAVVVWLYLWLW	0.585	0.666	1811.22
Gramicidin-Ile-A	L	IGALAVVVWLWLWLW	0.585	0.666	1825.25
Gramicidin-Val-B	L	VGALAVVVWLWLWLW	0.585	0.733	1772.18
Gramicidin-Ile-B	L	IGALAVVVWLFLWLW	0.585	0.733	1786.21
Gramicidin-Val-C	L	VGALAVVVWLYLWLW	0.585	0.666	1788.18
Gramicidin-Ile-C	L	IGALAVVVWLYLWLW	0.585	0.666	1802.21
Leu-Enkephalin	L	YGGFL	0.585	0.4	555.62
Leuprolide	L	PHWSYLLRP	2.585	0.22	1168.35
Somatostatin	C	AGCKNFFWKTFTSC	1.999	0.429	1639.9
Triptorelin	L	EHWSYWLRPG	2.578	0.1	1330.45

[2] https://ionsource.com/Card/protein/BovineSerumAlbumin.htm

A selection of peptides was afterwards done based on the resulting graph and availability on market where the main goal was to pick peptides with dissimilar properties. The selection resulted in the peptides given in Table 2 above.

The calculated data was then exported in csv-file format for PCA analysis of correlation. The objective was to validate that the chosen test set was large enough to assume that no correlation was to be found, this would otherwise had given obvious outcomes due to biased data.

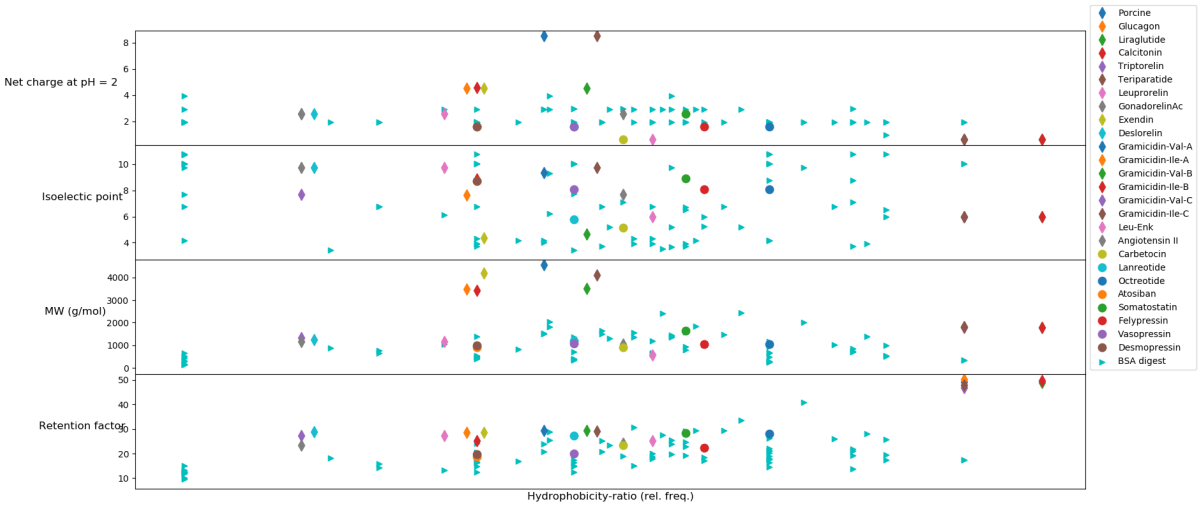

Figure 4. Calculated properties were plotted as function of the hydrophobicity ratio to assess the distribution in a chemical space.

4.2 Chemicals

Mobile phase for all experiments conducted on SFC instruments consisted of CO_2 (99.99%) from AGA Gas AB (Stenungsund, Sweden) and HPLC-grade MeOH. For solubility measurements and analysis of peptides and BSA-digest, a B-phase solvent was prepared using HPLC-grade MeOH with ultra-pure water with conductivity 18.2 MΩ from Merck Millipore (Darmstadt, Germany) and TFA (CAS# 76-05-1) from Sigma Aldrich (St. Louis, MO, USA) as additives. A make-up solvent for the mass spectrometer was prepared from MeOH and two wash solvents (weak/strong wash) were prepared using heptane, IPA and MeOH. Solvents for HPLC analysis were HPLC-grade ACN from VWR, ultra-pure water from Merck Millipore and TFA from Sigma Aldrich was used as additive. The diluents were prepared using the following solvents: t-BuOH (CAS# 75-65-0) from Honeywell Fluka (Seelze ,Germany) or Sigma Aldrich, HPLC grade MeOH, IPA, EtOH, Acetone (CAS# 67-64-1) Fischer Scientific (Loughborough, UK) and ACN (CAS# 75-05-8).

The two test compounds for solubility experiments aspirin (CAS# 50-78-2, >99% purity) and caffeine (CAS# 58-08-2) was provided by Sigma Aldrich. Acid purified SiO_2 sand (CAS# 60676-86-0) from Sigma Aldrich was used as stationary phase in packing the flow cell.

The peptide set used in retention analysis were Angiotensin II human (CAS# 4474-91-3, Bachem AG), Carbetocin (#CAS 37025-55-1, MedChemExpress LLC, Monmouth Junction NJ, USA), Desmopressin (CAS# 62288-83-9, MedChemExpress), Gramicidin A,B,C,D mixture (CAS# 1405-97-6, Sigma Aldrich), Leu Enkephalin (CAS# 58822-25-6, Bachem AG), Leuprolide (CAS# 53714-56-0, provider, country) and Triptorelin (CAS# 57773-63-4, provider, country). The BSA tryptic digest was purchased in 1 nmol amounts from Thermo Fisher (Waltham MA, USA). An additional cyclic peptide, Somatostatin (CAS# 38916-34-6) was synthesized on site at AstraZeneca. AgII and Leu-Enk were also included in solubility measurement experiments and had the same provider as above. The purity of the peptides was not found for all, however the data that was available is compiled in Appendix D D. An unknown crude extract of a peptide will here be mentioned as 'Peptide-A' and was included in the set of analytes. This was a non-purified API compound derived from AstraZeneca.

4.3 Instrumentation

Experiments were conducted on a total of five different instruments. For solubility measurements, a Waters Acquity UPC^2 instrument was used (Waters Corporation, Milford, MA, USA) installed with a PDA detector and two binary solvent managers. The method required a unique system setup following the one seen in Figure 5 in section 4.4.1 below. The setup required coupling of capillaries from binary solvent valves to two external six port, two position vici valves that were externally mounted and preinstalled. A Kromasil SFC-5-Diol column (#F05DIC15, Nouryon, Bohus, Sweden) with dimensions 150x3.0 mm and particle size 5 µm was used for experiments with aspirin and caffeine while a 150x3.0 mm Kromasil SFC-5-XT column (#F05XTC15), 5 µm particle size, was used for the peptide separation. The columns and flow cell were both mounted in a Waters Column Manager unit. The acquisition, analysis and exportation of data was managed in Empower 3 software suite by Waters Corporation.

The peptide analysis for SFC were carried out on a similar Waters Acquity UPC^2 installed with a temperature controlled autosampler with a 7.5 µL injection loop, a binary and isocratic solvent manager, a PDA detector with a detection range of 210-400 nm and a single-quadrupole mass spectrometer conducted in ESI positive and negative mode in range 120-802 m/z. The data was collected, analysed and exported using MassLynx software by Waters Inc. Separation was performed on two columns, Kromasil SFC-5-XT (#F05XTC15) and SFC-5-CN (#F05CNC15) namely. Both had dimensions 150x3.0 mm with 5 µm particle size and were kindly provided by Nouryon (Bohus, Sweden). The peptide analyses on HPLC was performed on an Agilent 1200 series (Agilent Technologies, Santa Clara, CA, USA) installed with an auto-sampler with a 100 µL injection loop, a binary solvent pump, a DAD UV-vis detector and a column thermostat. The column used was a Waters XBridge BEH C18 4.6x100 mm with 2.5 µm particle size (#186008986).

The analysis of BSA digest was carried out on two systems. First, an Acquity UPC^2 SFC instrument similar but not identical to the previous was used. It came installed with a binary and isocratic solvent manager, an auto-sampler with and a PDA detector. The system was coupled to a Synapt G2 High Definition MS system from Waters with quantitative time of flight and involved a mass splitter to enable both fraction collection and mass detection. Ionization was performed using ESI in positive mode in range 160-2000 m/z. For HPLC analysis it was chosen not to use the Agilent 1200 system due to the lack of a mass detector. Instead, an Acquity UPLC by Waters on AstraZeneca site was used installed with a binary solvent manager, a column manager with room for up to four columns, a PDA detector, an auto-sampler and a single-quadrupole mass spectrometer. The MS was run in ESI positive mode with a range of 160-2000 m/z. The separation was carried out on a Waters XBridge column with dimension 2.1x50 mm and 2.5 µm particle size (#186006029).

4.4 Method

4.4.1 Solubility method

The method presented in this work is based on the similar principle as the one proposed by Li et al. as it implements an SFC-instrument together with a small vessel for measuring solubility of organic compounds. An important part of this method is the setup. The coupling scheme represents this and is presented in Figure 5 below.

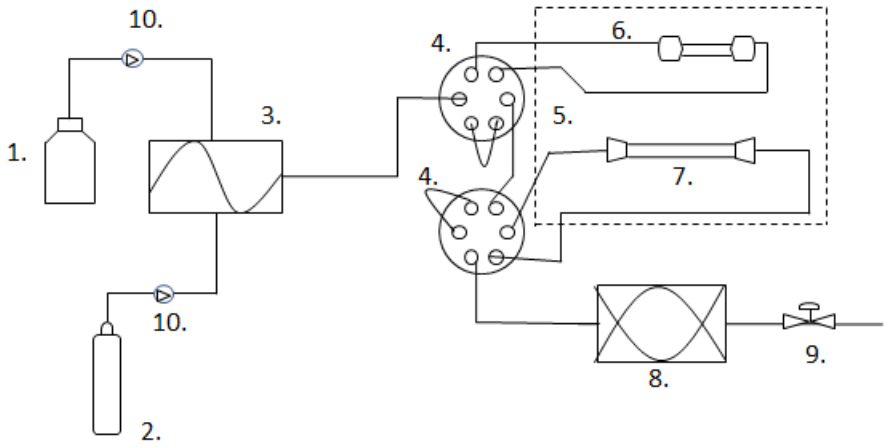

Figure 5. An overview of the coupling scheme used to carry out the solubility measurements. The enumerated species in the scheme are 1. Organic modifier solvent, 2. CO_2 gas cylinder, 3. Solvent mixer, 4. External 6-port-2-position vici-valves, 5. Column oven, 6. flow cell, 7. analytical column, 8. PDA detector, 9. BPR and 10. solvent pumps.

This setup involves the pumping and mixing of CO_2 and organic modifier to a defined mixture which make up the mobile phase. The mobile phase is directed through two six-port Vici valves that can be switched between two states. This enables direct control of flow path through flow cell, column or to waste outlet. Note that regardless of the two positions valve 2 can take, eluate will always finally be directed towards waste outlet.

The experiment was setup in Empower 3 software (Waters Inc.) according to Table 1 below. The method was designed to include a four-step procedure where three different combinations of two two-position valves were used. This method procedure is given in Table 1 and depicted in Figure 6 to facilitate explanation.

Table 3. Overview of the analysis sequence used for solubility measurements.

Valve position	Description	Amount modifier (v/v%)	Length (min)
1a	**Cell equilibration**	0.0	1.50
2	**Column equilibration**	2.0	2.00
1b	**Solubility analysis method**	*	**
3	**Termination of sequence**	2.0	2.00

* the amount of modifier was varied between experiments.
** the solubility method was varied in length and was analyte dependent.

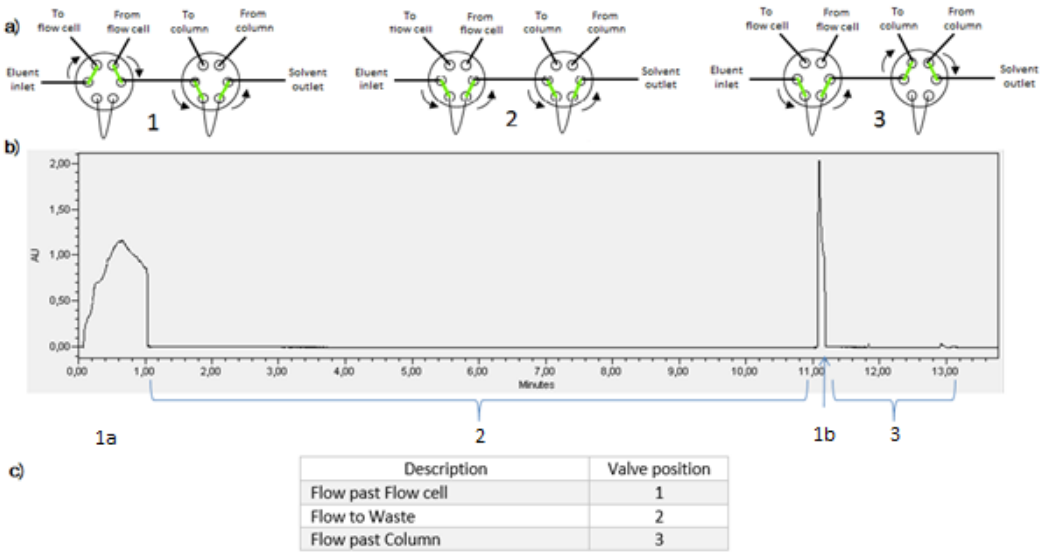

Figure 6. A chromatogram with valve positions depicting the characteristic look of the solubility method and effect of valve switching. The method includes the switching of vici valves to a total of four times into one of the three possible combinations shown above Green implies flow path of solvent. a) A detailed scheme of valve position and fluid flow direction for each condition. b) A characteristic elution diagram for this method with each valve position specified as interval or position. c) A verbal description of each of the valve positions and reference number corresponding to position in a).

In the first step, the cell is flushed with a solvent composition that you wish to evaluate solubility in. In the second, valves are switched to trap solvent in cell and direct newly flushed eluent towards waste. This step is referred to as equilibration step, where a static equilibrium is reached. In the third step, flow is again directed past the cell displacing the solvent which has been let statically equilibrated. This volumetric quantity, call it plug, is the part of analytical

interest and where the amount of analyte should have reached its limit of solubilization. The time constraint for this part determines how much of the dissolved, trapped analyte that is directed for analysis before valves switch to the third position which changes flow towards the column. The time constraints for each step was experimentally changed to evaluate which intervals that resembled into the highest detection of dissolved analyte, which thus would correspond to the solubility closest to the theoretical values. The following values were evaluated for several parameters, specified in Table 3 below.

Table 4. Values of method parameters assessed for this method.

Time [min] for valve position 2 (*equilibration*)	Time [min] for valve position 1b (*solvent transfer*)	System Pressure	Modifier (v/v% MeOH) for aspirin analyses
2.5 – 20	**0.02 – 0.25**	**120, 140, 160, 180**	**0, 2.4, 5.0, 8.0**

The time constraint for position 3 is not given in Table 3, but was experimentally determined empirically. The goal was to elute compound within 2-3 minutes after the valves were switched into position 3.

The analytes were loaded by dry-packing a 0.2 mL emptied 2.1x50 mm column with sand as stationary phase. The packing procedure commenced with analyte weighed in a 20 ml glass vial and then mixed with a predetermined amount of SiO_2 sand. A small, custom made funnel was used to facilitate the transfer of this mixture into the cell, before the end cap of column was screwed back on. The mass of each analyte loaded was empirically determined. The theoretical solubility for each compound for the actual conditions was used as guidance to avoid packing of too small amounts. For aspirin and caffeine, approximately 90 and 140 mg were loaded, respectively.

Two measurements were conducted with Leu-Enk, where the cell was packed with approximately 40 mg of Leu-Enk. For AgII, two experiments were conducted where two amounts was evaluated, 6 and 3 mg respectively. The unknown crude extract ‘Peptide-A’ was loaded in an amount of 4 mg.

4.4.2 Calibration Curves

Samples with 1 and 3 mg/ml concentration of caffeine were prepared by weighting the analyte powder in 2 ml glass vials before diluting with ethanol. The analytic peaks were generated by injecting a series of injection volumes from each sample ranging from 2 to 10 µL with an interval of 2 µL. Each injection was repeated to guarantee reproducibility. A gradient corresponding to 2-40 v/v% MeOH, 4 minutes was used which was sufficient to elute caffeine with an adequate peak retention and width. Peaks were integrated in Empower software and plotted against the theoretical masses for each injection.

Calibration curves were also generated for aspirin and Leu-Enk in the same manner as described for caffeine above. Sample concentrations for aspirin was set to 2 and 7 mg/ml and 0, 5 and 3 mg/ml for Leu-Enk. A gradient of 2-40 v/v% MeOH was run in 5 minutes for both analytes.
All calibration curves generated showed strong linear correlations which means that no mass injections were high enough to be found in the non-linear region. The coefficient of determination (R^2) was close to 1.00 for all regressions suggesting high degree of linear correlation. All calibration curves with respective coefficients have been included and can be found in Appendix B.

4.4.3 Analytical and overloaded study of Peptides in SFC and HPLC

Peptides were analyzed in SFC using a linear gradient of 20-70 v/v% co-solvent in 7 minutes, followed by 2 minutes of isocratic elution at 70 v/v% to elute any potential analytes still retained in column and 4.5 minutes isocratic elution at 20 v/v% to equilibrate before subsequent injection. The choice of using gradient was based on the report by Enmark et al. (2018) where a comparison between isocratic and gradient mode for separating Gramicidin isoforms showed greater robustness of separation when applying a co-solvent gradient elution. The study did also report greater robustness when using an increased gradient slope ($\geq$7% min^{-1}), which was used as basis for decision of gradient specification. A detailed description of chromatographic method for each technique is given in Appendix C. Injection volumes were set to 2 and 10 µL to evaluate both analytical and overloaded behavior of analyte and diluent. For HPLC measurements, a linear gradient with slope 2.5 v/v% min^{-1} ACN and 12 v/v% ACN was used for all peptides. The gradient was followed by a 2 min isocratic elution with 70 v/v% ACN and terminated with a 3 min 12 v/v% ACN equilibration step. Analytical injections were performed at 3 µL and the larger 100 µL sample loop allowed 20 µL overloaded injections. A 2 ml vial filled with ultrapure water was used as wash vial to reduce the risk of any carry-over by the needle between injections. For both SFC and HPLC, a 2-5 µL injection of ultrapure water was included prior to analysis in each sequence. This was also included within a sequence between analyses if diluents differed between samples. The main reason was to hamper any eventual carry-over from previous injections that would potentially affect peak shape and retention.

4.4.4 Analysis of BSA digest on SFC-QTof MS and LC-MS

The UV detector was set to measure between 200 and 350 nm and the mass spectrometer from 160 to 2000 m/z. The BSA digest was first analyzed in SFC with a gradient of 20-70 v/v% co-solvent in 7 minutes for collecting retention data of peptides to be included in the peptide set. A second run was carried out with a longer gradient using 20-80 v/v% co-solvent in 33 minutes to evaluate the separation performance of identified peptides. The mass data was worked with in MassLynx by comparing mass spectra with mass chromatograms of individual masses.

4.5 Mobile phase

For HPLC an A-phase solvent was prepared containing ultra-pure water and 0.1% TFA and a B-phase solvent consisting of ACN (VWR, Radnor, PA, USA) with 0.1% TFA. A magnetic

stirrer was used to blend both phases before being manually degassed using a water sonicator bath from VWR and a closed filtration funnel coupled to a vacuum inlet. The mobile phase composition for all SFC experiments consisted of HPLC-graded MeOH with 7% ultrapure water and 0.1% TFA. The weak and strong needle wash for BSA digest analysis were mixed 80/20 v/v% heptane/IPA and 50/50 v/v% MeOH/IPA, respectively. The SFC UPC^2 instruments were all equipped with degassers hence no need for manual degassing of these solvents. All mobile phases were prepared through gravimetric measure of amount solvent to increase measuring accuracy. All solvents were contained in 1L borosilicate flasks and stored in room temperature.

4.6 Preparation of diluents

Peptides were weighed in and diluted to a final concentration of 1 mg ml^{-1} in 2 ml glass vials, closed with cap and septa. Due to differences in polarity, some peptides precipitated in one condition whilst other were completely miscible. 1 µL of TFA was added for these cases to completely dissolve samples. All completed dissolutions were followed by sonication in 30 °C in a water sonication bath for a short amount of time. All samples were if not stowed in the auto-sampler stored in a +4 °C in a cooled storing room.

5 Results & Discussion

5.1 PCA results for therapeutic peptides

The score plot from the results of the PCA analysis is seen in Figure 7. The variance and distribution of peptide sequences was considered for the first three principle components to be sufficient to determine which peptides that should be included in the chromatographic analysis. Values for the variation for the first three principle components were t[1] = 0.384, t[2] = 0.208 and t[3] = 0.194 respectively. Occurrence of any eventual outliers was also assessed where the Hotelling's T^2 test was performed to define a useful boundary (given by the oval area in Figure 7). As seen, only five BSA fragments were found to be mild outliers together with gramicidin (not seen) that was considered as a strong outlier. Apart from the score plot in Figure 7, the three first principal components together with a summary of fit was examined to see how well

data was interspersed between the given descriptors. A loading scatter plot was also evaluated which suggested no certain correlations were to be found between descriptor variables.

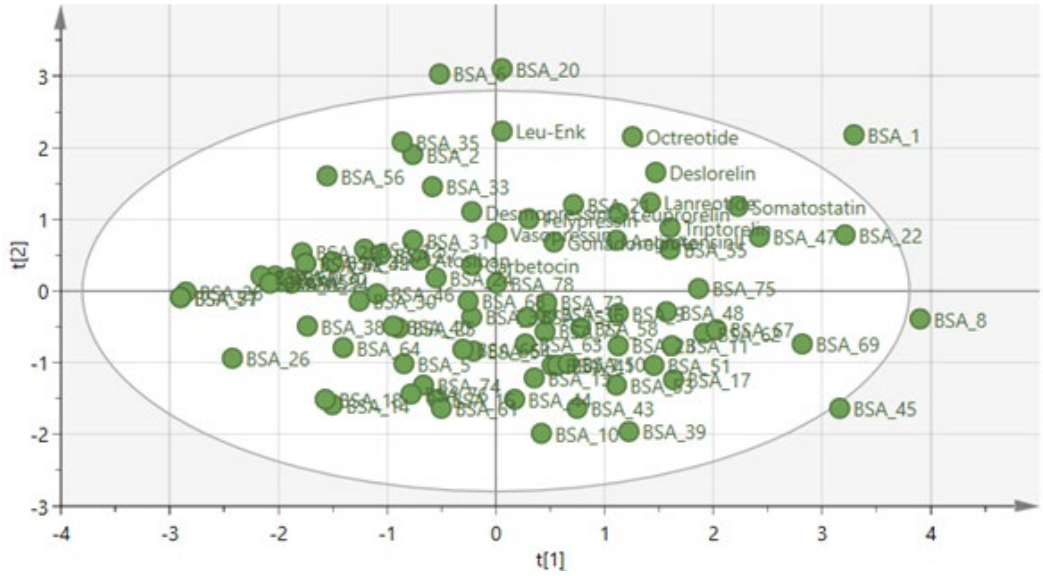

Figure 7. PCA analysis of the test set including BSA fragment data, where the oval-shaped region is defined by the Hotelling's T^2 with a pre-determined confidence interval. This is made to detect obvious outliers within data. The analysis was performed in Modde Simca® 15.

5.2 Solubility method feasibility

The method was evaluated for both pure SF-CO_2 and with organic modifier using the two different test compounds aspirin and caffeine. The method was finally assessed if it can be used to determine solubility of peptides. The choice of test compounds was mainly determined by the literature available. Studies have been conducted on extraction of caffeine from plant tissue since the 1990-ies (Mehr et al, 1996, Peker et al., 1992), caffeine has therefore been made as model compound for assessing solubility in SF-CO_2. The articles by Gahm et al. (2011) and Li et al. (2016) both involve a new technical approach of applying static solubility measurements, while using caffeine as model compound. The two reports however assess the solubility in pure and with organic modifier solvent, respectively. The experimental values from Li et al. were converted from percentage mol fraction into approximate mg ml^{-1} concentrations for the pressures 100 to 200 bar with 20 bar increments, and at constant temperature (40 °C). For measurements with organic modifier, caffeine was assessed however the results showed too high solubility which required substantial amounts of solute packed in flow cell. The pattern from chromatograms were between replicates showing uncertainty of amount dissolved, probably due to the larger impact of the equilibrium rate that is dependent on amount analyte present in column. Literature by Huang et al. (2005) and Vorobei et al. (2019) presented solubility measurements of Aspirin in SF-CO_2 with mol fraction organic modifier. Aspirin was therefore instead chosen to be analyzed for these conditions. Both works presented the amount of modifier in percentage mol fraction which required the conversion into concentration in mg ml^{-1}.

The initial experiments conducted on caffeine gave an important insight to which parameters had the most impact on the resulting output. The equilibrium time was evaluated in the range of 2.5 – 40 min which did not show any significant impact on solubility (not presented here). It was concluded to set equilibration time to either 10 or 20 minutes for the following experiments thus enabling sufficient equilibration. The time interval for flushing the flow cell (phase 1a) was also evaluated, which showed necessity for optimization. Too long interval resulted in the draining of analyte from flow cell, thus diminishing the subsequent solubility measurement due to too low amounts of test compound. Too short interval however meant that the flow cell was not completely filled with the solvent of choice. These conclusions resulted in an adequate time interval of 1 minute.

The overall method sequence was assessed where an initial cell equilibration in pure SF-CO_2 showed an increase in absorbance of analyte. This finding suggested that an initial flush of flow cell prepares sample by introducing the given pressure and temperature conditions. Hence, the analyte reaches equilibrium quicker in following solubility analysis steps.

When these initial parameters had been determined, it was evaluated if packing analyte with a stationary phase could enhance the dissolution through an increase of exposing surface. The packing of acid-purified SiO_2 showed improved acquisition of caffeine through increased detection and more robust measurements. The emptying of the flow cell after measurements showed aggregation of caffeine that was fairly difficult to rinse out. When sand was included in packing this issue was solved. It is reasonable to believe that the aggregation could have an impact on flow to and from flow cell as this can block the pores within the frits. An overall increase of system pressure was often encountered when not using sand as stationary phase.

The first few experiments with aspirin using 2.4% MeOH as co-solvent indicated that the measured solubility was highly sensitive to the chosen opening time and thus required optimization. This resulted in an attempt to introduce a simple way to calculate the solubility by taking the whole chromatogram of the analysis into consideration. It was concluded that the

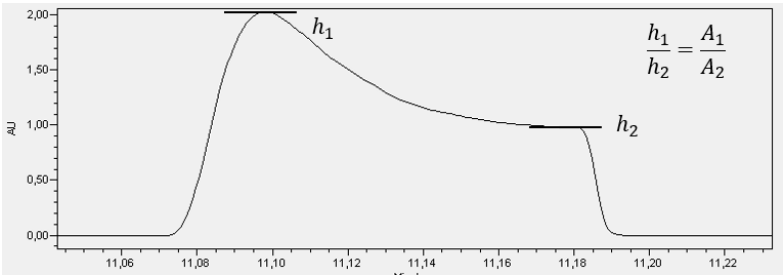

Figure 8. Measurements with modifier gave a characteristic elution profile, with a maximal peak, h_1 and a valley of height h_2. A mathematical relationship can be used to relate these two peaks to peak area. Chromatogram is taken from analysis of aspirin and 2.4% MeOH, measured at wavelength 295 nm.

introduction of modifier to the total mobile phase resulted in a characteristic shape of the elution profile that was generated from opening the valve from flow cell after equilibration. The shape is depicted in Figure 8 where the chromatogram was generated at wavelength 295 nm. The

elution behavior is explained by the difference in aspirin concentration over time. The solvent which has been equilibrated in the flow cell is passed through the detector which detects the highest concentration at h_1. The succeeding solvent has less aspirin dissolved over time until it plains out and the dynamic solubility is reached at h_2. The following analytic peak seen in phase 3, Figure 6, will be related to the solvent that is trapped at the end of this elution profile. By expressing the relationship between these two heights and the area of the following analytical peak A_2, the area of peak A_1, which relates to the highest concentration of aspirin present, can be calculated. This correction was evaluated for the experiments with aspirin and compared to non-corrected values.

The results now continue by dividing the results into three parts; solubility of caffeine in pure SF-CO_2, solubility of aspirin in SF-CO_2 and volume fraction MeOH and solubility evaluation of Leu-Enk and the unknown Peptide-A.

5.2.1 Solubility measurements of caffeine in pure SF-CO_2

Caffeine was measured both without and with SiO_2 sand. Packing with sand suggested an increase in dissolved amount of caffeine (Figure 9). Reference values from four studies are plotted with the measured solubilities from this work. The plots in Figure 9 suggests that measurements are within an acceptable solubility range and has a dependency with pressure. The major discrepancy between series of data is probably due to the method robustness in maintaining pressure. The study by Ramsey et al. (2016) together with this study are those two that show larger solubility values at each measured pressure. Surely, the method setup could induce higher pressures than is set by the BPR. The system pressure was monitored throughout the experiments and showed a 5 bar increase from the pressure set (a 3.5% error at most). If data was to be corrected to this deviation, it would still correlate well to reference data.

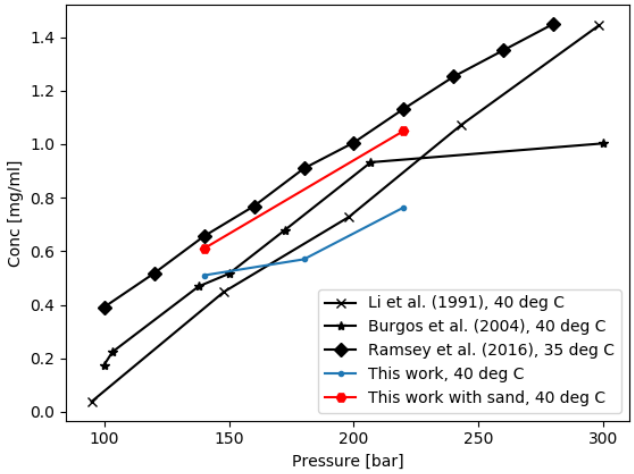

Figure 9. Solubility of caffeine as function of pressure of fluid. Reference values are here plotted in black. Measured values in this work are included in red for measurements with sand and in blue without sand.

Measurements using a fraction modifier in mobile phase was initially tested with caffeine. The compound was soluble to a high extent however the robustness was low as each replicate showed high variance. It was therefore concluded to use aspirin for assessing measurements with fraction modifier present.

5.2.2 Solubility of aspirin with MeOH as modifier

Following the 3 mol% of MeOH used in literature, initial experiments were conducted using approximated value of 2.4% MeOH. The results are seen in Figure 10 where measurements were made at the three pressures 140, 160 and 180 bar. The values were closest to Huang et al.'s results (2005). This could either implicate that these two methods either perform equally or are the closest to the theoretically true value. A deviation from the increasing trend of solubility as function of pressure is seen for the measurement at 160 bar. In contrast to the other measurements at 140 and 180 bar, no replicates were made at this state. It is therefore hard to conclude if the result suggests low robustness for these particular conditions or if the experimental setup was temporarily defective.

The impact on increasing fraction MeOH was evaluated briefly. Aspirin was measured with 5% MeOH at 120 bar which gave a solubility of 3.86 mg ml^{-1}. The value seems reasonable for being preliminary results when compared with Figure 10, but requires to be evaluated further at several pressures and fraction MeOH to be validated.

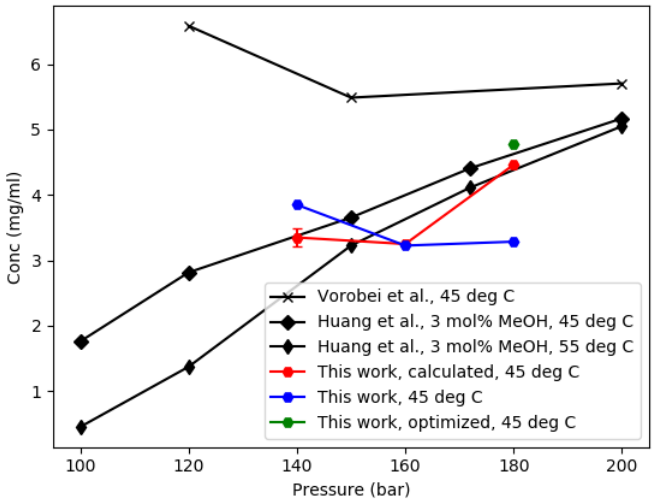

Figure 10. Solubility of aspirin and its dependency on pressure. The reference values from Vorobei et al. and Huang et al. are plotted in black. The measured solubilities from peak area are shown in blue and the calculated solubilities from elution profiles are shown in red.

5.2.3 Solubility measurements of peptides

Successful measurements carried out with caffeine and aspirin led to the investigation if the method could be applied to determine solubility of peptides. If this would prove to work, the method could be a valuable tool when separating in a preparative scale. This could be of particular use for assessing peptide separation in preparative scale SFC. The method parameters and setup were kept from the previous experiments with aspirin, apart from changing mobile and stationary phase. The approach was to adopt the conditions from peptide analysis subproject, by evaluating the solubility in the mobile phase composition at start of gradient, i.e. 20% MeOH with 7% water and 0.1% TFA. Leu-Enk showed to be a compatible compound for this method with a clear analytic peak. The measurement suggested a solubility of 1.90 mg ml^{-1} in 20 v/v% co-solvent. The solubility of Leu-Enk in pure water is 5 mg ml^{-1} [3], why this experimental value at least can be considered realistic. The solubilities of the peptides AgII and

[3] Data provided by Merck/Sigma Aldrich

the unknown “Peptide-A” could not be determined using 20% modifier, even when opening time was changed (Table 4). The results were both non-reproducible and showed precipitation in frits, implicating that the solubility in actual conditions was too low. An increase in volumetric fraction modifier in mobile phase should probably enhance solubility, as seen for both caffeine and aspirin in this study. Furthermore, an increased fraction water in modifier could potentially further increase solubility. This would in particular concern AgII since it has a five-fold larger solubility than Leu-Enk in water (Sigma-Aldrich), explained by its native presence in aqueous environments in the body.

Table 5. The resulting solubilities of three chosen peptides.

Peptide	Solubility [mg ml^{-1}]
Leu-Enk	1.90
AgII	n/a
“Peptide-A”	n/a

5.3 Diluent optimization study for SFC and HPLC

5.3.1 Evaluating fraction water in MeOH and t-BuOH

The volume fraction of water in diluent was assessed for both HPLC and SFC, where it was varied in the range between 10-50%. Since the distribution in physico-chemical properties were maximized within the peptide set, some were not soluble in this range. Therefore, a small amount of TFA was added to those samples that showed sign of precipitation. It should however be noted that this could potentially influence peak shape. Any diluent that was spiked with TFA will be denoted in the following section and further discussed. What is known about HPLC is that the best diluent is the one most similar in composition to the mobile phase. The results of Leu-Enk in ACN, MeOH and t-BuOH with different volumetric fraction of water is presented in Figure 11 below. The goal was to inject overloaded amounts to enhance any potential diluent effects. The injections correspond to 20 µL respectively.

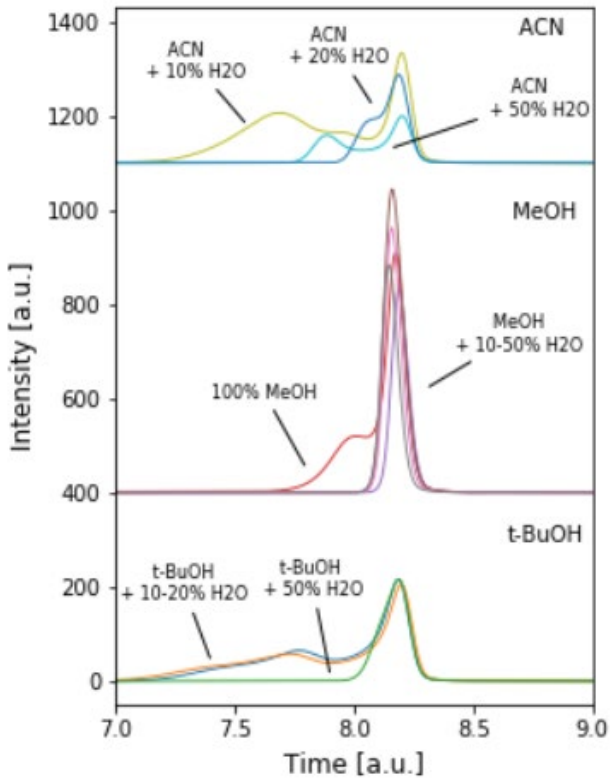

Figure 11. HPLC diluent experiment of Leu-Enk with different volumetric fraction of water. The experimental parameters were fixed to enable comparison between diluents where a 12-42% ACN gradient in 12 min was used. Injection volume is 20 µL from diluents with concentration of 1 mg ml^{-1}.

It can clearly be seen that diluents with 50% water show improved peak shape. Clear diluent effects can be seen for ACN and t-BuOH with 10 and 20% water and pure MeOH. The diluent effect is taken the form of a broad "hill-shaped" fronting phenomenon in the beginning of the peak. The results strongly suggest that MeOH with water is the best diluent of choice in terms of peak shape, however it is not possible to deduce an optimal volume fraction of water. A reservation should in this case be carried out as the results are specifically related to Leu-Enk and a result of multiple interactions with its intrinsic physico-chemical properties. Being a relatively small peptide, only five amino acids long and with no charged side groups, the peptide is relatively easy to dissolve in organic modifier. The similarity of amino acid composition and size of molecule should not result in any major dispersion that otherwise could cause peak broadening. Peak shape and elution behavior should however deviate for any of the other peptides in the set.

Similar experiments were carried out with SFC using the same set of diluents but with AgII instead of Leu-Enk. The main reason for evaluating AgII was because smeared peaks had been obtained in SFC after have performed quick screening experiments . Therefore, following experiments were assessed to try finding a diluent which could eventually narrow peaks. Experiments with ACN (not presented) were evaluated but showed heavy diluent effects and were therefore chosen not to be included in this section. Injection volumes of 2 and 10 µL for t-BuOH and MeOH are seen in Figure 12 and Figure 13, respectively. The analytic injections showed of none to low presence of diluent effects, as was expected. Both t-BuOH and MeOH show acceptable peak shapes for water fractions between 20-50% with similar peak widths (<0.5). Clear diluent effects are however observed for the 10 µL injections with water fractions of 30-50% for both solvents. The deviating peak shapes for t-BuOH with 10% water and MeOH with 20% water are probably caused by spiking of TFA, which was not necessary to add to the other diluents. The low volume fraction water caused AgII to precipitate, which is why these two were required to be spiked with TFA. These two peaks show no signs of diluent effects or peak split, but it comes at the price of increasing peak width. A probable explanation for the retention behavior of these is that TFA which is a strong acid protolyzes when being injected

into the diluent. The acid can after proteolysis ion-pair with charged residues of AgII, thus increasing overall hydrophobicity of the peptide. SFC is known for retaining polar compounds better than nonpolar, which is why the peak apex is found much earlier than in the other diluents. This does not however explain why all diluents showed identical tailing, but one explanation could be that the quantity of moles TFA does not equal to the moles of charged amino acid groups of AgII which would cause dispersion to some degree.

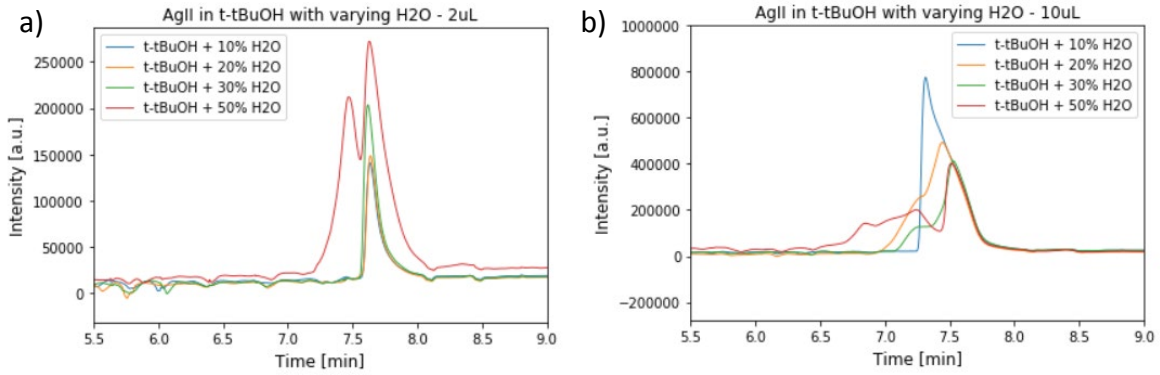

Figure 12. Injections of AgII with t-BuOH and varying fraction water in SFC. The chromatograms shown correspond to a) 2 and b) 10 µL injections.

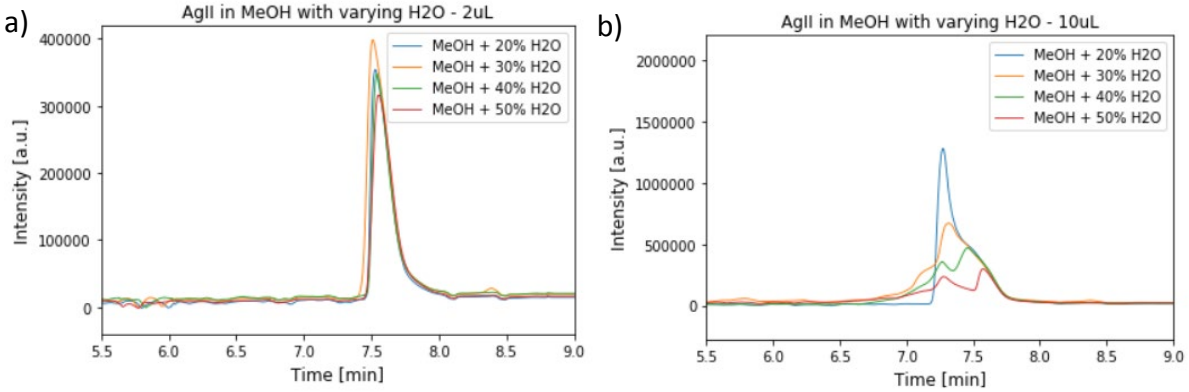

Figure 13. Injections of AgII with MeOH and varying fraction of water in SFC. The chromatograms shown correspond to a) 2 and b) 10 µL injections.

Apart from this, it can be concluded that 50% water increases diluent effects for both MeOH and t-BuOH which is why higher fraction water is probably not favorable for SFC. The complex relation in diluents is however that the fraction water should be adequate for peptides to dissolve. A volumetric fraction of 20 and 30% of water did demonstrate for both t-BuOH and MeOH some diluent effects, but is minor relative to what was seen for higher fractions.

5.3.2 Diluent effects on peak shape and retention in SFC

The results from Figure 12 and Figure 13 were taken into consideration when investigating diluent effect on a larger set of solvents. AgII was again chosen as analyte. After having performed a screening of several peptides in the set on SFC, AgII showed that it was prone to peak broadening even in analytical amounts. Each organic solvent was diluted with 30% water for two reasons. Mainly since results from section 5.3.1 showed acceptable peak shapes for volumetric fractions of 20-30% and second since this was the lowest amount of water that could be added for AgII to be miscible in all solvents. For analytical injections, no obvious diluent effects could be seen for any of the solvent samples. A deviation in signal intensity can be seen in Figure 14 where t-BuOH showed the least smallest peak area and IPA, ACN and acetone the most. For the 10 µL injections however, the effect of solvent becomes more obvious (Figure 15). Effects are seen for all diluents but to a different extent. It can most significantly be related to ACN, acetone and IPA while t-BuOH, MeOH and EtOH indicate acceptable peaks. Particularly, EtOH performs very well with a compact peak and with the least fronting. These signals show that IPA, EtOH and t-BuOH has good coherence of sample.

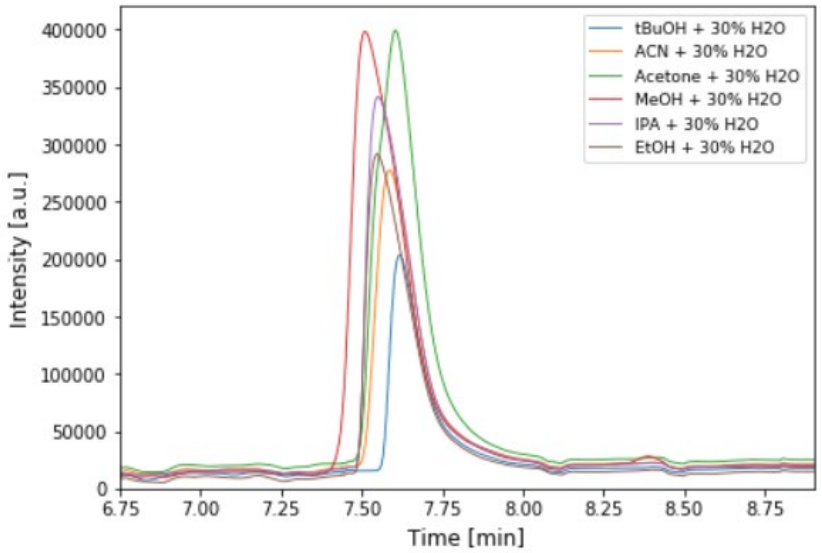

Figure 14. Analytical 2 µL injections of AgII in SFC with various solvents and 30% water. The experiments were performed on a Kromasil XT 3.0x150 mm.

A major implication for the diluent experiments presented here and in previous section 5.3.1 is that there are small differences in concentration of analyte present. For each diluent the mass of 0.5 mg of AgII was weighed using a scale with an accuracy of 0.1 mg. If the weighed mass would have an exact value of 0.45 the concentration would be 10% lower than presumed, thus resulting in at minimum 10% difference in peak area. As measurements have been performed on highly sensitive detection apparatus it is obvious that small errors will be obvious. No replicates were made for each diluent, however in hindsight this should have been considered

from the beginning. In the case of the experiments in section 5.3.1, no stock solution could be prepared either since the concentration would differ between samples.

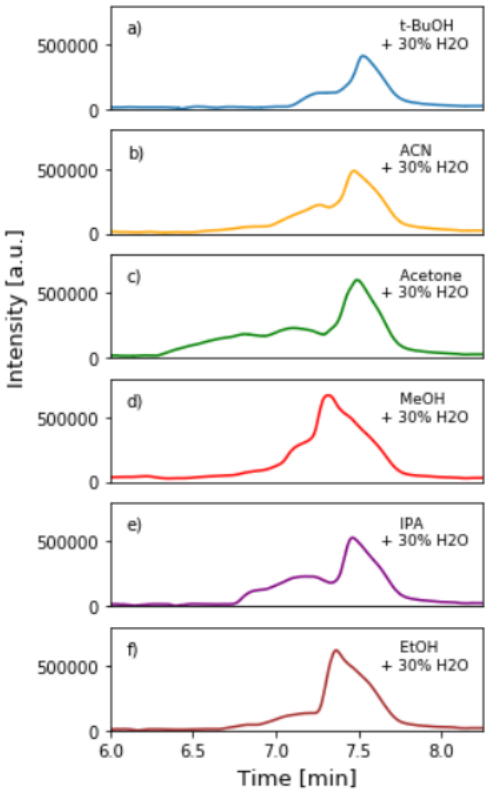

Figure 15. 10 µL injections of AgII in SFC with various solvents and 30% water. The experiments were performed on a Kromasil XT 3.0x150 mm. Mass-to-charge (m/z) signal is superpositioned over UV-chromatogram.

5.4 Selectivity and Retention in HPLC vs SFC

From section 5.3.2 it was decided to proceed with the diluents t-BuOH + 10% water and MeOH + 20% water as these gave adequate peak shape for AgII. Remaining peptides in the set were diluted with these two diluents to assess peptide selectivity for both HPLC and SFC. Here, the resulting elution profiles for each peptide are presented for respective technique. The expected retention behavior of RP-HPLC is that more hydrophobic compounds are retained longer while polar or charged compounds will be eluted earlier. The results seen in Figure 16 a) and b) suggests an elution pattern that does not follow this entirely. For instance, both Leuprolide and Triptorelin have hydrophobicity ratios of 0.22 and 0.1 respectively, which are the lowest values for this set of peptides (Appendix D). Interestingly, the two cyclic peptides Desmopressin and Carbetocin have similar molecular weights while Desmopressin has a notable larger net charge and higher hydrophobicity ratio. Still, the inverted retention behavior can be observed. This implicates that the retention cannot be described only by the intrinsic properties evaluated in this study.

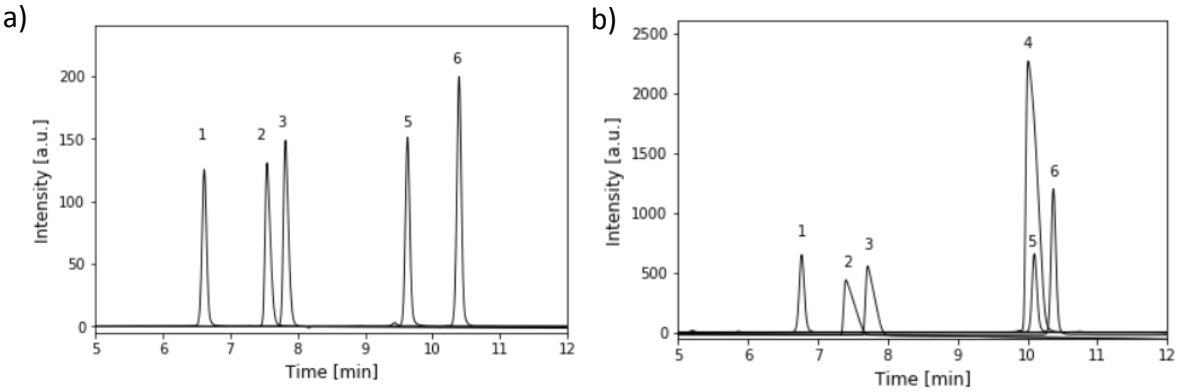

Figure 16. Separation of peptides in RP-HPLC with a) 3 µl and b) 20 µl injections, respectively. The enumerated peaks correspond to: 1. Desmopressin, 2. AgII, 3. Leu-Enk, 4. Leuprolide, 5. Triptorelin, 6. Carbetocin and 7. Gramicidin (not shown). Gramicidin was eluted after approximately 18 min. The separation was carried out on a Waters XBridge C18 column (4.6x100 mm), 2.5 µm particle size.

For analysis in SFC, the order of eluting peptides should be similar to that of NPLC, i.e. according to increasing polarity. Yet again, there is no particular correlation of hydrophobicity ratio to retention order (Figure 17).

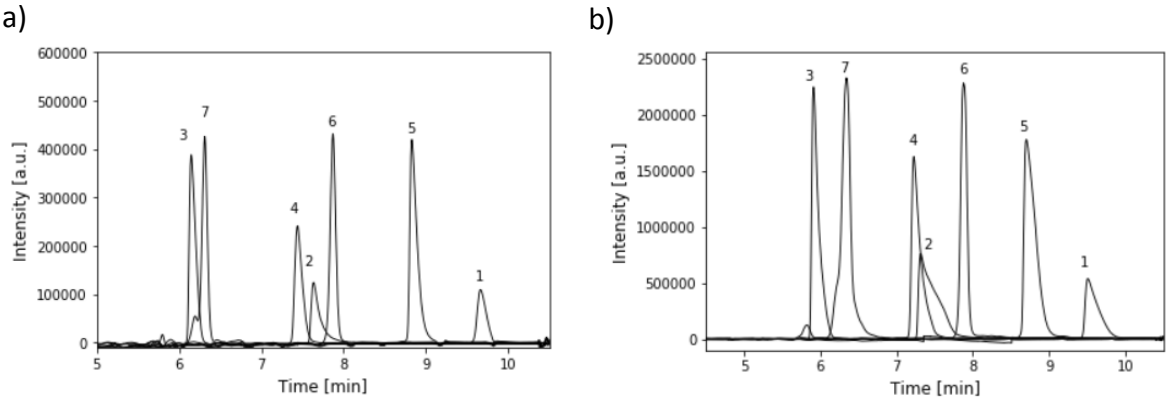

Figure 17. Separation of peptides in SFC with a) 2 µl and b) 10 µl injections, respectively. The enumerated peaks correspond to: 1. Desmopressin, 2. AgII, 3. Leu-Enk, 4. Leuprolide, 5. Triptorelin, 6. Carbetocin and 7. Gramicidin. The separation was carried out on a Kromasil SFC XT column (3.0x150 mm), 5 µm pore size.

The comparison of linear sequence data is not completely applicable, as it leaves many unanswered questions for the eluted order of the peptides. In Figure 18 the relation between three properties and the retention in SFC experiments (here in form of retention factor). The hydrophobicity ratio describes only the fraction of hydrophobic amino acids within the sequence, without taking into consideration how many that are exposed to the stationary phase and eluent due to conformation. Fundamentally, proteins have a hydrophobic core with polar residues directed from the center when surrounded by a polar environment. For shorter fragments like peptides, secondary and tertiary structures are not as commonly occurring as for larger proteins due to their constraint in their flexibility of peptide chain, however they do exist. Descriptors taking hydrophobic and hydrophilic surface area into consideration could potentially explain more about the results from a chromatographic process.

In terms of peptide separation however, SFC is for these specified conditions fully capable of separating the peptides within the set. Among these, Leuprolide and Triptorelin which were co-eluted for the HPLC procedure were fully separated in SFC. One difficulty that this setup might

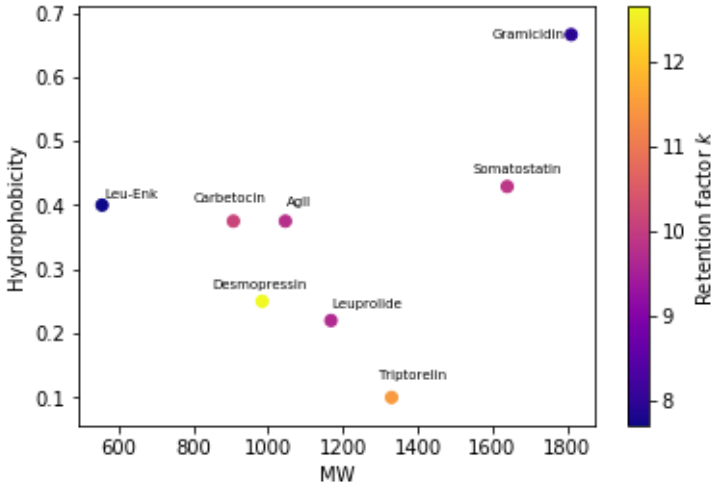

Figure 18. Spatial partion of peptides in a down-scaled chemical space with regard to their molecular weight (MW), hydrophobicity and retention factor in SFC. The most correlating property seems primarily to be the hydrophobicity ratio.

entail is the separation of AgII from Leuprolide and Somatostatin, which is further complicated by the broadening behavior of AgII. As the separation was performed on two different SFC column chemistries, it was also examined if the separation could be enhanced by change of stationary phase.

The selectivity for each pair of peptides was calculated for separations on both XT and CN column, and is shown in the two-dimensional surface plots below (Figure 19, calculated values are found in Appendix C). Interestingly, selectivity differed between the two columns with significant results for the two cyclic peptides Desmopressin and Carbetocin.

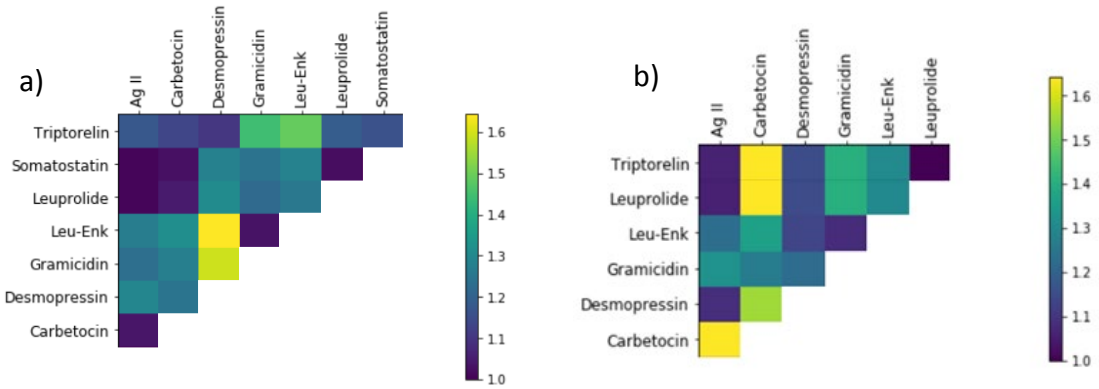

Figure 19. Comparison of selectivity of each pair of peptides between a) Kromasil XT and b) CN (3.0x150 mm) columns.

This result implies that in terms of selectivity, a CN or XT stationary phase can be used to promote small changes when separation needs to be enhanced. The differences in retention for the peptide set can also be seen in Figure 20. The spread of data points suggests that orthogonality between these two stationary phases exists to some extent, which further emphasizes that the choice between SFC column chemistries are important.

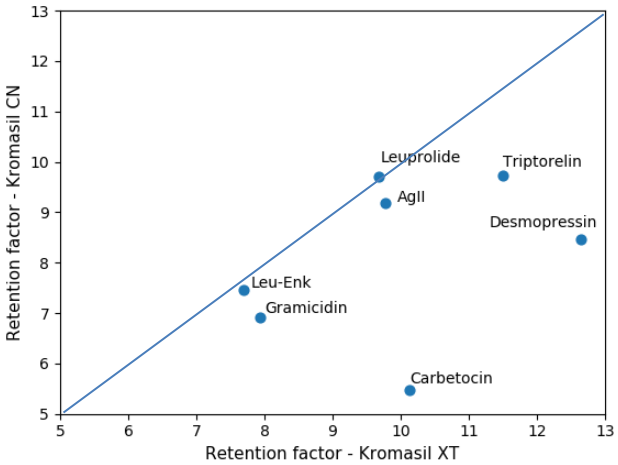

Figure 20. Difference in retention factor between the two columns XT and CN, using a MeOH, 7% water and 0.1% TFA mobile phase and gradient 20-70% in 7 minutes.

5.5 BSA analysis on SFC and HPLC

Analysis on BSA tryptic digest was carried out on SFC and HPLC. The experiment is interesting in two aspects, primarily to evaluate feasibility of SFC for separating a vast number of peptide fragments which could have applications in proteomics. The other aspect is to expand the set of peptides to link physico-chemical properties to retention behavior. A vast number of fragments were identified using both techniques, among these were ten distinct sequences able to be identified. Their retention masses and retention factors are summarized in Table 6.

The predicted retention factor was also calculated based on Gahm et al.s article (2011). The values derived from this model correlated well with those derived from the SFC experiment but deviated a lot from experiments with HPLC. The HPLC method gradient was scaled with regard to the column that was at that time available, a 2.1x50 mm Waters XBridge C18 to correlate with the one used for peptide analysis on the Agilent 1200 system. The length of method in time was also taken into consideration due to difference in column length.

Table 6. Isotopic masses and retention factors, *k* for HPLC and SFC for ten identified fragments in BSA digest.

Sequence	Isotopic mass	*k* - HPLC	*k* - SFC	*k* - predicted
FWGK	536,62	8,3	19,5	19,13
LVTDLTK	788,93	14,6	18,56	18,08
AEFVEVTK	922,03	9,98	20,59	17,46
YLYEIAR	927,06	13,68	19,74	19,7
DLGEEHFK	974,03	9,9	22,02	15,02
LVVSTQTALA	1 002,16	13,8	19,51	20,52
LVNELTEFAK	1 163,32	16,68	19,86	20,82
HLVDEPQNLIK	1 305,48	14,6	23,01	18,57
TVMENFVAFVDK	1 399,61	8,52	20,19	22,48
HPYFYAPELLYYANK	1 439,66	21,43	24,99	30

Without taking details into account, SFC show retention behavior similarly to RPLC. The performance in terms of separation of fragments was also evaluated for SFC using a lesser

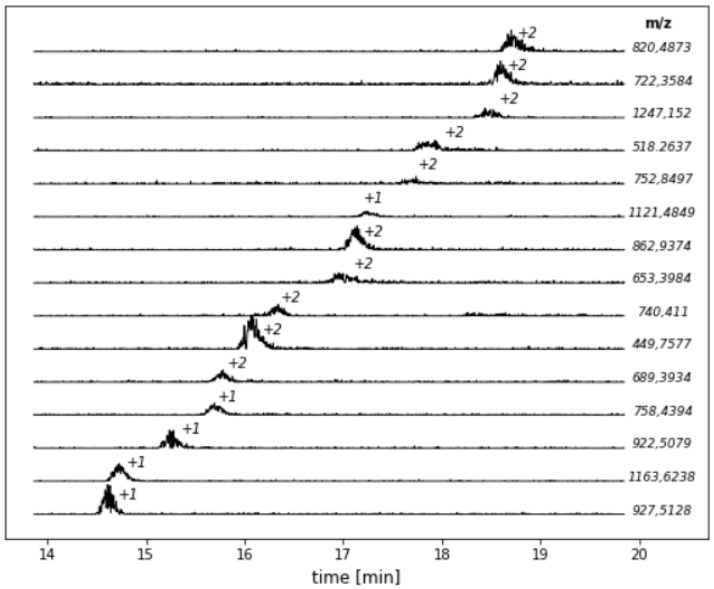

Figure 21. BSA tryptic digest analyzed with SFC. Mass chromatograms are here plotted for each mass-to-charge signal, given on the right. The degree of ionization is denoted for each peak.

gradient slope. Figure 21 depicts the retention of the strongest mass signals identified from the experiment, where selectivity can be considered acceptable. The resulting UV chromatogram from RP-HPLC is given in Figure 22 below.

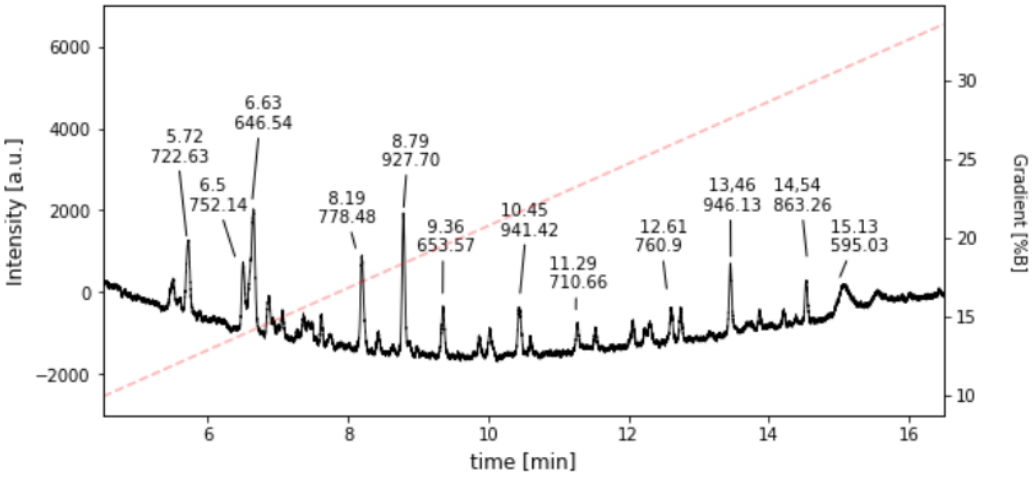

Figure 22. Separation of BSA tryptic digest using HPLC. The UV-chromatogram is annotated with retention time and m/z values for the largest peaks.

6 Conclusion

The experiments conducted on HPLC and SFC have shown that the volume fraction of water in diluents affect peak shape. This possibly has a large impact on separation in preparative scale. Two optimal volume percentages water in t-BuOH and MeOH were experimentally determined for diluents with 2 and 10 µL injections, corresponding to 20-30% water for SFC and 50% for HPLC. Diluent effects were not possible to avoid for SFC with overloaded injections for any of the diluents examined, though the results showed that the spiking of sample with TFA had an interesting impact which could potentially have important implications. The retention behavior of the peptides examined in this work showed that SFC performed well in terms of resolving power. While using SFC the peptides Leuprolide and Triptorelin were able to be well separated when using the XT organo-silane stationary phase ($\alpha \approx 1.19$). This was not possible when using conventional RP-HPLC. Retention behavior of SFC was in comparison with HPLC partly, but not completely, in the inverse order as expected and as concluded in other studies (Cheng et al. 2006, Shao et al. 2016). No correlation was however found between retention in RP-HPLC or SFC and the peptide sequence descriptors that were evaluated. The lack of relationship suggests that other properties rule the chromatographic behavior of peptides. The selectivity was assessed for two Kromasil SFC columns to evaluate any difference in separation performance. There were significant differences with overall better selectivity when using CN stationary phase. This report does also introduce a novel method for determining solubility in a supercritical fluid. Initial experiments suggest that the method is able to determine solubility for numerous substances including the peptide Leu-Enk. The method did provide a significant

increase in solubility with increasing pressure as expected. The solubilities for caffeine and aspirin were found to correlate to solubility data found in literature. Finally, this report presents the separation of a tryptic BSA digest on SFC which has to this date not been published in any scientific paper.

7 Acknowledgements

I would first and foremost like to thank my supervisors Martin Enmark (KAU) and Linda Thunberg (AZ). I am very thankful for all guidance, interesting and motivating talks and for everything that I have learned during this project. I would like to thank Prof. Torgny Fornstedt for welcoming me to the FSSG team and for including me in this collaboration within the KKS Synergy project. Thanks to Jörgen Samuelsson for great thoughts and input regarding the solubility method. Thanks to Annika Weinmann for showing me the workflow of preparative scale separation, for many valuable technical tips, especially for managing mass data in MassLynx. Thank you Hanna Leek for welcoming me to SSL labs and for giving me the opportunity to learn and to get involved with the work process at your division. I would like to thank Tomas Leek for guidance with Q-TOF analysis, articles on BSA tryptic digest separation and interesting talks about column peak capacity and Poppe plots.
I would also like to thank Fredrik Limé at Nouryon for providing this project with Kromasil SFC columns that made all peptide analyses possible. Finally I would like to thank all in Separation Science Laboratory (SSL) team for the great team spirit and the good fika-sessions.

This is a synergistic joint project between the Fundamental Separation Science Group (FSSG) at Karlstad University ad AstraZeneca made possible thanks to the Swedish Knowledge Foundation via the KKS SYNERGY project 2016 “BIO-QC: Quality Control and Purification for New Biological Drugs” (grant number 20170059).

References

2016. Mass Spectrometry Resource. WWW-dokument 2016-01-19: http://www.ionsource.com/. Hämtad 2019-06-02.

Baker DE, Levien TL. 2018. Angiotensin II. Hospital Pharmacy 53: 230–233.

Bjellqvist B, Ek K, Giorgio Righetti P, Gianazza E, Görg A, Westermeier R, Postel W. 1982. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. Journal of Biochemical and Biophysical Methods 6: 317–339.

Burgos-Solórzano GI, Brennecke JF, Stadtherr MA. 2004. Solubility measurements and modeling of molecules of biological and pharmaceutical interest with supercritical CO2. Fluid Phase Equilibria 220: 55–67.

Chen R. A Brief Review of Interfacing Supercritical Fluid Chromatography with Mass Spectrometry. Chromatography Today February/March 2009: 11–13.

Ciftci ON. 2017. Supercritical Fluid Technology: Application to Food Processing. Journal of Food Processing & Technology, doi 10.4172/2157-7110.1000e105.

De Pauw R, Choikhet K, Desmet G, Broeckhoven K. 2014. Occurrence of turbulent flow conditions in supercritical fluid chromatography. Journal of Chromatography A 1361: 277–285.

Enmark M, Glenne E, Leśko M, Langborg Weinmann A, Leek T, Kaczmarski K, Klarqvist M, Samuelsson J, Fornstedt T. 2018. Investigation of robustness for supercritical fluid chromatography separation of peptides: Isocratic vs gradient mode. Journal of Chromatography A 1568: 177–187.

Enmark M, Åsberg D, Shalliker A, Samuelsson J, Fornstedt T. 2015. A closer study of peak distortions in supercritical fluid chromatography as generated by the injection. Journal of Chromatography A 1400: 131–139.

Fekete S, Veuthey J-L, Guillarme D. 2012. New trends in reversed-phase liquid chromatographic separations of therapeutic peptides and proteins: Theory and applications. Journal of Pharmaceutical and Biomedical Analysis 69: 9–27.

Fosgerau K, Hoffmann T. 2015. Peptide therapeutics: current status and future directions. Drug Discovery Today 20: 122–128.

G. de la Torre B, Albericio F. 2019. The Pharmaceutical Industry in 2018. An Analysis of FDA Drug Approvals from the Perspective of Molecules. Molecules 24: 809.

Gahm KH, Huang K, Barnhart WW, Goetzinger W. 2011. Development of supercritical fluid extraction and supercritical fluid chromatography purification methods using rapid solubility screening with multiple solubility chambers. Chirality 23: E65–E73.

Gedela S, Medicherla NR. 2007. Chromatographic Techniques for the Separation of Peptides: Application to Proteomics. Chromatographia 65: 511–518.

Gilar M, Xie H, Jaworski A. 2010. Utility of Retention Prediction Model for Investigation of Peptide Separation Selectivity in Reversed-Phase Liquid Chromatography: Impact of Concentration of Trifluoroacetic Acid, Column Temperature, Gradient Slope and Type of Stationary Phase. Analytical Chemistry 82: 265–275.

Huang Z, Chiew YC, Lu W-D, Kawi S. 2005. Solubility of aspirin in supercritical carbon dioxide/alcohol mixtures. Fluid Phase Equilibria 237: 9–15.

Isidro-Llobet A, Kenworthy MN, Mukherjee S, Kopach ME, Wegner K, Gallou F, Smith AG, Roschangar F. 2019. Sustainability Challenges in Peptide Synthesis and Purification: From R&D to Production. The Journal of Organic Chemistry 84: 4615–4628.

Johannsen M, Brunner G. 1995. Measurements of Solubilities of Xanthines in Supercritical Carbon Dioxide + Methanol. Journal of Chemical & Engineering Data 40: 431–434.

Li B, Guo W, Song W, Ramsey ED. 2016. Interfacing supercritical fluid solubility apparatus with supercritical fluid chromatography operated with and without on-line post-column derivatization: Determining the solubility of caffeine and monensin in supercritical carbon dioxide. The Journal of Supercritical Fluids 115: 17–25.

Li S, Varadarajan GS, Hartland S. 1991. Solubilities of theobromine and caffeine in supercritical carbon dioxide: correlation with density-based models. Fluid Phase Equilibria 68: 263–280.

Lindskog MA, Nelander H, Jonson AC, Halvarsson T. 2014. Delivering the promise of SFC: a case study. Drug Discovery Today 19: 1607–1612.

Mehr CB, Biswal RN, Collins JL, Cochran HD. 1996. Supercritical carbon dioxide extraction of caffeine from guaraná. The Journal of Supercritical Fluids 9: 185–191.

Müller AT, Gabernet G, Hiss JA, Schneider G. 2017. modlAMP: Python for antimicrobial peptides. Bioinformatics (Oxford, England) 33: 2753–2755.

Nguyen LT, Haney EF, Vogel HJ. 2011. The expanding scope of antimicrobial peptide structures and their modes of action. Trends in Biotechnology 29: 464–472.

Peker H, Srinivasan MP, Smith JM, McCoy BJ. 1992. Caffeine extraction rates from coffee beans with supercritical carbon dioxide. AIChE Journal 38: 761–770.

Peschel A, Sahl H-G. 2006. The co-evolution of host cationic antimicrobial peptides and microbial resistance. Nature Reviews Microbiology 4: 529–536.

Reichardt C, Welton T. 2010. Solvents and Solvent Effects in Organic Chemistry: REICHARDT:SOLV.EFF. 4ED O-BK. doi 10.1002/9783527632220.

Saldaña MDA, Mohamed RS, Baer MG, Mazzafera P. 1999. Extraction of Purine Alkaloids from Maté (*Ilex paraguariensis)* Using Supercritical CO 2. Journal of Agricultural and Food Chemistry 47: 3804–3808.

Shao Y, Wang C, Apedo A, Mcconnell O. 2016. Rapid Separation of Five Cyclosporin Analogs by Supercritical Fluid Chromatography. Journal of Analytical Sciences, Methods and Instrumentation 06: 23.

Tognarelli D, Tsukamoto A, Caldwell J, Caldwell W. 2010. Rapid peptide separation by supercritical fluid chromatography. Bioanalysis 2: 5–7.

Vera CM, Shock D, Dennis GR, Samuelsson J, Enmark M, Fornstedt T, Shalliker RA. 2015. Contrasting selectivity between HPLC and SFC using phenyl-type stationary phases: A study on linear polynuclear aromatic hydrocarbons. Microchemical Journal 119: 40–43.

Vorobei AM, Pokrovskiy OI, Ustinovich KB, Parenago OO, Lunin VV. 2019. A method for measuring solubility in multi-component sub- and supercritical fluids using an online hyphenation of supercritical antisolvent precipitation and supercritical fluid chromatography. Journal of Molecular Liquids 280: 212–217.

Vorst GV der, Langenhove HV, Paep FD, Aelterman W, Dingenen J, Dewulf J. 2009. Exergetic life cycle analysis for the selection of chromatographic separation processes in the pharmaceutical industry: preparative HPLC versus preparative SFC. Green Chemistry 11: 1007–1012.

Webster GK. 2014. Supercritical Fluid Chromatography: Advances and Applications in Pharmaceutical Analysis. Pan Stanford, Singapore.

Wilm M. 2011. Principles of Electrospray Ionization. Molecular & Cellular Proteomics 10: M111.009407.

Wong MKS. 2016. Subchapter 29B - Angiotensin II. I: Takei Y, Ando H, Tsutsui K (red.). Handbook of Hormones, s. 258-e29B-4. Academic Press, San Diego.

Zhang G, Ueberheide BM, Waldemarson S, Myung S, Molloy K, Eriksson J, Chait BT, Neubert TA, Fenyö D. 2010. Protein Quantitation Using Mass Spectrometry. I: Fenyö D (red.). Computational Biology, s. 211–222. Humana Press, Totowa, NJ.

Zheng J, Pinkston JD, Zoutendam PH, Taylor LT. 2006. Feasibility of Supercritical Fluid Chromatography/Mass Spectrometry of Polypeptides with Up to 40-Mers. Analytical Chemistry 78: 1535–1545.

Softwares:

Agilent Technologies. Chemstation 3. Agilent. Santa Clara CA, USA.
URL: https://www.agilent.com/en/support/software-informatics/multiinstrumentsoftwarerev

Waters Inc. 20. Empower 3. Waters Corporation, Milford MA, USA.
URL: https://www.waters.com/waters/en_US/Empower-3-Chromatography-Data-Software/nav.htm?cid=513188&locale=en_US

Waters Inc. 2012. MassLynx V4.1. Waters Corporation, Milford MA, USA.
URL: https://www.waters.com/waters/en_US/MassLynx-Mass-Spectrometry-Software-/nav.htm?cid=513164&locale=en_US

U.S. Secretary of Commerce. 2018. Reference Fluid Thermodynamic and Transport Properties. National Institute of Standards and Technology, Gaithersburg MD USA.
URL: https://www.nist.gov/srd/refprop

Sartorius Stedim Data Analytics AB. 2017. Umetrics Simca 15 Multivariate Data Analytics Solution. Sartorius Stedim Biotech, Malmö, Sweden.
URL: https://umetrics.com/products/simca

Appendix A

Here some detailed information about provider and purity of peptides involved in this study together with a complete list of pharmaceutical peptides that were assessed are presented. Descriptors for all peptides were calculated using modlAMP package, however the charge specified at pH 2 was modified for cyclic peptides as charged residues at the N- and C-termini are neutralized when linked together.

Table 7. Information of CAS-number, provider and purity for each peptide.

Peptide	CAS#	Provider	Purity
Angiotensin II	4474-91-3	Bachem AG	≥ 93%
Carbetocin	37025-55-1	Medchemexpress	> 98%
Desmopressin	62288-83-9	Medchemexpress	n/a
Gramicidin A, B, C, D	1405-97-6	Sigma Aldrich	≥ 98%
Leu-Enkephalin	58822-25-6	Bachem AG	98.6%
Leuprolide	53714-56-0	Bachem AG	98.9%
Somatostatin	-	Synthesized on site at AZ	~96%
Triptorelin	57773-63-4	n/a	n/a

Table 8. Complete list of the primary peptide set. The descriptors are calculated using the modlAMP package.

Peptide	Cyclic/ Linear	Sequence	Charge (pH 2)	Hydrophobicity ratio	MW
Angiotensin II	L	DRVYIHPF	2.566	0.375	1046.18
Atosiban	C	YITNCPKG	-7.32E-07	0.25	895.04
Calcitonin	L	CSNLSTCVLGKLSQELHKLQ TYPRTNTGSGTP	3.993	0.25	3434.87
Carbetocin	C	YIQNCPLG	-7.32E-07	0.375	907.05
Deslorelin	L	PHWSYWLRP	2.585	0.11	1241.4

Desmopressin	C	YFQNCPRG	0.999	0.25	984.09
Exendin	L	HGEGTFTSDLSKQMEEEAV RLFIEWLKNGGPSSGAPPPS	4.531	0.256	4187.57
Felypressin	C	CFFQNCPKG	1.585	0.444	1043.22
Glucagon	L	HSQGTFTSDYSKYLDSRRA QDFVQWLMNT	4.528	0.241	3482.76
Gonadorelin	L	PHWSYGLRPG	2.585	0.1	1169.3
Gramicidin-Val-A	L	VGALAVVVWLYLWLW	0.585	0.666	1811.22
Gramicidin-Ile-A	L	IGALAVVVWLWLWLW	0.585	0.666	1825.25
Gramicidin-Val-B	L	VGALAVVVWLWLWLW	0.585	0.733	1772.18
Gramicidin-Ile-B	L	IGALAVVVWLFLWLW	0.585	0.733	1786.21
Gramicidin-Val-C	L	VGALAVVVWLYLWLW	0.585	0.666	1788.18
Gramicidin-Ile-C	L	IGALAVVVWLYLWLW	0.585	0.666	1802.21
Lanreotide	C	WCYDWKVCT	1.566	0.333	1203.39
Leu-Enkephalin	L	YGGFL	0.585	0.4	555.62
Leuprolide	L	PHWSYLLRP	2.585	0.22	1168.35
Liraglutide	L	HAEGTFTSDVSSYLEGQAA KEEFIAWLVRGRG	4.538	0.344	3512.8
Octreotide	C	FCFWKTCT	0.999	0.5	1035.24
Porcine	L	SYSMEHFRWGKPVGKKRRP VKVYPNGAEDELAEAFPLE F	8.531	0.308	4567.16
Somatostatin	C	AGCKNFFWKTFTSC	1.999	0.429	1639.9
Teriparatide	L	SVSEIQLMHNLGKHLNSME RVEWLRKKLQDVHNF	8.545	0.353	4117.73
Triptorelin	L	EHWSYWLRPG	2.578	0.1	1330.45
Vasopressin	C	CYFQNCPRG	0.999	0.333	1087.24

Appendix B

Calibration Curves with Regression Line and R^2

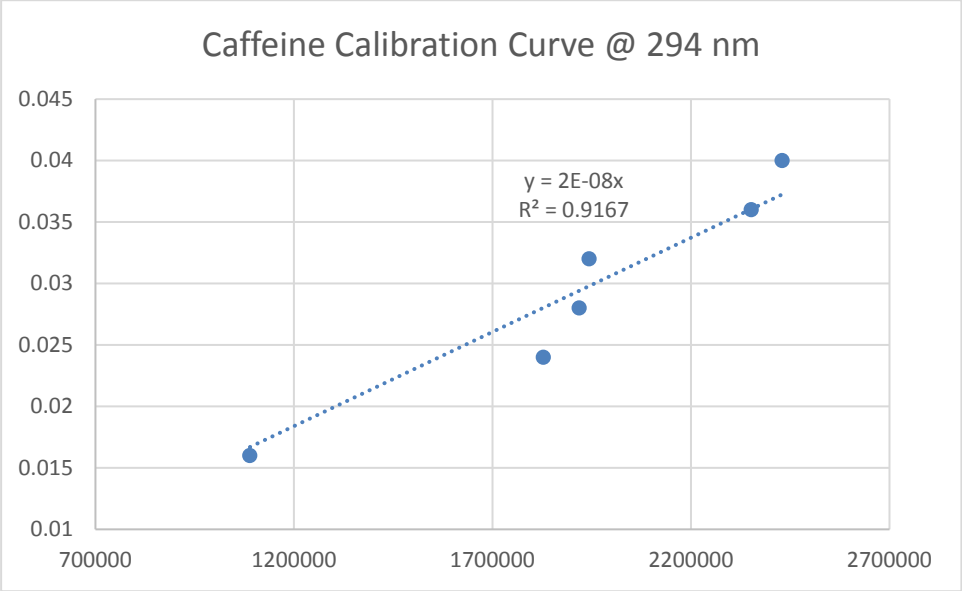

Figure 23. Calibration curve for caffeine with fitted line and the coefficient of determination (R^2). Peak areas are taken at wavelength 290 nm.

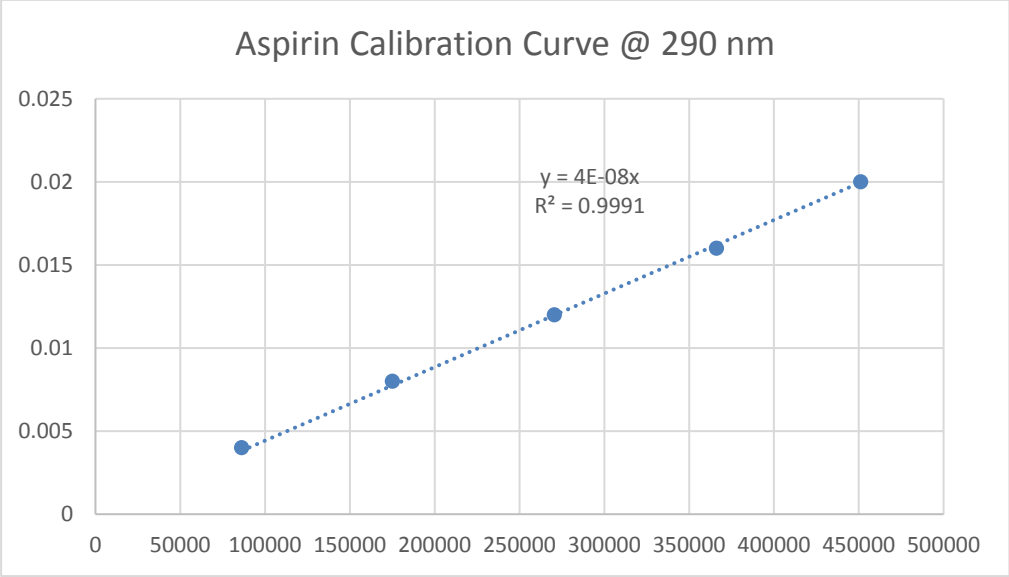

Figure 24. Calibration curve for aspirin with fitted line and the coefficient of determination (R^2). Peak areas are taken at wavelength 290 nm.

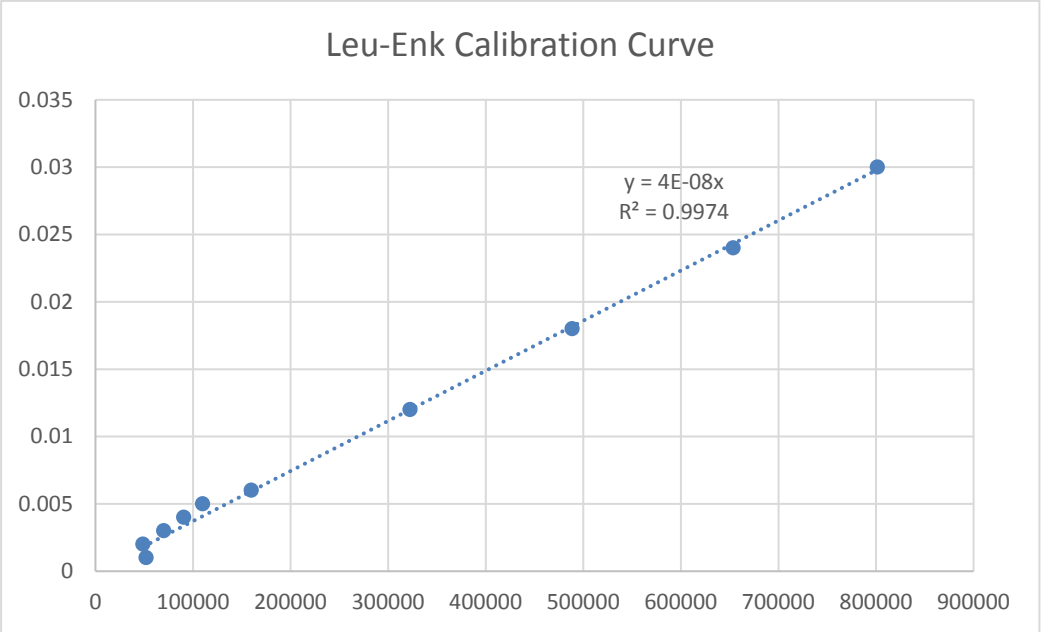

Figure 25. Calibration curve for Leu-Enk with fitted line and the coefficient of determination (R^2).

Appendix C

Details of SFC method designs.

Table 9. Method design for solubility analysis of aspirin and caffeine in Empower. In case of caffeine, volume percentage of phase B equaled to zero.

Time	%B Phase	Valve position (1,2 or 3)
Init	X	1
1.00	X	2
1.01	X	2
11.00	X	1
11.15	X	3
11.17	2	3
15.17	40	3

Table 10. Method design for solubility analysis of peptides in Empower.

Time	%B Phase	Valve position (1,2 or 3)
Init	20	1
1.00	20	2
1.01	20	2
11.00	20	1
11.15	20	3
11.17	20	3
18.17	70	3
20.17	70	3
23.17	20	3

Appendix D

Selectivity and retention data for peptides in SFC using both XT and CN columns. The selectivity was calculated according to Equation 3 given in theory section.

Table 11. Retention factor, *k*, for each peptide in the set. Data is generated from separation using Waters XBridge BEH 4.6x100 mm column.

	MeOH + 20% H_2O	**t-BuOH + 10% H_2O**	**MeOH + 50% H_2O**	**t-BuOH + 50% H_2O**
Angiotensin II	7,55	7,52	7,548	7,524
Carbetocin	10,398	10,39	10,394	10,396
Desmopressin	6,027	6,619	6,62	6,64
Gramicidin	22,17	22,15	-	22,25
Leu-Enkephalin	7,418	7,783	7,82	7,767
Leuprolide	10,174	10,206	10,116	10,176
Somatostatin	-	-	-	-
Triptorelin	9,615	9,618	9,634	9,423

Table 12. Retention factor, *k*, for each peptide in the set. Data is generated from separation using Kromasil XT 3.0x150 mm column.

	MeOH + 20% H_2O	**t-BuOH + 10% H_2O**	**MeOH + 50% H_2O**	**t-BuOH + 50% H_2O**
Angiotensin II	9,653650255	9,78098472	9,653650255	9,752688172
Carbetocin	10,13469157	10,13469157	10,12054329	10,12054329
Desmopressin	12,44086022	12,65308432	12,53989813	12,6672326
Gramicidin	7,941709111	7,941709111	-	7,941709111
Leu-Enkephalin	7,50311262	7,701188455	7,644595359	7,602150538
Leuprolide	9,20090549	9,681946802	9,116015846	9,285795133
Somatostatin	9,950764007	9,880022637	-	-
Triptorelin	11,52122241	11,49292586	11,52122241	11,50707414

Table 13. Retention factor, *k*, for each peptide in the set. Data is generated from separation using Kromasil CN 3.0x150 mm column.

	MeOH + 20% H_2O	t-BuOH + 10% H_2O	MeOH + 50% H_2O	t-BuOH + 50% H_2O
Angiotensin II	9,016977929	9,172608942	9,05942275	9,101867572
Carbetocin	5,465761177	5,465761177	5,451612903	5,451612903
Desmopressin	8,267119411	8,465195246	8,352009055	8,380305603
Gramicidin	6,937181664	6,92303339	-	6,894736842
Leu-Enkephalin	7,24844369	7,460667799	7,460667799	7,404074703
Leuprolide	9,20090549	9,71024335	9,116015846	9,285795133
Somatostatin	-	-	-	-
Triptorelin	9,696095076	9,724391624	9,696095076	9,667798529

Table 14. Selectivity factor for each pair of peptides evaluated in SFC. Separation was performed using Kromasil XT 3.0x150 mm column.

k2 \ k1	Angiotensin II	Carbetocin	Desmopressin	Gramicidin	Leu-Enkephalin	Leuprolide	Somatostatin	Triptorelin
Angiotensin II	1							
Carbetocin	1,036	1,000						
Desmopressin	1,294	1,248	1,000					
Gramicidin	1,232	1,276	1,593	1,000				
Leu-Enkephalin	1,270	1,316	1,643	1,031	1,000			
Leuprolide	1,010	1,047	1,307	1,219	1,257	1,000		
Somatostatin	1,010	1,026	1,281	1,244	1,283	1,020	1,000	
Triptorelin	1,175	1,134	1,101	1,447	1,492	1,187	1,163	1,000

Table 15. Selectivity factor for each pair of peptides evaluated in SFC. Separation was performed on Kromasil CN 3.0x150 mm column.

k2 \ k1	Angiotensin II	Carbetocin	Desmopressin	Gramicidin	Leu-Enkephalin	Leuprolide	Somatostatin	Triptorelin
Angiotensin II	1,000							
Carbetocin	1,678	1,000						
Desmopressin	1,084	1,549	1,000					
Gramicidin	1,325	1,267	1,223	1,000				
Leu-Enkephalin	1,229	1,365	1,135	1,078	1,000			
Leuprolide	1,059	1,777	1,147	1,403	1,302	1,000		
Somatostatin	0,000	0,000	0,000	0,000	0,000	0,000	1,000	
Triptorelin	1,060	1,779	1,149	1,405	1,303	1,001	0,000	1,000