Chromatographic Behavior of Peptides Containing Oxidized Methionine in Reversed-phase Chromatography: Application to Cyclolinopeptides in Flaxseed Oil and Linear Tryptic Peptides

By

Ying W. Lao

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Manitoba
Winnipeg

ABSTRACT

The thesis consists of two parts targeting the characterization of chromatographic behavior of linear tryptic and cyclic peptides containing oxidized methionine (Met) in reversed-phased chromatography. The retention order of methionine-containing peptide analogues was observed to be the same in both studies: Met oxide < Met dioxide < Met. For linear tryptic peptides, the magnitude of the retention time shift may vary dramatically: from –9 % to +0.36 % acetonitrile. Particularly, large negative retention time shifts are found mostly associated with methionine being in the hydrophobic face of an amphipathic helix. Contrary to previously reported observations, I demonstrate for the first time that methionine oxidation may increase peptide hydrophobicity, this occurs only when methionine is in the N3 position of the N-capping stabilization motif preceding an amphipathic helix. In the second study, the effect of peak splitting was observed for some Met oxide-containing cyclolinopeptides, which most likely appear due to diastereomerization.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Oleg Krokhin, from the Department of Internal Medicine, for his patient guidance, generosity and encouragement as well as providing me with all necessary facilities throughout the study. I am thankful to the committee members, Dr. Hélène Perreault and Dr. Joe O'Neil, from the Department of Chemistry, and Dr. Michel Aliani, from the Department of Human Nutritional Sciences, for their professional suggestions on the thesis. I am grateful to Dr. Kenneth Standing, from the Department of Physics and Astronomy, for providing lab space and equipment, as well as extended thank you to all staff in the lab for technical support. In particular, I appreciate Mr. Vic Spicer's technical support on processing large proteomic data, and occasional help on preparing manuscripts. Last, but not least, I would like to thank Dr. Lynda Donald, for helping me with thesis proofreading and editing.

I also would like to thank ShapeFoods Inc. (Brandon, Manitoba) for the collaboration opportunity and Natural Sciences and Engineering Research Council of Canada for providing financial support.

In addition, I would like to thank my family and friends for continuous encouragement, love and support.

FOREWORD

The thesis is a detailed investigation of the effect of methionine (Met) oxidation on peptide retention time in reversed-phase chromatography (RPC). Chapter 2 describes the chromatographic behavior of linear tryptic peptides containing oxidized methionine. All the data presented in this chapter were from part of a large study of bacterial and fungal proteomes. I reanalyzed the data collected, extracted pairs of Met/Mso-containing peptides and did further experiments with synthetic peptides. On the other hand, Chapter 3 is concerned with the onset of bitter taste in flaxseed oil, which is associated with methionine oxidation in major cyclolinopeptides. I developed a fast and reliable analytical method based on an HPLC procedure with UV detection and a phenyl-hexyl column for determination of the oxidation degree of major cyclolinopeptides. This will be used for an industrial application for quality control of flaxseed oil.

Chapter 2 will be a part of the manuscript, which is under preparation for submission to Analytical Chemistry: Lao, Y.W., Shamshurin, D., Spicer, V., Krokhin, O.V., Chromatographic behavior of peptides containing oxidized methionine in reversed-phase chromatography. In this work protein digestion was performed by Dmitriy Shamshurin; 2D LC-MS analysis of the resulting digest - by Oleg Krokhin at the Manitoba Centre for Proteomics and Systems Biology. Oxidation of synthetic peptides and tryptic digests using hydrogen peroxide was conducted by Ying W. Lao at Department of Physics & Astronomy. All data analysis was done by Ying W. Lao and Oleg Krokhin. Ying W. Lao contributed a significant part to the manuscript writing, which was edited by Oleg Krokhin.

Chapter 3 data was published as: Lao, Y.W., Mackenzie, K., Vincent, W., Krokhin, O.V., Characterization and complete separation of major cyclolinopeptides in flaxseed oil by reversed-phase chromatography. Journal of Separation Science 37 (2014) 1788. In this work, the production of flaxseed oil as well as accelerated aging of oil samples were performed by Kim Mackenzie and William Vincent from Shapefoods Inc. (Brandon, Manitoba). All instrumental work (peptide extraction, RP HPLC, MALDI MS) and data analysis were conducted by Ying W. Lao at the Department of Physics & Astronomy. Ying W. Lao wrote the manuscript, which was edited by Oleg Krokhin.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF COPYRIGHT CLEARANCES FOR ITEMS IN THE THESIS

Figure 1.1: Structure of an α-helix, and a β-sheet.
Copyright Sinauer Associates, Inc. Used with permission. p. 2

Figure 1.2: Intrinsic N-capping preference for all 20 amino acids, according to Richardson…
Image credit: Oleg Krokhin and Vic Spicer. Manitoba Centre for Proteomics and Systems Biology, Winnipeg, MB. Used with Permission. p. 4

Figure 1.3: Structure of a capping box located at the N-terminal end.
Image credit: Thomas E. Sladewski, Department of Molecular Physiology and Biophysics, University of Vermont. Proteopedia.org. Used with permission. p. 4

Table 1.2: Accuracy of retention time prediction models when applied to a collection of 3330…
Copyright Elsevier. Used with permission. p. 26

Chapter 3: Characterization and Complete Separation of Major Cyclolinopeptides in Flaxseed Oil by Reversed-phase Chromatography
Published as Lao, Y.W., Mackenzie, K., Vincent, W., Krokhin, O.V. (2014). Journal of Separation Science 37, 1788-1796. Copyright John Wiley and Sons. Used with permission from publisher and co-authors. pp. 57-88

LIST OF ABBREVIATIONS

ALA	α-Linolenic acid
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartic acid
CD	Circular dichroism
CID	Collision-induced dissociation
COFRADIC	Combined fractional diagonal chromatography
DNA	Deoxyribonucleic acid
ESI MS	Electrospray ionization mass spectrometry
FASP	Filter aided sample preparation
Glu (E)	Glutamic acid
HI	Hydrophobicity index
His (H)	Histidine
Ile (I)	Isoleucine
LA	Linoleic acid
LC-MS	Liquid chromatography – mass spectrometry
Lys (K)	Lysine
MALDI MS	Matrix-assisted laser desorption/ionization mass spectrometry
Met (M)	Methionine
MRM	Multiple reaction monitoring
MS/MS	Tandem mass spectrometry
Msn	Methionine *S*,*S*-dioxide
Mso	Methionine *S*-oxide

NMR	Nuclear magnetic resonance
Pro (P)	Proline
PTMs	Post-translational modifications
PUFAs	Polyunsaturated fatty acids
QqTOF	Quadrupole/time-of-flight
RNA	Ribonucleic acid
ROS	Reactive oxidative species
RP HPLC	Reversed-phase high performance liquid chromatography
RPC	Reversed-phase chromatography
Ser (S)	Serine
SRM	Selected reaction monitoring
SSRCalc	Sequence-specific retention calculator
SWATH	Sequential window acquisition of all theoretical mass spectra
TFA	Trifluoroacetic acid
Thr (T)	Threonine
Val (V)	Valine
VWD	Variable wavelength detector
UV	Ultraviolet
XIC	Extracted ion chromatogram

Chapter 1

Introduction

My thesis describes the application of modern LC-MS methodology to study the effect of methionine (Met) oxidation on peptide retention time in reversed-phase chromatography (RPC). The hybrid liquid chromatography – mass spectrometry (LC-MS) technique is based on the separation (analysis) of peptides by two independent techniques: reversed-phase high performance liquid chromatography (RP HPLC) and mass spectrometry (MS). Therefore, the introductory section describes the basic principles of both techniques and their application in proteomics. The majority of the topics covered in my thesis deal with the mechanism of peptide separation, the prediction of peptide retention times and the effect of a peptide's secondary structure on its interaction with the hydrophobic C18 phase. Therefore, this introduction also includes a basic discussion of protein/peptide structure, different approaches to predict peptide retention time in RPC and a comparison of the different models developed to date.

1.1 Basics of Protein Structure

Proteins are macromolecules, made of long chains of amino acids that fold into three-dimensional structures. The structural features of proteins are usually described at four different levels – primary, secondary, tertiary and quaternary structures. Primary structure refers to the linear amino acid sequence of the protein, which is determined by the DNA sequence. Adjacent

amino acids are linked together by peptide bonds, where water is removed during the condensation reaction. Secondary structure refers to the shape of a folded polypeptide chain due to repeated hydrogen bonding between its backbone amide and carbonyl groups, such as in α-helices and β-sheets (Figure 1.1). In an α-helix, the polypeptide is coiled into a helix (3.6 amino acids per turn) and held in place by hydrogen bonds between the H of amide groups and the O of carbonyl groups of every fourth amino acid along the chain. The amino acid side chains are directed outward from the helix axis. On the other hand, a β-sheet forms when polypeptide chains are arranged side by side with well-aligned hydrogen bonds between chains. Tertiary structure describes the overall specific three-dimensional shape of an entire protein molecule. This involves hydrophobic and hydrophilic interactions and cross links (such as disulfide bonds and hydrogen bonding) between different parts of the peptide chain. Quaternary structure is the combination of two or more tertiary units. Correct folding of a protein into its unique three-dimensional structure is crucial to its function in cells. [1]

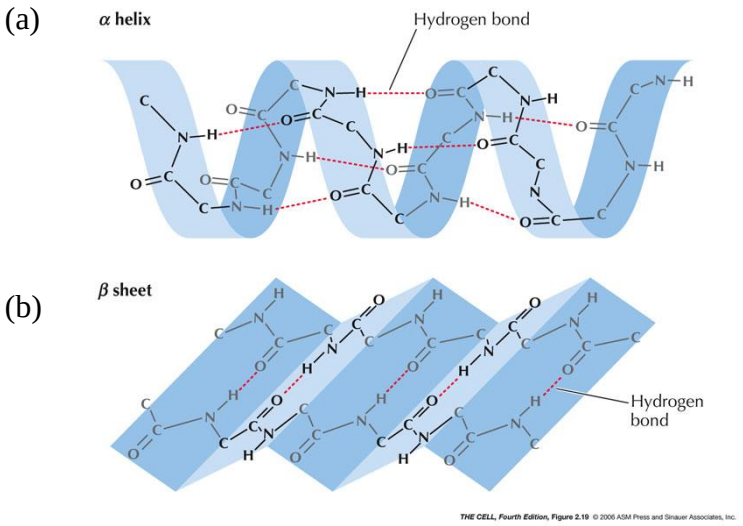


Figure 1.1 Structure of **(a)** an α-helix, and **(b)** a β-sheet. Red dashed lines represent hydrogen bonding interactions. Figure reproduced with permission from Sinauer Associates, Inc.

1.1.1 N-cap Residues in Helices

The α-helices are known as the most prevalent type of secondary structures in proteins [2]. However, only specific segments of a protein sequence exist in α-helical conformation. Therefore, residues in the first and last turn of the helices lack their respective hydrogen bonding partners. This led to the original hypothesis of Presta and Rose [3] that the signal sequences to "start" and "end" a helical stretch must be encoded in the primary sequence. The same year Richardson and Richardson [4] provided confirmation of this hypothesis. They analyzed the X-ray crystallography data of the structures of 215 helical sequences and determined amino acid preferences at the helical ends (Figure 1.2). They defined N-capping residues as boundary residues in the first turn of the helix which are capable of accepting hydrogen bonds from functional groups in amino acid side chains in order to stabilize the helical structures. The nomenclature for N-cap was proposed as follows: N''-N'-N_{cap}-Nl-N2-N3-N4; where N1-N4 belong to the helix and N-cap is located at the N-terminal end preceding the helix. Intrinsic preference of each amino acid at helical ends was confirmed using a large data set of known protein structures [5] and confirmed by extensive studies on synthetic peptides using circular dichroism (CD) spectroscopy [6]. It was established that asparagine (Asn, N), serine (Ser, S), aspartic acid (Asp, D), glycine (Gly, G) and threonine (Thr, T) have the highest N-cap preference (Figure 1.2), and that proline (Pro, P) is strongly favored at the N1 position. One of the most dominant N-capping motifs, the "capping box", features reciprocal hydrogen bonding interactions between the –NH backbone group of the N-cap residue (Asn, Ser, Thr, Asp) and the side chain of the N3 (glutamic acid, Glu, E) residue, and vice versa (Figure 1.3). The preference for hydrophobic residues is highest at the N4 position: the hydrophobic interaction between the

side-chains of the residues separated by three or four amino acids also contributes to helix stabilization [7].

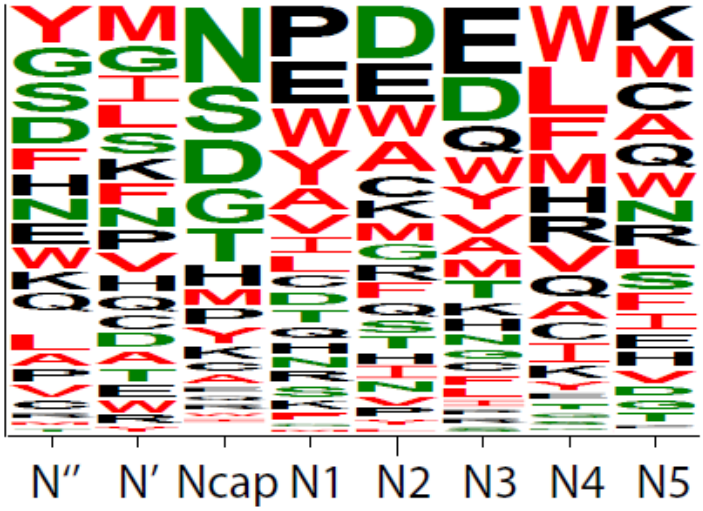
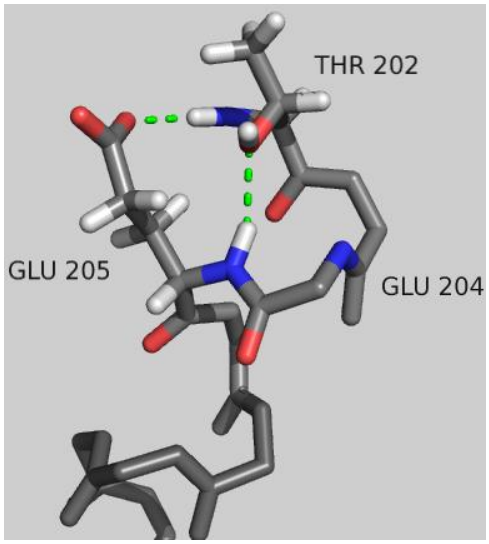


Figure 1.2 Intrinsic N-capping preference for all 20 amino acids (all are abbreviated using the one-letter designation), according to Richardson and Richardson [4]. Asn (N), Ser (S), Asp (D), Gly (G) and Thr (T) that were determined to be the preferred N-capping residues, are shown in green. Hydrophobic residues are shown in red and all others are in black. The height of the letters is proportional to the preference for that amino acid at that location near the N-capping residue of 215 helices used in that study. Figure was prepared by O. Krokhin and V. Spicer and used with permission.

Figure 1.3 Structure of a capping box located at the N-terminal end. The side chains of the N-cap (Thr) and N3 (Glu) residues are interacting with the –NH backbones from N3 and N-cap residues, respectively. The "capping box" hydrogen bonds are shown as green dashed lines. Image was kindly provided by Thomas E Sladewski. Image was created using atomic coordinates of Protein Data Bank accession no. 2HF4 [8]. Used with permission.



1.1.2 Helical Wheel Representation

The helical wheel projection is a simple way of determining the position of amino acids in an α-helical structure for any protein segment. Basically, a sequence of amino acids is plotted in a rotating manner down the axis of a helical wheel, with 3.6 residues per turn creating an arc of 100˚ between adjacent residues (Figure 1.4). Furthermore, the helical wheel can be used to predict whether a sequence of amino acids might form an amphiphatic α-helix, which is a special type of α-helix that contains hydrophilic residues almost exclusively on one side and hydrophobic residues on the opposite side along the axis of the helix. Figure 1.4 (a, b) shows two peptide sequences that are amphiphatic and non-amphiphatic, respectively. The hydrophobic residues are shown in red, thus permitting simple visual assignment of amphipathicity.

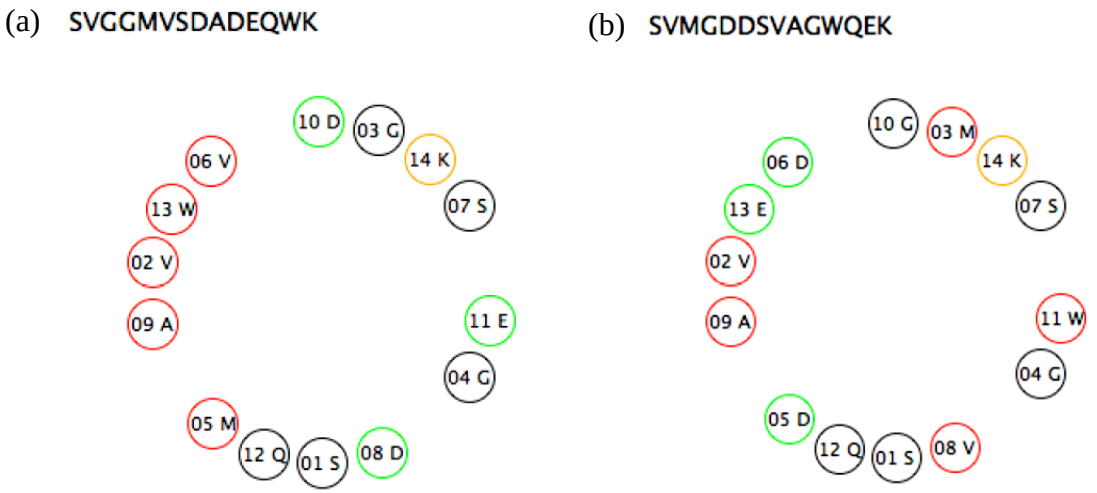


Figure 1.4 Helical wheel projection of **(a)** SVGGMVSDADEQWK and **(b)** SVMGDDSVAGWQEK. This displays the sequence in a helical representation as if looking down the axis of the helix. All amino acids are colored coded: the acidic residues are green, basic residues are yellow, hydrophobic residues are red and all others are black. The peptide **(a)** is an example of an amphiphatic helix, as all hydrophobic residues are located on one side of the helix, and the other side is dominated by hydrophilic residues. In contrast, the second peptide **(b)** would be an example of a non-amphiphatic helix, because its hydrophobic residues are located randomly in the helical wheel projection.

1.1.3 The Hydrophobic Moment

The hydrophobic moment (μ), introduced by Eisenberg et al. [9], can be used to quantify the degree of amphiphilicity of an α-helix. In this method, a given protein segment of n residues is projected in a helical wheel so that the side chains from consecutive amino acids emerge from the backbone at an angle $\delta = 100^{\circ}$ (assuming it is an α-helix). A vector of size proportional to the hydrophobicity of each corresponding amino acid (H_i) is drawn from the center of the helix and the hydrophobic moment calculated using the following equation [10]:

$$\mu = \left\{ \left(\sum_{i}^{i+n-1} H_i \sin \delta \right)^2 + \left(\sum_{i}^{i+n-1} H_i \cos \delta \right)^2 \right\}^{1/2}$$

It is important to realize that the hydrophobic moment computation does not predict the presence or absence of a given structural element. Rather, it is a numerical value that is indicative of peptide amphipathicity. In practice, a cutoff value can be derived from known amphipathic helices. At the same time, the hydrophobic moment value is determined by the hydrophobicity of individual amino acids (H_i), which can be determined by various methods and may vary in different systems. Hydrophobic moment also does not take into account the propensity of different residues to assume specific angles between chemical bonds (the characteristic of a helical structure).

1.1.4 Experimental Measurement of Helicity Using Circular Dichroism Spectroscopy

Circular Dichroism (CD) is a spectroscopic technique widely used for the study of the secondary structure of proteins and the conformation of peptides. CD uses a source of circularly

polarized light, in which the electromagnetic light vector oscillates rotationally to the right or to the left, forming a helix around the axis of propagation. If a molecule is optically active, it absorbs left- and right-handed circularly polarized light to different extents. The difference in absorption (ΔA) of the left- and right-components (A_L and A_R) can be measured and it obeys the Beer-Lambert Law:

$$\Delta A = A_L - A_R = \varepsilon_L \times L \times C - \varepsilon_R \times L \times C = \Delta\varepsilon \times L \times C$$

where ε is the molar extinction coefficient in $M^{-1}cm^{-1}$, L is the optical path length in cm and C is the concentration in M. It is worth noting that a CD signal can be positive or negative, depending on whether left-handed circularly polarized light is absorbed to a greater extent than right-handed circularly polarized light (CD signal positive) or to a lesser extent (CD signal negative).

CD probes the secondary structure of proteins because the peptide bond is asymmetric and molecules without a plane of symmetry show the phenomenon of circular dichroism. Right-handed α-helical elements in proteins typically have a circular dichroism band with positive rotational strength at 190 nm and bands with negative rotational strengths near 208 and 222 nm, whereas β-sheets have a positive band near 195 nm and a single negative band at 215 nm. Random coils have a very different spectrum, with a negative band of great magnitude at around 200 nm. [11] CD measurements were used extensively for studies of peptide helicity in solution [6]. Subsequently, similar measurements were applied to understand the connection between peptide amphipathic helicity and increased chromatographic retention in RP systems. Blondelle et al. [12] performed CD measurements on a series of synthetic peptides absorbed on silica slides

with chemically attached aliphatic groups. They found a very good correlation between peptide helicity on these surfaces and peptide behavior in chromatographic systems.

1.2 Proteomics: Techniques and Challenges in Protein Identification and Characterization

Proteomics can be defined as a combination of tools to study all the proteins expressed in a particular organism (proteome). Major tasks in proteomic experiments include identification of all proteins in complex mixtures, along with their post-translational modifications (PTMs), detection of interacting partners as well as protein quantitation across different samples. One of the most popular methods in modern protein's analytical chemistry, particularly for protein identification and peptide sequencing, is referred to as bottom-up proteomics. In general, this workflow starts with protein digestion by a proteolytic enzyme (split a given protein into smaller pieces, called peptides), separation of the peptide mixture by HPLC with MS detection and finally peptide fragmentation and identification by tandem mass spectrometry (MS/MS).

1.2.1 Reversed-phased Chromatography

In proteomics, reversed-phased chromatography (RPC) is often considered the method of choice for peptide separation due to the unique aspects of the peptide structures and the compatibility of buffers with MS detection systems. The mechanism of RP chromatographic separation is based on differences in the hydrophobicity of the analytes, which provides distinct selectivity different from other modes of separation based on size or charge. In RPC, retention of peptides strongly depends on the concentration of the organic solvent and gradient elution is

usually required. The principle of gradient elution is based on varying the composition of the mobile phase from low to high eluting strength. First, a polar mobile phase with low elution strength (such as water) is employed; this allows the injected peptides to be absorbed onto the column through hydrophobic binding. Slowly, the concentration of a solvent with high elution strength (such as methanol and acetonitrile, also known as an organic modifier) is increased to permit desorption and elution of peptides from the column. Note that any peptide with a specific sequence only elutes when the organic modifier rises to a critical concentration. The retention order in RPC is generalized as follows: solutes elute in increasing order of hydrophobicity or decreasing order of net charge, degree of ionization and ability to participate in hydrogen bonding.

In RPC, the stationary phase is made with porous spherical silica beads (2-5 micron), the surface of which is covered with –Si-OH silanol groups. These silanol groups serve as attachment sites for hydrophobic ligands (such as C4, phenyl, C8, and C18). The traditional types of resin are made porous to increase surface area available for binding. Most interactions between the analyte and the particle occur at the hydrophobic surface of the particle. Smaller particles generally provide more surface area and better separations but at the expense of higher backpressure, which is not compatible with most conventional HPLC equipment. Typical column dimensions in analytical RPC are 2.1-4.6 mm inner diameter and 30-250 mm length. The column sizes in typical proteomic RPC applications are much smaller. The size of the column is determined by the necessity to achieve maximal absolute sensitivity. Most of the proteomic samples are expensive to prepare and limited in quantity. Therefore detection at attomole and femtomole levels is required. This can be achieved only using extremely low flow rates.

Practically, with the current minimal size of tubings (~20 micron), the reasonable flow rate ranges between 200-500 nL/min. The size of the column should be scaled respectively to provide reasonable “dead” time for the chromatographic system. Therefore, typical column dimensions in proteomic applications range between 75 and 150 micron inner diameter and 5-40 cm in length.

The resolving power of RPC is remarkable. Peptides with similar chemical structures can be resolved using relatively short gradients. C8 and C18 ligands are the most commonly used for peptide analyses and usually the best initial choice for method development; however, little difference in selectivity between C18 and C8 is observed. Unlike alkyl-silica columns, columns with aromatic functionality can provide additional selectivity due to pi-pi interactions (non-covalent interaction that involves direct attraction between arene rings) for peptides containing a different number of aromatic amino acids. The main energetic contribution to these pi-pi interactions are Van der Waals dispersion and electrostatics.

1.2.1.1 Ion Pairing Agents in Reversed-phase Chromatography

As mentioned earlier, silanols are the binding site for alkyl groups to form the support matrix. Bulky groups such as C18 may prevent all of the exposed silanols from reacting due to steric hindrance, and thus unreacted silanols with acidic properties can induce unwanted interactions with basic residues in peptides. One way to minimize or eliminate the interactions of peptides with silanols is to operate at low pH (<4) where the silanols are not charged. When separation is performed under acidic conditions (addition of formic, acetic or trifluoroacetic acid (TFA)), basic amino acids (lysine (Lys), arginine (Arg), histidine (His)) are positively charged

and very hydrophilic. Counter-anions from acidic modifiers (formate, acetate, or trifluoroacetate) form ion-pairs with the ionic groups of a peptide and sometimes enhance its binding to the hydrophobic stationary phase due to formation of hydrophobic ion-pairs. Due to the different hydrophobicity of ion-pairing modifiers, they increase peptide retention in different manners. Clearly, the use of ion pairing agents can alter the ionic or hydrophobic characteristics of the solute. With regard to the ion-pair chromatography, two models of ion-pairing have been proposed:

Model 1: mobile phase ion-pairing formation: first the ion-pairing agent binds to the charged solute in the mobile phase to form a neutral ion-pair which subsequently binds to the reversed-phase matrix because of its increased hydrophobicity.

Model 2: dynamic ion exchange: first, the ion-pairing agent is proposed to be absorbed onto the surface of the stationary phase through hydrophobic interaction. This results in the ionic portion of the ion-pairing agent extending into the mobile phase and serving as a point for possible ion-exchange interaction with the charged solute.

It is conceivable that both models can be used to explain influences of ion-pairing agents on the retention of charged solutes. In fact, the ion-pairing interaction may be a combination of both mechanisms.

1.2.1.2 Conformation of Peptides in Reversed-phase Chromatography

Peptides tend to form different secondary structures (i.e. α-helices and β-sheets) to maximize hydrogen bonding in order to obtain the most stable conformation. The specific conformation adopted by a given peptide depends not only on its primary sequence but also its

surrounding environment. For instance, a non-polar environment (such as the hydrophobic stationary phase in RPC) can induce α-helical conformation in potential helical peptides due to binding of preferred hydrophobic domains to the stationary phase [12]. Thus, the RPC column is often found as a reasonable mimic for the natural environment of membranes when studying peptide/protein interactions with lipids. This phenomenon often occurs with amphiphatic α-helices. It was suggested [13] that the hydrophilic and hydrophobic moieties of such molecules have strong affinities for polar and non-polar environments, respectively. It is worth noting that, in general, conformations adopted by peptides in RPC can be broadly divided into two groups: amphiphatic α-helices and random coils. The presence of the preferred binding domain (to the hydrophobic stationary phase), such as the hydrophobic face of an amphipathic helix, leads to increased retention of the peptide in RPC. As mentioned above, Blondelle et al. [12] showed a correlation between the peptide helicity and increased retention time compared to the predicted value. The same group used synthetic amphipathic peptides with sequential substitution of all residues in 18 amino acids – long peptides with methionine [14]. After methionine oxidation, peptides containing methionine inside the hydrophobic face of the helices showed dramatically reduced retention times, suggesting the existence of a preferred binding domain to the hydrophobic stationary phase in RPC. Finally Sereda et al. [15] showed that disruption of the helical configuration results from substitution of D-amino acids in amphipathic helical peptides.

1.2.2 Principles of Mass spectrometry and Tandem Mass Spectrometry

Mass spectrometry is one of the leading methods for protein characterization. Mass spectrometers are often coupled with RP-LC to allow for analysis of complex protein/peptide

mixtures. The basic principle of mass spectrometry is to first generate ions upon ionization of the analytes and subsequently separate these ions in a mass analyzer by their mass-to-charge ratio (*m/z*). Accurate mass measurement of the digested proteins can allow one to identify unknown proteins by high-accuracy peptide mass mapping and database searching. This quick, simple and cheap method is known as peptide mass fingerprinting.

However, peptide mass fingerprinting does not allow for unequivocal identification of the proteins (especially in complex mixtures) as it does not provide information regarding the amino acid sequence of the peptide. Tandem mass spectrometry (MS/MS) can provide such information and has been used extensively for peptide sequencing in proteomic studies. There are two stages of mass analysis during a typical MS/MS experiment: first the mass analyzer is used to isolate the precursor ion, which then undergoes fragmentation via collision with an inert gas, such as helium, nitrogen or argon (this process is known as collision-induced dissociation (CID)), to yield daughter ions; the second analyzer is used to separate these daughter ions according to their masses. The accepted nomenclature of the daughter ions generated by CID was first established in the 1980s [16]. N- and C-terminal fragments are named b- and y-ions, respectively. The designations correspond to cleavage at the amide bond. The subscript number, n (in b_n or y_n) indicates the residue number relative to the terminus that the fragment contains. [16]

The calculation of mass difference between peaks in the fragmentation spectrum (it is also a mass difference between consecutive daughter ions belonging to the same ion series), allows the determination of peptide's primary sequence. This approach is called *de novo* sequencing. It is particularly important for protein identification in organisms with unsequenced genomes. However, *de novo* sequencing can be very time-consuming and almost impossible for

analysis of a complex proteome. Alternatively, large-scale peptide sequencing can be performed through database searching – an experimental fragmentation spectrum would be searched against a database of predicted fragmentation spectra from the genome. For instance, X!Tandem is one of the most commonly used search engines that use MS/MS fragmentation data to identify protein from primary sequence databases. A unique peptide "hit" is found when the similarities between the predicted and observed fragmentation match to a significantly greater degree than any other peptide candidates. The overall scheme of operation of mass-spectrometers in modern proteomics experiments include: mass measurement of peptides in MS mode; selection of parent ions for fragmentation; MS/MS fragmentation and mass measurement of fragments. Depending on quality and principle of operation for different mass spectrometers, the time required to perform these operations (duty cycle) and number of MS/MS in this cycle can be different. As an example, the TripleTOF 5600 (ABSciex) mass spectrometer used in my work is capable of performing one MS/MS acquisition in 50-100 ms. A typical duty cycle includes one 250 ms MS acquisition and up to 20 MS/MS acquisitions (100 ms each). Therefore the total length of each duty cycle can take up to 2.25 s. The output information from these acquisitions contains data on the *m/z* of the parent ion, its charge, and the *m/z* of daughter ions. Each MS/MS entry contains information on time of acquisition, which is equivalent to the retention time – a measure of peptides' hydrophobicity in RP HPLC systems.

1.3 Post-translational Modification of Proteins

Potentially, any protein in a proteome can be post-translationally modified. A post-translational modification (PTM) of a protein is a chemical modification, which occurs after the

protein has been synthesized by translation of the messenger RNA and results in changing the size and structure of a protein [17]. These modifications include phosphorylation, oxidation, glycosylation, ubiquitylation, nitrosylation, methylation, acetylation, lipidation, proteolysis [18] and others. Although the biological significance of many of these modifications remains difficult to predict, there is ample evidence showing that protein properties including but not limited to protein stability [19], protein function [20,21], and protein localization [22] can all be affected by such modifications. Therefore, monitoring *in vivo* PTM has been an essential part of modern systems biology studies. Often, PTMs are envisioned as a change made to the structure of an amino acid side chain, which in turn produces a corresponding change in the molecular weight. This is the basis of detection and characterization of PTMs by mass spectrometry.

Over the past decade, mass spectrometry instruments have significantly improved to provide higher mass accuracy and resolution to enable large-scale analysis of PTMs as well as continuous discovery of novel PTMs. Presently, about 500 modifications are known and reported [23]. Despite these advances in mass spectrometry technology, there are problems pertinent to PTMs in complex proteomic analyses that have yet to be resolved. Despite bottom-up proteomics being the most widely used technique in peptide/protein identification, the main obstacle encountered in large-scale bottom-up proteomics is an inadequate depth of analysis due to the wide dynamic range of protein concentrations [24] and presence of post-translationally/ chemically modified peptides. Chemical instability of the analytes also complicates targeted (i.e. SRM, selected reaction monitoring) and data independent (i.e. SWATH, sequential window acquisition of all theoretical mass spectra) quantitative techniques. SRM is a data-dependent acquisition method that utilizes a triple quadrupole instrument, where a precursor ion of interest

is selected initially for MS/MS fragmentation and subsequently, a daughter ion from the fragmentation reaction of the precursor ion is selected for further monitoring and quantification. Thus, the peptides with potential modification sites are typically excluded from SRM assays [25]. On the other hand, SWATH is a data-independent acquisition method that attempts to collect MS/MS spectra for all detectable analytes in a sample; data acquisition cycles repeatedly through a number of consecutive predefined-precursor mass ranges (also refers to as a SWATH window). The assignment of a potentially modified peptide in SWATH can be ambiguous as it is often affected by the modified counterpart which usually lands in the same SWATH window [26]. In this case, groups of non-shared fragment ions from modified and unmodified peptides would have to be extracted in order to determine which fragment ions correspond to which precursor ions. Introduction of a peptide retention time prediction for modified peptides could eliminate these complications. To date, there are no retention time prediction models which can consistently deal with the large number of post-translationally modified peptides observed in proteomic experiments.

In particular, there is a growing interest of applying structure-based peptide RP-LC retention time prediction as a secondary constraint in peptide/protein identification (including PTM carrying peptides) [27]. Peptide retention time can be considered as a structure-dependent parameter that is constant for given chromatographic separation conditions (mobile phase composition, stationary phase, temperature, pH). Typically, falsely identified peptides can be filtered out based on large discrepancies between observed and predicted retention times on a correlation plot. The literature [28] indicates that retention time prediction for a given peptide could help to improve the confidence of peptide identification based solely on peptide

fragmentation by MS/MS. Furthermore, predicting peptide retention time is also of great interest to targeted proteomics. For instance, multiple reaction monitoring (MRM) is a method for targeted quantitative measurement of proteins using mass spectrometry. Typically, if the retention time of a specific peptide of interest is known, acquisition can be scheduled to monitor continuously the specific ion in order to achieve high sensitivity.

1.4 Peptide Retention Time Prediction in Reversed-phase Chromatography

Understanding and predicting peptide retention times in RP chromatography has a long history: The first publication on this subject appeared in 1980 [29]. At the time, it was assumed that retention time of any peptide is proportional to its hydrophobicity and determined by the hydrophobicity of all its residues. Thus, the first basic model for peptide retention time prediction was known as the additive model, in which the retention time for a given peptide was equal to the sum of the hydrophobicity (retention) coefficients of each residue and terminal group. Two different RP HPLC approaches were used to derive hydrophobicity coefficients of amino acids. The first approach utilizes sets of designed synthetic peptides with substitution at a particular position; the second uses collections of retention time data for random sets of peptides. The first hydrophobicity scales using synthetic peptides were reported by Guo et al. [30] in 1986. The authors used Ac-Gly-X-X-Leu-Leu-Leu-Lys-Lys-amide peptides substituted in the X position with all 20 amino acids to derive retention values. In 2006, Kovacs et al. [31] reported one of the best hydrophobicity scales derived from a set of twenty synthetic peptides using the sequence motif Ac-XGAKGAGVGL-Amide (where X is the target amino acid residue and Ac refers to an acetylated N-terminus). The peptide was designed to eliminate any secondary

structure and nearest neighbor effects on the intrinsic hydrophobicity of the substituted residue in a random coil conformation. Hydrophobicity coefficients were then determined by measuring their effect on retention time of the model synthetic peptide during an RP HPLC separation. With the random peptide collection approach, on the other hand, retention coefficient values are determined from optimization of the model that correlates observed and predicted retention times (also known as linear regression analysis). The first example of this model was demonstrated by Meek [29] based on optimization for 24 natural peptides separated using RP HPLC. Later, a similar approach was used by many researchers for different chromatographic conditions [32-34]. A sufficient number of random peptides (>100) are required for proper evaluation of retention coefficients using regression analysis. Most of the models described in the 1980's were optimized for a very low number of peptides and produced very poor correlation with the data obtained by the synthetic peptide approach [31]. Recently, Shamshurin et al. [35] conducted a comparative study of retention coefficients reported by Kovacs et al. [31] and the ones obtained during an optimization of the Sequence-Specific Retention Calculator (SSRCalc) model. The authors concluded that there is virtually no difference between the scales obtained, except for the determination of hydrophobicity of charged groups (Arg, His and Lys) at pH 2. Clearly, all charged residues in a peptide sequence are involved in ion pairing interactions under acidic conditions in RP HPLC – it is convincing that the hydrophobicity of charged residues is strongly affected by the type and concentration of the ion pairing modifier. It was also demonstrated by Shamshurin et al. [35] that the hydrophobicity of basic amino acids is strongly affected by the overall hydrophobicity of the peptide, rather than its position within the sequence.

Retention time prediction of peptides based solely on summation of retention coefficients of 20 amino acids (or the additive model) cannot provide correlation better than 0.85-0.9, even though the optimization of retention coefficients utilized a reasonable-size training set. Furthermore, it was realized that corrections, which account for important peptide structural properties should be introduced to the additive model in order to facilitate a more accurate assignment of retention coefficients. In the next section, I will briefly review sequence-specific corrections used in the SSRCalc model.

1.5 Sequence-Specific Retention Time Calculator

Sequence-specific retention time calculator (SSRCalc) is one of the most sophisticated peptide retention time prediction models that take into account important peptide structural features. Currently, a family of SSRCalc models were optimized for different RP HPLC separation systems, such as a C18 column (100Å pore size) with 0.1% TFA in the eluent [36], a C18 column (300Å pore size) with 0.1% TFA [36], a C18 column (100Å pore size) with 0.1% formic acid [37] and a C18 column (100Å pore size) at pH 10 [37]. There have been significant improvements to the SSRCalc model by introduction of new sequence-specific features and a gradual increase in the size of the optimization data set.

The first version of the SSRCalc model was developed in 2004 [38]. A training set of ~350 peptides was collected using offline HPLC-MALDI MS and a correlation of R^2~0.94 between observed and predicted peptide retention times was produced. At the time, it was an additive model with correction factors related to the peptide length and the shielding effect for

the amino acids located at the N-terminus due to ion pairing. A brief explanation of each correction is provided as follows:

1. Length correction. The retention times of smaller peptides are easier to predict, as all their residues are readily available for interaction with a stationary phase. On the other hand, for longer peptides the retention contribution of each amino acid decreases due to the overall size of the peptide. This correction was first introduced by Mant et al. [39] and applied to most of the models developed to date.

2. Ion-pairing shielding effect observed for N-terminal residues. The shielding effect can be envisioned as the positively charged N-terminal surrounded by a cloud of counterions (formate, trifluoroacetate) due to ion-pairing. This directly reduces the retention contribution of the nearby residues (particularly hydrophobic residues in positions 1, 2 and 3). Different types of ion-pairing modifiers (hydrophobic vs. hydrophilic) and their concentration also affect the retention behavior of peptides. Corrections related to ion-pairing at the N-terminus were first proposed by Krokhin et al. [38], and subsequently used by other research groups in their models [40].

In the second version of the SSRCalc [36], a much larger data set was used (~2000 tryptic peptides) to characterize additional structural properties of peptides. This includes additional corrections for the nearest neighbor effects, the influence of isoelectric point (pI) of a peptide, the effect of overall peptide hydrophobicity on retention contribution of the charged residues (His, Arg, and Lys), the presence of a cluster of hydrophobic residues and Pro residue(s) as well as the influence of peptide propensity to form helical structures. A brief explanation of each correction is provided as follows:

1. Nearest neighbor effects. All amino acids adjacent to positively charged residues (His, Arg, or Lys) inside the peptide chain will be affected by ion-pairing. This is similar to the shielding effect observed for the N-terminal residues as explained earlier.

2. Influence of pI of a peptide. The isoelectric point of a given peptide is the pH where the molecule carries zero net charge. When LC separation operates under acidic conditions (pH~2.5), basic amino acids (His, Arg, and Lys) are the only residues that carry a positive charge and all acidic amino acid residues remain neutral. Observation reveals that low pI favors retention of relatively hydrophobic peptides and decreases retention for hydrophilic ones. The reverse is also true; high pI is the reason for preferential retention of relatively hydrophilic compared to hydrophobic peptides. Such a phenomenon can be explained from the point of view of ion-pair formation as well.

3. The presence of a cluster of hydrophobic residues will decrease the apparent hydrophobicity of the peptide. This observation can be explained from the point of view of steric hindrance that prevents neighboring amino acids from simultaneous interaction with the C18 phase.

4. The presence of a proline residue in a peptide sequence can potentially prevent a peptide from folding, making the rest of the amino acids available for hydrophobic interactions due to a rigid structure (side chain is covalently linked with nitrogen in the peptide chain). It was also found that the peptides carrying proline repeats typically show negative deviation from the predicted retention time [36].

5. The effect of overall peptide hydrophobicity. It was found that peptides with different hydrophobicities exhibited different prediction errors on average. The second version of the

SSRCalc contained an approximate correction related to these effects. Later it was established that these effects are related to the intrinsic properties of the charged amino acids, rather than the whole peptides [35]. Each tryptic peptide in the training dataset contains at least one charged residue. This determined the overall dependence of retention time shifts on peptide hydrophobicity.

6. The influence of peptide propensity to form helical structures. The influence of this parameter was established by monitoring peptides with substantial positive deviations from the predicted retention time. Correction was only introduced to account for short stretches of helical structure that show the following patterns, XXOX, XXOXX, XXOOXX, where X represents the hydrophobic amino acids and O for all others. Each possible combination was assigned with an optimized value of a correction factor. The incorporation of corrections for multi-turn helices in SSRCalc is still under development.

The third version of the model included all features of the second one, but it was re-optimized for an even larger dataset containing 5,000-10,000 peptides. At the same time two new separation conditions were introduced: a C18 column (100Å pore size) with 0.1% formic acid and a C18 column (100Å pore size) at pH 10 [37]. Subsequent development of SSRCalc included introduction of sliding hydrophobicity scales (a weighted correction for the retention coefficient of each amino acid is added to the calculation of peptide retention time depending on the overall hydrophobicity of the particular peptide) [35] and improvements in predicting helical contributions. Alpha-helices are major elements of protein and peptide secondary structure. Improvements in the understanding of how helical peptides interact with the hydrophobic C18 phase will be critical for further progress in the field. The size of retention time data sets is

growing continuously. Currently, the SSRCalc model is being enhanced using a ~280,000 peptide collection.

The long-term goal of the research group is to construct an enhanced peptide retention time prediction algorithm, which will encompass various chemically and biologically relevant PTMs. Following development of the core SSRCalc model [36], Krokhin et al. began systematic studies of modified peptides with N-terminal cyclization at glutamine and carbamidomethyl-cysteine [41], and with five different cysteine-alkylating agents [42]. These studies highlighted the importance of amphipathic helicity in correct functioning of prediction models [43]. Yet this feature remains unaccounted for in current peptide retention time prediction models. The hydrophobic character of methionine, its abundance inside amphipathic helices and significant differentiation in retention time shifts upon oxidation depending on secondary structure [14] make this modification a key in better understanding the effects of amphipathic helicity on RP HPLC peptide retention. Therefore the first study presented in the thesis focused on characterization of the chromatographic behavior of linear tryptic peptides containing oxidized methionine in RPC.

1.5.1 Comparison of Different Retention Time Prediction Models

Rapid development in proteomics brought renewed interest in peptide retention time modeling. These developments were also accelerated by the growing size of the peptide training sets. More than one dozen models were reported from 2003 to 2014 [36-38,40,44-51]. Comparison of different models is a very complicated task as all of them were optimized for different systems and retention datasets. Proper comparison must be based on identical data: the

best model should provide a better correlation. Another way to compare different models is to understand which parameters are taken into account. Thus, there are a number of algorithms, developed in the proteomic era, which do not take into account N-terminal effects due to ion-pairing. These algorithms belong to a category of additive models [49-51] and cannot be compared to more sophisticated approaches. Table 1.1 shows the features of different models. The SSRCalc has more features than any other models. At the same time, it should be noted, that some models are based on different algorithmic approaches and authors do not always reveal all details. Thus, the analytical neural network (ANN) algorithm [47,48] uses so-called "hidden nodes" in the optimization process.

Reimer et al. [43] attempted a comparison of different models using the same optimization dataset of ~3,330 peptides, separated using C18 300Å sorbent with 0.1 % TFA as the ion-pairing modifier. Authors of respective algorithms were contacted to perform the calculations [40,48,49], while an on-line version of BioLCCC (Liquid Chromatography of Biomolecules at Critical Conditions) [45] was used for calculations. Table 1.2 shows the comparison of reported model accuracies and the actual R^2-values for the dataset under consideration. The SSRCalc demonstrated the best performance among all tested algorithms; ANN approach and Elude algorithm showed similar results; the additive model by Gilar [49] was the distant fourth; and BioLCCC demonstrated an extremely poor correlation – significantly lower reported values. Validity of the SSRCalc approach is also confirmed by the fact that the Elude model is based on the same principles and correction factors.

Table 1.1 Retention time prediction models developed in the proteomic era: correction factors.

Model	Sum of retention coefficients	N-terminal hydrophobicity	C-terminal hydrophobicity	Peptide length	Nearest neighbor effects	Effect of Pro in the sequence	Clusters of hydrophobic residues	Helicity
Analytical Neural Network [47,48] Includes undisclosed "hidden nodes" of ANN	+	+	+	+	-	-	-	-
Analytical Neural Network [46]	+	-	-	+	-	-	-	-
BioLCCC [45]	+	Undisclosed model to predict peptide's secondary structure						
SSRCalc [36-38]	+	+	+	+	+	+	+	+
Additive model [49]	+	-	-	+	-	-	-	-
ELUDE [40]	+	+	+	+	+	-	+	+
QSRR – quantitative structure-retention relationship [44]	+	Undisclosed collection of "molecular descriptors"						

Note that a "+" sign indicates the respective algorithm features the corresponding structural correction, whereas a "-" sign indicates that correction has not yet been taken into account in the algorithm.

It should be noted that the comparison shown in Table 1.2 was performed based on the version of respective algorithms available in 2010. The newest version of SSRCalc was introduced with sliding hydrophobicity scales for charged residues and better prediction of helicity.

Table 1.2 Accuracy of retention time prediction models when applied to a collection of 3330 tryptic peptides [43]. Used with permission from Elsevier.

Model	Reported accuracy for training datasets	Test set of 3330 tryptic peptides	
		Model accuracy for all peptides	Model accuracy excluding cysteine-containing peptides
SSRCalc 300Å – TFA* [36]	~0.98	0.977	0.977
ANN predictor [48]	~0.98	0.967	0.969
BioLCCC** [45]	~0.9	0.862	0.859
ELUDE [40]	~0.94	0.965	0.966
Additive model with length correction by Gilar et al. [49]	~0.96-0.97	0.927	0.925

* - http://hs2.proteome.ca/SSRCalc/SSRCalcX.html; ** - http://theorchromo.ru

1.6 Polyunsaturated Fatty Acid and Flaxseed Oil

Polyunsaturated fatty acids (PUFAs) are long-chain fatty acids that have two or more carbon-carbon double bonds in their hydrocarbon structure. They are mainly classified into four families designated n-3, n-6, n-7 and n-9, where n- denotes the position of the last double bond from the methyl end of fatty acids. Extensive studies have shown that the n-3 and n-6 families of fatty acids are physiologically important for many vital functions in biological membranes and as precursors of a variety of lipid regulators of cellular metabolism. The commonly known α-linolenic acid (ALA, 18:3n-3) and linoleic acid (LA, 18:2n-6) act as the parent fatty acid for the

n-3 and n-6 families, respectively. They are also considered as essential fatty acids that cannot be synthesized by humans, as humans lack the desaturase enzymes required for their production. Therefore, LA and ALA have to be obtained from the diet, and further modified into longer PUFAs. Dietary sources of LA and ALA include fish oil, flaxseed, soybean, walnuts and the oil of flaxseed, rapeseed and soybean.

Polyunsaturated fatty acids are readily susceptible to chemical modifications because their double bonds are unstable. The oxidation of unsaturated fats produces a variety of compounds that smell and taste rancid. Thus, oxidation is of serious concern in the food industry. In this instance, I conducted analyses on flaxseed (linseed) oil, which contains high levels of polyunsaturated fatty acids and cyclic methionine-containing peptides that are prone to oxidation [52,53]. This work was done through collaboration with the industrial partner, ShapeFoods Inc. (Brandon, Manitoba). While the freshly pressed flaxseed oil has a delicate nutty flavor, an unpleasant bitter taste starts to develop after only one day of storage due to methionine oxidation in its major cyclolinopeptides [52]. A recent study by Bruhl et al. [52] had identified cyclolinopeptide E (cyclo-MsoLVFPLFI) – CLE, (where Mso represents methionine oxide) as responsible for the onset of bitter taste associated with flaxseed oil aging. Hydroperoxide produced from lipid oxidation is believed to be the main source of reactive oxidative species (ROS), which act on methionine-containing cyclolinopeptides [54]. Quantitation of CLE by reversed-phase HPLC was used as an indicator of oil deterioration in the assessment of shelf-life of this product [55]. A series of samples prepared by ShapeFoods during accelerated shelf-life trial experiments were analyzed to determine stability of the commercial product during storage. None of the previous studies demonstrated complete separation of major cyclolinopeptides,

possibly due to their similar chromatographic behavior. Application of modern core-shell separation technology with different reversed-phase functionality was selected as one of the possible approaches to address this problem. The goal of the second part of the thesis was to develop an efficient chromatographic procedure to monitor the oxidation of major cyclolinopeptides in flaxseed oil, as well as characterize major cyclolinopeptides by RPC and mass spectrometry.

1.7 Core-shell Reversed-phased Technology

The construction of high-efficiency LC columns for fast separation is of great technological importance – not only to increase separation efficiency but also to shorten analysis time and increase throughput. A recent innovation in column particle technology is the invention of core-shell phases. These columns are packed with uniformly distributed particles that are made of solid silica cores and covered with a homogenous spherical porous shell. Separation efficiency of core-shell columns has improved over traditional fully porous chromatographic materials, and most importantly, they operate at pressure compatible with conventional HPLC equipment. To understand the principle of core-shell columns, the concept of the height equivalent to a theoretical plate (H) and the theory of band broadening are briefly discussed in the following section 1.7.1.

1.7.1 Band Broadening Theory

Column efficiency expressed as number of theoretical plates (N) is a mathematical concept that measures the broadening of a peak during passage through the LC column:

$$N = 5.545\left(\frac{t_R}{w_{0.5}}\right)^2$$

where t_R is the retention time of a chosen peak and $w_{0.5}$ is the peak width at its half height [56]. Typically, columns with higher plate numbers are considered to be more efficient, hence capable of better separation due to narrower peaks. The height equivalent to a theoretical plate (H) is generally used as a measure of efficiency when comparing different columns. It relates the plate number (N) to the column length (L) as follows:

$$H = \frac{L}{N}$$

However, this simplistic plate theory did not seem to address the contributions of peak broadening in chromatography and it has thus been expanded to include key aspects of modern HPLC methodology, including particle size, diffusion, flow dynamics and resultant effects on the mass transfer of a solute between the mobile and stationary phases of the column. Therefore, an equation that combined all critical bandspreading contributions was introduced by van Deemter in 1956:

$$H = A + \frac{B}{\mu} + C\mu$$

where μ is the linear flow velocity of the mobile phase and A, B, and C are constants. The A term, Eddy diffusion, describes multiple diffusion paths taken by an analyte through the column packing. It is dependent on particle size and the homogeneity of the packed bed. Smaller particles as well as uniform particle size distribution reduce the A-term and therefore improve efficiency. The B term, the longitudinal diffusion constant, quantifies band broadening due to the

longitudinal diffusion of the solute. Typically, the higher the eluent velocity, the lower the diffusion effect on the band broadening. Moreover, molecular diffusion in the liquid phase is about five orders of magnitude lower than that in the gas phase and so this effect is almost negligible at the standard HPLC flow rate but more significant in gas chromatography. The C term, resistance to mass transfer, is proportional to linear velocity of the mobile phase through the column and it has two contributors: resistance to the mass transfer in the stationary phase C_s and mobile phase C_m. The effect of the former one is only relevant if thick bonded phases are used. The latter one, also known as stagnant mobile phase, is kinetic hindrance due to penetration of the solutes into the pores of the support packings. Typically, the slower the velocity, the more uniformly analyte molecules may penetrate inside the particle, and a decreased effect of different penetration on the efficiency. In contrast, faster flow rates will lead to high elution distance between molecules with different penetration depth.

Core-shell particle morphology addresses two of the most critical sources of band broadening: A term (the Eddy diffusion) and C term (resistance to mass transfer) in the van Deemter equation. First, by narrowing the particle-size distribution and optimizing packing, band broadening effects due to analyte diffusion (Eddy diffusion) is dramatically diminished and leads to sharper peaks independent of the mobile phase velocity. Second, since analyte interaction with the stationary phase and LC separation occur only in the outer shell of the core-shell particles, the analyte spends less time diffusing in and out of the pore media, thus reducing the dispersive effect known as resistance to mass transfer. As a result, minimizing the A and C terms leads to an overall increase in column performance. On the other hand, the thin porous shell of core-shell particles reduces the sample loading capacity and can be considered a disadvantage compared to

the traditional columns packed with fully porous particles. Thus, preparative separations that require high loading capacity would be better suited for the traditional porous sorbents. Analytical applications with restricted sample load fit perfectly with core-shell technology.

Chapter 2

Chromatographic Behavior of Linear Tryptic Peptides Containing Oxidized Methionine

2.1 Literature Review

2.1.1 Methionine Oxidation in Proteomic Studies

Oxidation is often encountered in proteomic analyses. This chemical modification can readily occur when peptides/proteins are exposed to air during sample preparation. Sulfur-containing amino acids such as cysteine and methionine are generally prone to oxidation. Typically, reduction and alkylation of cysteine are commonly used to prepare proteins for enzymatic digestion in proteomic experiments. The latter step mainly prevents the formation of disulfide bonds, yet helps to reduce the chance of exposing reactive free cysteine residues to air oxidation. On the other hand, methionine oxidation is one of the most frequently observed PTMs on peptides/proteins, and is present up to 10% of the total non-redundant peptide sequences identified in typical bottom-up LC-MS experiments [57]. Unintentional oxidation occurs primarily during sample preparation and handling, and it was reported that electrospray ionization (ESI) may also induce oxidation [58]. Usually, an in-source oxidation can easily be differentiated based on identical elution profiles of modified and non-modified peptides [59].

The first and the main product of methionine oxidation is methionine *S*-oxide (methionine sulfoxide, Mso, +16 Da adduct), while intense oxidation can cause methionine *S*-

oxide to react further to produce a methionine *S*,*S*-dioxide derivative (methionine sulfone, Msn, +32 Da adduct) [60]. The latter product is very rarely observed in proteomic experiments because the reaction is much slower. Sources of reactive oxygen species (ROS) may be endogenous (i.e. naturally generated superoxide anions and hydrogen peroxide as a result of oxidative phosphorylation [61]) or exogenous (i.e. pollutants, chemicals and toxins taken in from the environment). Furthermore, external sources such as sunlight and other forms of radiation can generate endogenous ROS, which can lead to a number of diseases. Ghesquiere et al. [62] investigated the oxidation degree of different methionines within the same protein, and found unequal oxidation susceptibility among them. Interestingly, methionine oxidation occurs preferentially near polar neighboring amino acids, most likely because Met becomes more surface-oriented within a polar environment and thus turn itself into an oxidation target. For instance, methionines surrounded by acidic amino acids were surprisingly more susceptible to oxidation than those surrounded by basic amino acids at physiological pH. In addition, methionine residues were more likely to be oxidized when close to alanines, threonines and serines.

The reaction mechanism for methionine oxidation in peptides to produce methionine *S*-oxide is rather difficult to characterize even if the source of oxidation is known and the reaction system is relatively simple [63]. Schoneich et al. [63] demonstrated several possible oxidation pathways that can be encountered during the oxidation of peptides, such as oxidation by free hydroxyl radicals, or a metal-catalyzed oxidation with or without a neighboring group effect. Most interestingly, however, the formation of methionine *S*-oxide leads to a mixture of RS- and

SS-diastereomers, since the sulfur of the methionine thioether is a prochiral center (Figure 2.1).

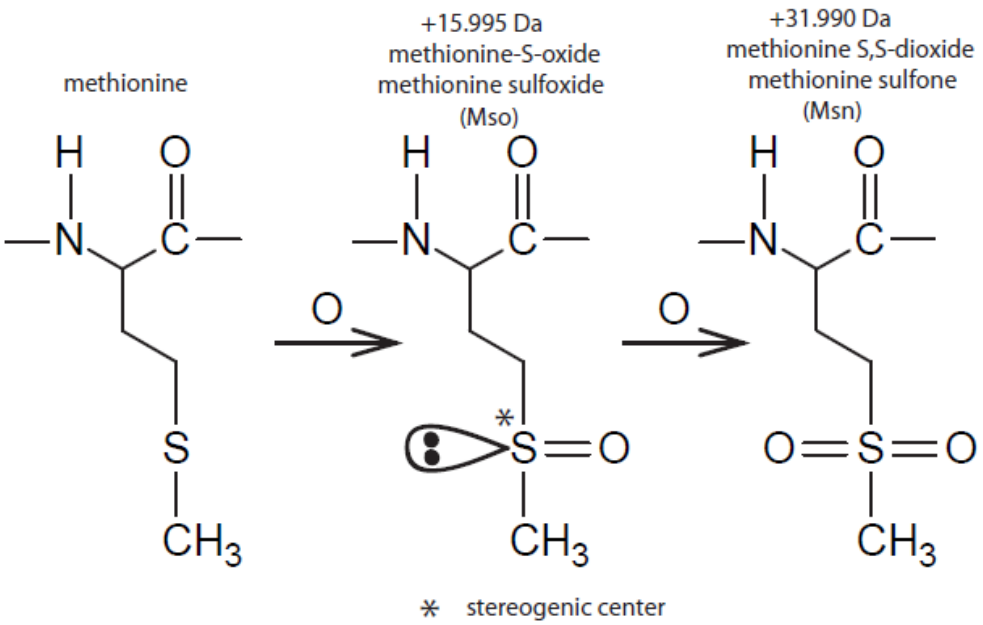


Figure 2.1 Oxidation of methionine to methionine *S*-oxide (Mso) and methionine *S,S*-dioxide (Msn). The stereogenic center at the S atom of Mso is labeled with *.

Peptides carrying Mso diastereomers can be easily separated using conventional (achiral) RP HPLC methods only when there are effects of secondary structure. Blondelle et al. [14] reported that a doublet peak was observed for amphipathic helical peptides containing methionine *S*-oxide that are located in the hydrophobic face of the helix. The same positioning caused significantly larger negative retention time shift after oxidation. Peak splitting for Mso-containing peptides is sometimes observed in proteomic experiments [64]. Recently, Khor et al. [65] demonstrated successful identification and differentiation of the RS- and SS-diastereomers using methionine sulfoxide reductase enzymes MsrA and MsrB, respectively. In most biological systems, methionine sulfoxide reductase is responsible for catalyzing the reduction of methionine *S*-oxide

back to methionine [66,67]. Thus, methionine that scavenges ROS is often considered as a catalytic antioxidant, which protects both proteins and other macromolecules [68]. The oxidation and subsequent reduction of methionine residues is of great interest for understanding different aspects of how oxidative stress affects function and other aspects of proteins. Nevertheless, repair by methionine sulfoxide reductase is inefficient. Many age-related diseases are known to be associated with the accumulation of Mso, such as Alzheimer's, Parkinson's, emphysema, bronchitis, rheumatoid arthritis, inflammatory disorders, and cataract formation [69].

2.1.2 Use of Methionine Oxidation in COFRADIC Technology

The use of modified peptides, which exhibit a different affinity for the stationary phase in reversed-phase chromatography, has opened up a new avenue of exploration in peptide-centric proteomics [70]. The combined fractional diagonal chromatography (COFRADIC) approach aims to reduce the overall peptide mixture complexity to a more manageable level by isolating peptide sets from the digest mixture for further MS/MS analysis [71]. This method utilizes two-dimensional RP-LC separation under identical conditions. Fractions collected in the first dimension are subjected to a chemical modification in order to alter the peptides' chromatographic shifts in the second dimension run. Currently, the COFRADIC approach has been designed for selective isolation of various types of modified peptides [72], such as methionine *S*-oxide-containing peptides. In this instance, H_2O_2 oxidation was introduced to all fractions collected in the first dimension. Knowing that methionine oxidation reduces retention time allowed for the prediction of the expected elution window for modified peptides in the

second dimension. Due to the relatively low abundance of methionine this led to a higher dynamic range for the analysis of complex peptide mixtures.

The recent COFRADIC approach was modified slightly by reducing isolated methionine *S*-oxide-containing peptides with recombinant methionine sulfoxide reductases between two COFRADIC RP HPLC separations [61]. As a result, Msr-affected peptides that become more hydrophobic are isolated from the bulk of peptides for further analysis, such as determination of oxidation degree of protein-bound methionines in a proteome sample, and correction for any artificial methionine oxidation that may occur during sample handling [61]. Both studies required isotope labeling on the respective reference and *in vivo* stressed cell cultures with different methionine variants. In the first study, the reference trypsin-digested cell culture is initially oxidized *in vitro* to generate all possible methionine *S*-oxide-containing peptides, and then, combined with the *in vivo* stressed trypsin-digested cell culture. This mixture is then subjected to COFRADIC and the MS readout allows for determination of the degree of *in vivo* oxidation. To correct for artificial oxidation occurring during sample handling, one can combine the reference cell culture with the *in vivo* stressed one, digested with trypsin, then subject the mixture to COFRADIC. It is important to realize that artificial methionine oxidation is expected to occur to the same extent in both cell cultures, providing the nature of the sample handling was the same. Thus, peptides that were *in vitro* oxidized in both instances would appear in the MS spectra as doublet ions with similar intensities, whiline peptides originally oxidized *in vivo* would be seen as single peaks.

2.1.3 Chromatographic Behavior of Peptides Containing Oxidized Methionine

Systematic large-scale observations on chromatographic behavior of PTM/CM peptides including methionine oxidation appeared following rapid developments in LC-MS proteomic applications. The literature indicates that peptides containing oxidized methionine are more hydrophilic under RP HPLC conditions than their unmodified analogs. Zybailov et al. [57] summarized the effects of 15 different PTMs on peptide elution times and found that retention time increases in the order Mso < Met. Lengqvist et al. [58] investigated the effect of common modifications on retention and isoelectric point values and arrived at the same conclusion. Both studies indicated that "dioxidation" (+32 Da) reduces retention time even more. These observations were made, however, for peptides carrying two Mso residues rather than one methionine *S*,*S*-dioxide (Msn). Recently, Moruz et al. [27] applied their peptide retention time prediction model Elude to five different PTMs. The authors collected retention information from proteomic analyses of 12 different organisms and modified the prediction model to incorporate PTMs. Some of these runs were performed using the COFRADIC approach, thus the authors encountered oxidation to both Mso and Msn. Optimization was performed for ~ 1,500 peptides for both Mso and Msn-containing peptides and showed increase in retention coefficients as follows: Mso < Msn < Met. Despite ample evidence of significant variation in retention time shift upon oxidation provided by Houghten and co-workers 20 years ago [73], these effects were never explored in detail. Assumption of identical effects of methionine oxidation (independent of its positional and secondary structure effects) on the chromatographic behavior of any peptides, the accuracy of such peptide retention time prediction model reaches its potential maxima at R^2 = 0.93-0.94 (based on the correlation between predicted and actual retention times).

The goal of this study was to expand the application of the SSRCalc model in order to process peptides with methionine oxidation. Priority was given to ubiquitous methionine *S*-oxide-containing peptides routinely observed in LC-MS experiments. I attempted analysis of retention time shifts for a few thousand pairs of Met-Mso containing tryptic peptides observed in three 2D LC-MS runs. Retention behavior of Met-Mso-Msn triads was investigated on a smaller scale by 1D LC-MS analysis of tryptic digests oxidized with hydrogen peroxide (H_2O_2). The observations were also supported by the analysis of designed synthetic peptides. Overarching objectives of this study included using methionine oxidation as a probing tool to better understand effects of helicity on peptide interaction with hydrophobic surfaces – one of the key mechanisms determining peptide interaction in biological systems.

2.2 Experimental

2.2.1 Materials

All chemicals were analytical grade and sourced from Sigma-Aldrich (St-Louis, MO), unless noted otherwise. Synthetic peptides were purchased from BioSynthesis Inc. (Lewisville, TX) and JPT (Berlin, Germany), in mg and Spike Tide formats (50 ng per peptide) respectively. HPLC-grade acetonitrile and de-ionized water from Fisher Scientific were used for the preparation of eluents. Sequencing-grade modified trypsin (Promega, Madison, WI) was used for digestion.

2.2.2 Protein Digestion and Sample Preparation

Tryptic digests of whole cell lysates of *Neurospora crassa*, *Thermoanaerobacter thermohydrosulfuricus* WC1, and *Clostridium butyricum* were prepared using the FASP (filter aided sample preparation) protocol [74]. This digestion protocol utilized a filter with a particular molecular weight cutoff to collect proteins that have been denatured in sodium dodecyl sulfate (SDS), reduced (by dithiothreitol) and alkylated (by iodoacetamide). SDS is then washed away using urea and followed by an on-filter digestion. Finally, the resulting digested peptides were collected for LC-MS analysis. Digests were acidified with TFA and purified by RP HPLC. Approximately 100 µg of each digest (according to NanoDrop 2000, Thermo Scientific) were used for 2D LC-MS analysis. Samples of synthetic peptides were prepared by dilution in Buffer A (0.1% formic acid in water) to provide injected amounts of ~200-400 fmole per injection. To produce samples with a higher degree of oxidation the samples were subjected to oxidation with 1% H_2O_2. Depending on the desired oxidation state and degree of oxidation in the original samples the time of exposure to H_2O_2 can be varied. Synthetic peptides and the tryptic digest of *T. thermohydrosulfuricus* WC1 were oxidized with 1% H_2O_2 for six days to produce sufficient amounts of Msn-containing species. Methionine-containing Spike Tide peptides from JPT usually exhibit ~50-60% oxidation degree to Mso without any treatment. All individual samples and fractions collected in the first dimension of the 2D LC procedure were spiked with a mixture of six standard peptides P1-P6 [64] for retention time alignment purposes (400-600 fmole per injection each).

2.2.3 Chromatographic Conditions

First dimension separation was performed using RP HPLC at pH 10, using a method first described by Gilar et al. [75] and modified as described elsewhere [37]. Aliquots of digests containing ~100 µg of peptide mixtures were fractionated. Forty 1-minute fractions were collected (roughly 2.5 µg per fraction). Fractions were concatenated pair-wise to ensure optimal coverage of separation space in the second separation dimension. One third of each fraction (~1.5 – 2 µg) was subjected to a 1 hr second dimension LC-MS analysis.

Both 1D and 2D analysis procedures employed identical nano-flow RP HPLC instrument settings on the final stage of separation prior to the MS acquisition. A splitless nano-flow 2D-Ultra system (Eksigent/ABSciex, Dublin, CA) with 10 µL sample injection through a 300µm × 5mm PepMap100 (Thermo Scientific) trap-column and a 100 µm × 200mm analytical column packed with 5µm Luna C18(2) (Phenomenex, Torrance, CA) was used. Both eluents A (water) and B (acetonitrile) contained 0.1 % formic acid as the ion-pairing modifier. Linear gradients of 0.6 – 0.8 % acetonitrile per minute (0.5 – 30 % B) were used for LC-MS acquisitions in 2D analyses. A 0.33 % acetonitrile per minute linear gradient (0.5 – 30 % B) was used for peptide elution for 1D LC-MS. All gradient runs were completed with 5 min of 80 % phase B and 7 min of equilibration with starting conditions of 0.5 % B. Total LC-MS analysis time was ~1 hr per fraction in 2D and ~2 hrs per sample in 1D, respectively.

2.2.4 Mass Spectrometry

A TripleTOF^{TM}5600 mass spectrometer (ABSciex, Concord, ON) was used in standard MS/MS data-dependent acquisition mode. Survey MS spectra 250 ms were collected (*m/z* 400-

1500) and followed by up to 20 MS/MS measurements on the most intense parent ions (300 counts/sec threshold, +2 - +4 charge state, *m/z* 100-1500 mass range for MS/MS, 100 ms each). Previously targeted parent ions within a 50 mDa window were excluded from repetitive MS/MS acquisition for 12 sec. The following search parameters for identification with X!Tandem (MS/MS database searches for protein identification) were used: ±10 ppm and ±30 mDa mass tolerance for parent and fragment ions, respectively; constant modification of Cys with iodoacetamide; variable modifications +15.995 and +31.990 at Met; expectation value cut-off of Log(e) < –3 and < -1 for non-modified and oxidized peptides, respectively.

2.2.5 Hydrophobicity Expression and Calculations

Retention times for the standard P1-P6 peptides [64] and tryptic digests were assigned to each non-redundant identification as the intensity weighted time average for the two most intense MS/MS spectra. Peptide retention time was expressed in Hydrophobicity Index (HI) units (% acetonitrile) [76] on a least square linear regression fit using the HI values previously assigned to each member of the P1-P6 mixture. An in-house version of the SSRCalc algorithm written in Perl was used for the HI calculations [36].

2.3 Results and Discussion

2.3.1 Impact of Methionine Oxidation on the Output of LC-MS Proteomic Analyses

Table 2.1 shows the summary of protein and peptide identification for one- and two-dimensional LC-MS analyses of the digests used in this study including detection of peptides containing methionine *S*-oxide. Coverage of the respective proteomes (60-70%) in 2

Table 2.1 Detection of methionine oxidation for typical 1D and 2D LC-MS bottom-up proteomic experiments.

Sample origin	Number of spectra/ Number of identified peptides[a]	Identified proteins[b]	Unique identifications for non-modified peptides	Unique Met-containing peptides	Unique Mso-containing peptides	Oxidation rate[c]
1D LC-MS						
C. butyricum	34182/ 20001	912	6565	2204	222	3.4%
N. crassa	15398/ 10712	1153	5979	1419	745	12.5%
T. thermohydrosulfuricus WC1	35886/ 20939	1101	6969	1957	224	3.2%
2D LC-MS						
C. butyricum	301661/ 164756	2174	25027	8800	2880	11.5%
N. crassa	306150/ 180347	4288	41504	11693	7441	17.9%
T. thermohydrosulfuricus WC1	393105/ 202969	1998	26633	8567	2733	10.3%

[a] Expectation value cut-off of Log(e) < -1; [b] Expectation value cut-off of Log(e) < -3; [c] Oxidation rate, I computed the ratio of the total number of Mso-containing peptides identified to the total number of non-redundant peptides.

dimensional runs with ~20 hrs of MS acquisition is typical for the TripleTOF^{TM}5600 platform. The percentages of Mso-containing peptides for *C. butyricum* and *T. thermohydrosulfuricus* WC1 digests found for 1D and 2D analyses performed in my laboratory were ~3 and ~10%, respectively. As a comparison, Zybailov et al. [57] reported ~10% of peptides carried methionine *S*-oxide in the analysis of leaf proteome from *Arabidopsis thaliana* wild type. Their 1D LC-MS

runs, however, were performed for in-gel digested proteins, which usually results in additional methionine oxidation. A higher degree of oxidation in the *N. crassa* digest in this case was probably due to differences in sample preparation. In all cases two-dimensional separation increases dynamic range of MS detection and makes it accessible to low-abundance species. This results in better detection of Mso-containing peptides for all three samples. An additional source of this increase includes oxidation of peptides after collecting the fractions in the first dimension as they were pairwise concatenated and lyophilized.

It should be noted that I did not detect Msn-containing peptides in the samples shown in Table 2.1. Zybailov et al. [57] found the "Dioxidation (M) [+31.99]" modification, which corresponds to oxidation of two methionines in the same peptide, rather than formation of methionine *S,S*-dioxide. Msn-containing species are observed when intentional oxidation is employed as a part of sample preparation protocol. Therefore, investigation of the chromatographic behavior of Msn-containing peptides is relevant to the method development of specialized proteomic protocols such as methionine – COFRADIC whereas better understanding of the behavior of Mso-containing molecules would benefit most proteomic applications.

2.3.2 Peak Splitting for Mso-containing Diastereomers: Confirmation of Literature Data Using Synthetic Peptides

Addition of an oxygen atom (Figure 2.1), as Met is oxidized to Mso, results in increased polarity of the side chain leading to decreased RP HPLC retention time of the modified peptide. According to Blondelle et al. [14] a decrease in retention time is more profound when the modified residue is situated in the peptide preferred binding domain, the hydrophobic face of an

amphipathic helix to the hydrophobic stationary phase of the column. The same positioning leads to peak splitting for Mso-containing peptides. Addition of the second oxygen (Figure 2.1) eliminates the stereogenic center, which should lead to elimination of peak splitting. To verify this, I used oxidation with H_2O_2 to prepare mixtures containing sufficient amounts of all three versions (Met, Mso and Msn – containing analogs) of the synthetic amphipathic peptides ASGVSSVMSSVVGK and ASGVSMVVSSVVGK. Both peptides contain comparable numbers of hydrophobic residues and exhibit a well-defined hydrophobic face as shown by the respective helical projections in Figure 2.2. These two identical features were introduced to generate sequences with comparable hydrophobicity and to highlight the influence of methionine positioning within an amphipathic structure. The following observations can be made based on the corresponding extracted ion chromatogram (XIC) shown in Figure 2.2:

a) Retention time (hydrophobicity of a side chain) increases in the following order: Mso < Msn < Met, which agrees with Moruz et al. [27] data;

b) A much larger retention time shift is observed when Mso was positioned in the hydrophobic face in the amphipathic α-helical structure as compared to that in the hydrophilic face: -5.6 and -1.8 % acetonitrile for ASGVSSVMSSVVGK and ASGVSMVVSSVVGK, respectively;

c) Peak splitting only occurs when Mso is positioned in the hydrophobic face of an amphipathic helix;

d) Further oxidation to methionine *S,S*-dioxide results in elimination of peak splitting.

Manual examination of the data from real proteomic runs shown in Table 2.1 confirms these findings (not shown) as it pertains to the Mso-Met pair: a larger negative retention time

shift is associated with peak splitting. Due to the small difference in retention time between pairs of peptides containing Mso diastereomers, these effects are largely ignored in large-scale proteomic runs. Both Mso-containing peptides have identical MS/MS spectra and are usually reported as a single detectable feature.

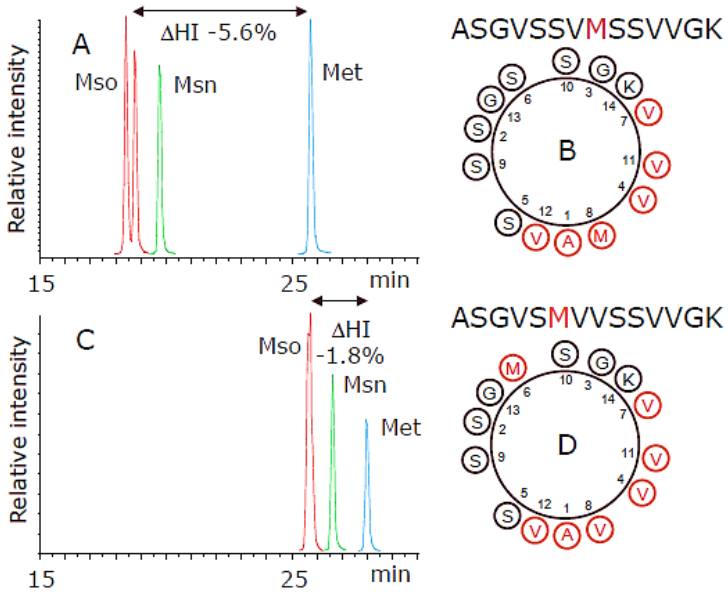


Figure 2.2 Retention time shifts associated with methionine oxidation to Mso and Msn depending on their position within the amphipathic helical structure. **(A, B)** – XICs and axial helical projection for ASGVSSVMSSVVGK (Met is located in the hydrophobic face of the helix); **(C, D)** – XICs and axial helical projection for ASGVSMVVSSVVGK (Met is located in the hydrophilic face). The slope of the gradient for these separations was 0.78 % acetonitrile per minute. Therefore a retention time shift of 7.16 minutes in case of ASGVSSVMSSVVGK vs ASGVSSVMsoSSVVGK is equivalent to 5.6 % acetonitrile. Similarly, a retention time shift of 2.29 minutes is equivalent to 1.8 % acetonitrile for ASGVSMVVSSVVGK vs ASGVSMsoVVSSVVGK.

To rationalize the observations in Figure 2.2, it is important to recognize that, depending on the peptide sequence, hydrophobic methionines could play different roles upon interaction with a hydrophobic C18 surface. When in a preferred binding domain (hydrophobic face of an

amphipathic helix to the stationary phase of the column), methionine provides additional helix stabilization due to the hydrophobic character of its side chain. Consequently, when Met is oxidized to Mso, the more hydrophilic side chain disrupts this stabilization (Figure 2.2 a, b). Methionine oxidation occurring outside of the preferred binding domain (to the hydrophobic stationary phase) seems to affect peptide retention time to a lesser extent (Figure 2.2 c, d). Due to close "intimate" [14] contact of the preferred binding domain with a hydrophobic surface in RPC, even small alterations in peptide chemistry in this domain affect the interaction energy. In the case of a Mso diastereomeric pair this results in chiral separation using an achiral separation system.

2.3.3 Impact of Methionine Oxidation on Peptides' RP-LC Retention Time: Positional and Secondary Structure Effects

Data sets of 6421 pairs of non-modified – Mso containing peptides were compiled using 2D LC-MS identifications from Table 2.1. For the purpose of peptide retention time modeling, this collection of all modified species can be used directly for model optimization independently of their non-modified counterparts as shown by Moruz et al. [27]. Nevertheless I have chosen to analyze retention time shifts for each pair: this might reveal additional information required for proper algorithm implementation.

Figure 2.3 a and b show the dependence of retention time shifts (expressed in acetonitrile %, HI units) on peptide molecular weight (a) and hydrophobicity (b). The horizontal line at zero can be considered as the reference point for the retention time of any non-modified

methionine-containing peptide. Expressing retention time properties in acetonitrile percentage units allows for direct visualization of the chromatographic outcome of individual LC-MS

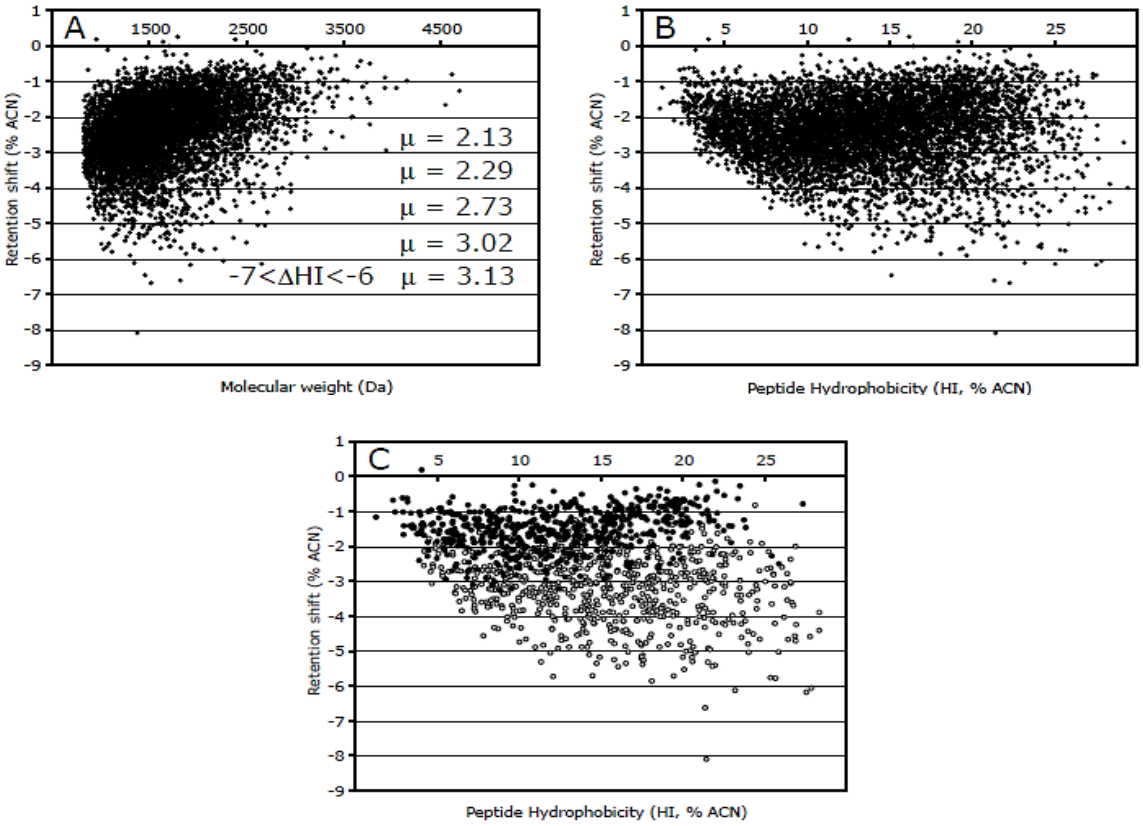


Figure 2.3 Effect of the size and hydrophobicity of peptide on retention time shifts ($\Delta HI = HI_{Mso} - HI_{Met}$) associated with oxidation of a single Met to Mso. **(a)** – ΔHI vs. molecular weight; **(b)** – ΔHI vs. peptide hydrophobicity expressed in HI units; **(c)** – retention time shifts (ΔHI) for groups (a) and (c) of peptides shown in Table 2.2. Filled circles – peptides carrying oxidized methionine in the N-terminal position; empty circles – amphipathic helical peptides with a methionine residue on the hydrophobic face of the helix. In **(a)**, average values of Eisenberg's hydrophobic moment (μ) were determined for groups of peptides showing the following retention shifts: 6-7, 5-6, 4-5, 3-4, and 2-3% acetonitrile.

acquisitions where the gradients could be different. Addition of standard peptides was used for alignment of the data from three proteomic runs.

Several trends are worth noting from these dependencies:

Table 2.2 Examples of Mso-containing peptides with similar structural characteristics and the corresponding retention time shifts.

Peptide origin	Sequence	Retention time shift ΔHI = (HI_{Mso} – HI_{Met}) (% acetonitrile)
(a) N-terminal Met: small negative retention time shift		
T. thermohydrosulfuricus WC1	MAGSHTDITER	-0.66
N. crassa	MLCSDWANIKPDIITLGK	-0.60
C. butyricum	MAQSENPDIIFGTDPDCDR	-0.47
(b) Met adjacent to positively charged residue: small (moderate) negative retention time shift		
N. crassa	IEEHVMKPSR	-0.85
C. butyricum	ALFIPNGTWLPDEMKDAPR	-0.84
C. butyricum	ALVMKPQASQEDQNEAWDYLVNDVTAK	-0.60
(c) Met in hydrophobic face of amphipathic helix: large negative retention time shift, associated with separation of diastereomeric pair of Mso-containing peptides		
Synthetic	ASGVSSVMSSVVGK	-5.60
Synthetic	GSIVEEMEEVIGEGER	-9.05
T. thermohydrosulfuricus WC1	SETIMEAVK	-3.52
T. thermohydrosulfuricus WC1	EGYETLLNTDMELELDNLAR	-3.39
C. butyricum	EAITMVVND**M**ELLSAVTK	-4.17
C. butyricum	EIIEEMEAFLK	-8.08
T. thermohydrosulfuricus WC1	VGENLAEDIEVMEAIFEVTK	-6.05
C. butyricum	LFLEDMESFLK	-5.11
d) Met in hydrophilic face of amphipathic helix: small negative retention time shift		
Synthetic	ASGVSMVVSSVVGK	-1.79
C. butyricum	EAIT**M**VVNDMELLSAVTK	-1.28
N. crassa	GVDMFIESLAR	-0.71
C. butyricum	SLNEALEEIGMVVEGVK	-0.72
T. thermohydrosulfuricus WC1	VYIIDEVHMLSTGAFNALLK	-0.80
e) Met in N3 position of N-capping box prior to amphipathic helix: small positive retention time shift		
Synthetic	GSIVMEIEEVIGEGER	0.36
N. crassa	SLEMELESLQGNAVR	0.01
N. crassa	TLQMLITELSPDTR	0.13
T. thermohydrosulfuricus WC1	TLEMILADTESEQYK	0.30

1) Retention time shifts caused by methionine oxidation range from -8 to 0.3 % acetonitrile, -2.37 % on average (Figure 2.3 a, b). In most cases, oxidation of Met to Mso reduces peptide hydrophobicity (retention in RP HPLC). However, I found five instances of positive retention time shifts, these data points appear above the zero line;

2) Methionine oxidation shows smaller effects for large peptides (Figure 2.3 a), which is consistent with that shown for other modifications [41,42]: modification of a single residue induces smaller changes in larger peptides;

3) Extremely large retention time shifts (5-8 % acetonitrile) were found only for relatively hydrophobic peptides (HI > 10 % acetonitrile in Figure 2.3 b).

That large retention time shifts were found associated with relatively hydrophobic peptides provides additional insight into the effect of peptide amphipathic helicity on the peptides' RP HPLC behavior. The presence of methionine in the hydrophobic face of an amphipathic helix is a prerequisite for a significant retention time shift. Therefore these peptides must exhibit a well-defined hydrophobic face, e.g. contain several hydrophobic residues in their sequences. Because of that, relatively hydrophilic species with HI < 10 simply do not fit this criterion as they contain relatively low numbers of hydrophobic amino acids. Figure 2.3 a also shows average values of Eisenberg's hydrophobic moment (μ) for groups of peptides depending on their retention shifts (6-7, 5-6, 4-5 % acetonitrile, etc.). Clearly, amphipathicity increases for peptides with large negative retention time shifts.

To further elucidate positional and structural effects of methionine oxidation I analyzed retention time shifts for five different groups of peptides from the data sets, examples of which are shown in Table 2.2. The variation in retention time shifts was explained based on the current

understanding of the separation mechanism according to the SSRCalc model and new findings associated with effects of secondary structure:

a) Small negative retention shifts are often associated with Met being in the N-terminal position as shown by filled circles in Figure 2.3 c. Part (a) in Table 2.2 lists several examples of those peptides selected from the data sets, such as MAGSHTDITER from *T. thermohydrosulfuricus* WC1. This effect is caused by ion-pairing of the peptide in RP HPLC. The positively charged N-terminal amino group interacts with hydrophilic formate counter ions, which provide a "shielding" effect from neighboring amino acids. This leads to a significant reduction of apparent hydrophobicity of the residues at the peptide N-terminus and results in a smaller retention time difference between Met (hydrophobic) and Mso (hydrophilic) containing peptide analogs.

b) Similar ion-pairing nearest neighbor effects are responsible for relatively low decreases in retention times for methionine's neighboring positively charged residues (Lys, Arg, His), such as IEEHVMKPSR from *N. crassa* shown in Table 2.2, part (b).

c) Extremely large negative retention time shifts were observed when methionine is a part of the hydrophobic face of an amphipathic helix. I confirmed this effect by selecting a group of peptides with extremely high hydrophobic moment values ($\mu > 3$), all of which contain methionine in the hydrophobic face of the helix (empty circles in Figure 2.3 c). See part (c) in Table 2.2 for examples.

d) Relatively small negative retention shifts are observed for methionine residues located in the hydrophilic face of an amphipathic helix (similar to Figure 2.3 c, d). For example, EAITMVVNDMELLSAVTK (*C. butyricum*) with two methionines additionally highlights the

difference between groups (c) and (d) in Table 2.2. Met_5 is located in the hydrophilic and Met_{10} in the hydrophobic face of the amphipathic helix. Retention time shifts upon oxidation of the same parent peptide were found to be –1.28 and –4.17 % acetonitrile for these two positions, respectively. This example indicates that chromatographic behavior may clearly distinguish between peptides with different positions of a modified residue.

e) Peptides exhibiting a small positive retention time shift contain a methionine residue in the N3 position of an N-capping box stabilization motif preceding the amphipathic helix. This group of peptides found in real samples is very small: three cases from 6421 pairs. However, they highlight the extreme importance of structural features for understanding effects of peptide helicity. I further characterized this unexpected behavior using designed synthetic sequences (see section 2.3.4).

Detailed inspection of oxidized / non-oxidized peptide pairs may reveal additional features related to the effects of secondary structure. For example, the amphipathic peptide QMFQQMQQDMQR selected from *T. thermohydrosulfuricus* WC1 carrying three Met residues was found in all three different oxidation states. All the methionines are located in the hydrophobic face of the helix, however exhibiting different shifts upon oxidation: -3.4, -5.0 and -3.7 % acetonitrile for positions 2, 6, and 10, respectively. This implies that residues located in the middle of a helical stretch (e.g. Met_6) are more involved in stabilization of the molecule upon interaction with a hydrophobic surface compared to terminal ones.

2.3.4 Positive Retention Time Shift upon Oxidation of Methionine in the N3 Position of the N-cap Box Motif: Confirmation Using Synthetic Peptides

It has been shown that, similar to helices in proteins, N-cap box motifs preceding amphipathic helical structures significantly increase chromatographic retention time, possibly due to helix stabilization [77]. This stabilization occurs due to reciprocal hydrogen bonding interactions of the side chain of the N-cap residue (according to Richardson and Richardson terminology [4]) with the backbone –NH group of an N3 residue, and vice-versa; the side chain of N3 residue with the –NH backbone group of the N-cap amino acid. This motif consists of four residues with each position showing different preferences for individual residues. First position (N-cap) is preferable for Asn, Asp, Ser, and Thr. Second (N1) and third (N2) positions favor hydrophobic residues. Fourth (N3) position is predominantly occupied by Glu, Gln and Asp.

Three out of the five peptides (SLEMELESLQGNAVR, TLQMLITELSPDTR, TLEMILADTESEQYK) with positive retention time shifts shown in Figure 2.3 (a) and (b) are amphipathic and contain a common sequence motif X-Leu-Z-Met, where X is represented by Ser or Thr and Z by Glu or Gln. These three peptides start with typical N-cap residues (Ser, Thr), followed by hydrophobic Leu (N1) and have a Met residue in position 4 (N3): SLEM, TLQM, TLEM. I believe that positive retention time shifts observed after methionine oxidation in these three natural peptides are caused by the better ability of the Mso residue to form a hydrogen bond with the N-cap –NH group compared to non-modified methionine. This would further stabilize an amphipathic helix.

The above hypothesis was confirmed by analyzing retention behavior of oxidized/non-oxidized pairs for four synthetic peptides carrying predicted N-cap motifs with a Met residue

placed in the N-cap (hydrophilic face), N1 (hydrophobic face), N3 (hydrophilic face) or in the middle of an amphipathic helical stretch (hydrophobic face of the helix). The peptide GS**IV**EE**I**EE**VI**GEGER with an optimal composition of N-cap box motif (SIVE) preceding two turns of amphipathic sequence (**IV**EE**I**EE**VI**) was chosen as a template. To additionally increase helical interactions, amino acids with small compact aliphatic chains were used to construct the hydrophobic face of the helix (isoleucine (Ile, I) and valine (Val, V)), these single letter amino acid codes are shown in bold in the sequence. Figure 2.4 shows respective extracted ion chromatograms and axial helical projections for these sequences.

GMIVEEIEEVIGEGER has a Met residue in the N-cap position, placed in the hydrophilic face of the helix (Figure 2.4 a). This explains the relatively low (-1.63 % acetonitrile) decrease in retention upon oxidation to Mso and no peak splitting for Mso diastereomers.

GSMVEEIEEVIGEGER sequence has a Met residue shifted into the next position compared to the previous molecule. However it was enough to move it into the hydrophobic face of the helix (Figure 2.4 b), causing diastereomers' peak splitting and a significantly larger decrease in peptide hydrophobicity upon oxidation (-5.51 % acetonitrile).

GSIVMEIEEVIGEGER contains a Met residue in the N3 position of the N-cap box motif (Figure 2.4 c). Hydrophobicity of the oxidized peptide is higher (+0.36 %), which confirms the findings made based on the analysis of natural tryptic digests in Table 2.2, part (e).

GSIVEEMEEVIGEGER features a Met residue halfway along the linear sequence of an amphipathic sequence (IVEEMEEVI) and it is located in the middle of the hydrophobic face in axial helical projection. This position plays the most important role in the stabilization of an amphipathic helix on a C18 surface. Replacement of Met for the more hydrophilic Mso

decreases retention time dramatically (-9.05 % acetonitrile) and gives baseline resolution of Mso diastereomers (Figure 2.4 d).

These four examples show how knowing the rules of helices interactions with the hydrophobic C18 surface allowed us to design peptides exhibiting extreme values of retention shifts, both positive and negative.

This section gives an explanation for three Mso-containing peptides that show positive retention time shifts out of five, whereas the reason for similar behavior for other two peptides still remains unclear.

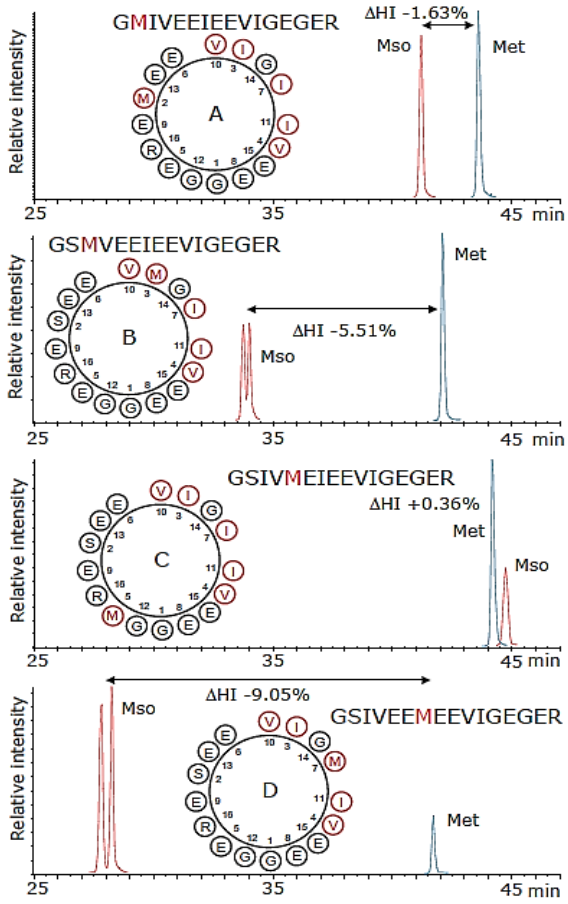


Figure 2.4 Retention time shifts associated with methionine oxidation to Mso depending on the position of Met within N-cap box motif and inside amphipathic stretch of the GSIVEEEIEEVIGEGER model peptide. Extracted ion chromatograms and axial helical projection for **(a)** – GMIVEEIEEVIGEGER; **(b)** – GSMVEEIEEVIGEGER; **(c)** – GSIVMEIEEVIGEGER; **(d)** –

2.4 Conclusions

For the first time I attempted a large-scale systematic study of the RP HPLC behavior of peptides carrying oxidized methionine residues. In addition to the commonly accepted phenomenon of lowering the hydrophobicity of the methionine side chain upon oxidation, I focused on the retention mechanism and effects of peptide secondary structure. Oxidation of Met-containing peptides to methionine *S*-oxide (+16 Da) is the most typical modification, which results in –2.37 % acetonitrile retention time shifts on average, compared to the non-modified Met analogs. Further oxidation of the methionine *S*-oxide to methionine *S,S*-dioxide (+32 Da) increases retention time slightly with an average retention time shift of -2.1 % (compared to the Met analogs). The former represents the most typical modification observed in bottom-up proteomics and was the major focus of this study. The latter is relatively rare, and typically observed when oxidation is used in sample preparation protocols.

A detailed overview of the chromatographic behavior of peptides carrying oxidized methionine revealed an extremely complex picture in which multiple interactions (hydrophobic, electrostatic, hydrogen bond) affect the change of the peptide hydrophobicity. The amplitude of this retention time shift may vary dramatically: from –9 % to +0.36 % acetonitrile. Stabilization of amphipathic helical structures upon interaction with the hydrophobic C18 surface (i.e. secondary structure effect) is the main reason for such variation. Methionine is a moderately hydrophobic residue, which may be positioned on the hydrophobic face of the amphipathic helix – preferred binding domain of an amphipathic helical molecule to the C18 surface. Any alteration (Met oxidation in this case) in this domain causes significant change in a peptide's retention time. Due to the presence of a stereogenic center in the Mso structure, this also leads to

partial or complete resolution of peaks corresponding to Mso diastereomers. Moreover, for the first time, I demonstrate that methionine oxidation may increase peptide retention time. This phenomenon is also related to amphipathic helicity. When placed in the N3 position of the N-cap box motif, the side chain of Mso serves as a better partner for hydrogen-bond interactions, which additionally stabilizes helical structure.

Specific properties of methionine (e.g. hydrophobic, prone to oxidation) make it a key in understanding the retention mechanism and developing better peptide retention time prediction models. Collection of representative data sets for modeling the retention times of peptides with Mso residues was the major focus of this part of thesis. Incorporation of established effects into the SSRCalc model is on-going and will be presented in a manuscript under preparation.

Chapter 3

Characterization and Complete Separation of Major Cyclolinopeptides in Flaxseed Oil by Reversed-phase Chromatography*

*Published as Lao, Y.W., Mackenzie, K., Vincent, W., Krokhin, O.V. (2014). Journal of Separation Science 37, 1788-1796.

3.1 Literature Review

3.1.1 Protein Oxidation in Food Products

Oxidation is of serious concern in the food industry due to the fact that manufacturing processes and storage conditions facilitate oxidation. Protein oxidation in foods often leads to loss of digestibility and biological value [78]. In particular, the interaction of proteins with oxidizing lipids is one of the most important oxidative pathways encountered in food products. Schaich [79] reviewed the chemistry of lipid oxidation and the transfer of free radicals to proteins from oxidizing lipids. Roubal and Tappel [80] also reported substantial losses in amino acids when proteins were exposed to lipid hydroperoxides which are products of lipid oxidation. They observed that methionine, histidine, cysteine and lysine were the most vulnerable to damage [80]. Particularly, Tannenbaum et al. [81] reported that methionine residues in proteins were oxidized to methionine *S*-oxide in the presence of oxidizing lipid.

3.1.2 Bitter Taste in Flaxseed Oil is linked to Methionine Oxidation of Major Cyclolinopeptides

Canada is a main producer and exporter of flax, accounting for approximately 40 percent of the world's production [82]. The oil obtained from flaxseed contains high levels of polyunsaturated fatty acids, 40-68% of which is α-linolenic acid (ALA, a particular form of plant based essential omega-3 fatty acids) [83]. The high content of ALA and overall ratio of omega-3 to omega-6 fatty acids makes flaxseed oil a very attractive dietary product containing essential fatty acids [53]. Flaxseed oil has been widely used as an additive to formulated supplements, as well as a stand-alone dietary product.

Fresh flaxseed oil has a smooth buttery taste. However, during storage the dietary quality of the product begins to deteriorate, developing an unpleasant bitter taste [52]. A recent study by Bruhl et al. [52] had identified cyclolinopeptide E (cyclo-MsoLVFPLFI) – CLE, (where Mso represents methionine oxide) as a principal bitter component in stored cold pressed linseed oil. Oxidation of flaxseed oil is a major commercial concern, and is likely caused by the chemical degradation process of oxidative rancidification, occurring when unsaturated fatty acids auto-oxidize via free radical reactions into hydroperoxides, cyclic peroxides, and epoxides [54]. These primary products either degrade into secondary products (such as saturated or unsaturated aldehydes and ketones) [84] or react further with protein-bound amino acids. In the case of flaxseed oil, the presence of highly reactive double bonds in the omega-3 fatty acids makes the interaction of oxygen with the unsaturated fatty acids highly probable. The amino acid methionine is the most likely to be oxidized by the primary products. Initial work suggests that oxidation stability depends on temperature, storage time and exposure to ambient light.

Quantitation of CLE by reversed-phase HPLC was used as an indicator of oil deterioration in assessment of shelf-life of this product [55]. These findings also prompted further studies on the effect of cyclolinopeptides on oxidative stability of flaxseed oil [85].

3.1.3 Overview of Cyclolinopeptides Discovery

Cyclolinopeptides (CLPs) are a group of naturally occurring hydrophobic cyclic peptides derived from flaxseed oil (*Linum usitatissimum* L.). They consist of eight or nine amino acid residues with molecular weight 960-1100 Da (Table 3.1 (p. 86), Figure 3.1). The isolation and structural elucidation of cyclolinopeptide A was first reported in late 1950's by Kaufman and Tobschirbel [86]. Subsequent isolation and characterization work led to the discovery of additional members of the cyclolinopeptide family, with an alphabetical naming convention based on the order of discovery [87-91]. Figure 3.1 shows sequences and oxidative pathways of all major cyclolinopeptides containing non-modified methionine and methionine *S*-oxide. The most abundant CLA does not contain methionine, CLB-CLL-CLM – one methionine residue, and CLN-CLO – two methionine residues. The second group of peptides, upon oxidation, yields a single product – CLC-CLE-CLD, respectively. Peptides with two methionines could be converted into two positional isomers each with one methionine *S*-oxide (CLT, CLI, CLH and CLP), and finally into peptides with two methionine *S*-oxide residues (CLF, CLG).

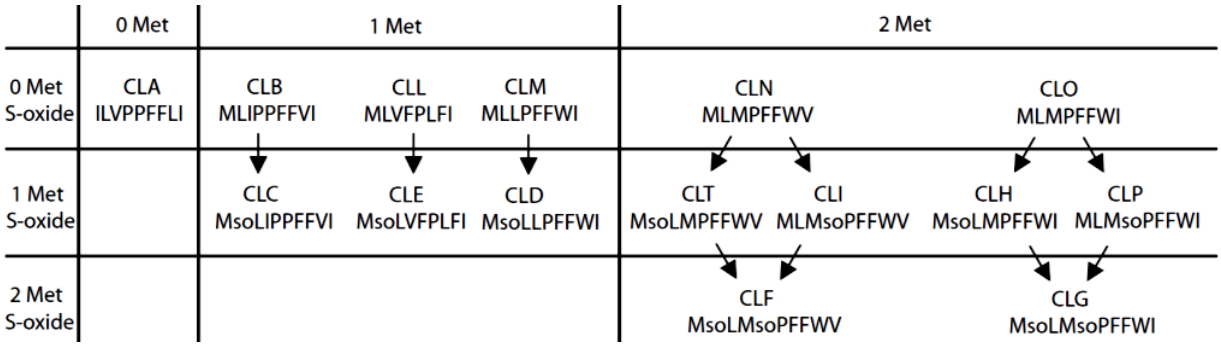


Figure 3.1 Amino acid sequences and oxidative pathways of cyclolinopeptides.

The order of CLPs discovery was driven in part by their abundance and chemical stability. CLA is the most abundant and only stable component due to lack of methionine in its amino acid sequence. CLB – CLI were subsequently isolated and structurally characterized using NMR and MS methods between 1997 and 2001 [87-90]. CLB is the only component with a non-modified methionine among them, and is the second most abundant cyclolinopeptide [87]. The existence of non-oxidized counterparts (L, M, N, O) in fresh flaxseed oil was hypothesized, and confirmed by direct MS and MS/MS measurements [90,91]. Two peptides carrying methionine *S,S*-dioxide, Msn (products of further oxidation of CLC and CLE) were characterized recently in detail [92]. These peptides demonstrate chemical stability [92] due to completion of methionine oxidation. The same study has shown that cyclolinopeptides containing methionine *S*-oxide can be chemically transformed into peptides containing methionine or methionine *S,S*-dioxide through chemical reduction or oxidation, respectively [92]. Discovery of new members of the cyclolinopeptide family continues: in early 2014 Okinyo-Owiti et al. [93] reported characterization of the CLQ (cyclo-MLKPFFFWI), CLR (cyclo-MsoLKPFFFWI), and CLS (cyclo-GIPPFWLTL).

The natural function of cyclolinopeptides in plants is still unknown. However, *in vitro* studies demonstrated that these peptides have biological activity [94] with immunosuppressive and anticancer effects, indicating that these compounds have potential pharmaceutical applications. CLA possesses immunosuppressive activity similar to cyclosporine A [95]. In addition, CLA, CLC and CLE can induce apoptosis in a human lung epithelial cancer cell line [96]. These findings prompted significant interest towards the preparative isolation of stable forms of CLPs. Reaney and co-workers have published extensively on the analytical characterization of these peptides, their distribution within the seed, and quantitation across different flax cultivars [97,98].

3.1.4 Literature Review of RP HPLC Analysis of Cyclolinopeptides

The literature indicates that the content of CLPs extracts can differ significantly depending on the sample preparation protocols. Gui et al. [97] employed solid phase extraction to isolate the peptide-containing fractions on silica-gel using an acetone-based eluent. RP HPLC analysis of these extracts showed the presence of CLF, CLG, CLC, CLE and CLA (listed in the order of their retention times on C18 sorbent). Lower amounts of CLB eluting prior to CLA were found for some isolates. Olivia et al. [99] used water-methanol extracts from flaxseed oil and reported RP HPLC detection of the same compounds plus CLD. Bruhl et al. [52] used RP HPLC for the analysis of pre-fractionated water-methanol extracts from flaxseed oil and found five major peaks on the chromatogram corresponding to CLF, CLG, CLC, CLE and CLA. All these studies have detected oxidized species and stable CLA as major components of mixtures. At the same time, MS and MS/MS analysis of methanol extracts from freshly ground flaxseed by

Stefanowicz [91] showed an almost complete set of CLPs except for CLF, CLG, CLI and CLT. The absence of doubly-oxidized CLF and CLG, and the presence of methionine-containing peptides indicate that the original oxidation state of the CLPs was better preserved in this study. Recently Aladedunye et al. [100] used ultrasonic methanol extraction of CLPs from a sample of cold-pressed flaxseed oil and the high-efficiency core-shell Kinetex^{TM} C18 column to show separation of 14 components including all non-oxidized CLPs.

CLE was determined [52] as the major bitter component in stored flaxseed oil, thus detection and quantitation of CLE and its non-oxidized counterpart CLL is critical for determination of the stability of flaxseed oil. Development of a rapid, reliable and inexpensive RP HPLC method using ultraviolet (UV) detection would provide a significant advantage in an industrial process control setting. Major CLPs are not fully resolved under typical reversed-phase LC conditions: C18 chemistry and water/acetonitrile gradients. Thus, Jadhav et al. [92] concluded that CLL, a major target in the study of oxidation, always overlaps with the very abundant CLA. Recent work by Aladedunye et al. [100] used a high-efficiency core-shell Kinetex^{TM} C18 column and water/acetonitrile gradient with the addition of isopropanol. They achieved partial resolution of the CLA-CLL pair and reported separation of 14 major components.

In this study, the goal was to apply high-efficiency core-shell reversed-phase sorbents of different chemistry (C8, C18, Phenyl-Hexyl) to achieve the separation of the major cyclolinopeptides responsible for the onset of bitter taste in flaxseed oil during storage. I also used MALDI MS and electrospray ionization mass spectrometry (ESI MS) measurements

coupled with RP HPLC to confirm the sequences of previously reported identifications of major cyclolinopeptides.

3.2 Experimental

3.2.1 Materials

Deionized water, HPLC-grade acetonitrile and methanol (EMD, Mississauga, ON) were used for the preparation of the eluents and extraction solvents. Trifluoroacetic acid (TFA), formic acid and dithiothreitol were sourced from Sigma Aldrich (St-Louis, MO). The cold-pressed flaxseed oil was provided by ShapeFoods Inc. (Brandon, Manitoba, Canada), with each batch packaged into 250-mL volume dark-glass bottles with an argon-filled headspace and stored for one week at 4 ˚C prior to analysis.

3.2.2 Extraction of Cyclolinopeptides

An aliquot (500 μL) of the flaxseed oil was added to 500 uL of 3:2 methanol/water containing 10 mM dithiothreitol to prevent oxidation of the cyclolinopeptides during the extraction process. This mixture was shaken in a vortex mixer for 20 minutes, followed by five minutes centrifugation at 4000 g. The methanol/water supernatant was extracted for immediate MALDI MS or HPLC analysis.

3.2.3 Matrix-Assisted Laser Desorption/Ionization – MS Measurements

The cyclolinopeptide extracts and RP HPLC fractions were analyzed using matrix-assisted laser desorption/ionization (MALDI) quadrupole/TOF (QqTOF) prototype mass

spectrometer (University of Manitoba/ABSciex, ON, Canada) [101]. Samples were mixed 1:1 with 2, 5-dihydroxybenzoic acid MALDI matrix solution (150 mg/ml in 1:1 water/acetonitrile) and deposited onto a stainless steel target. Peptide sequence identity was confirmed by high-accuracy mass measurements (10 ppm) in both MS and MS/MS modes.

3.2.4 Chromatographic Conditions

A micro-Agilent 1100 Series system (Agilent Technologies, Wilmington, DE) was used with a manual injector and VWD G1314A detector set to 214 nm. All separations were carried out at room temperature (22 - 24 ˚C). Preliminary experiments for RP HPLC separations of oil extracts (10 µL injection) were performed using a Luna^{TM} C18(2) column (100 × 1 mm, 5 µm, 100 Å, (Phenomenex, Torrance, CA)) at a 150 µL/min flow rate. A binary gradient of 1 % acetonitrile per minute (25 - 75% in 50 minutes) was applied with both eluents A (water) and B (acetonitrile) containing 0.1 % TFA as ion-pairing modifier.

High-resolution gradient RP HPLC separations of the oil extract were subsequently performed using a 50 µL injection and 0.5 mL/min flow rate. I employed four different separation columns: core-shell Kinetex^{TM} C18, C8 and Phenyl-Hexyl columns (150 × 4.6 mm, 2.6 µm, 100 Å, Phenomenex) and Luna^{TM} C18 (2) column (150 × 4.6 mm, 3 µm, 100 Å). A linear gradient slope (30 – 90 %) of 3 % acetonitrile per minute was applied with both solvents containing 0.1 % TFA.

A nano-flow 2D LC Ultra^{TM} system (Eksigent - ABSciex, Dublin, CA) with 10 µL sample injection via a 300 µm × 5 mm PepMap100 (Thermo Fisher Scientific, Rockford, IL) trap-column and a 100 µm × 200 mm analytical column packed with 5 µm Luna^{TM} C18(2) was

used for the HPLC-ESI MS experiments. Both eluents A (water) and B (acetonitrile) contained 0.1 % formic acid as an ion pair modifier. A linear gradient of 2 % acetonitrile per minute was used (0 – 80 % in 40 minutes). Cyclolinopeptide extracts were diluted by a factor of ~ 100 with 50 % methanol prior to sample injection.

3.2.5 Electrospray Ionization Mass Spectrometry Measurements

A TripleTOF^{TM}5600 mass spectrometer (ABSciex) with NanoSprayIII source was used in standard MS/MS data-dependent acquisition mode. 250 ms survey MS spectra were collected (*m/z* 400 – 1500) and followed by up to 20 MS/MS measurements on the most intense parent ions (300 counts/sec threshold, +1 – +2 charge state, *m/z* 100 – 1500 mass range for MS/MS, 100 ms each). Previously targeted parent ions were excluded from repetitive MS/MS acquisition for 12 seconds.

3.2.6 Accelerated Aging (Oxidation) of Flaxseed Oil Samples

A fresh sample of flaxseed oil was subjected to accelerated aging using the Schaal oven storage stability test methodology [102]. One-day storage at 68 ˚C (Isotemp oven Model 737F, Fisher Scientific) is the equivalent of 1 month on the shelf at room temperature for flaxseed oils. Several aliquots were taken (1 – 36 days) to monitor the oxidation degree of cyclolinopeptides, which corresponds to up to 3 years of storage at room temperature.

3.2.7 Oxidation of Fractionated Peptide Samples to Confirm the Presence of Diastereomeric Pairs of Peaks

Fractions containing cyclolinopeptides G, P, P* and A were collected during an RP HPLC separation using a Phenyl-Hexyl column (150 × 4.6 mm, 2.6 μm, 100 Å, Phenomenex). A linear gradient slope (30 – 90 %) of 3 % acetonitrile per minute was applied with both eluents A (water) and B (acetonitrile) containing 0.1 % TFA. 1.5 mL plastic vials containing CLG, CLP and CLP* were open to air at room temperature for three weeks. 200 μL of 7:3 acetonitrile:water were added to each vial daily to replace the evaporated solvent and keep the peptides in solution. Aliquots of these samples were collected every week, and analyzed by LC-ESI MS (~1/20 dilution with 1:1 methanol:water). These experiments were performed using the nano-LC-ESI MS system (as used in section 3.2.4 – 3.2.5) with a 30 min 30 – 90 % acetonitrile gradient. Non-oxidized CLA was added to each sample as an internal retention time standard.

3.3 Results and Discussion

3.3.1 Detection of Major Cyclolinopeptides in Flaxseed Oil Extracts Using MALDI MS

My initial elucidation of cyclolinopeptide structures started with direct MALDI-MS analysis of the extracts. Figure 3.2 a and b shows the MALDI spectra corresponding to extracts from a fresh and moderately aged (5 days storage at 68 ˚C) flaxseed oil samples. Peaks corresponding to CLL and CLE differ by a nominal mass of +16 Da, indicating that CLE is indeed an oxidized form of CLL. This is also the case for the pair of CLB and CLC, where an oxidation in CLB also yields a mass shift of +16 Da to produce CLC. Comparison of relative peak intensities indicates an increase in oxidized species (CLE, CLC) and corresponding

decrease in original CLL and CLB species as a function of storage time (Figure 3.2 a, b). These variations can be also verified by comparing peak intensities against CLA, which contains no oxidizable Met residues.

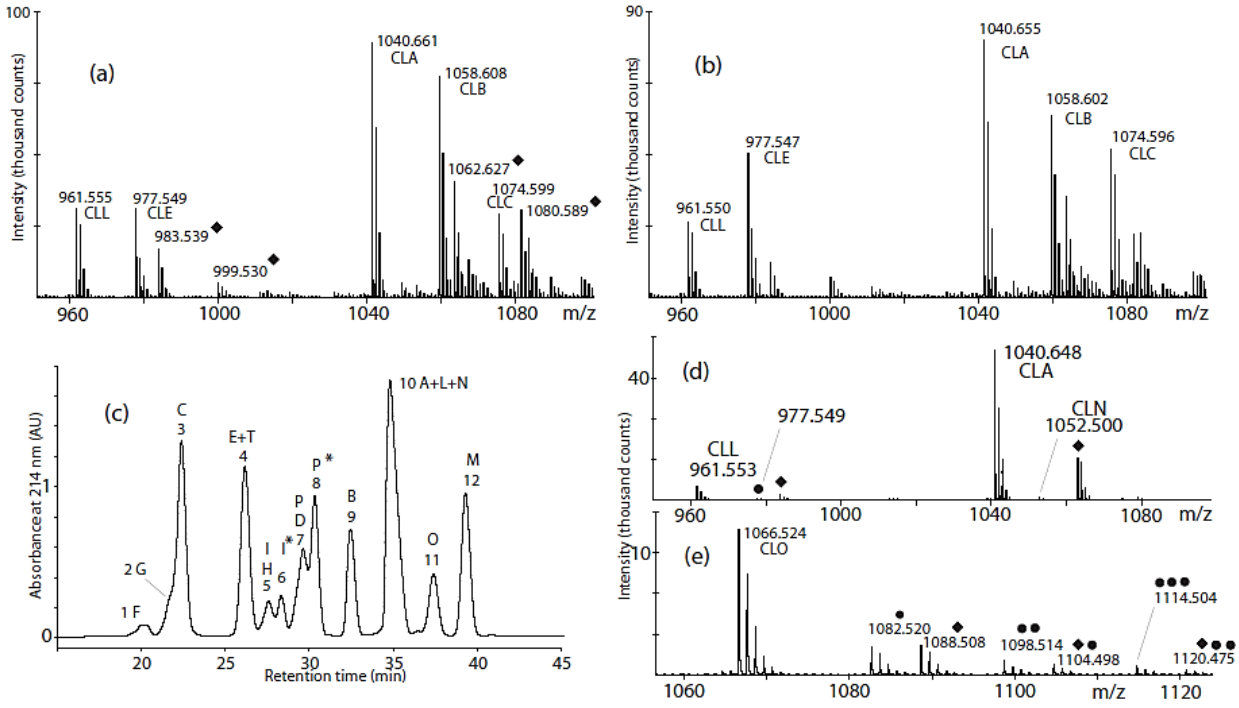


Figure 3.2 MALDI MS and RP HPLC-MALDI MS analysis of flaxseed oil extracts. (a) and (b) are MALDI spectra corresponding to the extracts from fresh and moderately aged flaxseed oil samples, respectively. (c) RP-LC chromatogram of a fresh oil extract using a Luna C18 column (100×1mm, 5mm, 100Å). 12 fractions corresponding to major peaks were collected and analyzed by MALDI MS. The second peaks in diastereomeric pairs of CLP and CLI are labeled with an asterisk. (d) and (e) are MALDI MS spectra corresponding to Fractions 10 and 11. Oxidation products of (+16 Da) and sodium adducts (+22 Da) are labeled with a solid circle and a square, respectively.

3.3.2 Off-line Reversed-phase HPLC/MALDI MS Analysis

Detection of the remaining minor CLPs (1048-1100 Da, Table 3.1 (p. 86)) using direct MALDI MS is problematic due to the overlap of isotopic peaks from abundant components and their sodium adducts. These interferences can be eliminated using RP HPLC pre-fractionation followed by MALDI MS of individual fractions. Figure 3.2 (c) shows the initial RP-LC separation of cyclolinopeptides from an oil extract using a Luna^{TM} C18 column (100×1mm, 5μm, 100Å) and a 1 % acetonitrile per minute gradient. Each numbered fraction on the chromatogram was analyzed by MALDI MS (examples shown in Figure 3.2 d, e). The peaks from various CLPs (Table 3.1, p. 86) were initially identified by accurate mass measurement, and their sequence identity confirmed by MS/MS. Figure 3.2 (c) and Table 3.1 (p. 86) show the peak assignments, confirming the identification of all previously reported CLPs.

Overall, the initial RP separation confirmed the previously reported elution ordering of major CLPs on C18 columns: F < G < C < E < B < A [52,99,100]. Cyclolinopeptides can also be divided into three groups based on their elution order: CLF and CLG with two Mso residues elute first, peptides with one oxidation elute second, and non-modified ones represent the last group of analytes.

3.3.3 Chromatographic Behavior of CLP-CLH and CLI-CLT Pairs: Separation of Positional Isomers and Diastereomers

Two components of the peptide mixture, CLN and CLO, contain two Met residues (Figure 3.1). One could expect each to yield two oxidized products of identical mass with different positions of the Mso residue: CLO will be converted to CLP (cyclo-MLMsoPFFWI)

and CLH (cyclo-MsoLMPFFWI) – both at *m/z* 1082.52; CLN to CLI (cyclo-MLMsoPFFWV) and CLT (cyclo-MsoLMPFFWV) – at *m/z* 1068.51. An identification of CLT has not been reported to date, so I chose to test the hypothesis of its formation upon oxidation of flaxseed oil. Aladedunye et al. [100] reported separation of three out of these four components with elution order of CLI < CLH < CLP. The RP HPLC – MALDI MS analysis detected both 1082.52 (CLH, CLP) and 1068.51 ions (CLI and CLT) in three different fractions (Table 3.1 (p. 86), Figure 3.2 c). MS/MS analysis of these fractions (not shown) confirmed that CLH is present as a minor component in Fraction 5, and CLP is a major component in Fractions 7 and 8. Similarly, CLT was found in Fraction 4, while CLI was a major component in Fractions 5 and 6. The second peaks of CLP and CLI are labeled with * in Figure 3.2 c. Keeping in mind that the observed peak splitting could be a product of inconsistent fraction collection and peak overlap, I proceeded with an on-line HPLC-ESI MS measurement providing both continuous extracted ion chromatograms (XIC) profiles and simultaneous MS/MS based sequence confirmation.

Figure 3.3 a shows the total ion ESI MS chromatogram of the cyclolinopeptide extract with retention order of the major components consistent with both my and others' [52,99,100] findings: F < G < C < E < B < A. XICs of both the 1068.51 and 1082.52 ions showed three peaks: a low abundance one eluting prior to the partially resolved pair of peaks (Figure 3.3 b, c). Additional proof of the correctness of my peak assignment from the previous section comes from the MS/MS analysis. MS/MS spectra provided in Figure 3.4 unequivocally identified the first 1082.52 peak in Figure 3.3 c as CLH (cyclo-MsoLMPFFWI), while the second and third are CLP/CLP* (cyclo-MLMsoPFFWI). Similar assignments were made for the 1068.51 ion (Figure

3.3 b): CLT (cyclo-MsoLMPFFWV) was followed by two split peaks of CLI/CLI* (cyclo-MLMsoPFFWV).

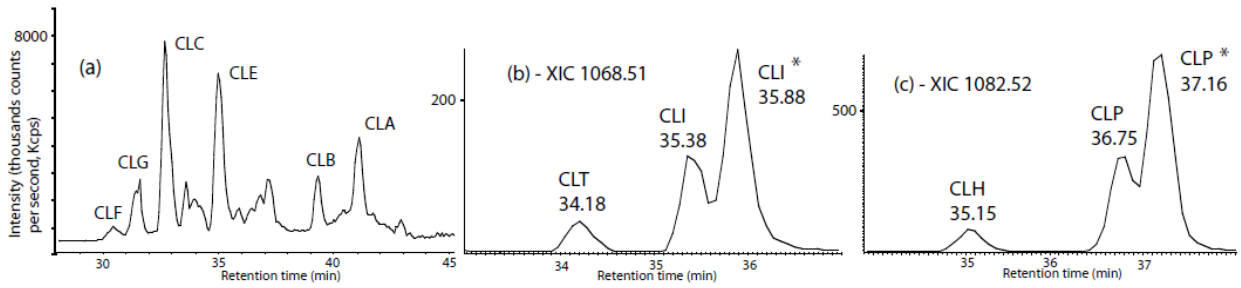


Figure 3.3 RP HPLC analysis of CLPs using ESI MS detection. **(a)** – total ion chromatogram of flaxseed oil extract obtained using LC-ESI MS; **(b)** – extracted ion chromatogram for *m/z* 1068.51 molecular ion. Sequence confirmation by MS/MS revealed that these peaks correspond to peptides derived from CLN: CLT, CLI and CLI*; **(c)** – extracted ion chromatogram for *m/z* 1082.52 ion. MS/MS analysis confirmed that these peaks correspond to oxidation products of CLO: CLH, CLP and CLP* (Figure 3.4).

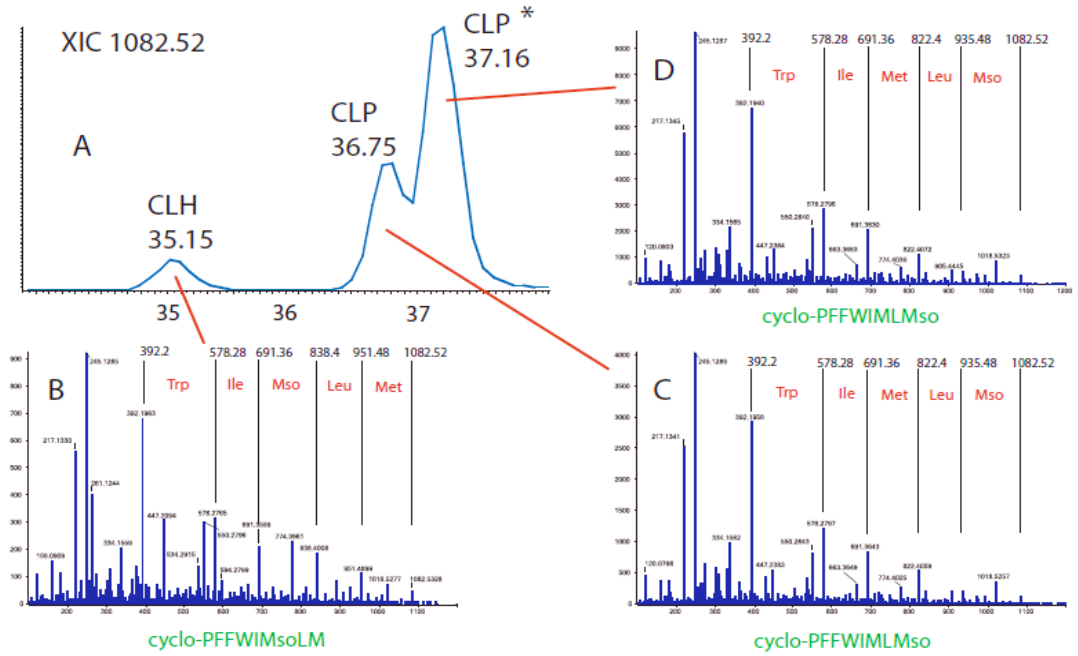


Figure 3.4 LC-ESI MS and MS/MS analysis of positional isomers and diastereomers for oxidation products of CLO: CLH and CLP. **(A)** extracted ion chromatogram for the singly charged molecular ion of *m/z* 1082.52. **(B)** MS/MS spectrum of CLH at 35 min. **(C)** MS/MS spectrum of CLP at 36.7 min. **(D)** MS/MS spectrum of CLP* at 37.1 min.

While the peak splitting upon methionine oxidation is common for Mso-containing amphipathic α-helices in RP HPLC (Chapter 2 of this thesis) [14], I believe the observation of it is the first example of such behavior for cyclic peptides. At this point the assignment is made based on MS/MS measurements, and requires additional confirmation by using alternative techniques.

3.3.4 Reversed-phase HPLC – Electrospray Ionization MS

On-line HPLC-ESI MS identified the entire repertoire of cyclolinopeptides (chromatogram in Figure 3.3 a, Table 3.1 (p. 86)). Modern LC-MS techniques allow for simultaneous identification by both accurate mass and MS/MS of all components of the cyclolinopeptide extracts. Figure 3.5 shows the MS/MS spectra from singly and doubly charged CLA ions in comparison with tandem mass spectra acquired during a LC-MALDI MS experiment. I also found a significant variation in ratios of intensities between 2+ and 1+ molecular ions for different cyclolinopeptides (Table 3.1, p. 86). Clearly, more hydrophilic peptides (CLF, CLG) exhibit much higher intensity of doubly charged ions compared to late eluting ones.

The high sensitivity of HPLC ESI-MS allowed us to confirm the identification of CLK and CLJ peptides containing methionine *S,S*-dioxide [92]. These two peptides are the products of further oxidation of CLC and CLE, and were retained longer than their singly-oxidized parents [92].

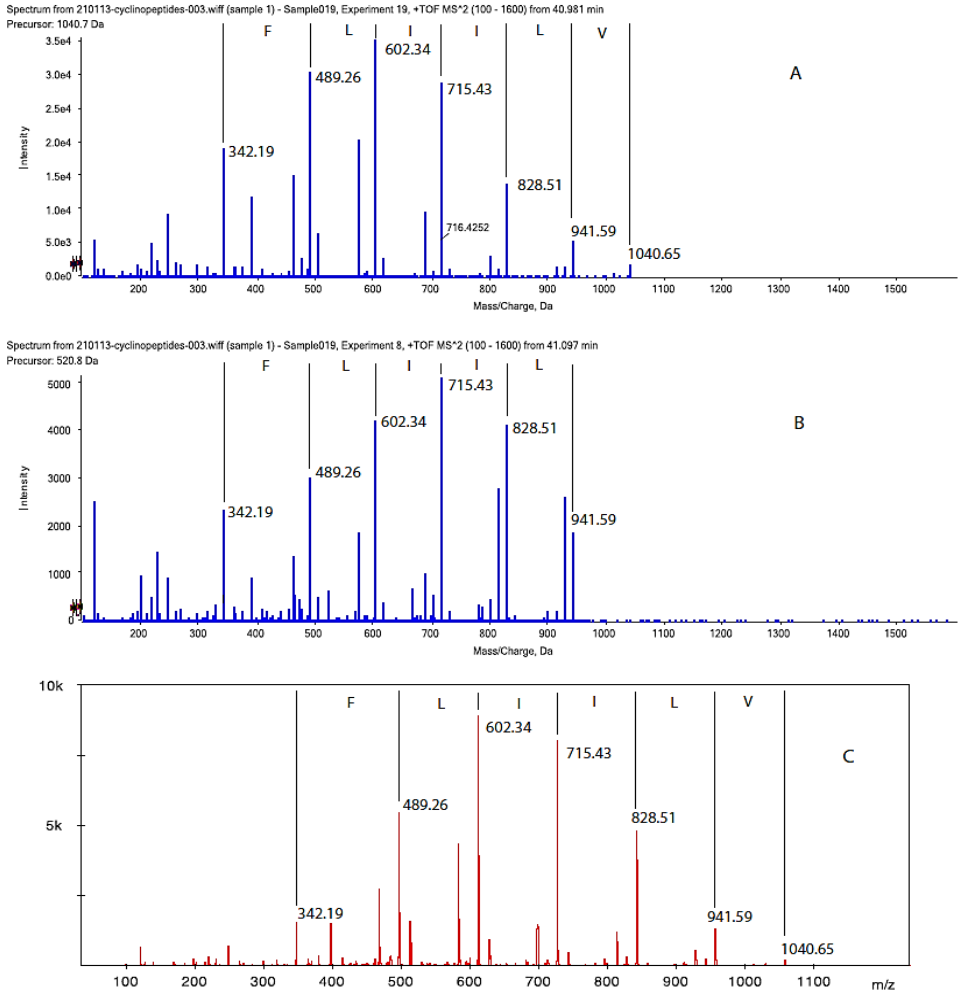


Figure 3.5 MS/MS spectra of cyclolinopeptide A using ESI and MALDI ionization techniques. **(A)** – LC-ESI MS/MS on +1 ion; **(B)** – LC-ESI MS/MS on +2 ion; **(C)** – LC-MALDI MS/MS on +1 ion.

3.3.5 Comparison of CLPs' Separation Selectivity on Different Reversed-phase Sorbents

Because of the high operational cost of modern mass spectrometers, my goal was the design of a RP HPLC procedure for the separation of major CLPs, allowing simple UV detection to determine the degree of peptide degradation. The variation in charge distribution of molecular ions during ESI MS detection (shown in Table 3.1, p. 86) is another argument in favor of UV detection. I have chosen to use core-shell Kinetex^{TM} 2.6 µm sorbents of different chemistry (C18,

C8, Phenyl-Hexyl) as well as standard porous C18 silica. Core-shell technology offers significant improvements in separation efficiency compared to commonly used C18 material, improving the chances to resolve closely eluting components. This work was performed in the summer of 2012, prior to the report of Aladedunye et al. [100] on CLPs separation on Kinetex™ C18 column. The authors employed a water-acetonitrile-isopropanol gradient and demonstrated partial resolution of the CLA – CLL pair, critical for monitoring the degree of CLL – CLE oxidation.

Figure 3.6 a-d show a comparison of separation selectivity between four columns of identical size using the same 3% per minute linear water-acetonitrile gradient. Major CLPs are labeled, except for the less abundant overlapping CLP, CLI, CLD, CLH and CLT. I can make several observations based on these data:

a) Application of core-shell technology significantly improves the separation efficiency of CLPs. The average peak widths at half maximum height for well-resolved CLC and CLB components were found to be 0.11, 0.08, 0.07 and 0.07 minutes for Luna™ C18, Kinetex™ C18, C8 and Phenyl-Hexyl, respectively;

b) Kinetex™ Phenyl-Hexyl is the only packing material exhibiting baseline separation of CLA-CLL peaks using water/acetonitrile gradients and capable of monitoring the degree of CLL-CLE conversion;

c) Overall peptide retention time is lower on core-shell sorbents due to a smaller carbon load compared to porous Luna™ C18(2) material;

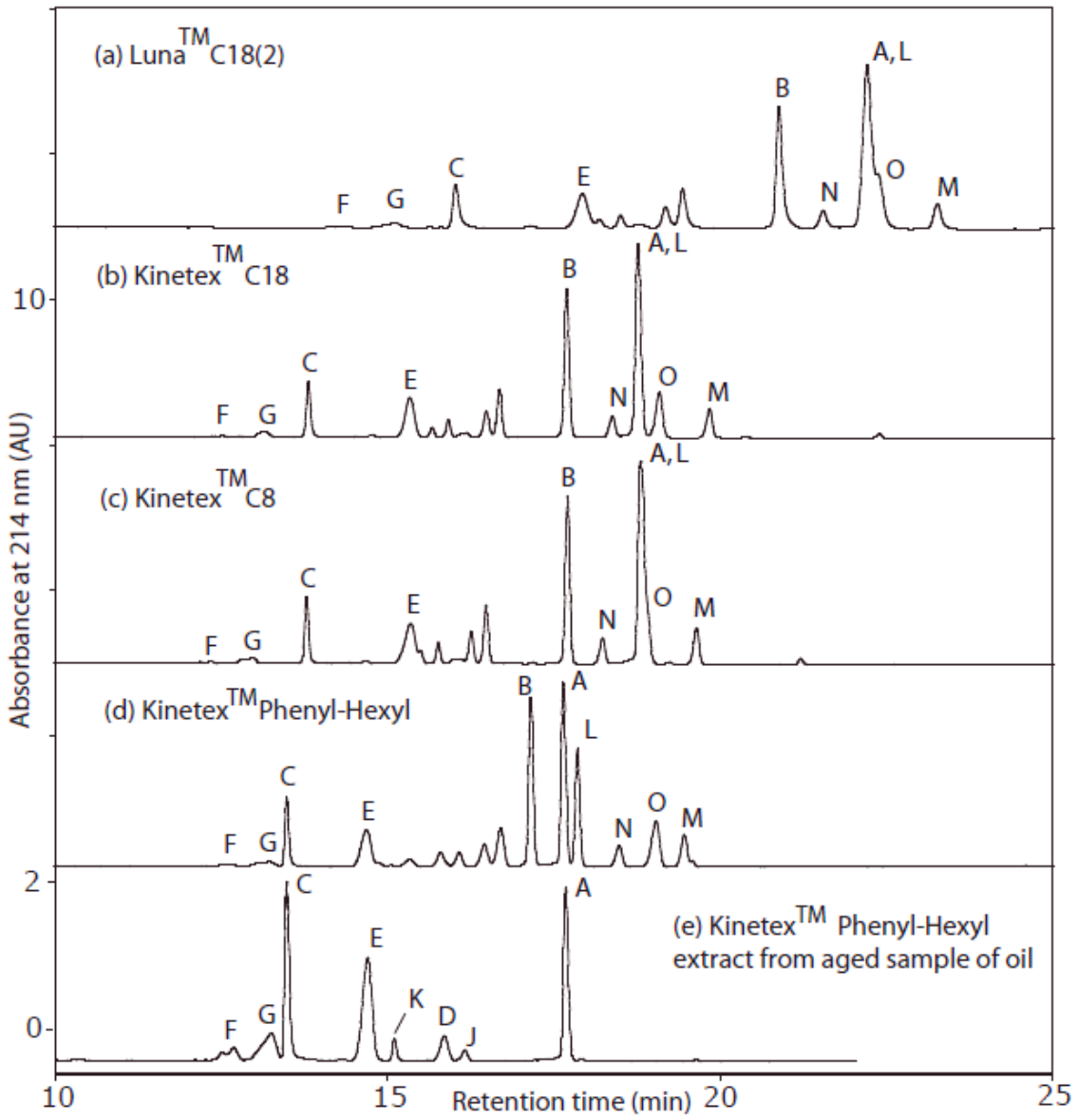


Figure 3.6 RP HPLC separations of CLPs using **(a)** Luna™ C18, **(b)** Kinetex™ C18, **(c)** Kinetex™ C8, and **(d,e)** Kinetex™ Phenyl-Hexyl columns with identical gradient settings: 0.5 ml/min flow rate; 3% per minute acetonitrile gradient. **(d and e)** – separations of extracts from fresh and aged flaxseed oil samples, respectively. Note complete oxidation of all Met-containing peptides **(e)** and the appearance of CLK and CLJ containing methionine *S*,*S*-dioxide - products of continuous oxidation of CLC and CLE.

d) Kinetex™ C8 and C18 exhibit very similar separation selectivity, while the Phenyl-Hexyl phase shows a higher affinity to peptides with larger numbers of aromatic residues. All non-oxidized CLPs carry two Phe residues, while only CLM, CLN and CLO have an additional Trp. This explains the relatively higher retention time of the latter and their oxidized derivatives on Kinetex™ Phenyl-Hexyl stationary phase. Clearly, the additional π-π interactions between the analyte molecules and phenyl functionality produce an alternative selectivity compared to alkyl phases;

e) Separations on all four sorbents show patterns corresponding to peak splitting for CLP and CLI.

I obtained further confirmation of the peak assignments shown in the previous sections by off-line RP HPLC MALDI MS analysis of CLPs extracts on the Phenyl-Hexyl column (Figure 3.7, Table 3.1 (p. 86)). Similar to the first separation on Luna™ C18 (2) (Figure 3.2 c), fractions were collected manually and analyzed by MALDI MS and MS/MS. The relative abundance of CLM was found to be higher in this batch of oil. Therefore the peak corresponding to its singly oxidized product CLD is also more intense, overlapping with CLI. It is also worth noting that the separation speed can be increased by adjusting the initial acetonitrile concentration and the slope of the gradient. Figure 3.8 shows a 10-minute separation of CLPs on Kinetex™ Phenyl-Hexyl column using 6% per minute gradient starting from 52% acetonitrile.

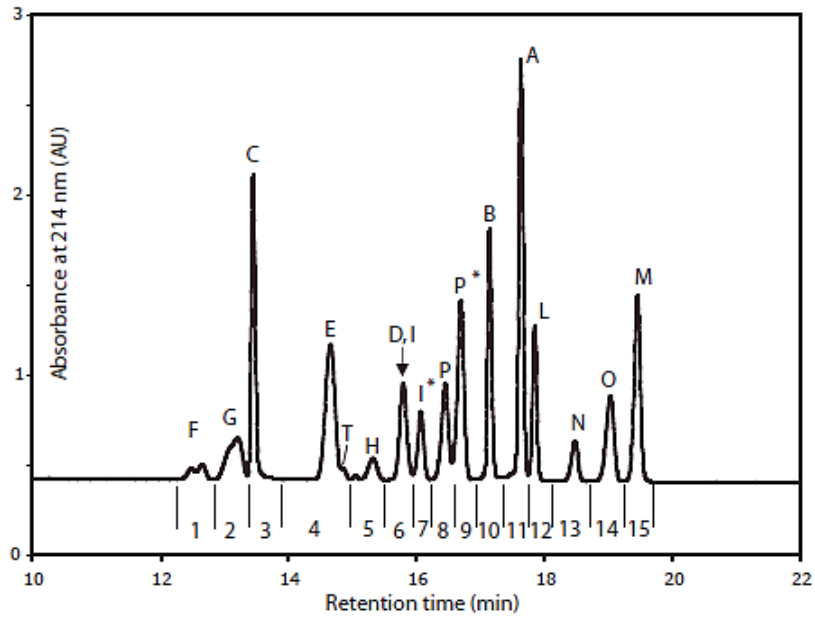


Figure 3.7 RP-LC MALDI MS analysis of flaxseed oil extract using a Kinetex™ Phenyl-Hexyl column. Separation conditions are identical to those in Figure 3.6. Peak assignment is also shown in Table 3.1, p. 86.

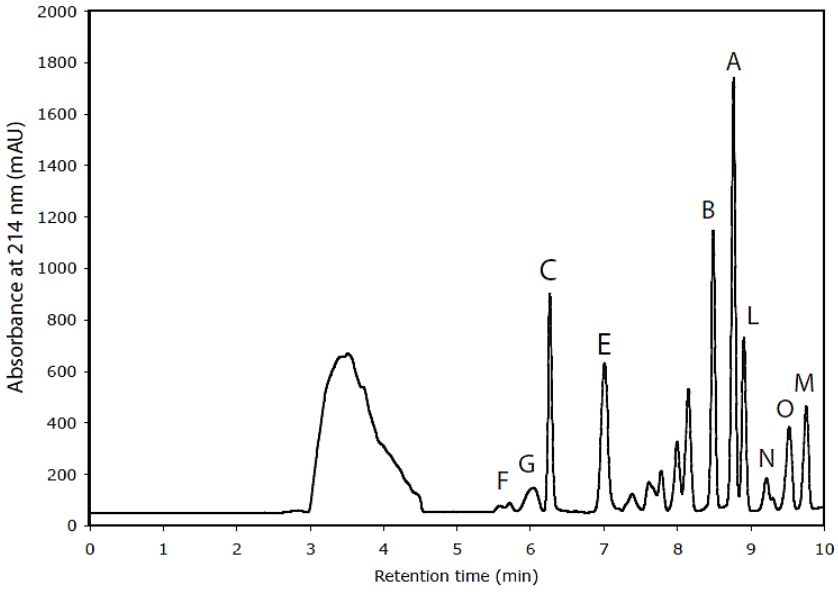


Figure 3.8 Fast separation of cyclolinopeptides using a Kinetex™ Phenyl-Hexyl column. Kinetex™ Phenyl-Hexyl column (150×4.6 mm, 2.6 mm, 100 Å, Phenomenex); linear water/acetonitrile gradient of 6% per minute (52-90% acetonitrile); 0.5 ml/min flow rate.

3.3.6 Analysis of Aged Flaxseed Oil and Estimation of Oxidation Degree of CLL and CLB

Analysis of the degree of oxidation in cyclolinopeptides is a key element in assessing flaxseed oil quality [55,100]. A sample of fresh oil was stored at 68 ˚C to monitor peptide oxidation. Oxidation manifests as a gradual decrease of non-oxidized Met-containing peptides, and an increase in CLC, CLE, CLD, CLG and CLF. Incomplete peptide oxidation, where only one of two potential Met is oxidized, occurs in samples with moderate periods of oxygen exposure, but is absent from aged samples. Figure 3.6 d and e show two chromatograms corresponding to fresh and oxidized oil (36 days storage at 68 ˚C). The latter contains CLF, CLG, CLC, CLE, CLD and CLA as well as two new peaks, which were identified as CLC and CLE analogs with further oxidation to methionine *S,S*-dioxide. This finding is consistent with a recent report by Jadhav et al.[92], which showed their higher retention time compared to Mso analogs. The same species (CLK and CLJ) were identified in the LC-ESI MS experiment with retention times higher than their Mso-containing parents (Table 3.1, p. 86). Based on these results, one would expect the formation of a CLD + 16Da analog, eluting close to the CLA peak, but I have not been able to detect it in fractions collected during the separation shown in Figure 3.6 e.

My results confirm that CLA, CLB and CLL are by far the most abundant species among six non-modified cyclolinopeptides. They are encoded by one gene as three separate peptides in a ratio of 1:1:1 [97]. Oxidation of the latter two yields CLC and CLE, respectively, which were found to be the most abundant components in oxidized oils, after stable CLA [98]. An earlier study by Bruhl et al. [52] identified CLE as a major contributor to the bitter taste of oxidized flaxseed oil. It is conceivable that other Mso-containing peptides may also contribute to bitterness of the oil; the high abundance of CLC makes it a good first choice to test this

hypothesis. Since its isolation can be achieved through RP HPLC separation, verification will require additional sensory evaluation experiments.

A batch of 24 bottles of flaxseed oil samples was subjected to an accelerated shelf life trial. According to the Schaal oven storage stability test methodology, one day of storage at 68 ˚C is equivalent to one month of accelerated shelf life at room temperature [102]. I estimated the oxidation status of these flaxseed oil samples by monitoring the ratios of peak areas for CLE/CLL and CLC/CLB pairs. RP HPLC separation using a Kinetex^{TM} Phenyl-Hexyl column allows for the complete resolution of these four components and, therefore, enables direct measurement using UV detection. The sum of peak areas $S_{CLC+CLB}$, $S_{CLE+CLL}$ represents the total amount of components in the respective pairs. Thus the degree of oxidation is simply:

"Oxidation degree CLL" = $S_{CLE}/(S_{CLE+CLL})$; (1)

"Oxidation degree CLB" = $S_{CLC}/(S_{CLC+CLB})$ (2)

I determined these values across 24 bottles of flaxseed oil samples stored for different periods of time, and found a very good (0.998 R^2-value) correlation between them, with the oxidation degree of the CLL/CLE pair found to be ~1.2 times higher than that for the CLB/CLC pair. Duplicate analysis was performed for each bottle to ensure reproducibility. The time dependence of oxidation degrees of CLL and CLB throughout 24 days of accelerated shelf life trial is shown in Figure 3.9. It is very important to note that these flaxseed oil samples were prepared and processed under the same conditions, for instance, the origin of the seeds, temperature setting during the oil press, as well as minimal light exposure to each bottle during the packaging process. Since flaxseed oil is a complex system containing many ingredients, varying any conditions mentioned above can directly modify the unsaturated fatty acids which in

turn affect the final amount of CLE and CLC through lipid oxidation. As an example, we found an unanticipated technological source of oxidation due to difference in the packaging process – variable headspace above the oil at the bottleneck. We established that the size of the air pocket above the oil correlates with the oxidation degrees of CLE and CLB to an extent. This may cause fluctuation in their oxidation values shown in Figure 3.9. Nonetheless, the oxidation degrees of CLE and CLB appeared to reach a plateau (around 80% oxidation) after five days of accelerated shelf life trial (Figure 3.9). In addition, oxidation degrees of CLL and CLB could decrease after long exposure to high temperature due to further oxidation into CLJ and CLK, respectively. Interestingly, Bruhl et al. [55] conducted a sensory test on flaxseed oil samples that were stored for different periods of time and found that bitterness continued to increase even after the content of CLE stabilized. This finding merits further investigation.

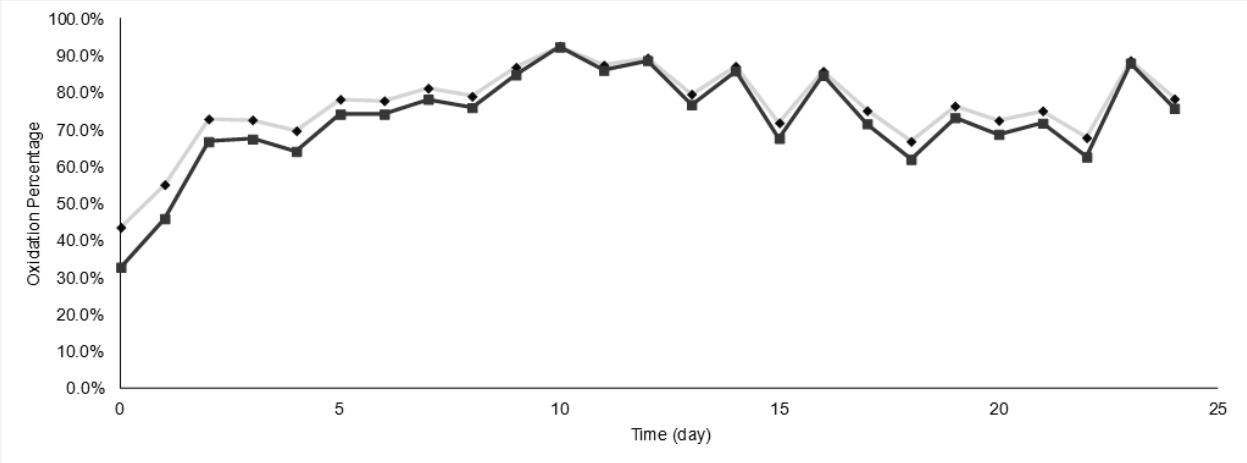


Figure 3.9 Oxidation degree of CLL and CLB during the accelerated shelf life experiment: storage at 68˚C for up to 24 days. Diamonds and squares represent the oxidation degree of CLL and CLB, respectively.

To further elucidate the reproducibility of the method, I performed 10 replicate analyses (both extraction and separation) of the same sample. The measurement of peak areas for CLL, CLE, CLB, CLC and CLA showed ~ 6 % standard deviation, whereas the relative computations of oxidation degree showed ~1% standard deviation.

3.3.7 The Effect of Peak Splitting on Separation Efficiency of Cyclolinopeptides

MALDI MS acquisition on the fractions from the Kinetex^{TM} Phenyl-Hexyl separation (Figure 3.7) confirmed peak splitting of CLI-CLI* and CLP-CLP* pairs. Moreover, their analogs carrying two Mso residues CLF and CLG exhibit the same feature, manifesting as a distorted peak shape. Aladedunye et al. [100] observed similar results for the latter pair, however the authors did not provide any explanation. All four different columns in this study showed similar peak shapes of CLF and CLG (Figure 3.6). This characteristic feature may also impact the separation efficiency of CLE; all four chromatograms in Figure 3.6 show a wider peak for this component.

The previous literature on the separation of CLPs does not discuss these effects, as most of these studies were performed for flaxseed oil with a higher degree of oxidation, and using C18 sorbents of lower separation efficiency. Cyclolinopeptides F, G, C, E, and A are normally observed under these conditions, and changes in peak shape are not always visible in low efficiency separations [52]. Similarly, it is hard to notice an increase in peak width of the CLE peptide when I perform the separation using a 10 cm long column packed with 5 μm sorbent (Figure 3.2 c). Olivia et al. employed high efficiency monolithic columns to optimize the separation of CLPs [99]; a detailed inspection of a chromatogram from a Chromolith SpeedROD

column shows an increase in the peak widths for CLF, CLG and CLE compared to CLC and CLA. Application of core-shell sorbents and high-quality fresh flaxseed oil in my work, and by Aladedunye et al. [100] provides deeper insights into the separation of CLPs. Fresh oil contains significant amounts of the CLI-CLI* and CLP-CLP* pairs, which show significant changes in retention times and can be separated even using conventional C18 columns (Figure 3.6 a). The higher separation efficiency of core-shell sorbents provides the resolution to observe variations in peak shapes across various CLPs. Thus, chromatograms obtained using Kinetex^{TM} columns clearly indicate the unique characteristics of CLF, CLG and CLE.

The MS/MS measurements on CLI-CLI* and CLP-CLP* pairs suggest the presence of two potential diastereomers. Taking the identical (except for Met oxidation) amino acid compositions of CLL, CLE and CLJ as an example (Figure 3.6 d, e), diastereomerization is the most likely mechanism leading to peak widening for CLE. It is visually apparent that both CLL (Met) and CLJ (Msn) peaks exhibit good separation efficiency: 0.08 and 0.09 min peak width at half maximum, respectively. CLE is an intermediate component, the only one containing a (Mso) residue capable of diastereomerization, and exhibits poor peak shape (0.16 min peak width at half maximum).

I performed additional experiments to confirm the effect of diastereomerization through continuous oxidation of individual members of the CLO oxidation cascade (CLP, CLP* and CLG). Fractions containing these components were collected during a separation similar to that shown in Figure 3.7, and subjected to oxidation with atmospheric oxygen at room temperature. All of these components feature an oxidized second methionine (Mso) in the original CLO sequence (cyclo-MLMPFFWI). I expected that as soon as second methionine in the original CLO

sequence (cyclo-MLMPFFWI) was further oxidized to methionine *S*,*S*-dioxide, peak splitting would disappear and the final product of oxidation would be identical for all individual fractions containing CLP, CLP* and CLG. The results shown in Figure 3.10 unequivocally confirmed this hypothesis.

a) I found that one week of oxidation of individual CLP and CLP* (cyclo-MLMsoPFFWI) for each gives rise to a different peak (Figure 3.10 b, c). The retention time difference between these peaks is ~0.15 min, while MS/MS confirmed they have the same sequence: CLG (cyclo-MsoLMsoPFFWI). This explains the distorted peak shape of CLG originally observed in all previous chromatograms, as it contains two individual products (assigned as CLG and CLG* in Figure 3.10 a-c). Figure 3.10 a shows the extracted ion chromatograms for CLP and CLG from the separation of a non-fractionated extract, which resemble the superposition of two chromatograms from panels b and c.

b) Further exposure of CLP/CLG to oxygen in a water/acetonitrile mixture results in oxidation of the Trp residue in the first instance, rather than a Mso to Msn conversion. In fact, one-week of oxidation yields ~50% conversion of tryptophan to kynurenine (Trp +3.997 Da adduct). Kynurenine-containing analogs of CLP and CLP* elute prior to their parent compounds. Two weeks of oxidation exhibits an almost complete Trp – kynurenine conversion. Therefore all oxidation products for a two-week (or more) exposure will contain kynurenine (W+4) instead of Trp.

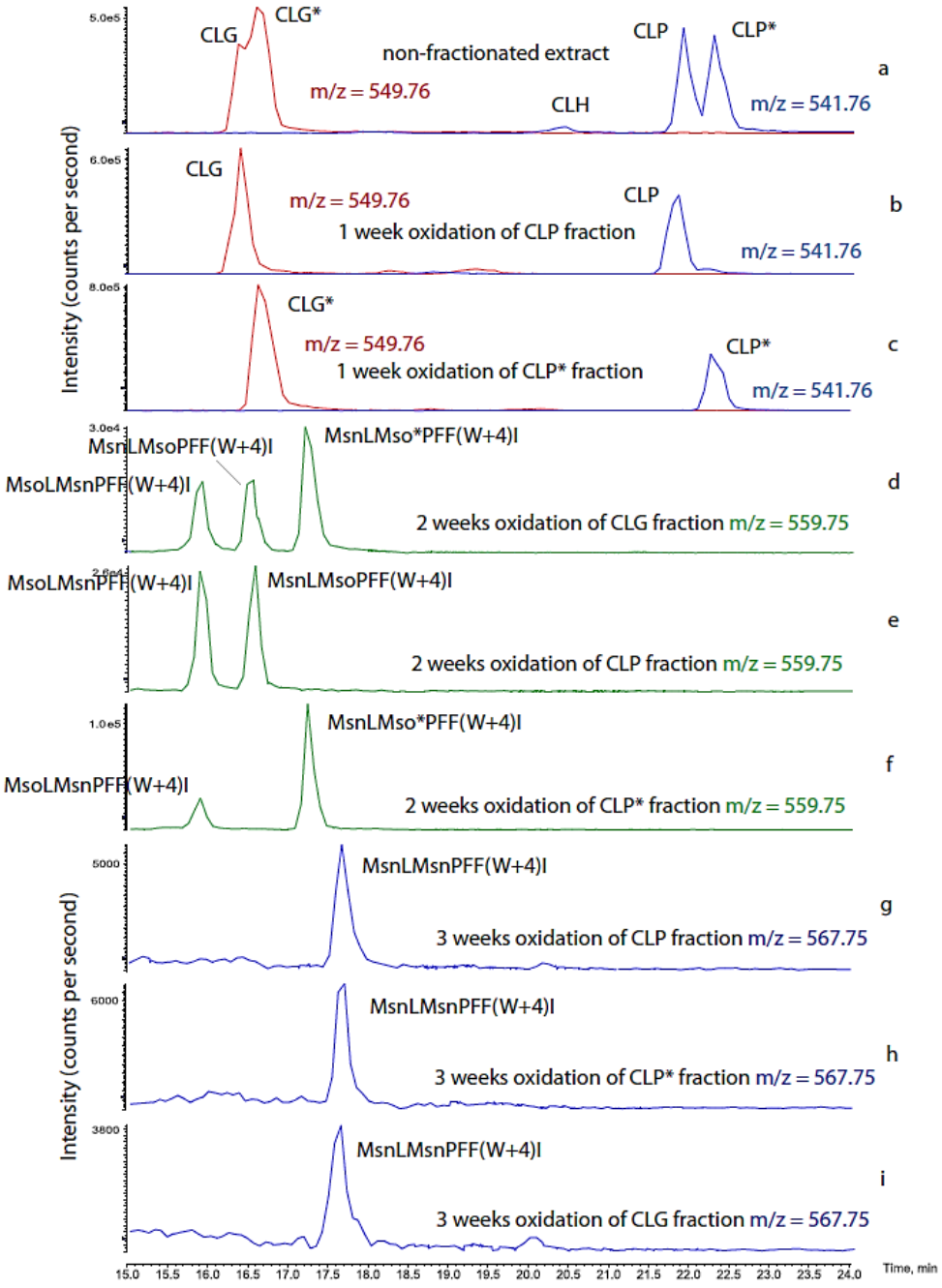


Figure 3.10 Extracted ion chromatograms for 2+ molecular ions for different members of the CLO oxidation pathway. **(a)** – non-fractionated extract; **(b)** – one week oxidation of CLP fraction; **(c)** – one week oxidation of CLP* fraction; **(d)** – two weeks oxidation of CLG fraction; **(e)** – two weeks oxidation of CLP fraction; **(f)** – two weeks oxidation of CLP* fraction; **(g)** – three weeks oxidation of CLP fraction; **(h)** – three weeks oxidation of CLP* fraction; **(i)** – three weeks oxidation of CLG fraction. Nano-LC ESI MS using 30 min 30 – 90 % acetonitrile gradient.

c) Two weeks of oxidation results in the appearance of the products containing one Mso and one Msn residue. Similar to the CLP-CLH pair, one can expect the formation of positional isomers. MS/MS fragmentation of chromatographic peaks with identical *m/z* = 559.75 for 2+ ions (Figure 3.10 d-f) confirmed that as soon as the second Met residue is oxidized to Msn, both CLP and CLP* fractions show the presence of the same compound – cyclo-MsoLMsnPFF(W+4)I. Oxidation of the CLG fraction, which contained both diastereomeric components, results in three peaks: two diastereomers cyclo-MsnLMsoPFF(W+4)I and one positional isomer cyclo-MsoLMsnPFF(W+4)I with no peak splitting.

d) Three weeks of oxidation shows the presence of the final product - cyclo-MsnLMsnPFF(W+4)I with identical retention times and MS/MS spectra for all three original fractions (CLG, CLP, CLP* in Figure 3.10 g-i). This compound is a final product of oxidation, similar to its CLO parent, exhibiting no diastereomeric peaks due to elimination of the chiral center at the sulfur atom.

Peak splitting for short peptides under conditions of RP HPLC may occur due to the existence of diastereomers, enantiomers, conformers, and isopeptide isomers [103]. For example, non-enzymatic deamidation of Asn residues leads to formation of two products containing asparatic and iso-asparatic acid residues with different chromatographic properties [104]. Slow conversion between cis- and trans-conformers for proline-containing peptides resulted in peak widening during fast RP HPLC separations [105]. Peak splitting for linear methionine *S*-oxide-containing peptides was reported by Blondelle et al [14]. As shown in Chapter 2, this is also routinely observed in proteomic experiments. The separation of diastereoisomeric pairs of linear peptides occurs only when the methionine *S*-oxide residue is located in the hydrophobic face of

amphipathic α-helical structures. Due to close “intimate” interaction of the hydrophobic face of the helix with C18 sorbents, even a small variation in its structure will result in different interaction energies and peak splitting. The fact that different CLPs exhibit a different degree of peak splitting may be related to their differences in orientation upon interaction with the stationary phase. Thus, CLI-CLI* and CLP-CLP* appear as baseline-separated pairs of peaks, but CLC shows no peak splitting and good separation efficiency, while the intermediate CLE manifests as a wider peak.

Table 3.1 Cyclolinopeptides in flaxseed oil extract detected using off-line LC-MALDI MS and on-line LC-ESI MS.

Abbreviation	Sequence[a]	$[M+H]^+$	Off-line LC-MALDI MS analysis (Figure 3.2 c)	On-line LC-ESI MS analysis (Figure 3.3 a)		Off-line LC-MALDI MS analysis (Figure 3.7)	Reference[b]
			Fraction #	Retention time (min)	Intensity ratio (2+/1+ ions)	Fraction #	
Peptides with two Mso residues							
CLF	cyclo-MsoLMsoPFFWV	1084.501	1	30.44	4.68	1	[89]
CLG	cyclo-MsoLMsoPFFWI	1098.514	2	31.56	2.03	2	[89]
Peptides with one Mso residue							
CLC	cyclo-MsoLIPPFFVI	1074.605	3	32.69	0.69	3	[88]
CLT	cyclo-MsoLMPFFWV	1068.503	4	34.18	0.80	4	This work
CLE	cyclo-MsoLVFPLFI	977.552	4	34.99	0.40	4	[88]
CLH	cyclo-MsoLMPFFWI	1082.520	5	35.15	0.50	5	[89]
CLI	cyclo-MLMsoPFFWV	1068.503	5	35.38 (split)[c]	1.10	6	[89]
CLI*	cyclo-MLMsoPFFWV	1068.503	6	35.88 (split)[c]	1.22	7	
CLD	cyclo-MsoLLPFFWI	1064.564	7	36.41	1.31	6	[88]
CLP	cyclo-MLMsoPFFWI	1082.520	7	36.75 (split)[c]	1.02	8	[90,91]
CLP*	cyclo-MLMsoPFFWI	1082.520	8	37.16 (split)[c]	0.69	9	
Non-oxidized peptides							
CLB	cyclo-MLIPPFFVI	1058.609	9	39.30	0.46	10	[87,88]
CLN	cyclo-MLMPFFWV	1052.508	10	40.40	0.15	13	[90,91]
CLA	cyclo-ILVPPFFLI	1040.654	10	41.07	0.21	11	[86]
CLL	cyclo-MLVFPLFI	961.557	10	41.14	0.14	12	[90,91]
CLO	cyclo-MLMPFFWI	1066.523	11	41.63	0.21	14	[90,91]
CLM	cyclo-MLLPFFWI	1048.568	12	42.93	0.21	15	[90,91]
Msn-containing peptides							
CLK[d]	cyclo-MsnLIPPFFVI	1090.594	-	35.22	1.12	-	[92]
CLJ[d]	cyclo-MsnLVFPLFI	993.530	-	37.36	0.54	-	[92]

a) The order of amino acids according to flax genome sequence [106]; b) First report on identification of individual components; c) peak splitting observed for pairs of CLI/CLI* and CLP/CLP* peaks; d) identified in the sample of aged flaxseed oil (Figure 3.6 e).

3.4 Conclusions

I used mass spectrometry in an on-line and off-line combination with reversed-phase HPLC to characterize cyclolinopeptides in freshly prepared and aged flaxseed oil. LC- ESI MS analysis using a TripleTOF^{TM}5600 mass spectrometer allows for the simultaneous detection of all cyclolinopeptides previously reported in the literature. I also identified a new cyclolinopeptide T (CLT, cyclo-MsoLMPFFWV), which is a positional isomer of the singly oxidized CLI. This completes the discovery of all possible oxidation products of the five original CLPs containing methionine *S*-oxide (Mso). Further oxidation of these components into methionine *S,S*-dioxide-containing peptides will generate another set of the species containing all possible permutations of Mso-Msn oxidation states. I also confirmed a prior report on the identification of two members of this group – methionine *S,S*-dioxide-containing analogs of CLC and CLE. Retention time for all three variants of Met increases in the order similar to the one found for the linear peptides: Mso < Msn < Met.

I have also observed a new feature in the RP HPLC behavior of cyclic peptides containing methionine *S*-oxide: the complete or partial resolution of pairs of peaks for CLP, CLI, CLE, CLF and CLG. This drastically affects the separation efficiency for these components using high-efficiency RP HPLC sorbents.

Four different stationary phases were compared, including porous C18 and core-shell particles with C8, C18 and Phenyl-Hexyl functionalities. Introducing an aromatic surface chemistry alters separation selectivity due to stronger interaction with peptides carrying an additional aromatic Trp residue. The Kinetex^{TM} Phenyl-Hexyl was the only column providing a complete separation of all major cyclolinopeptides. This allowed development of a

straightforward UV-detection based procedure for the determination of the degree of oxidation in major cyclolinopeptides contributing to bitter taste in flaxseed oil.

Chapter 4

General Conclusions

Oxidation is one of the most frequently encountered PTMs in a number of analytical techniques targeting peptides and proteins. Here I aimed specifically to characterize the chromatographic behavior of linear tryptic peptides from proteomic experiments and cyclic peptides from flaxseed oil that contain oxidized methionine derivatives (methionine *S*-oxide and methionine *S,S*-dioxide) in RPC. Understanding of this behavior and the ability to predict retention time shifts for modified peptides brings a new element to method development in a wide range of bioanalytical techniques that employ RPC. Chromatography and mass-spectrometry are both separation techniques. Principles of separation based on mass to charge ratio are well established, while measurements in the chromatographic dimension still require establishment of better metrics and the ability to calculate expected retention times. My results will help with respective improvements for a very large group of peptides with PTMs – those peptides carrying oxidized methionine residues.

Peptides containing methionine *S*-oxide are generally anticipated to have lower retention times in RP-LC compared to non-modified analogs. This thesis is a detailed inspection of a large collection of pairs of oxidized and non-oxidized linear tryptic peptides for the purpose of getting a more complete picture. The amplitude of the retention time shift may vary dramatically between -9% to +0.36% acetonitrile. In particular, the effect of methionine position within

amphipathic helices is very important in determining the retention time shift. Thus I concluded that destabilization of amphipathic helical structures upon interaction with a C18 stationary phase is the main reason for large negative retention time shifts. This is a further advance to understanding the effect of amphipathic helicity on RP HPLC peptide retention time. I also report for the first time that positive retention time shifts can be observed. These happen because Mso residues are better than non-modified methionine to form hydrogen bonds with the -NH group of the N-cap amino acid and therefore stabilize the amphipathic helix. These important findings are critical for accurate predictions of peptide retention time.

A number of comparable observations were made on both linear and cyclic peptides: first, retention times of methionine-containing peptide analogues increases in the order: Mso < Msn < Met; second, peak splitting happens when there is a chiral sulfur in the Mso side chain (demonstrated for the first time for cyclic peptides). It is conceivable that peak splitting for diastereomers only occurs when Mso is positioned in the hydrophobic face of an amphipathic helix of a linear peptide, and that the same positioning leads to a significantly larger negative retention time shift. Strong interaction within a binding domain between Mso and the stationary phase could be the reason for peak splitting in cyclic peptides. Although peak splitting for cyclic peptides did not correlate with the largest retention time shifts, this merits further study on the conformation of the cyclic peptide upon contact with a hydrophobic surface. The largest retention time shift for cyclolinopeptides caused by Met oxidation to Mso is –14.6% acetonitrile. The most extreme cases observed in the study of linear peptides and similar ones conducted by Blondelle et al. [14] were –9% and –9.3% acetonitrile, respectively.

Understanding the chromatographic behavior of compounds of interest is useful in various applications. The study of linear peptides has served as a foundation for the development of a peptide retention time prediction model, which was aimed to aid proteomic research in both targeted and bottom-up discovery methods. In the second part of the study, I developed a straightforward UV-detection based procedure that benefits industrial applications for quality control of flaxseed oil. This rapid and reliable analytical method can not only be used for determination of the oxidation degree of major cyclolinopeptides, but also for the isolation of individual major cyclolinopeptides for further sensory tests to identify new components contributing to bitter taste.

LIST OF REFERENCES

[1] J.B. Reece, M.R. Taylor, E.J. Simon, J.L. Dickey, N.A. Campbell, K.G.E. Scott, Pearson Education, Inc., Upper Saddle River, New Jersey, USA, 2012, p. 48.

[2] M. Roberts, M.J. Reiss, G. Monger, Nelson Thornes, Delta Place, Cheltenham, United Kingdom, 2000, p. 28.

[3] L.G. Presta, G.D. Rose, Science 240 (1988) 1632.

[4] J.S. Richardson, D.C. Richardson, Science 240 (1988) 1648.

[5] R. Aurora, G.D. Rose, Protein Sci 7 (1998) 21.

[6] A.J. Doig, R.L. Baldwin, Protein Sci 4 (1995) 1325.

[7] V. Munoz, L. Serrano, J Mol Biol 245 (1995) 275.

[8] M.A. Rould, Q. Wan, P.B. Joel, S. Lowey, K.M. Trybus, J Biol Chem 281 (2006) 31909.

[9] D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, J Mol Biol 179 (1984) 125.

[10] R. Brasseur, Molecular Description of Biological Membrane Components by Computer Aided Conformational Analysis, CRC Press, 1990, p. 192.

[11] W.W. Parson, Springer, New York, USA, 2009, p. 328.

[12] S.E. Blondelle, J.M. Ostresh, R.A. Houghten, E. Perez-Paya, Biophys J 68 (1995) 351.

[13] M. Jafari, B. Zargar, M. Soltani, D.N. Karunaratne, B. Ingalls, P. Chen, in: A. Tiwari, M. Ramalingam, H. Kobayashi, A.P.F. Turner (Eds.), Biomedical Materials and Diagnostic Devices, John Wiley & Sons, Hoboken, NJ, USA., 2012, p. 482.

[14] S.E. Blondelle, E. Perez-Paya, G. Allicotti, B. Forood, R.A. Houghten, Biophys J 69 (1995) 604.

[15] T.J. Sereda, C.T. Mant, F.D. Sonnichsen, R.S. Hodges, J Chromatogr A 676 (1994) 139.

[16] R. Soares, E. Pires, A.M. Almeida, R. Santos, R. Gomes, K. Koči, C.F. Franco, A.V. Coelho, in: J. Prasain (Ed.), Tandem Mass Spectrometry - Applications and Principles, InTech, 2012, p. 35.

[17] M. Kinter, N.E. Sherman, Protein Sequencing and Identification Using Tandem Mass Spectrometry, Wiley, 2005, p. 270.

[18] R.S.B. Clark, H. Bayir, L.W. Jenkins, Crit Care Med 33 (2005) S407.

[19] C.S. Hwang, A. Shemorry, A. Varshavsky, Science 327 (2010) 973.

[20] H. Lin, J. Du, H. Jiang, T.P. Begley, Wiley Encyclopedia of Chemical Biology, John Wiley & Sons, Inc., 2007, p. 1.

[21] J. Seo, K.J. Lee, J Biochem Mol Biol 37 (2004) 35.

[22] G.M. Forte, M.R. Pool, C.J. Stirling, PLoS Biol 9 (2011) e1001073.

[23] E. Ahrne, M. Muller, F. Lisacek, Proteomics 10 (2010) 671.

[24] A. Michalski, J. Cox, M. Mann, J Proteome Res 10 (2011) 1785.

[25] P. Picotti, O. Rinner, R. Stallmach, F. Dautel, T. Farrah, B. Domon, H. Wenschuh, R. Aebersold, Nat Methods 7 (2010) 43.

[26] L.C. Gillet, P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Mol Cell Proteomics 11 (2012) O111.016717.

[27] L. Moruz, A. Staes, J.M. Foster, M. Hatzou, E. Timmerman, L. Martens, L. Kall, Proteomics 12 (2012) 1151.

[28] T. Baczek, R. Kaliszan, Proteomics 9 (2009) 835.

[29] J.L. Meek, Proc Natl Acad Sci U S A 77 (1980) 1632.

[30] D. Guo, C.T. Mant, A.K. Taneja, J.M.R. Parker, R.S. Rodges, J Chromatogr A 359 (1986) 499.

[31] J.M. Kovacs, C.T. Mant, R.S. Hodges, Biopolymers 84 (2006) 283.

[32] C.A. Browne, H.P. Bennett, S. Solomon, Anal Biochem 124 (1982) 201.

[33] S.-J. Su, B. Grego, B. Niven, M.T.W. Hearn, J Liq Chromatogr 4 (1981) 1745.

[34] T. Sasagawa, T. Okuyama, D.C. Teller, J Chromatogr A 240 (1982) 329.

[35] D. Shamshurin, V. Spicer, O.V. Krokhin, J Chromatogr A 1218 (2011) 6348.

[36] O.V. Krokhin, Anal Chem 78 (2006) 7785.

[37] R.C. Dwivedi, V. Spicer, M. Harder, M. Antonovici, W. Ens, K.G. Standing, J.A. Wilkins, O.V. Krokhin, Anal Chem 80 (2008) 7036.

[38] O.V. Krokhin, R. Craig, V. Spicer, W. Ens, K.G. Standing, R.C. Beavis, J.A. Wilkins, Mol Cell Proteomics 3 (2004) 908.

[39] C.T. Mant, T.W. Burke, J.A. Black, R.S. Hodges, J Chromatogr A 458 (1988) 193.

[40] L. Moruz, D. Tomazela, L. Kall, J Proteome Res 9 (2010) 5209.

[41] J. Reimer, D. Shamshurin, M. Harder, A. Yamchuk, V. Spicer, O.V. Krokhin, J Chromatogr A 1218 (2011) 5101.

[42] X. Jiang, D. Shamshurin, V. Spicer, O.V. Krokhin, J Chromatogr B Analyt Technol Biomed Life Sci 915-916 (2013) 57.

[43] J. Reimer, V. Spicer, O.V. Krokhin, J Chromatogr A 1256 (2012) 160.

[44] T. Baczek, P. Wiczling, M. Marszall, Y.V. Heyden, R. Kaliszan, J Proteome Res 4 (2005) 555.

[45] A.V. Gorshkov, I.A. Tarasova, V.V. Evreinov, M.M. Savitski, M.L. Nielsen, R.A. Zubarev, M.V. Gorshkov, Anal Chem 78 (2006) 7770.

[46] K. Shinoda, M. Sugimoto, N. Yachie, N. Sugiyama, T. Masuda, M. Robert, T. Soga, M. Tomita, J Proteome Res 5 (2006) 3312.

[47] K. Petritis, L.J. Kangas, P.L. Ferguson, G.A. Anderson, L. Pasa-Tolic, M.S. Lipton, K.J. Auberry, E.F. Strittmatter, Y. Shen, R. Zhao, R.D. Smith, Anal Chem 75 (2003) 1039.

[48] K. Petritis, L.J. Kangas, B. Yan, M.E. Monroe, E.F. Strittmatter, W.J. Qian, J.N. Adkins, R.J. Moore, Y. Xu, M.S. Lipton, D.G. Camp II, R.D. Smith, Anal Chem 78 (2006) 5026.

[49] M. Gilar, H. Xie, A. Jaworski, Anal Chem 82 (2010) 265.

[50] N. Pfeifer, A. Leinenbach, C.G. Huber, O. Kohlbacher, BMC Bioinformatics 8 (2007) 468.

[51] A.A. Klammer, X. Yi, M.J. MacCoss, W.S. Noble, Anal Chem 79 (2007) 6111.

[52] L. Bruhl, B. Matthaus, E. Fehling, B. Wiege, B. Lehmann, H. Lehmann, K. Bergander, K. Quiroga, A. Scheipers, O. Frank, T. Hofmann, J Agric Food Chem 55 (2007) 7864.

[53] B.J. Soto-Cerda, S. Duguid, H. Booker, G. Rowland, A. Diederichsen, S. Cloutier, Theor Appl Genet 127 (2014) 881.

[54] H.K. Nielsen, J. Loliger, R.F. Hurrell, Br J Nutr 53 (1985) 61.

[55] L. Bruhl, B. Matthaus, A. Scheipers, T. Hofmann, Eur J Lipid Sci Technol 110 (2008) 625.

[56] B.A. Bidlingmeyer, Practical HPLC Methodology and Applications, John Wiley & Sons, 1992, p. 215.

[57] B. Zybailov, Q. Sun, K.J. van Wijk, Anal Chem 81 (2009) 8015.

[58] J. Lengqvist, H. Eriksson, M. Gry, K. Uhlen, C. Bjorklund, B. Bjellqvist, P.J. Jakobsson, J. Lehtio, Amino Acids 40 (2011) 697.

[59] M. Berg, A. Parbel, H. Pettersen, D. Fenyo, L. Bjorkesten, Rapid Commun Mass Spectrom 20 (2006) 1558.

[60] W. Vogt, Free Radic Biol Med 18 (1995) 93.

[61] B. Ghesquiere, K. Gevaert, Mass Spectrom Rev 33 (2014) 147.

[62] B. Ghesquiere, V. Jonckheere, N. Colaert, J. Van Durme, E. Timmerman, M. Goethals, J. Schymkowitz, F. Rousseau, J. Vandekerckhove, K. Gevaert, Mol Cell Proteomics 10 (2011) M110.006866.

[63] C. Schöneich, F. Zhao, J. Yang, L. Miller Brian, Therapeutic Protein and Peptide Formulation and Delivery, American Chemical Society, 1997, p. 79.

[64] M. Grigoryan, D. Shamshurin, V. Spicer, O.V. Krokhin, Anal Chem 85 (2013) 10878.

[65] H.K. Khor, M.E. Jacoby, T.C. Squier, G.C. Chu, D. Chelius, mAbs 2 (2010) 299.

[66] R. Grimaud, B. Ezraty, J.K. Mitchell, D. Lafitte, C. Briand, P.J. Derrick, F. Barras, J Biol Chem 276 (2001) 48915.

[67] A. Hansel, S.H. Heinemann, T. Hoshi, Biochim Biophys Acta 1703 (2005) 239.

[68] S. Luo, R.L. Levine, FASEB J 23 (2009) 464.

[69] J.H. Baxter, C.S. Lai, R.R. Phillips, L. Dowlati, J.J. Chio, S.T. Luebbers, S.R. Dimler, P.W. Johns, J Chromatogr A 1157 (2007) 10.

[70] M.W. Duncan, R. Aebersold, R.M. Caprioli, Nat Biotechnol 28 (2010) 659.

[71] K. Gevaert, J. Van Damme, M. Goethals, G.R. Thomas, B. Hoorelbeke, H. Demol, L. Martens, M. Puype, A. Staes, J. Vandekerckhove, Mol Cell Proteomics 1 (2002) 896.

[72] K. Gevaert, J. Vandekerckhove, in: B.D. Hames, C.D. O'Connor (Eds.), Proteomics: Methods Express, Scion Publishing Ltd. , Bloxham, 2007, p. 61.

[73] R.A. Houghten, S.T. DeGraw, J Chromatogr 386 (1987) 223.

[74] J.R. Wisniewski, A. Zougman, N. Nagaraj, M. Mann, Nat Methods 6 (2009) 359.

[75] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Anal Chem 77 (2005) 6426.

[76] O.V. Krokhin, V. Spicer, Anal Chem 81 (2009) 9522.

[77] V. Spicer, Y. Lao, D. Shamshurin, P. Ezzati, J. Wilkins, O. Krokhin, Submitted to Anal Chem, 2014.

[78] T. Horigome, T. Yanagida, M. Miura, J Agric Chem Soc Jpn 48 (1974) 195.

[79] K.M. Schaich, Crit Rev Food Sci Nutr 13 (1980) 131.

[80] W.T. Roubal, A.L. Tappel, Arch Biochem Biophys 113 (1966) 5.

[81] S.R. Tannenbaum, H. Barth, J.P. Leroux, J Agric Food Chem 17 (1969) 1353.

[82] C. Viju, M.T. Yeung, W.A. Kerr, World Economy 37 (2014) 137.

[83] R. Przybylski, in: F. Shahidi (Ed.), Bailey's Industrial Oil and Fat Products, John Wiley & Sons, Inc., 2005, p. 281.

[84] E.N. Frankel, Prog Lipid Res 22 (1983) 1.

[85] O. Sharav, Y.Y. Shim, D.P. Okinyo-Owiti, R. Sammynaiken, M.J.T. Reaney, J Agric Food Chem 62 (2014) 88.

[86] H.P. Kaufmann, A. Tobschirbel, Chem Ber Recl 92 (1959) 2805.

[87] H. Morita, A. Shishido, T. Matsumoto, K. Takeya, H. Itokawa, T. Hirano, K. Oka, Bioorg Med Chem Lett 7 (1997) 1269.

[88] H. Morita, A. Shishido, T. Matsumoto, H. Itokawa, K. Takeya, Tetrahedron 55 (1999) 967.

[89] T. Matsumoto, A. Shishido, H. Morita, H. Itokawa, K. Takeya, Phytochemistry 57 (2001) 251.

[90] P. Stefanowicz, Acta Biochim Pol 48 (2001) 1125.

[91] P. Stefanowicz, Eur J Mass Spectrom 10 (2004) 665.

[92] P.D. Jadhav, D.P. Okinyo-Owiti, P.W.K. Ahiahonu, M.J.T. Reaney, Food Chem 138 (2013) 1757.

[93] D.P. Okinyo-Owiti, L. Young, P.G.G. Burnett, M.J.T. Reaney, Biopolymers 102 (2014) 168.

[94] B. Picur, M. Cebrat, J. Zabrocki, I.Z. Siemion, J Pept Sci 12 (2006) 569.

[95] T.J. Gaymes, M. Cebrat, I.Z. Siemion, J.E. Kay, Febs Letters 418 (1997) 224.

[96] M.J. Reaney, Y. Jia, J. Shen, C. Schock, N. Tyler, J. Elder, S. Singh, WO2009079792 A1, W.I.P.O. (WIPO), 2009.

[97] B. Gui, Y.Y. Shim, R.S.S. Datla, P.S. Covello, S.L. Stone, M.J.T. Reaney, J Agric Food Chem 60 (2012) 8571.

[98] B. Gui, Y.Y. Shim, M.J.T. Reaney, J Agric Food Chem 60 (2012) 8580.

[99] C.M. Olivia, P.G.G. Burnett, D.P. Okinyo-Owiti, J.H. Shen, M.J.T. Reaney, J Chromatogr B Analyt Technol Biomed Life Sci 904 (2012) 128.

[100] F. Aladedunye, E. Sosinska, R. Przybylski, J Am Oil Chem Soc 90 (2013) 419.

[101] A.V. Loboda, A.N. Krutchinsky, M. Bromirski, W. Ens, K.G. Standing, Rapid Commun Mass Spectrom 14 (2000) 1047.

[102] N.T. Joyner, J.E. McIntyre, Oil & Soap 15 (1938) 184.

[103] D. Winter, R. Pipkorn, W.D. Lehmann, J Sep Sci 32 (2009) 1111.

[104] L. Huang, J. Lu, V.J. Wroblewski, J.M. Beals, R.M. Riggin, Anal Chem 77 (2005) 1432.

[105] W.R. Melander, J. Jacobson, C. Horvath, J Chromatogr 234 (1982) 269.

[106] P.G. Arnison, M.J. Bibb, G. Bierbaum, A.A. Bowers, T.S. Bugni, G. Bulaj, J.A. Camarero, D.J. Campopiano, G.L. Challis, J. Clardy, P.D. Cotter, D.J. Craik, M. Dawson, E. Dittmann, S. Donadio, P.C. Dorrestein, K.D. Entian, M.A. Fischbach, J.S. Garavelli, U. Goransson, C.W. Gruber, D.H. Haft, T.K. Hemscheidt, C. Hertweck, C. Hill, A.R. Horswill, M. Jaspars, W.L. Kelly, J.P. Klinman, O.P. Kuipers, A.J. Link, W. Liu, M.A. Marahiel, D.A. Mitchell, G.N. Moll, B.S. Moore, R. Muller, S.K. Nair, I.F. Nes, G.E. Norris, B.M. Olivera, H. Onaka, M.L. Patchett, J. Piel, M.J.T. Reaney, S. Rebuffat, R.P. Ross, H.G. Sahl, E.W. Schmidt, M.E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R.D. Sussmuth, J.R. Tagg, G.L. Tang, A.W. Truman, J.C. Vederas, C.T. Walsh, J.D. Walton, S.C. Wenzel, J.M. Willey, W.A. van der Donk, Nat Prod Rep 30 (2013) 108.