Mass Spectrometric Studies Of Keap1-Nrf2 Binding Interactions

BY

ZHIYUAN SUN
B.S, China Pharmaceutical University, China, 2006
M.S., University of Kentucky, Lexington, Kentucky, 2009

THESIS

Submitted as partial fulfillment of the requirements
For the degree of Master of Science in Medicinal Chemistry
In the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Richard van Breemen, Chair and Advisor
John Fitzloff,
Michael Federle

This thesis is dedicated to my parents, Rui Sun and Min Fu, as well as my relatives and friends for their endless support, encouragement and providing the strength, to complete my studies.

ACKNOWLEDGMENTS

I want to express my sincere gratitude and appreciation to all the people who helped me and support me throughout my work. In particular, I would like to thank my major advisor, Dr. Richard B. van Breemen for his constant support and guidance. I further want to extend my gratitude to the other members of my committee, which include Dr. John Fitzloff and Dr. Michael Federle.

In addition to my committee I would extend my appreciation to Dr. Jinghu Li, Dr. Shunyan Mo, Dr. Yang Song, Linlin Dong, Yang Yuan and other van Breemen lab members. They not only gave me a lot of instructions and suggestions, but also to share ideas with me about my projects.

I thank all the Biogen Idec Analytical Biochemistry Group, including Dr Dingyi Wen, Dr Sheng Gu, Dr Yan Wang and Mrs Yaping Sun for their guidance and help during my summer internship at Biogen Idec.

In addition to the technical and valuable assistance above, I would like to thank my family and friends for their equally important assistance. My parents, I thank for your support, understanding, and patience with me, I would like to extend my sincerest gratitude and love for your support.

I appreciate the Department of Medicinal Chemistry and Pharmacognosy for financial support. Major financial assistance for this work was supported by grants by NIH grants P01 CA048112 and P01 CA048112-S1from the National Cancer Institute and 1 S10 RR14686 from the National Center for Research Resources.

ZS

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF ABBREVIATIONS

2-D	Two-dimensional
ARE	Antioxidant response element
CHCA	alpha-cyano-4-sinapinic acid
CID	Collision-induced dissociation
COX2	Cyclooxygenase-2
DHB	2,5-dihydroxybenzoic acid
DTT	1,4-dithio-D,L-threitol
ECD	Electron capture dissociation
EGCG	Epigallocatechin gallate
ER	Estrogen receptors
ESI	Electrospray
ETD	Electron transfer dissociation
FA	Ferulic acid
H/D	Hydrogen/deuterium
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	High performance liquid chromatography
HXMS	Hydrogen/deuterium exchange mass spectrometry
IVR	Intervening region
KEAP1	Kelch-like ECH associating protein 1
LC-MS	High performance liquid chromatography coupled with mass spectrometry
MALDI	Matrix-assisted laser desorption ionization
MS	Mass spectrometry
MS-MS	Tandem mass spectrometry
NTR	*N*-terminal region
ROS	Reactive oxygen species
SA	Sinapinic acid
SDS	Sodium dodecyl sulfate
TCEP	Tris (2-carboxyethyl) phosphine hydrochloride
TFA	Trifluoroacetic acid
TOF	Time-of-flight
TOFMS	Time of flight mass spectrometry
UF	Ultrafiltration

SUMMARY

Chemoprevention is an active cancer preventive strategy to inhibit, delay or reverse human carcinogenesis using naturally occurring or synthetic chemical agents. In recent years, it has emerged as one of the major approaches for reducing cancer. Nrf2 is a member of the NF-E2 family of nuclear basic leucine zipper transcription factors. In the cell nucleus, Nrf2 binds to the 5' upstream regulatory antioxidant response element (ARE) regions of phase II genes and accelerates their transcription. Under basal conditions, Nrf2 concentration in the cell is low because of continuous ubiquitination by the Culln3-based (Cul3) E3 ligase ubiquitination complex and degradation (Zhang 2006). Therefore, phase II gene expression is suppressed under basal conditions. Keap1 serves as a bridge between Nrf2 and Cul3 and binds both proteins. When cells are under oxidative stress, signals are transmitted to the Keap1-Nrf2 complex, which interrupts ubiquitination of Nrf2 and increases its concentration in cells, including the nucleus. Consequently, protective phase II enzymes are up-regulated by the increasing level Nrf2 in the nucleus (Yamamoto, Itoh *et al.* 2003).

The Keap1-Nrf2 system has been shown to play an important role in the chemoprevention process. However, how the Keap1 protein interacts with transcription factor Nrf2 is still unclear. In this thesis, an ultrafiltration mass spectrometry screening assay was developed for characterizing protein-protein interactions. To minimize the interference from trypsin acting on the protein target during the binding assay, a trypsin deactivation step was included in the assay. This step significantly reduced the activity of trypsin remaining in the peptide mixture used in the binding incubation. This

ultrafiltration mass spectrometry method should be suitable for use as an alternative to H/D exchange or protein footprinting to study protein-protein interactions. In this investigation, the Keap1 tryptic peptide ^{337}QSLSYLEAYNPSDGTWLR^{354} was identified as having the highest affinity for the Neh2 domain of Nrf2. These results support the view that the human Kelch domain (amino acid 327-amino acid 609) is at least in part responsible for association between Keap1 and Nrf2.

1. INTRODUCTION

1.1 Proteomics and mass spectrometry

Proteomics is a powerful and indispensable technology widely used in biology. It can be used to identify the components of small protein complexes and large scale organelles, to determine post-translational modifications and may be used in sophisticated functional screens (Steen and Mann 2004). Page *et al*. defined proteomics as the, "the direct qualitative and quantitative analysis of the broad complement — or subset thereof — of the proteins present in an organism, tissue or cell under a given set of physiological or environmental conditions" (Smith, Page *et al.* 2004). It is believed that proteomics will provide valuable information for understanding how complex biological processes occur at a molecular level, how they differ in different cell types and how they are changed in disease states.

In the early years of proteomics, two-dimensional (2-D) gel electrophoresis was the major tool for obtaining a global picture of expression levels of the whole proteome. In this method, the proteins were first separated in one direction by isoelectric focusing and in the orthogonal direction by molecular mass using electrophoresis in a gel containing sodium dodecyl sulfate (SDS) (O'Farrell, Goodman *et al.* 1977). Using 2-D gel electrophoresis, tens of thousands of proteins can be separated in a single gel. However, this method has many drawbacks: it has poor dynamic range; it is biased toward soluble and abundant proteins; and it cannot identify the proteins that have been resolved.

During the 1990s, mass spectrometry (MS), in which biomolecules are ionized and their masses are measured by following their trajectories in the vacuum system, became widely used in proteomics. Mass spectrometry coupled with protein separation methods have evolved as the

primary tool in protein identification and protein complex deconvolution. The key improvements were the application of the time-of-flight (TOF) mass spectrometer and relatively nondestructive methods to convert proteins and peptides into gas-phase ions. Two kinds of soft ionization methods, matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI), have made it possible to analyze large biomolecules, such as proteins and peptides, using mass spectrometry (Yates 1998).

1.1.1 Ionization methods

In an initial method that is still in use, protein mixtures are first separated using 2-D gel electrophoresis followed by the excision of protein spots from the gel. The next step is digestion of the protein using sequence-specific protease such as Lys C or trypsin, and then the peptides from the digestion are analyzed by MS. When MALDI is used, the samples are in the solid phase within an acidified matrix, which absorbs energy from a laser pulse and dissipates the energy thermally. The rapid transferred energy generates a vaporized plume of matrix and thereby simultaneously ejecting the analytes into the gas phase. A strong electric field between the MALDI target and the entrance of the MS analyzer accelerates the ions from the ion source into the TOF tube where the ions are separated based on their different velocities which are related to their mass-to-charge (*m/z*) ratios. One of the most significant advantages of MALDI-MS and MALDI –tandem mass spectrometry (MS-MS) is the high throughput with which protein or peptide identification can be carried out (using MALDI targets containing 96 or 384 samples at one time). MALDI-MS-MS provides a fast way to identify proteins when a fully decoded genome is available because the deduced masses of the resolved analytes can be compared to those calculated for the predicted products of all of the genes in the genomes of an organism (Zhu, Bilgin *et al.* 2003).

Electrospray is another widely used method to introduce mixtures of large biomolecules into the mass spectrometer. The unique feature of ESI is that at atmospheric pressure it allows the rapid transfer of analytes from the liquid phase to the gas phase (Moran, Figeys *et al.* 2001). During ESI, the sample solution is passed through a conductively coated needle. A high voltage is applied to the needle and results in charges being added to a spray of droplets emerging from the needle. As the solvent evaporates, the sample can be positively or negatively charged. However, positive mode is used more frequently in protein and peptide analysis. A significant advantage of ESI is it can generate multiple charged ions from large biomolecules. Depending on the size and amino acid composition, peptides typically are detected as doubly or triply charged ions. The formation of multiply charged ions enables quadrupole mass spectrometer or ion traps, which have modest mass ranges, to measure peptides and proteins with masses far beyond their nominal *m/z* range.

Because ionization of analytes in complex mixtures can be a competitive process during ESI and MALDI, not all species in complex mixtures can be observed. Therefore, protein or peptide mixtures are not usually introduced into the mass spectrometer all at once. Instead, these mixtures are simplified using high performance liquid chromatography (HPLC) either prior to mass spectrometric analysis using MALDI or on-line using ESI in a process called LC-MS. Note that ESI but not MALDI MS may be coupled directly with HPLC (Figure 1.1).

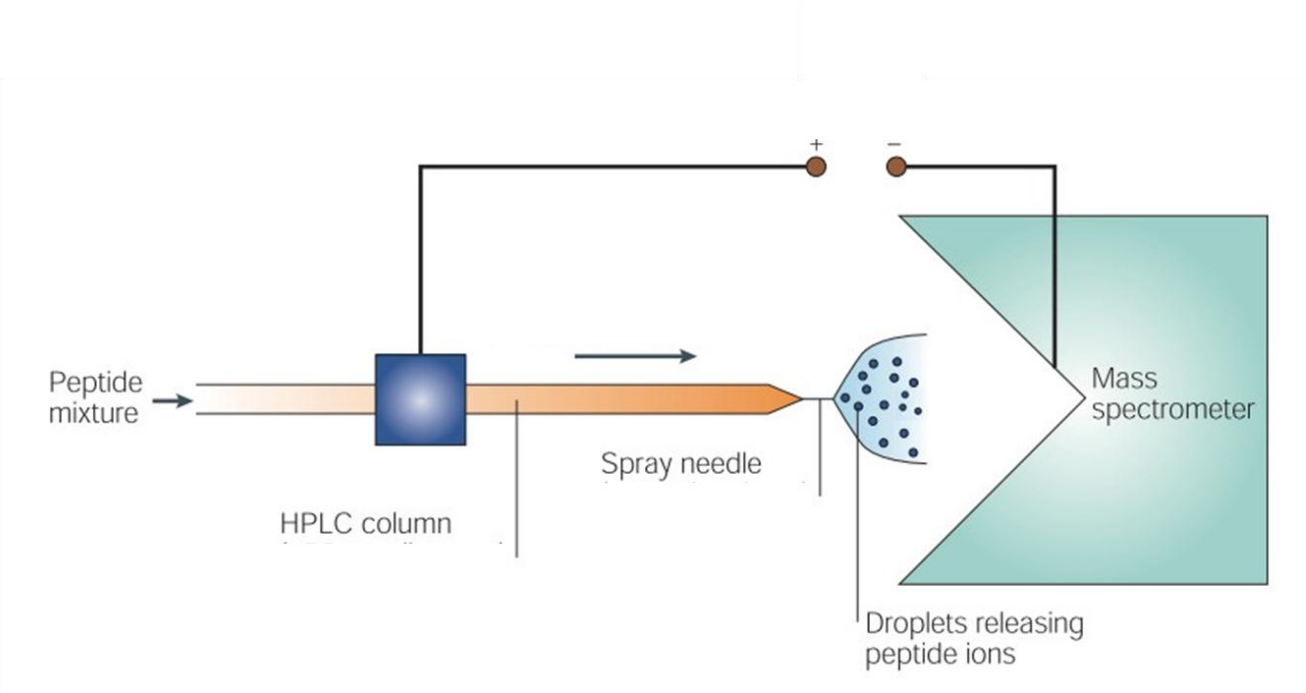


Figure 1.1 Schematic representation of a nanoelectrospray LC-MS apparatus using microcapillary liquid chromatography (adapted from Steen and Mann 2004). Note that when analytical HPLC is used with flow rates up to 300 µl/min, this process is called electrospray LC-MS.

Peptides and small proteins are usually separated using reversed phase HPLC. During reversed phase separations, a solvent gradient is used from water containing a weak ion pair agent such as formic acid to an organic solvent (usually acetonitrile). Different species are eluted sequentially in increasing order of their hydrophobicity. However, very hydrophilic peptides may be poorly retained on the column and elute immediately. Extremely hydrophobic peptides might not elute from the column using standard gradients, and more hydrophobic solvents must be used such as isopropanol.

After entering the mass spectrometer, the mixture of charged peptides is separated in the first MS according to their *m/z* ratios, and the masses of the peptides are recorded in the mass spectrum. In a second analysis, specific peptide ions (usually doubly charged) are selected based on the *m/z* values and instead of being detected directly, these species are fragmented using collision-induced dissociation or electron capture dissociation to generate fragment or "product" (Figure 2). With appropriate collision energy, fragmentation occurs predominantly at the peptide bonds such that a ladder of fragment ions, each of which differs by the mass of a single amino acid, is generated. The product ions are separated according to their *m/z* ratio in a second stage of mass spectrometry, and the sequence of the peptide can be deduced from the resulting fragments (Godovac-Zimmermann and Brown 2001; Moran, Figeys *et al.* 2001).

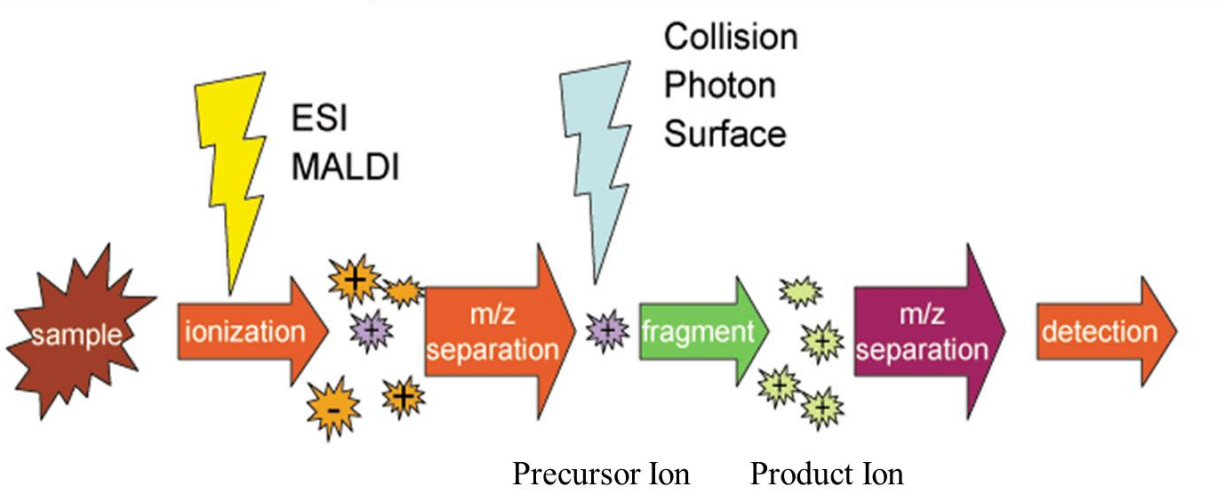


Figure 1.2 Schematic representation of a tandem mass spectrometry experiment. For peptide and protein analysis, electrospray (ESI) or matrix-assisted laser desorption ionization (MALDI) are usually used.

1.1.2 Protein identification by mass spectrometry

Before mass spectrometry became the primary tool for proteomics, Edman degradation was used for peptide sequencing. This technique was developed by Peer Edman. During the Edman degradation, a peptide is chemically cleaved, one amino acid at a time, from the amino terminus of the protein. Therefore, a free amino terminus is required. If the protein is acetylated at its amino terminus or is otherwise blocked, the method will fail. During the 1990s, mass spectrometry replaced Edman degradation as the most popular method of protein identification and sequencing, because it is much more sensitive, is not limited to peptides containing a free terminal amino group, and can fragment the peptides in seconds instead of hours or days (Wilm, Shevchenko *et al*. 1996). Additionally, mass spectrometry can sequence peptides containing modified amino acids and does not require pure peptides like Edman degradation. Today, mass spectrometry has become the definitive tool to study the amino acid sequence of proteins including the identification of sites of alkylation and post-translational modification.

There are two main approaches employed in protein analysis by mass spectrometry: top-down and bottom-up (Smith and Bogdanov 2005). Top-down strategy mainly focuses on the intact protein while bottom-up strategy is used to study the peptides generated from enzymatic digestion. In the top-down strategy, an intact, multiply-charged protein is analyzed directly by using electrospray mass spectrometry. Once trapped in the gas phase in a Fourier transform ion cyclotron resonance mass spectrometer, the multiply-charged ion is fragmented using electron capture dissociation, and the fragment ions are measured using MS/MS. Then the resulting tandem mass spectra are database searched to get information about protein identification and

characterization. When an Orbitrap mass spectrometer is used in place of the FT ICR instrument, electron transfer dissociation is used to fragment multiply charged peptides and proteins.

The bottom-up strategy is much easier than the top-down strategy while identifying and characterizing proteins, in terms of both biochemistry and instrumentation. First of all, the enzymatically digested peptides have higher water solubility compared with intact protein; therefore a wider range of proteins can be studied using the bottom-up strategy. Additionally, the sensitivity of the mass spectrometer is much lower for intact protein compared with peptides. Furthermore, tandem mass spectrometry is more efficient for obtaining sequence information from peptides in the range of 5 to 20 residues than from proteins. Most importantly, sequence information for just a few peptides might provide enough information for protein identification.

A typical bottom-up experiment consists of four stages. In stage 1, the proteins from biological samples are separated by gel electrophoresis or liquid chromatography, so that the sample contains no more than a few proteins. In stage 2, proteins are enzymatically digested. Usually, an endopeptidase such as trypsin is used to cleave the protein at the carboxyl side of the amino acids lysine (K) and arginine (R), except when either is followed by proline. In stage 3, the peptide mixture generated from the enzymatic digestion is analyzed by MS and MS-MS. Finally, the peptide masses obtained from MS scanning and their sequence data are searched against a protein database using database search programs such Mascot and Sequest.

Mass fingerprinting is one of the major methods used for bottom-up protein identification (Henzel, Billeci *et al.* 1993). In this method, proteins are often purified by one- or two-dimensional gel electrophoresis, enzymatically digested in the gel, and then mass spectrometry is carried out to analyze the peptides released from the gel bands or spots. The peptides extracted

from the gel spots or bands are analyzed using MALDI or electrospray mass spectrometry. For identification, all of the proteins in selected databases are "in silico" digested based on the specificity of the endopeptidase used, and the resulting peptides masses are calculated. Then the enzymatically digested peptide masses are matched with the calculated theoretical peptides masses for each protein in the database. The protein is identified based on the closest match between the measured peptides masses and the theoretical peptides masses. Although MALDI fingerprinting works well in many cases, peptide sequencing provides additional data that enable protein identification with greater confidence (Steen and Mann 2004). In this method, peptides are sequenced using MALDI MS-MS or else separated by HPLC and then analyzed on-line using MS/MS for protein identification. In this case, reversed HPLC coupled with electrospray tandem mass spectrometry is often used. During the LC-MS process, the most abundant peptide ion in the mass spectrum is selected automatically for further fragmentation, and then product ion tandem mass spectra are acquired. After that, the instrument resumes MS mode and the next cycle starts. The collected masses of precursor ions and the fragment ions are searched against database using algorithms such as Mascot (http://www.matrixscience.com) or Sequest (Thermo Finnigan, San Jose, CA). Confident identification is achieved by matching the experimentally measured precursor and fragment ion masses to the theoretical masses from the protein database.

During MS/MS, peptide ions are selected and then fragmented. Figure 1.3 shows the Roepstorff-Fohlmann-Biemann nomenclature of the peptide fragment ions (Steen and Mann 2004). If the ions contain the amino terminus of the peptide, they are labeled consecutively from the terminus a_m, b_m and c_m, in which the m represents the number of amino acids these ions contain. Peptide fragment ions containing the carboxylic acid terminus are labeled consecutively from the carboxyl terminus as $z_{(n-m)}$, $y_{(n-m)}$ and $x_{(n-m)}$.

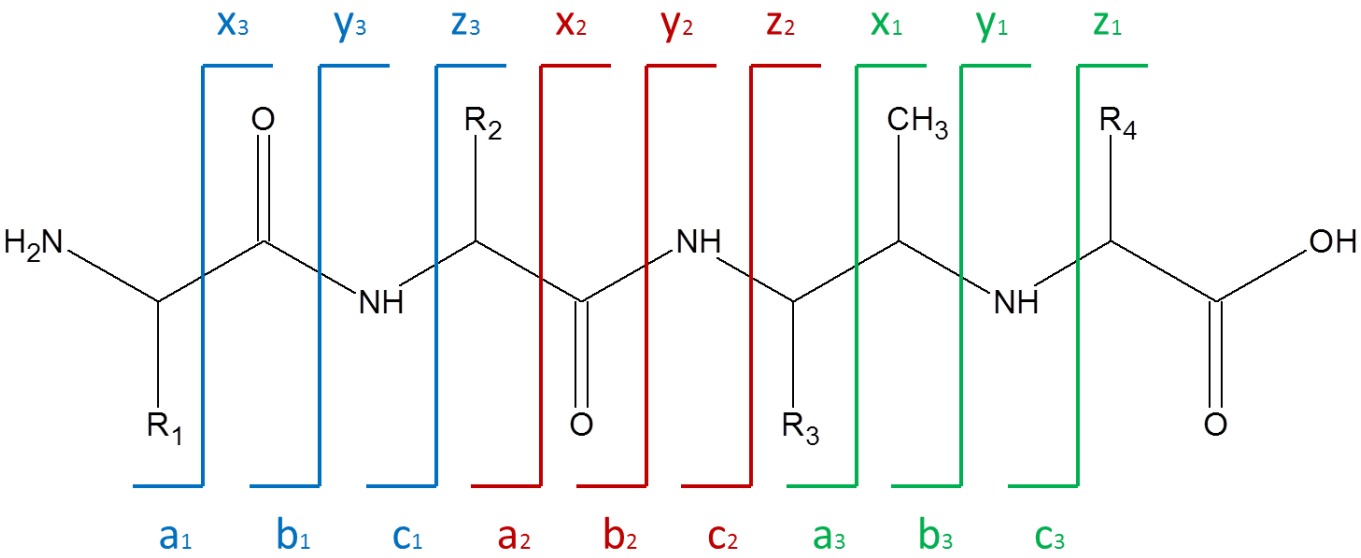


Figure 1.3 Roepstorff-Fohlmann-Biemann nomenclature of peptide fragment ions.

Many dissociation techniques have been used for peptide fragmentation. These techniques includes surface-induced dissociation (Ijames and Wilkins 1990), collision-induced dissociation (CID) (Wells and McLuckey 2005), infrared multi-photon dissociation (Zubarev, Witt *et al.* 2005), sustained off-resonance irradiation (Hofstadler, Wahl *et al.* 1994), electron capture dissociation (ECD) (Jensen, Stensballe *et al.* 2000), and electron transfer dissociation (Hunt, Syka *et al.* 2004). The most common fragmentation methods are CID and ECD. In CID, the molecular ions collide with neutral molecules such as helium, nitrogen or argon, and then break into smaller fragments. The main peptide fragment ions formed during CID are b-ions and y-ions. During ECD, the precursor ion interacts with a free electron to form an odd-electron ion (radical cation), and the liberation of the electric potential energy results in fragmentation to

form predominately c- and z-ions. One of the advantages of ECD is it can retain post-translational modifications during fragmentation, which are often lost during CID. However, ECD still has its drawbacks, for example the low efficiency of fragmentation and the requirement that the precursor ion must be multiply charged.

1.1.3 The protein database search algorithm Mascot

Database searching is a key step in mass spectrometry based proteomics. Mascot is one of the most widely used algorithms, and it was developed from one of the first programs for identifying proteins by peptide mass fingerprinting, MOWSE. Mascot became commercially available in early 1999 through Matrix Science (Boston, MA).

Mascot is database search software designed to identify the characteristics and post-translational modifications of peptides and proteins. The results of a protein identification analysis are presented with probability scores, and the probability-based scoring algorithm involves calculating the theoretically predicted fragments for all the peptides in the database. The predicted fragments are matched to the experimental fragments in a top-down fashion, starting from the most intense b- and y-ions (Figure 1.4). The probability that the number of fragment matches is random is calculated, and the negative logarithm of this number (multiply by 10) is the identification score (Steen and Mann 2004).

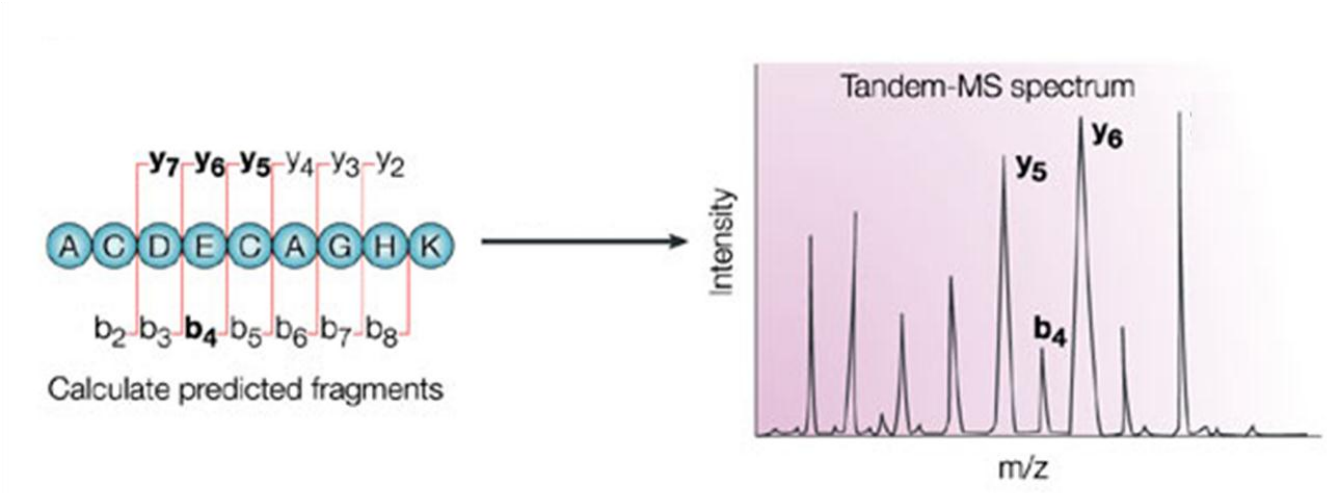


Figure 1.4 The algorithm of protein database search engine Mascot (adapted from Steen and Mann 2004).

1.2 Matrix-assisted laser desorption ionization time of flight mass spectrometry

1.2.1 Matrix-assisted laser desorption ionization

MALDI was first introduced in the 1980s by two key mass spectrometrists, Koichi Tanaka and Franz Hillenkamp, following many years research on the use of lasers for ionization and time-of-flight mass spectrometers for analysis of biomolecules (Karas and Hillenkamp 1988; Tanaka, Waki *et al.* 1988). The principle of MALDI is based on mixing applying nonvolatile analytes with excess amounts of small organic molecules or matrix in a solution of water and a volatile organic solvent (normally acetonitrile, acetone or ethanol). The solvent facilitates mixing of analyte and matrix and dilutes and isolates the analytes to prevent aggregation or formation of

clusters. (Chait and Kent 1992). Several microliters of mixture solution are spotted on a MALDI target plate, and then evaporation will result in a bed of analyte mix/co-crystallized matrix.

The most commonly used matrixes for detection of proteins and peptides in MALDI systems utilizing a UV laser include alpha-cyano-4-sinapinic acid (CHCA); 2,5-dihydroxybenzoic acid (DHB); sinapinic acid (SA); and ferulic acid (FA) (Fenselau and Demirev 2001; Zhang, Zhang *et al.* 2010). If an IR laser is used, then matrix selection includes compounds that absorb strongly in the infrared region such as glycerol. The target plate is placed in a high vacuum source region of mass spectrometer and is irradiated with a pulsed laser beam (usually 337 nm for UV lasers) (Kicman, Parkin *et al.* 2007). The matrix is selected so that it will absorb most of the laser energy since its concentration is much higher than the analytes. Then, sublimation of the matrix forms a dense gas cloud, expanding the matrix and analyte into the gas phase. This cloud expands supersonically, and some of the laser energy is transferred from matrix to analyte. During this desorption step (Figure 1.5), the analytes become protonated (or deprotonated) by a process believed to be the direct result of collision between the neutral analyte and the energetic matrix ions (Gimon, Preston *et al.* 1992). The protonated molecules formed during MALDI are primarily singly charged.

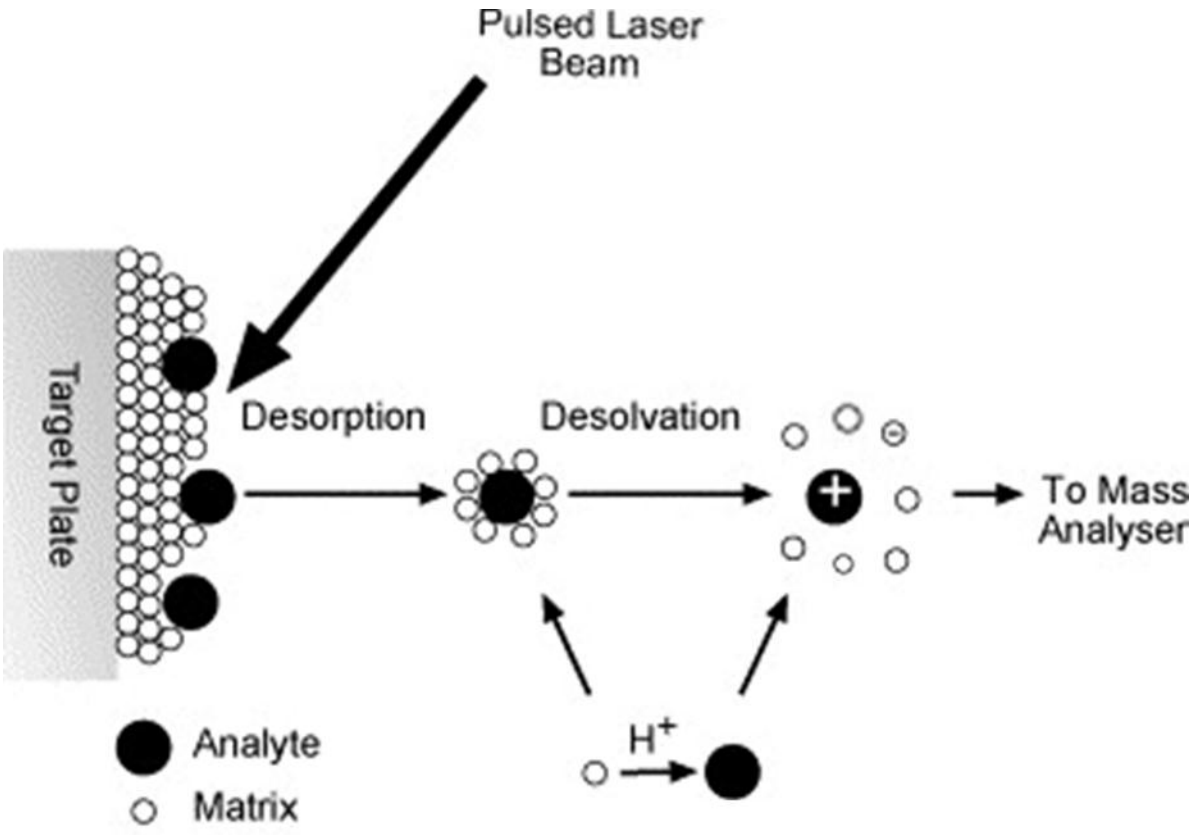


Figure 1.5 Ionization of analytes by MALDI. The co-crystal of matrix and sample is irradiated by a laser pulse causing a desorption process followed by desolvation and then introduction into the mass spectrometer. (adapted from Kicman, Parkin *et al.* 2007)

1.2.2 Time of flight mass spectrometry

Time of flight mass spectrometry (TOFMS) was first reported by Cameron *et al.* in 1948 (Mirsaleh-Kohan, Robertson *et al.* 2008). TOFMS is able to provide high resolution and sensitivity, which has enabled it to become one of the most popular instruments used in proteomics. The TOFMS works by measuring the time required for ions generated in the ion source at an equal energy level to traverse the analyzer and strike an impact detector at intervals

related to their *m/z* values. The principle is based on an ion of mass *m* leaving the ion source with a charge *z* and the accelerating potential *V*, therefore having the energy *zV* equal to the kinetic energy of ion:

$$K = zV = \frac{mv2}{2} \quad (1)$$

If the time taken, *t*, for the ion to fly over the distance *d* of the flight tube at velocity *v* is given by:

$$t = \frac{d}{v} \quad (2)$$

Substituting (2) into (1) gives:

$$t^2 = \frac{m}{z}\left(\frac{d^2}{2V}\right)$$

Since the terms in parentheses (related to a fixed distance and accelerating potential) remain constant, the *m/z* value can be determined by t^2. Since all of the ions have the same kinetic energy, singly charged ions which have a larger mass will move with a lower velocity compared with those ions of smaller mass. An advantage of TOFMS is that it has in theory no upper mass limit, making it ideal for coupling with MALDI for the analysis of large intact biomolecules such as proteins (Kicman, Parkin *et al.* 2007).

TOFMS can be used to measure molecules in a wide mass range from small molecules to large intact proteins. The wide range of applications for TOF analyzers has been further extended by the development of the tandem TOF (TOF/TOF) mass spectrometers (Medzihradszky, Campbell *et al.* 2000). In a TOF/TOF MS, the first TOF mass spectrometer isolates a precursor ion using a velocity filter, and after fragmentation, the second TOF-MS analyzes the fragment

ions. This method allows for specific ion selection and then fragmentation to obtain more structural information about the compound.

1.2.3 MALDI-TOF-TOF: fragmentation of molecules

In the MALDI TOF mass spectrometer, the rate of spontaneous fragmentation of ionized molecules is very low. Use of high laser power can induce molecular fragmentation either in the ion source region (called in-source decay) (Demeure, Gabelica *et al.* 2010) or in the field-free region of a TOFMS (referred to as post-source decay) (Talbo, Suckau *et al*. 2001). Although the above methods have been widely used for peptide and protein sequencing, they still have a major drawback: lack of sensitivity due to the inefficient fragmentation (Medzihradszky, Campbell *et al.* 2000). More efficient methods have been developed, including some that are suitable for TOF MS mass analyzers (Volmer and Sleno 2004). High and low energy collision-induced dissociation (CID) are suitable for TOF MS but not electron capture dissociation (ECD) (McLafferty, Zubarev *et al.* 2000) or electron transfer dissociation (ETD) (Hunt, Syka *et al.* 2004). CID is the most widely used method in the area of mass spectrometry of biomolecules, because of its good compatibility with those most frequently used mass spectrometers including TOF/TOF MS.

In CID, protonated peptides are excited to a higher energy level upon collision with molecules of collision gas, for example nitrogen or argon. Each collision event increases the internal energy of the precursor ion, which finally results in the breakage of a chemical bond (Wells and McLuckey 2005) (Figure 1.6). Peptide ester bonds, for example, are prone to cleavage under CID conditions. This process can be used to sequence peptides based on the measurement of different product ions generated by the fragmentation of one precursor ion

(Biemann 1990; Steen and Mann 2004; Medzihradszky 2005). Additionally, CID of precursor ions can generate a compound-specific fragment ion pattern, which is highly reproducible under same CID conditions. (Kafka, Kleffmann *et al.* 2011).

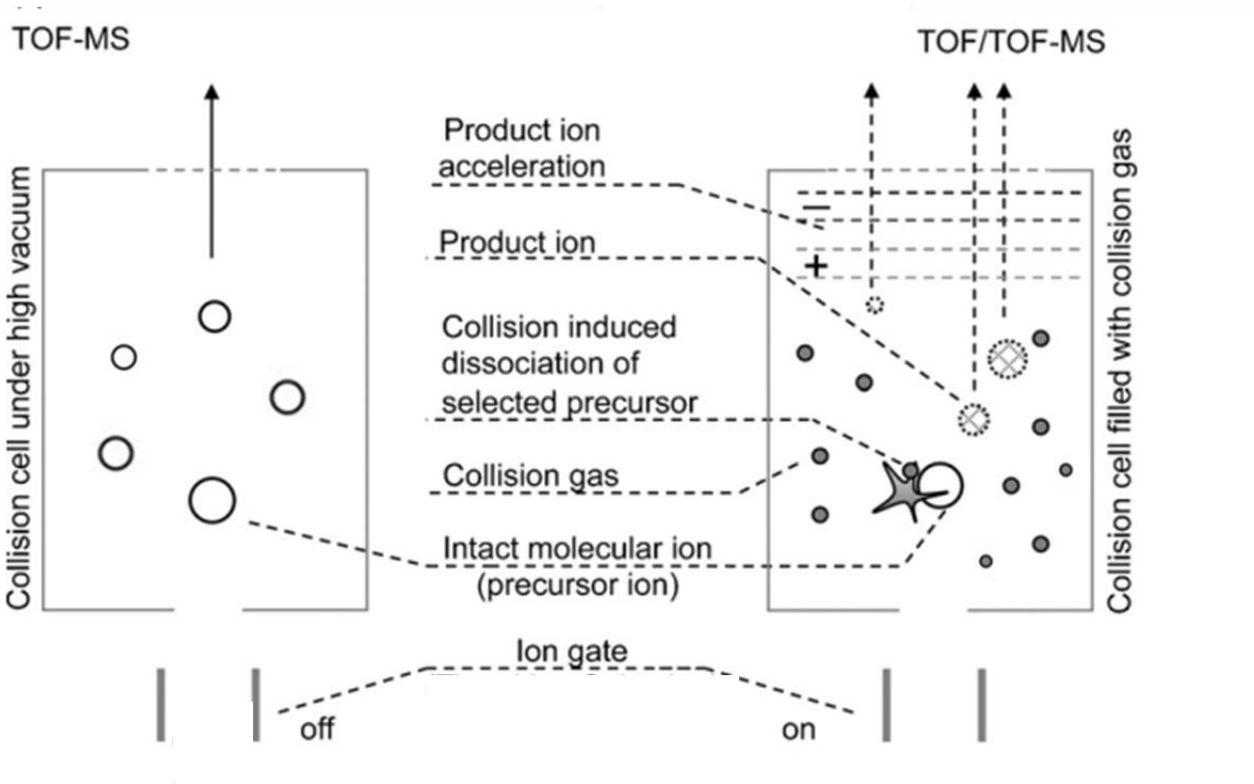


Figure 1.6 Collision induced dissociation (CID) in a tandem time-of-flight (TOF/TOF) mass spectrometer (adapted from Kafka, Kleffmann *et al.* 2011).

1.3 Mass spectrometric studies of protein-protein interactions

The study of protein-protein interactions has provided great insight in biology. Protein-protein interactions are important because they constitute the metabolic and signaling pathways that control the growth and development, structure, operation, replication and selective elimination of cells. Mass spectrometry is a powerful tool for identifying protein-protein interactions in multi-protein complexes. This is in part due to the tremendous advantages that

mass spectrometry offers over other techniques in terms of unambiguous identification of proteins and the accurate measurement of peptide and protein masses (Figeys, McBroom *et al*. 2001). There are several mass spectrometry-based techniques that have rapidly evolved and become used by many research groups, including hydrogen/deuterium exchange and protein footprinting.

1.3.1 Hydrogen/deuterium exchange

Amide proton hydrogen exchange was first developed as a structural technique at the Carlsberg laboratories by Linderstrom-Lang in the 1950s (Hvidt and Linderstrom-Lang 1954; Hvidt and Nielsen 1966). Exchange of an amide proton with a solvent deuteron, referred to as hydrogen/deuterium (H/D) exchange, needs both solvent and the breaking of the amide hydrogen bond (Englander, Mayne *et al*. 1997). Although usually applied to small globular proteins, H/D exchange mass spectrometry (HXMS) experiments may be extended to include larger proteins or protein complexes. The basis of HXMS lies in the fact that proteins and peptides can be ionized and measured with enough mass resolution to detect 1Da mass increments, which results from the exchange of a proton for a deuteron (Lanman and Prevelige 2004).

The basic design of an H/D exchange experiment is to initiate labeling of the protein of interest by performing a 10- or 20 fold dilution of a concentrated stock solution into D_2O. The exchange reaction is sampled at several time points, and the H/D exchange is quenched by lowering the pH to 2.5 and the temperature to 4 °C. After quenching slows the deuterium exchange rate, pepsin is added into the solution to digest the protein. Then the peptide mixture is analyzed using mass spectrometry (Figure 1.7).

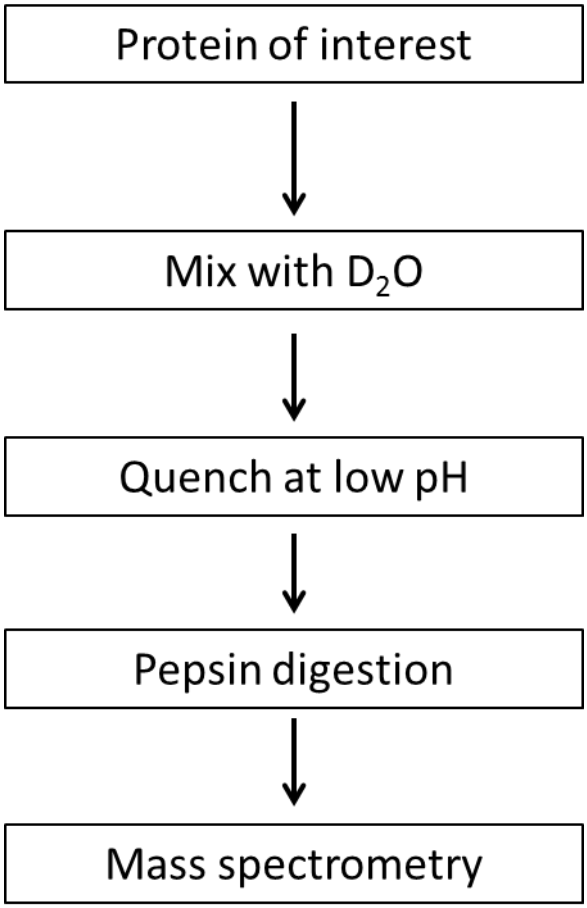


Figure 1.7 Flow chart illustrating the steps involved in performing an H/D exchange mass spectrometry experiment.

1.3.2 Protein Footprinting

The term ’footprinting’ refers to experiments that examine macromolecular structural changes by determining the solvent accessibility of the peptide backbone, bases or side chains of macromolecules based on their sensitivity to chemical or enzymatic cleavage or modification (Guan and Chance 2006; Takamoto and Chance 2006). This approach was first applied to the analysis of the binding of protein to DNA. In the protein/DNA binding studies, footprinting

provided a sensitive and specific structure probe that used a nonspecific enzymatic nuclease to cleave the DNA backbone, then the labeled cleavage fragments were analyzed using gel electrophoresis (Galas and Schmitz 1978; Schmitz and Galas 1979). This approach was then extended to proteins.

In early protein footprinting studies, end-labeled macromolecules were subjected to enzymatic cleavage, and the cleavage products were examined by sodium dodecyl sulfate (SDS) gel electrophoresis to find ligand-dependent changes in backbone conformation (Heyduk and Heyduk 1994; Zhong, Lin *et al*. 1995; Heyduk, Heyduk *et al*. 1997; Shea, Sorensen *et al*. 2000). Nowadays, gentle ionization methods for mass spectrometry, including electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), have been coupled with sensitive detection methods to provide high resolution tools for mass analysis of proteins and peptides at femtomole level (Aebersold and Mann 2003; Ferguson and Smith 2003). Protein footprinting approaches have been developed which use oxidation of proteins by hydroxyl radicals, and the oxidized proteins are then analyzed by proteolysis and mass spectrometric analysis (Goldsmith, Guan *et al*. 2001; Maleknia, Ralston *et al*. 2001; Guan, Vorobiev *et al*. 2002). The reactivity of side chains with •OH can be determined by this analysis (Guan and Chance 2005).

The overall protein footprinting experiment is summarized in Figure 1.8. Proteins are first exposed to hydroxyl radicals that can react with side chain sites which are solvent accessible. Then, proteolysis is utilized to generate well-defined peptide species using site-specific proteases. High performance liquid chromatography (HPLC) coupled with mass spectrometry (LC-MS) is used to separate the peptide mixture and to determine the masses of digested peptides. Within the HPLC chromatogram, the peak areas corresponding to the mass-to-charge values of the modified and unmodified peptides in the total ion chromatogram can be extracted and integrated. The peak

area ratio of unmodified peptides compared with the sum of modified and unmodified peptides is able to provide a quantitative measurement of oxidation extent expressed as fraction unmodified, where the value for the unexposed protein is assigned to 1.0. Multiple experiments that gradually increase the dose of radicals are carried out, and a plot of the unmodified fraction versus time of exposure gives a dose-response curve. This curve is fit to a first order function to provide the modification rate. The protein is then examined under alternative conditions (for example, denaturing conditions) in which conformation changes are expected. Peptides that are seen to experience changes in the observed oxidation rate have side chain residues that change conformation in response to the change in solution conditions or ligand. The specific oxidation sites representing the structural probes are further identified by MS/MS sequencing of the oxidized peptides (Guan and Chance 2005).

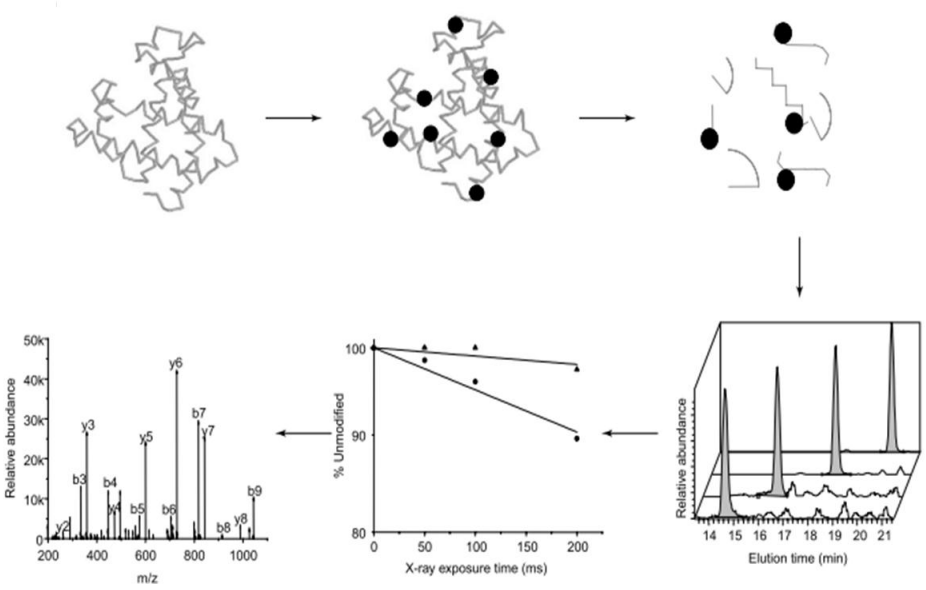


Figure 1.8 Schematic representation of protein footprinting using hydroxyl radicals and mass spectrometry. (adapted from Guan and Chance 2006)

1.4 Cancer chemoprevention

1.4.1 Cancer and carcinogenesis

Cancer is a large, heterogeneous class of diseases in which a group of cells display uncontrolled growth, invasion that intrudes upon and destroys adjacent tissues, and often metastasizes, wherein the tumor cells spread to other locations in the body via the lymphatic system or through the bloodstream. Cancer is the world's second biggest killer after cardiovascular disease, but one of the most preventable noncommunicable chronic diseases. Cancer killed 7.6 million people in 2005, three quarters of whom were in low- and middle-income countries. By 2015, that number is expected to rise to 9 million and increase to 11.5 million in 2030 (WHO 2007). There are several kinds of cancers. Carcinoma is the cancer derived from the epithelial cells that line or cover internal organs. Sarcoma is the cancer arising from the connecting tissues, such as bone, cartilage, fat, muscle and blood vessel. Leukemia is a cancer that starts from the hematopoietic (blood-forming) cells in the bone marrow, and results in an abnormal increase of immature white blood cells. Germ cell tumor is origin from pluripotent cells and often presenting in the testicle or the ovary.

A carcinogen is an agent that has the ability to damage the genome or to disrupt cellular metabolic processes leading to cancer. Carcinogens include chemical compounds, viruses or ionizing radiation. This term is often used more narrowly to mean chemical carcinogens only. Carcinogenic compounds, such as electrophiles and reactive oxygen species (ROS), contribute to DNA damage and the development of malignancy (Heidelbe 1970; Sims, Grover et al. 1974; Ames 1983).

Carcinogenesis is a multistage process that consists of three independent but closely linked processes: initiation; promotion and progression. Initiation is defined as irreversible modification of the target cell DNA, while promotion stands for the expansion of initially damaged or mutated cells to form a clone of an actively proliferating multicellular premalignant/benign tumor cell population, which eventually progresses to the malignancy characterized by increased invasiveness and metastatic potential (Surh 2002).

1.4.2 Chemoprevention and chemopreventive agents

Although there is no 'magic bullet' that can completely conquer cancer, cancer risk still can be reduced by eliminating exposure to identified carcinogens. Up to 40% all cancer deaths can be avoided by reducing tobacco use, improving diets and physical activity, lowering alcohol consumption, eliminating workplace carcinogens, and immunizing against hepatitis B virus and the human papillomavirus (WHO 2007). The concept of chemoprevention, as first defined by Sporn in 1976, uses natural, synthetic, or biologic chemical agents to reverse, suppress or prevent carcinogenic progression (Sporn 1976). After Wattenberg articulated the concept of selective inhibition of carcinogenesis during initiation, promotion or progression phases (Wattenberg 1990), the field of cancer chemoprevention has expanded.

It has been estimated that more than two-thirds of human cancer would be preventable by appropriate lifestyle modification. Recently, a high level of awareness and interest has developed concerning the chemopreventive activities of dietary fruits and vegetables. Doll and Petro reported that 10% to 70% (average 35%) of human cancer mortality could be attributed to diet (Doll and Peto 1981). Cancer cell lines and animal models have been developed and are being used to evaluate the chemopreventive effects of dietary compounds. These effects have yielded

great success in the discovery of novel chemopreventive agents such as polyphenols in green tea and flavonoids from soybeans. These compounds can effectively protect cells against damage caused by electrophilic or ROS by inducing cytoprotective genes which could enhance the expression level of phase II antioxidant and detoxification enzymes (Talalay, Fahey *et al.* 1995).

1.4.2.1 Natural products in chemoprevention

Throughout history, many naturally occurring agents have been used for the prevention and treatment of cancer and other disease. However, mechanisms of diet-derived chemoprevention were not understood until the late twentieth century. Since the National Cancer Institute initiated a diet and cancer program in the 1970s, studies have shown that phytochemicals in vegetables and fruits contribute significantly to the prevention of chronic diseases, such as cancer, diabetes and hypercholesteremia (Colic and Pavelic 2002; Mehta, Murillo *et al.* 2010).

1.4.2.2 Existing chemopreventive agents from natural products

Curcumin, from turmeric, is reported to suppress tumor promotion of skin carcinogenesis by reducing the gene expression of cyclooxygenase-2 (COX2) (Plummer, Holloway *et al.* 1999). Epigallocatechin gallate (EGCG), originally from green tea, has been shown to suppress malignant transformation in the mouse epidermal JB6 cell line by blocking the activation of Ap1 (Dong, Ma *et al.* 1997; Dong, Nomura *et al.* 2000) and NF- κ B (Yang, Chung *et al.* 1999). Genistein, a soy-derived isoflavone, inhibits the hydrogen peroxide- or TNF- α -induced activation of NF- κ B in human prostate cancer cell lines (Sarkar, Davis *et al.* 1999), and resveratrol, a phytoalexin in grapes, has been shown to induce apoptosis in both rat and human

pancreatic carcinoma cell lines (Gukovskaya, Mouria *et al.* 2002). Other common chemopreventive agents and their dietary sources are shown in Figure 1.9.

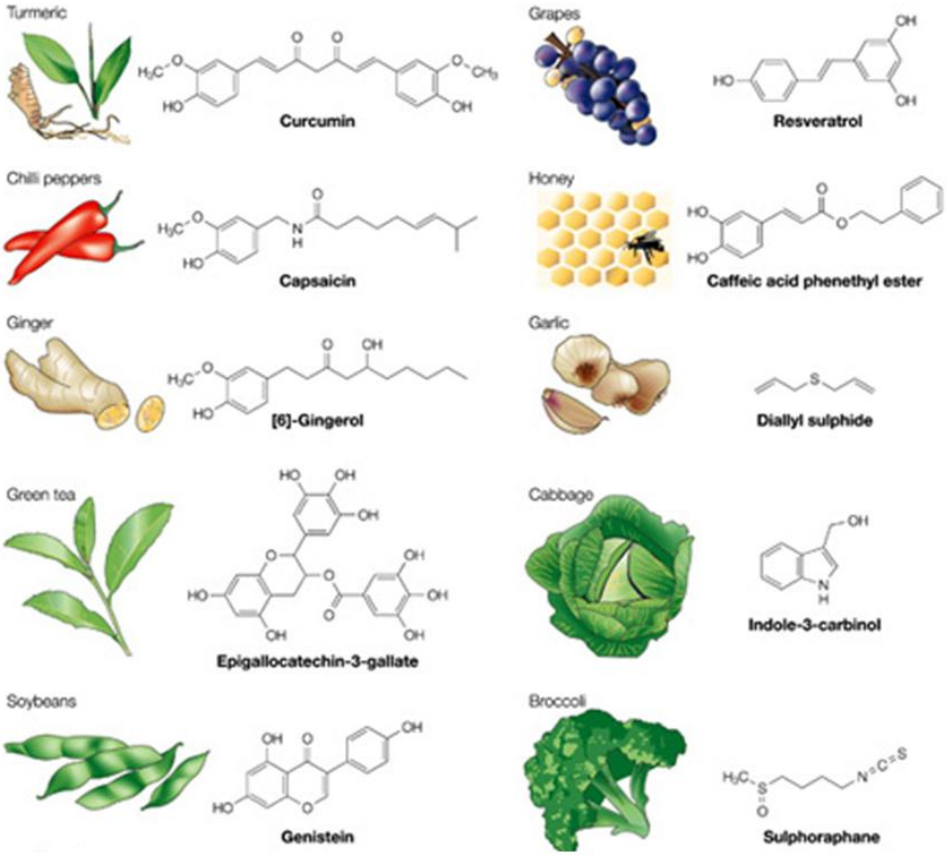


Figure 1.9 Representative chemopreventive agents and their dietary sources (adapted from Surh 2003).

1.5 Keap1-Nrf2 system

Nrf2 (Figure 1.10a) is a member of the NF-E2 family of nuclear basic leucine zipper transcription factors. In the cell nucleus, Nrf2 binds to the 5' upstream regulatory antioxidant response element (ARE) regions of phase 2 genes and accelerates their transcription. These enzymes provide efficient cytoprotection by regulating the intracellular redox state (Alam, Gong *et al.* 2002; Itoh, Mochizuki *et al.* 2004). Many natural antioxidants and potential chemopreventive agents have been shown to enhance the expression level of Nrf2 and induce phase II gene expression (Balogun, Hoque *et al*. 2003; Jemal, Murray *et al*. 2005). The Neh2 *N*-terminal domain of Nrf2 plays an important role in redox-dependent regulation of protein stability due to its interaction with Keap1 (Kelch-like ECH associating protein 1) and conjugation with ubiquitin (Zhang 2006). Domains Neh3, Neh4 and Neh5 mediate Nrf2 transactivating effects by binding to histone acetyltransferase (Pickett, Nioi *et al*. 2005; Osada, Ohta *et al.* 2007; Zhang, Hosoya *et al.* 2007). Human Keap1 protein consists of five domains (Figure 1.8b): 1) *N*-terminal region (NTR); 2) BTB domain responsible for dimerization and interaction with Cul3-E3-ligase (Zipper and Mulcahy 2002; Yamamoto, Kobayashi *et al*. 2004; Tanaka, Tong *et al*. 2006); 3) IVR (Intervening region) contains cysteine residues that is sensitive to attack by oxidative species (Talalay, Dinkova-Kostova *et al*. 2002); 4) Kelch-domain consists of six repeats (KR1-KR6) and possessing the structure of a six-bladed β-propeller that mediates association of Keap1 with Nrf2 and cytoskeleton proteins actin and/or myolin VIIa (Zhang 2006; Kong and Li 2009); and 5) the C-terminal region.

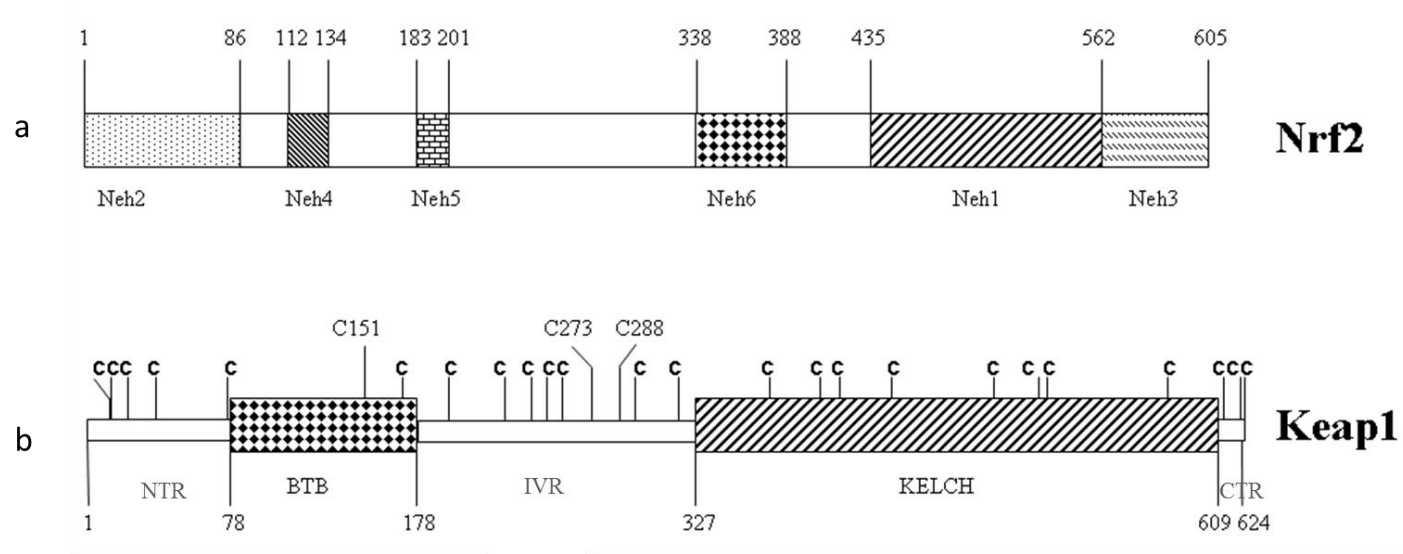


Figure 1.10 Schematic representations of different regions of human Nrf2 (a) and Keap1 (b) proteins. (adapted from Zhang 2006)

Under basal conditions, Nrf2 expression in the cell is low because of continuous ubiquitination by the Culln3-based (Cul3) E3 ligase ubiquitination complex and degradation (Zhang 2006). Therefore, phase II gene expression is suppressed under basal conditions. Keap1 (Figure 1.10b) serves as a bridge between Nrf2 and Cul3 and binds both proteins. When cells are under oxidative stress, signals are transmitted to the Keap1-Nrf2 complex, which interrupts ubiquitination of Nrf2 and increases its concentration in cells, including the nucleus. Consequently, the protective phase II enzymes are up-regulated by the increasing level Nrf2 in the nucleus (Yamamoto, Itoh *et al.* 2003) (Figure 1.11)

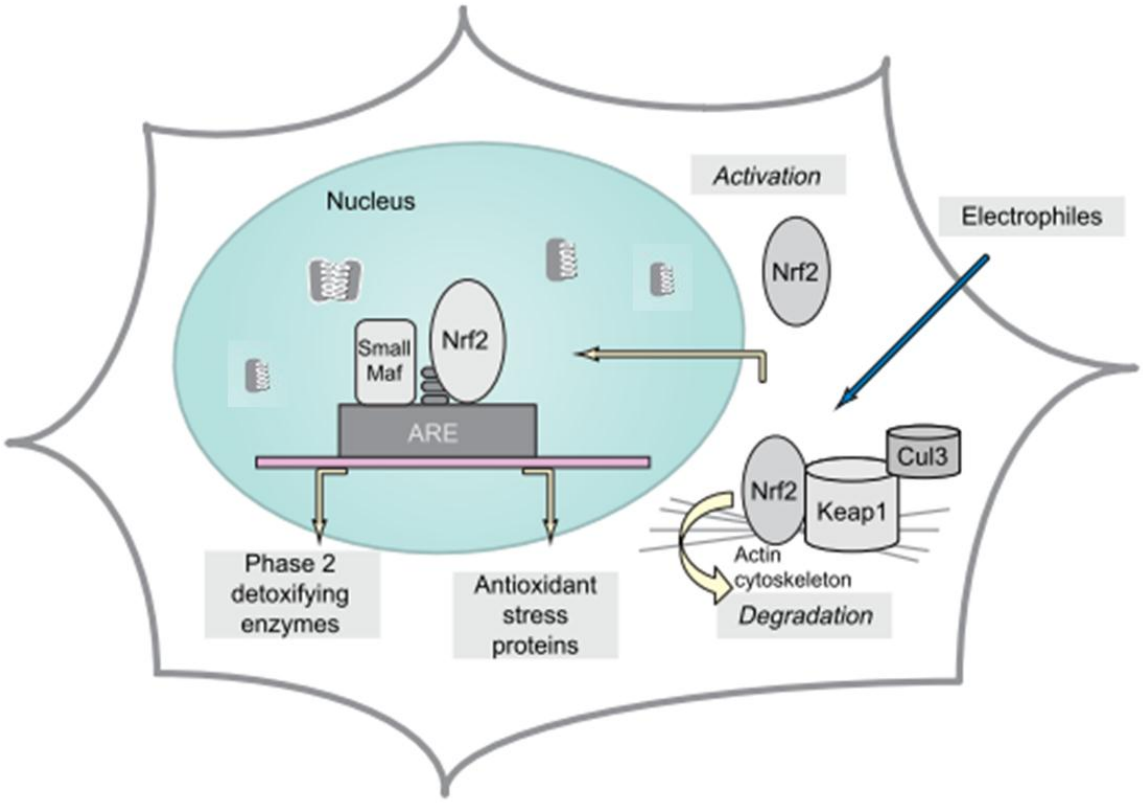


Figure 1.11 Schematic illustration of the molecular mechanisms of the Nrf2–Keap1 regulatory system and its endogenous activating signals(adapted from Motohashi and Yamamoto 2004).

The mechanism of how cancer chemopreventive agents activate Nrf2 pathway is still under investigation. Recent studies proposed several hypotheses to explain the regulatory mechanisms involving Nrf2-Keap1. These studies include direct and indirect signal transductions of the Keap1-Nrf2 pathway.

One possible chemoprevention mechanism is disruption of the Keap1-Nrf2 complex. Itoh *et al.* first hypothesized that the Nrf2 is released from the Keap1-Nrf2 complex in the cytosol following exposure to electrophiles such as ROS that would covalently modify one or more of the 27 cysteine sulfhydryl groups of Keap1. Then, Nrf2 would be free to translocate to the

nucleus and activate the ARE-response genes. However, this mechanism was disproved by Mesecar and coworkers (Mesecar, Eggler, *et al.* 2005) who showed that modification of Keap1 cysteine residues did not alter binding between Keap1 and Nrf2. Mass spectrometry and site specific mutagenesis studies found certain cysteine residues of Keap1, including C257, C273, C288 and C297 play important role in the regulation of Nrf2 degradation (Zhang, Krutchinsky *et al*. 2003; Dinkova-Kostova, Holtzclaw *et al.* 2005). Subsequently, our group and others showed that the most important site of Keap1 modification by electrophiles leading to up-regulation of the ARE is C151 (Luo, Eggler, *et al.* 2007).

Another possible mechanism that has been proposed for the regulation of Nrf2 levels by Keap1 is to regulate the ubiquitination and degradation of Nrf2. Under basal conditions, Nrf2 has a short half-life of ~15 to 45 min depending on cell types. The degradation of Nrf2 is mediated primarily by the ubiquitin-26S proteasome pathway, and Keap1 binds a Cul3-based E3 ligase adaptor protein that targets Nrf2 for ubiquitination and degradation (Hannink and Zhang 2003; Diehl, Cullinan *et al*. 2004). Therefore, modification of Keap1 by chemopreventive agents might prevent ubiquitination of Nrf2 by Cul3 and thereby increase the half-life and concentration of Nrf2.

Phosphorylation of Nrf2 has also been proposed to be an Nrf2-Keap1 signaling mechanism. Several protein kinases including phosphatidylinositol-3-kinase, protein kinase C, mitogen-activated protein kinases, and RNA-dependent protein kinase-like endoplasmic reticulum kinase have been reported to participate in the phosphorylation of Nrf2 (Kobayashi and Yamamoto 2005). Activation of certain protein kinase, such as protein kinase C-δ, can promote the stabilization and nuclear localization of Nrf2, thereby enhancing the induction of ARE-dependent phase II genes in mammalian cell lines (Niture, Jain *et al.* 2010).

The mechanism whereby the ARE is regulated via Keap1-Nrf2 signaling is an area of active investigation. A better understanding of this cellular signaling pathway induced by chemoprevention agents might provide valuable new targets and facilitate the design of new therapeutic agents against cancer.

1.6 Ultrafiltration mass spectrometry

Ultrafiltration (UF) was originally developed as a size-based separation method for macromolecules such as proteins and polymers. Now the ultrafiltration has become a filtrate selection method with a wide range of biomedical and clinical applications (Leegsma-Vogt, Janle *et al.* 2003). During ultrafiltration, a solution containing solutes of different sizes is passed through a semipermeable membrane. The molecular cut-off of the semipermeable filtration membrane determines the maximum size of molecule that can pass through the membrane. The pore size of ultrafiltration membrane ranges from 5 to 100 nm which retains molecules in the range of molecular weight of 10 kD to 1 MD. The applicability of ultrafiltration has been demonstrated in several biomedical and clinical applications, including ion dynamics (Linhares and Kissinger 1993); mineral metabolism (Janle and Sojka 2000); drug kinetics (Linhares and Kissinger 1993), and protein studies (Schneiderheinze and Hogan 1996).

Ultrafiltration mass spectrometry was invented and developed in the van Breemen laboratory for screening combinatorial library mixtures and natural product extracts in order to identify ligands of macromolecular targets, such as adenosine deaminase (van Breemen, Huang *et al.* 1997; Zhao, vanBreemen *et al.* 1997); dihydrofolate reductase (van Breemen and Nikolic 1998); cyclooxygenase-2 (van Breemen, Nikolic *et al*. 2000); serum albumin (van Breemen, Gu

et al. 1999) and estrogen receptors (ER) (van Breemen, Sun *et al.* 2005). During ultrafiltration mass spectrometric screening, ligands in the mixture are able to bind to the specific target protein receptor, and then ultrafiltration is used to separate the receptor-ligand complex from unbound low mass compounds. The ligands can be released from the receptor-ligand complex by changing pH value, adding organic solvent or other destabilizing methods. Finally, the ligand can be analyzed in a mass spectrometer. Control assays are carried out using inactive receptor or no receptor to test for non-specific binding to the biomolecules or the ultrafiltration membrane. This screening process is summarized in Figure 1.12.

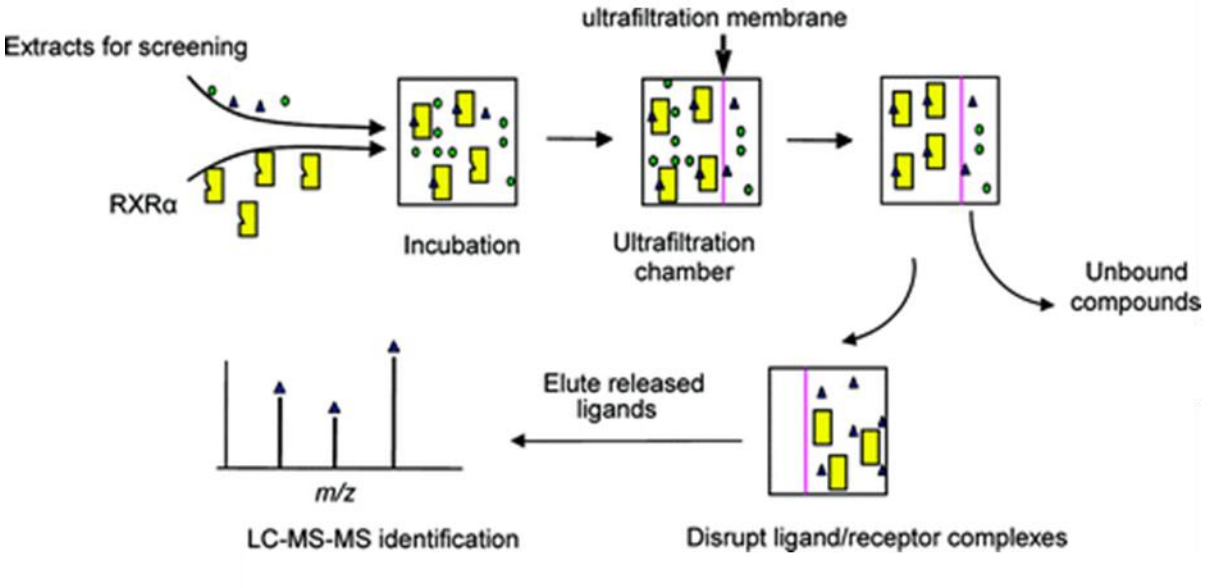


Figure 1.12 Experimental design of ultrafiltration LC−MS/MS screening of solutions for ligands to human RXRα (adapted from Liu, Guo *et al*. 2007)

2 INVESTIGATION OF KEAP1-NRF2 BINDING USING ULTRAFILTRATION MASS SPECTROMETRY

2.1 Introduction

Due to the recent enhancements in biomedical mass spectrometry that have made protein characterization and pharmaceutical analysis easier and faster, mass spectrometry is becoming integrated into the discovery of protein-protein interactions. Recombinant protein techniques and human genome sequencing also provide excellent tools to study protein interactions. By combining the power of mass spectrometry and recombinant proteins, new opportunities are available for the identification of protein interaction binding sites. As a simpler alternative to H/D exchange and mass footprinting, which are mass spectrometry-based methods for investigating protein-protein interactions, a new mass spectrometry-based method, based on ultrafiltration LC-MS, was developed and tested for the discovery of the protein interaction sites of Keap1 and Nrf2.

The binding mechanism of Keap1-Nrf2 has been studied for a long time, but the exact protein-protein interaction binding sites remain unknown. Keap1 has five distinct domains: the *N*-terminal domain (amino acids 1-60); the BTB domain (amino acids 61-178); a central linker domain (amino acids 179-321); the Kelch repeat domain (amino acids 322-608), and a C-terminal domain (amino acids 609-625). The BTB domain not only facilitates the dimerization of Keap1 (Zipper and Mulcahy 2002), but also binds to the adaptor protein in the Cul3-dependant ubiquitination system (Dinkova-Kostova, Holtzclaw *et al.* 2005). Although Itoh *et al.* proposed that the Kelch repeat domain binds to Nrf2 based on their electrophilic agent activation experiments (Yamamoto, Itoh *et al.* 1999), there is still no direct evidence indicating which

amino acids or peptide sequences are involved in specific binding of Keap1 to Nrf2. Protein crystallography has been ineffective in addressing this issue, since crystals of Keap1/Nrf2 have not yet been formed. In this investigation, a method based on ultrafiltration and MALDI-TOF/TOF mass spectrometry was developed to facilitate the characterization of binding sites of Keap1 that interact with Nrf2.

2.2 Experimental section

2.2.1 Protein expression and preparation

Recombinant human Keap1 protein was provided by collaborators Dr. A.D. Mescar and Dr. A.L. Eggler of Purdue University (West Lafayette, IN). The recombinant Keap1 containing a histidine-tag was expressed in and purified from *Escherichia coli*. The details of the cloning, expression, and purification of Keap1 have been published previously (Mesecar, Eggler *et al.* 2005). Keap1 cDNA from *Homo sapiens* skin cells was obtained from the American Type Culture Collection (ATCC, Manassas, VA). The gene encoding full-length Keap1 was directionally cloned using PCR into the pET15b vector (Novagen, San Diego, CA) for expression of an *N*-terminal-$(His)_6$-tagged version of the protein. The entire gene sequence in the expression vectors was verified by dideoxy sequencing. The Keap1 protein was expressed in *Escherichia coli* Rosetta (DE3) pLysS cells (Novagen, San Diego, CA) by induction with isopropyl β-D-thiogalactoside. Cells were harvested by centrifugation, and the Keap1 culture was raised to pH 7.8 with 1M Tris buffer before harvesting. Cell pellets were frozen at -80 ℃.

For purification, the cell pellet (15 g) was resuspended in 50 mL of Buffer A (50mM Tris pH 8.0, 500 mM NaCl, 10mM imidazole) which also contained 7.5 mM $MgSO_4$, 500 µg/mL

lysozyme, 3 mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP), a Roche Complete EDTA-free protease inhibitor tablet, and a small amount of DNase. The cells were lysed by sonication on ice and centrifuged at 40,000×*g*. The cell lysate was loaded into a 5 mL HiTrap column (Amersham Biosciences, Piscataway, NJ) bound with Ni^{2+} and equilibrated with Buffer A. The column was washed with 5% Buffer B consisting of 50 mM Tris pH 8.0, 500 mM NaCl and1M imidazole. Fractions were collected, and fraction purity was analyzed by SDS-PAGE. Pooled fractions were dialyzed against Buffer C (10 mM NaCl, 20 mM Tris pH 8.0, and 2 mM 1,4-dithio-D,L-threitol (DTT)) at 4℃ under N_2. Keap1 was then loaded onto a Mono Q column (Amersham Biosciences) equilibrated with Buffer C, washed with Buffer C, and eluted with a 0-40% gradient of Buffer D (1 M NaCl, 20 mM Tris pH 8.0, 5.5 mM DTT). Pure fractions were pooled, buffer was exchanged with 4 mM TCEP, 21 mM Tris pH 8.0 by multiple concentrations steps using a Centricon 20 ultrafiltration centrifuge tube with a 10,000 nominal molecular weight cut-off membrane (Amicon, Billerica, MA). After determining the protein concentration, an equal volume of freezing buffer (40% (v/v) glycerol, 29 mM Tris, pH 8.0) was added, and the protein aliquots were flash frozen on dry ice and stored at -80℃. During purification and subsequent experiments, Keap1 was kept under 4℃.

The gene of human Neh2 domain (amino acids 1-97) of Nrf2 was cloned by using PCR into pET15b vector. After expressing in *Escherichia coli*, the protein was purified by using a HiTrap-Ni^{2+}-chelate affinity column followed by a MonoQ anion-exchange column. The details of the cloning, expression, and purification of the Neh2 domain of Nrf2 have been published previously (Mesecar, Eggler *et al.* 2005).

2.2.2 Tryptic digestion and deactivation of trypsin

Keap1 (50 µg) was heated to 95 ℃ for 5 min and then cooled to room temperature. The denatured Keap1 was incubated with 1 µg trypsin in 25 mM ammonium bicarbonate buffer (pH 8.0) at 37 ℃. After 3 h, another 1 µg trypsin was added, and the solution was incubated at 30 ℃ overnight. After overnight digestion of Keap1, the remaining trypsin was deactivated by incubation at 55 ℃ for 30 min (Mária L. Simon 2001). The solution was stored at -80 ℃ until use.

2.2.3 Binding of Keap1 tryptic peptides to the Neh2 domain and ultrafiltration

In preparation for ultrafiltration screening, 20 µL Keap1 tryptic digest and 5 µL Neh2 (4.2 µM in binding buffer) were mixed with 100 µL of binding buffer, which consisted of 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 8) and 1 mM DTT. The mixture was incubated for 10 min on ice. Since the molecular mass of Neh2 is 12 kD, 10,000 molecular mass cutoff filters were used for ultrafiltration. After incubation, each mixture was filtered through a Microcon (Millipore, Bedford, MA) YM-10 centrifugal filter containing a regenerated cellulose ultrafiltration membrane with a 10,000 MW cutoff at 13,000 *g* for 15 min at 4 ℃. The Neh2/tryptic Keap1 peptide mixture was washed three times with 200 µL aliquots of 50 mM HEPES solution (pH 8.0) followed each time by centrifugation at 13,000*g* for 15 min at 4 ℃ to remove unbound peptides. The washed solution was treated with 200 µL 90% methanol in deionized water to disrupt the non-covalent binding between Neh2 and Keap1 peptides. Then, the released peptides were isolated from the denatured Neh2 domain by centrifugation at 13,000 *g* for 15 min.

The ultrafiltrate containing released peptides was desalted using Millipore Ziptip C_{18} pipette tips. Before use each Ziptip pipette tip was equilibrated with three portions of 10 µL equilibration buffer consisting of 0.1% trifluoroacetic acid (TFA) in deionized water. Then, the peptides were bonded to the Ziptip pipette tips by fully depressing the pipette plunger and aspirating and dispensing the sample for 10 cycles. The peptides were eluted by aspirating and dispensing 10 µL elution buffer (methanol/deionized water/TFA; 50:49.9:0.1, v/v/v) through the Ziptip pipette tip at least three times without introducing any air. The eluents were frozen at -80 ℃ until analysis using mass spectrometry.

2.2.4 MALDI spot preparation

In preparation for analysis using MALDI-TOF mass spectrometry, a 2.0 µL aliquot of a matrix solution consisting α-cyano-4-hydroxycinnamic acid at a concentration of 10 mg/mL in acetonitrile/water/ TFA (50:49.9:0.1, v/v/v) was mixed with 2 µL of 5 µM tryptic digest. Immediately before analysis, 1 µL of the mixture was spotted on the MALDI sample plate and air-dried.

2.2.5 MALDI-TOF and MALDI-TOF-TOF mass spectrometry

Positive ion MALDI-TOF mass spectra were acquired over the range of *m/z* 700-4,000 for peptides using an Applied Biosystems (Foster City, CA) Voyager DE-Pro MALDI-TOF mass spectrometer, which was operated in linear mode for proteins and in reflector mode for peptides.

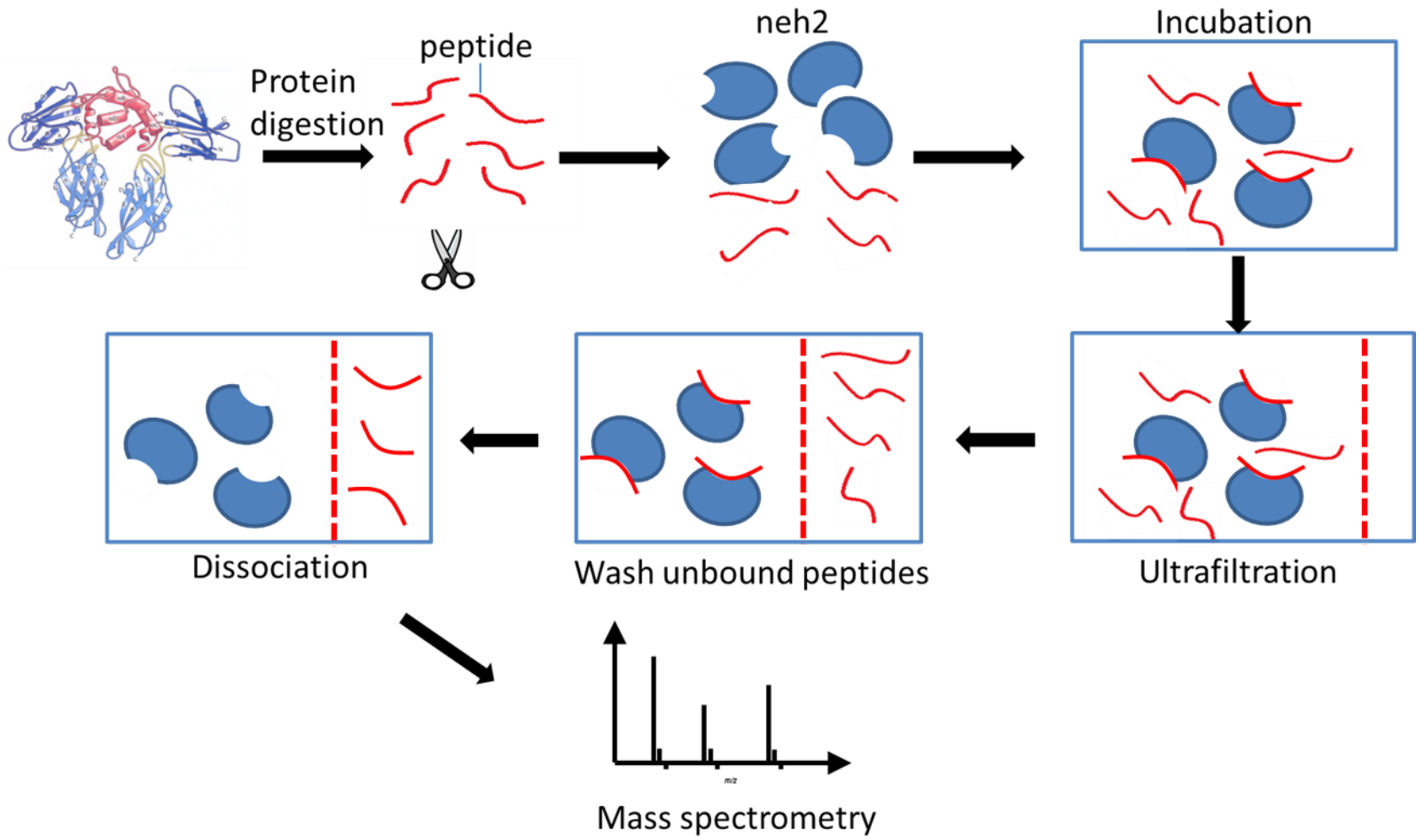
After time-delayed extraction, the ions were accelerated to 20 kV for TOF mass spectrometric analysis.

MS/MS analysis was carried out using an Applied Biosystems 4700 Proteomics analyzer TOF-TOF mass spectrometer using linear mode. A nitrogen laser (λ=337 nm) was utilized, and the laser intensity was set just above the threshold for ion production. The mass spectrometer was used in the positive ion mode with an acceleration voltage of 15 kV. Mass spectra were obtained over the range up to *m/z* 2,100. The instrument was calibrated externally every 20 samples. For each mass spectrum, 50 laser shots (0.5 ns each) at up to 40 locations per sample (a total of up to 2000 laser shots per mass spectrum) were acquired and summed. The TOF measurements were converted to *m/z* values using Applied Biosystems Data Explorer software. These mass spectra were exported as peak lists for Mascot database searching.

2.3 Results and Discussion

2.3.1 Method development of ultrafiltration mass spectrometric screening for protein-protein interaction studies

An ultrafiltration mass spectrometric screening assay has been developed for the studies of protein-protein interactions. This method was optimized to minimize potential interference due to non-specific binding of tryptic peptides to the ultrafiltration membrane. Also, interference from trypsin acting on the protein target during the binding assay was minimized by including a trypsin deactivation step. This step significantly reduced the activity of trypsin remaining in the peptide mixture used in the binding incubation. Overall, ultrafiltration mass spectrometry is a powerful new method for the study of protein-protein interaction (Figure 2.1)



.

Figure 2.1 Schematic representation of a new application of ultrafiltration mass spectrometry developed during this thesis research.

2.3.2 Mass spectrometric mapping of Keap1 tryptic peptides

The purified human Keap1 protein was digested with trypsin overnight. The MALDI mass spectrum of the tryptic digest was acquired and shown in Figure 2.2. The peptide mass fingerprint was compared with the *in silico* digestion results, and the measured peptides covered over 94% of the sequence of human Keap1 (Figure 2.3).

2.3.3 Ultrafiltration mass spectrometric screening for binding of Keap1 tryptic peptides to Neh2

To determine if any Keap1 peptides bind to the Neh2 domain of Nrf2, the peptides from the Keap1 tryptic digest were assayed using ultrafiltration mass spectrometry for binding to Neh2. Neh2 domain locates at the N terminus of Nrf2 and interacts with the cytoplasmic protein Keap1 (Itoh, Wakabayashi *et al*. 1999). The positive ion MALDI-TOF mass spectrum of the ultrafiltrate containing peptides that bound to Neh2 is shown in Figure 2.4. As a control, another ultrafiltration assay was carried out that was identical except that Neh2 was denatured by incubation at 55 ℃ for 30 min, and the MALDI mass spectrum of the control ultrafiltrate is shown in Figure 2.5. After comparing the mass spectrum obtained using ultrafiltration mass spectrometry with Neh2 and the Keap1 tryptic digest to that obtained in the control containing denatured Neh2, only the ion of *m/z* 2100 was found to be significantly enhanced. Therefore, the ion of *m/z* 2100 likely corresponds to a Keap1 peptide that binds to Neh2.

In the control mass spectrum (Figure 2.5), the ion of *m/z* 2100 was not abundant. Instead, the most abundant ion in the control mass spectrum was detected at *m/z* 1028 which corresponded to a tryptic peptide from Neh2. To minimize the interference of trypsin, the tryptic

digested Keap1 peptides had been incubated at 55℃ for 30 min to deactivate trypsin. Nevertheless, some tryptic digestion of Neh2 was still observed.

To confirm that the ion of *m/z* 2100 corresponds to a human Keap1 tryptic peptide, MS-MS sequencing was used to characterize the ion. The MALDI-TOF-TOF spectrum (Figure 2.6) was analyzed by database searching using Mascot. As shown in Figure 2.7, the product ion tandem mass spectrum of *m/z* 2100 corresponded with high confidence to a Keap1 peptide with sequence ^{337}QSLSYLEAYNPSDGTWLR^{354}. The ion score of this MS/MS ion search result is 113, which indicates identity or extensive homology ($p<0.05$). These results suggest that the Keap1 tryptic peptide ^{337}QSLSYLEAYNPSDGTWLR^{354} corresponds to a region of Keap1 that binds to the Neh2 domain of Nrf2.

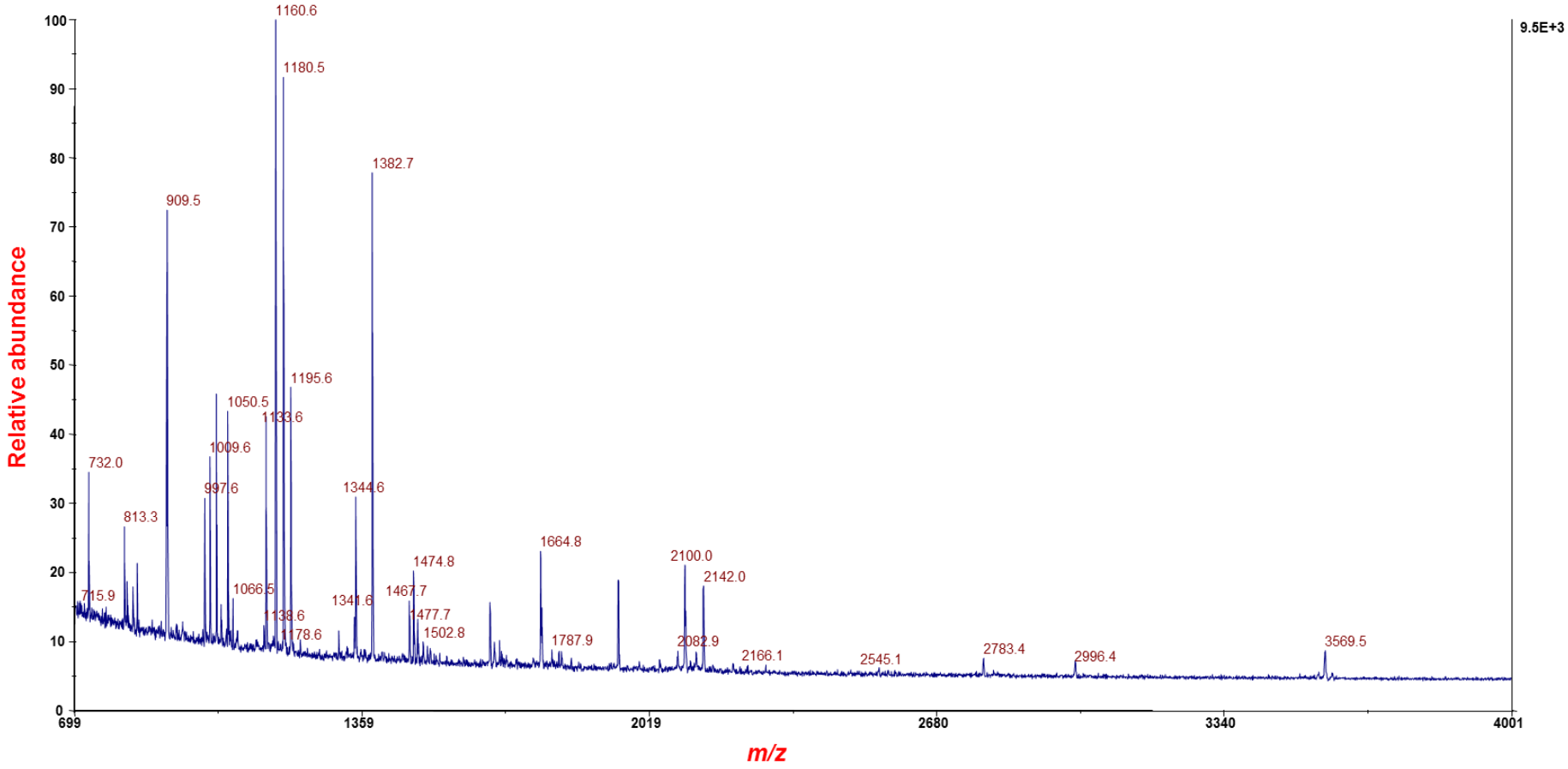


Figure 2.2 Positive ion MALDI-TOF mass spectrum of peptide map of human Keap1 digested by trypsin.

His-tag

1 MGSSHHHHHH SSGLVPRGSH MQPDPRPSGA GACCRFLPLQ SQCPEGAGDA VMYASTECKA EVTPSQHGNR TFSYTLEDHT

81 KQAFGIMNEL RLSQQLCDVT LQVKYQDAPA AQFMAHKVVL ASSSPVFKAM FTNGLREQGM EVVSIEGIHP KVMERLIEFA

161 YTASISMGEK CVLHVMNGAV MYQIDSVVRA CSDFLVQQLD PSNAIGIANF AEQIGCVELH QRAREYIYMH FGEVAKQEEF

241 FNLSHCQLVT LISRDDLNVR CESEVFHACI NWVKYDCEQR RFYVQALLRA VRCHSLTPNF LQMQLQKCEI LQSDSRCKDY

321 LVKIFEELTL HKPTQVMPCR APKVGRLIYT AGGYFRQSLS YLEAYNPSDG TWLRLADLQV PRSGLAGCVV GGLLYAVGGR

401 NNSPDGNTDS SALDCYNPMT NQWSPCAPMS VPRNRIGVGV IDGHIYAVGG SHGCIHHNSV ERYEPERDEW HLVAPMLTRR

481 IGVGVAVLNR LLYAVGGFDG TNRLNSAECY YPERNEWRMI TAMNTIRSGA GVCVLHNCIY AAGGYDGQDQ LNSVERYDVE

561 TETWTFVAPM KHRRSALGIT VHQGRIYVLG GYDGHTFLDS VECYDPDTDT WSEVTRMTSG RSGVGVAVTM EPCRKQIDQQ

641 NCTC

Figure 2.3 Amino acid sequence of Keap1. The tryptic peptides that were detected (94%) are highlighted.

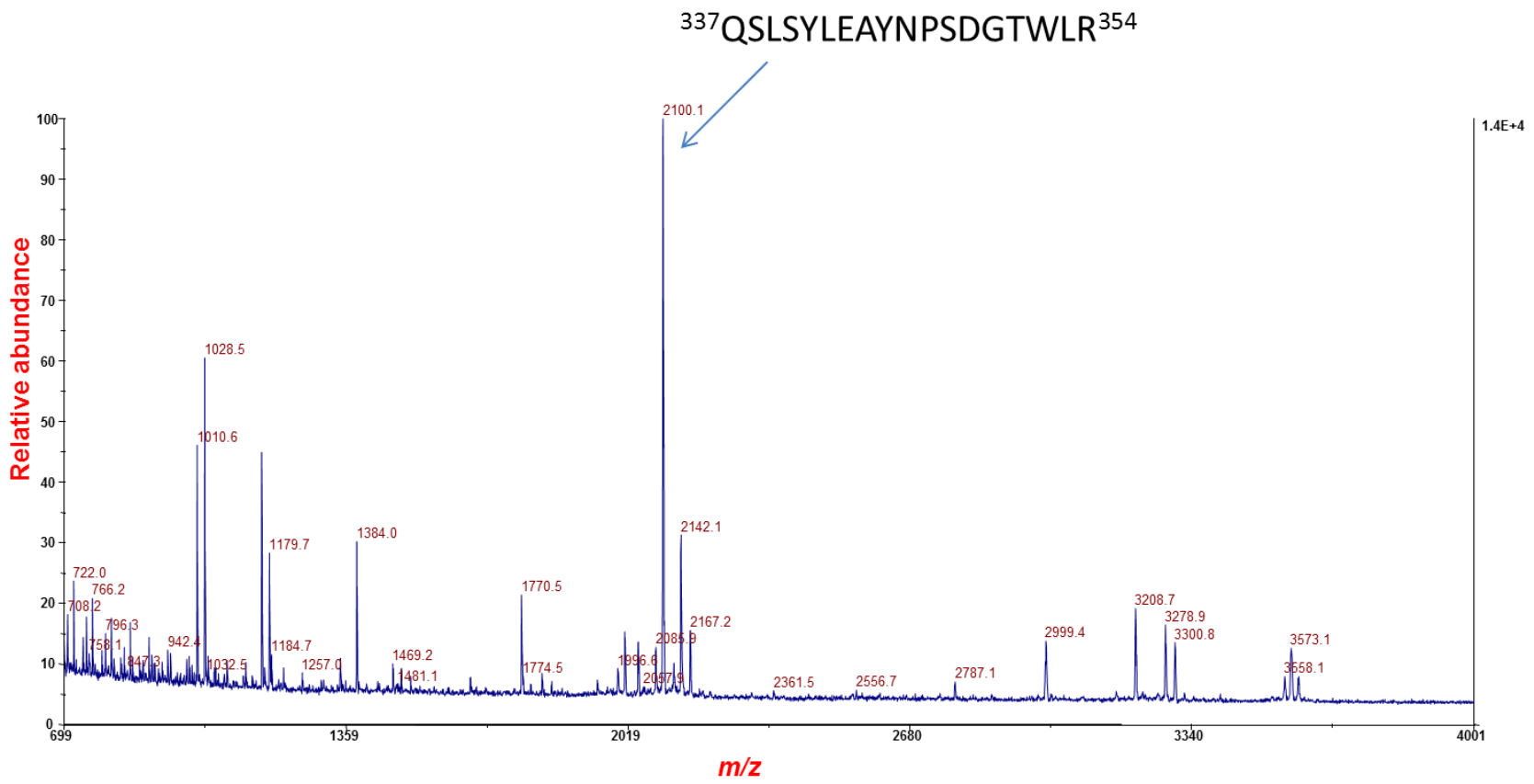


Figure 2.4 Positive ion MALDI-TOF mass spectrum of the ultrafiltrate after mass spectrometric screening of Keap1 tryptic peptides for binding to Neh2.

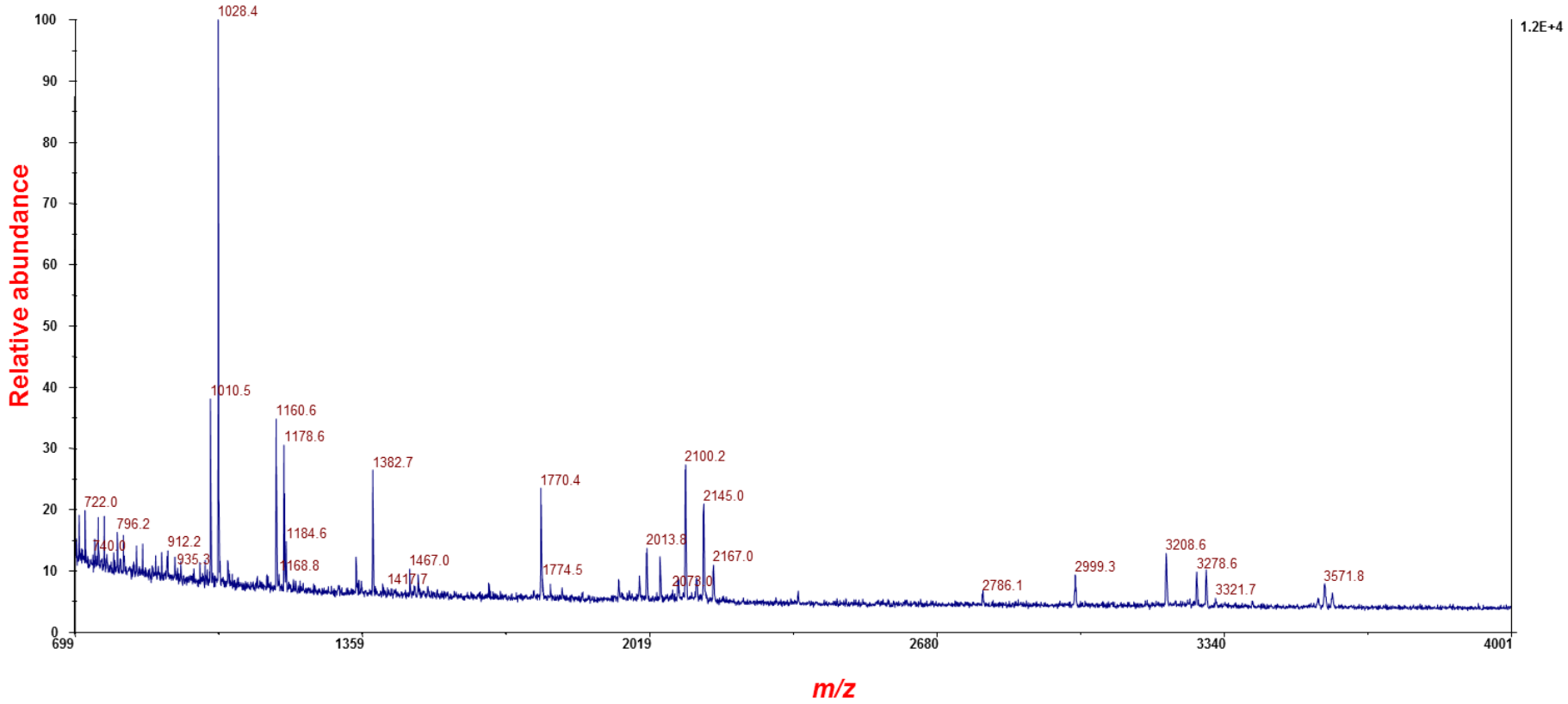


Figure 2.5 Positive ion MALDI-TOF mass spectrum of the control ultrafiltrate obtained after incubation of the Keap1 tryptic digest with denatured Neh2.

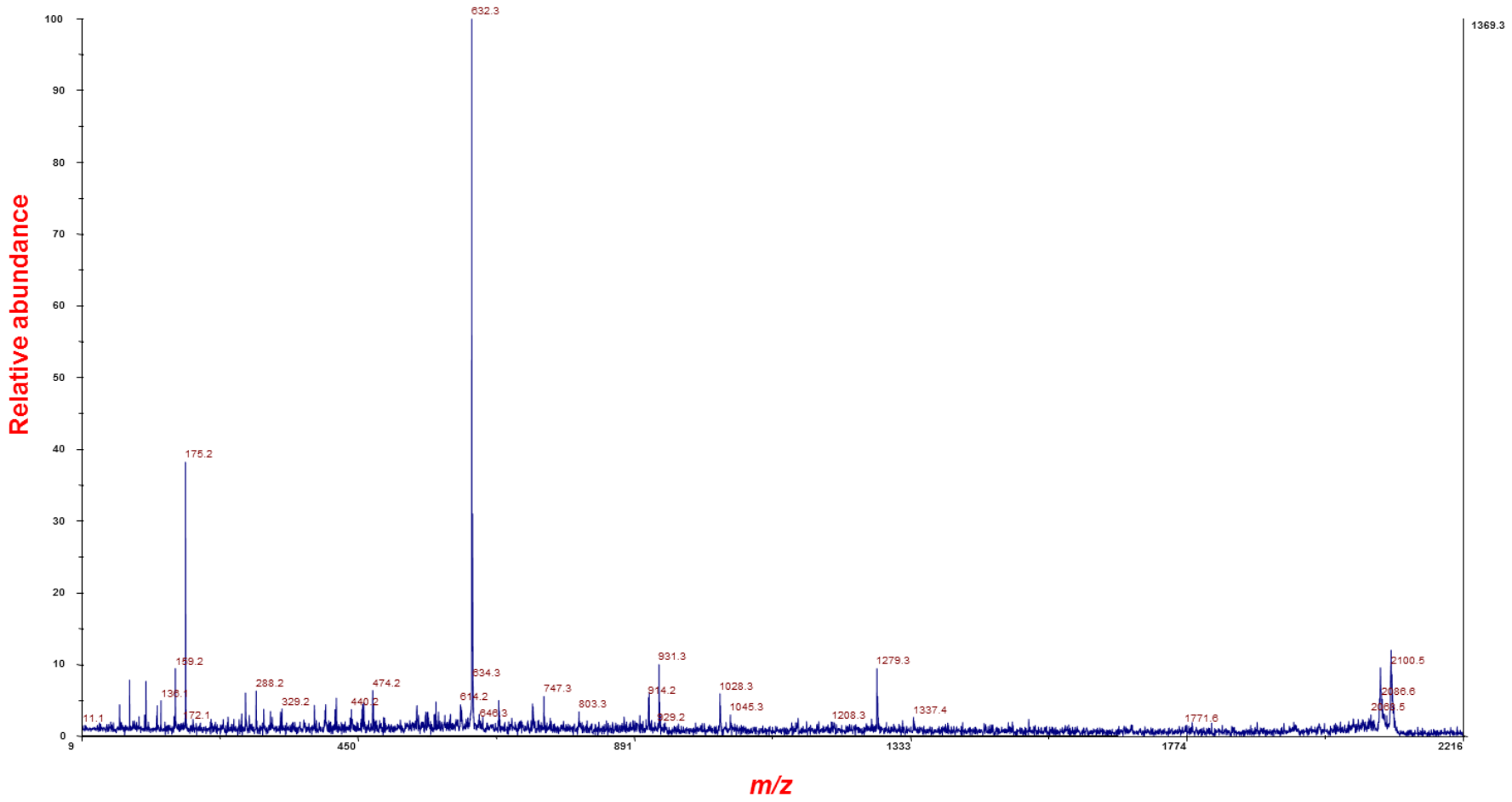


Figure 2.6 Positive ion MALDI-TOF-TOF product ion tandem mass spectrum of the unknown peptide of *m/z* 2100.

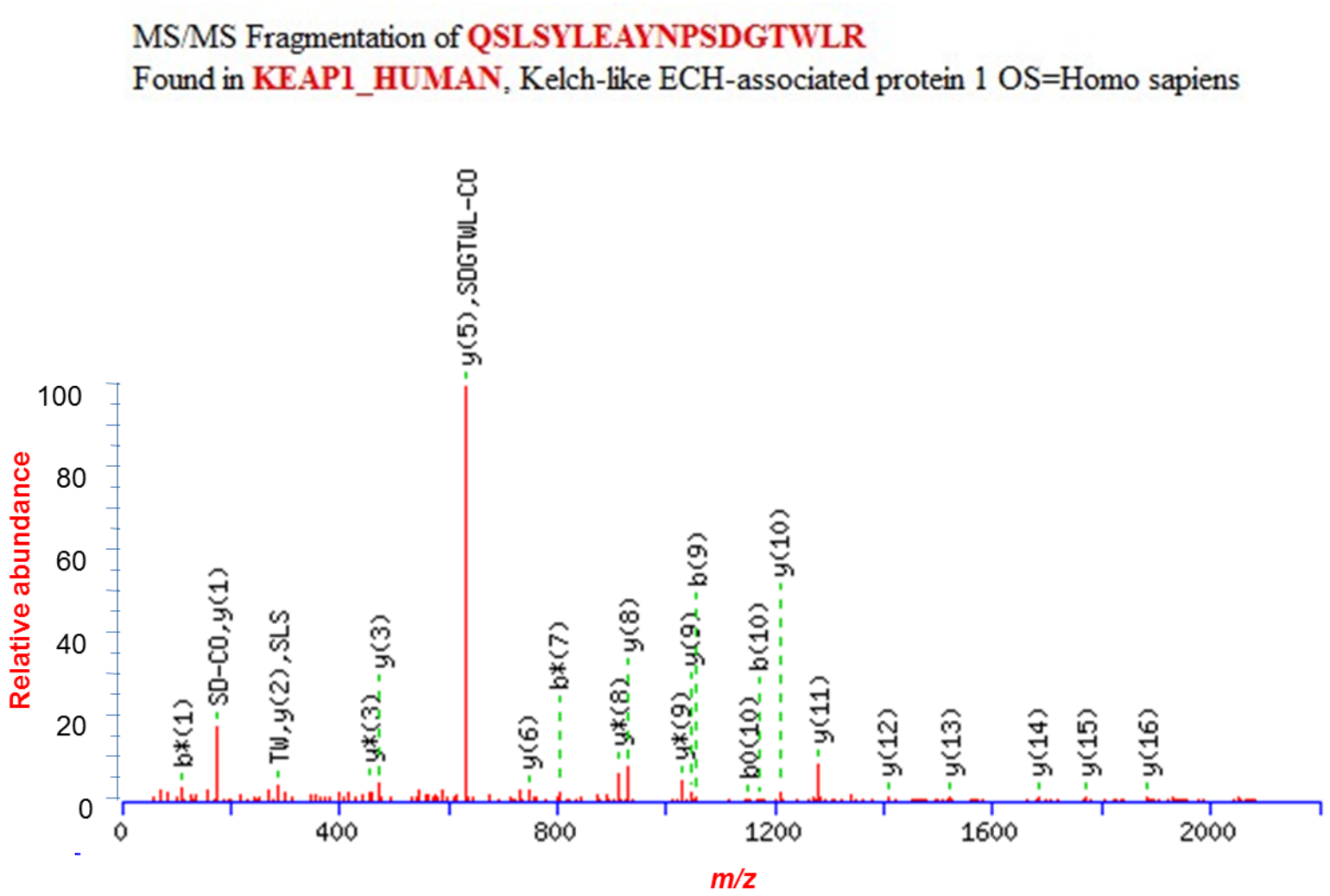


Figure 2.7 Mascot protein database search results and the peptide sequencing annotation for the tandem mass spectrum shown in Figure 2.6.

3. CONCLUSIONS

An ultrafiltration mass spectrometry screening assay was developed for characterizing protein-protein interactions. This new assay was tested by investigating the binding of the chemoprevention target Keap1 to the Neh2 domain of its binding partner Nrf2. This method was optimized to minimize potential interference due to non-specific binding of tryptic peptides to the ultrafiltration membrane. Also, interference from trypsin acting on the protein target during the binding assay was minimized by including a trypsin deactivation step. This step significantly reduced the activity of trypsin remaining in the peptide mixture used in the binding incubation. This ultrafiltration mass spectrometry method should be suitable for use as an alternative to H/D labeling or protein footprinting to study protein-protein interaction.

Mass spectrometry has become a powerful tool in the field of proteomics. This ultrafiltration mass spectrometry assay was developed and employed for the discovery of potential binding peptides in protein-protein interactions. Identification of the Keap1 peptides that bind to Neh2 should help contribute to the understanding of Keap1-Nrf2 binding interactions.

I hypothesis that even after trypsinization, Peptides from Keap1 corresponding to the region of the protein that interacts with Neh2 will still bind to Neh2. Therefore, an ultrafiltration mass spectrometry method was developed to characterize the affinity of human Keap1 tryptic peptides toward the Neh2 domain. This ultrafiltration method was optimized to minimize non-specific binding and trypsin digestion of the protein binding partner of the peptides. In this investigation, the Keap1 tryptic peptide ^{337}QSLSYLEAYNPSDGTWLR^{354} was identified as having the highest affinity for the Neh2 domain of Nrf2. These results support the view that the

human Kelch domain (amino acid 327-amino acid 609) is at least in part responsible for association between Keap1 and Nrf2.

LITERATURE CITED

Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.

Alam, J., P. F. Gong, et al. (2002). "Activation of the mouse heme oxygenase-1 gene by 15-deoxy-Delta(12,14)-prostaglandin J(2) is mediated by the stress response elements and transcription factor Nrf2." Antioxidants & Redox Signaling **4**(2): 249-257.

Ames, B. N. (1983). "Dietary Carcinogens and Anticarcinogens - Oxygen Radicals and Degenerative Diseases." Science **221**(4617): 1256-1264.

Balogun, E., M. Hoque, et al. (2003). "Curcumin activates the haem oxygenase-1 gene via regulation of Nrf2 and the antioxidant-responsive element." Biochemical Journal **371**(Pt 3): 887-895.

Biemann, K. (1990). "Nomenclature for Peptide Fragment Ions (Positive-Ions)." Methods in Enzymology **193**: 886-887.

Chait, B. T. and S. B. H. Kent (1992). "Weighing Naked Proteins - Practical, High-Accuracy Mass Measurement of Peptides and Proteins." Science **257**(5078): 1885-1894.

Colic, M. and K. Pavelic (2002). "Molecular, cellular and medical aspects of the action of nutraceuticals and small molecules therapeutics: From chemoprevention to new drug development." Drugs under Experimental and Clinical Research **28**(5): 169-175.

Demeure, K., V. Gabelica, et al. (2010). "New Advances in the Understanding of the In-Source Decay Fragmentation of Peptides in MALDI-TOF-MS." Journal of the American Society for Mass Spectrometry **21**(11): 1906-1917.

Diehl, J. A., S. B. Cullinan, et al. (2004). "The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: Oxidative stress sensing by a Cul3-Keap1 ligase." Molecular and Cellular Biology **24**(19): 8477-8486.

Dinkova-Kostova, A. T., W. D. Holtzclaw, et al. (2005). "The role of Keap1 in cellular protective responses." Chemical Research in Toxicology **18**(12): 1779-1791.

Doll, R. and R. Peto (1981). "The Causes of Cancer - Quantitative Estimates of Avoidable Risks of Cancer in the United-States Today." Journal of the National Cancer Institute **66**(6): 1191-&.

Dong, Z. G., W. Y. Ma, et al. (1997). "Inhibition of tumor promoter-induced activator protein 1 activation and cell transformation by tea polyphenols, (-)-epigallocatechin gallate, and theaflavins." Cancer Research **57**(19): 4414-4419.

Dong, Z. G., M. Nomura, et al. (2000). "Inhibition of ultraviolet B-induced AP-1 activation by theaflavins from black tea." Molecular Carcinogenesis **28**(3): 148-155.

Englander, S. W., L. Mayne, et al. (1997). "Hydrogen exchange: the modern legacy of Linderstrom-Lang." Protein Science **6**(5): 1101-1109.

Fenselau, C. and P. A. Demirev (2001). "Characterization of intact microorganisms by MALDI mass spectrometry." Mass Spectrometry Reviews **20**(4): 157-171.

Ferguson, P. L. and R. D. Smith (2003). "Proteome analysis by mass spectrometry." Annual Review of Biophysics **32**: 399-424.

Figeys, D., L. D. McBroom, et al. (2001). "Mass spectrometry for the study of protein-protein interactions." Methods **24**(3): 230-239.

Galas, D. J. and A. Schmitz (1978). "DNAse footprinting: a simple method for the detection of protein-DNA binding specificity." Nucleic Acids Research **5**(9): 3157-3170.

Gimon, M. E., L. M. Preston, et al. (1992). "Are Proton-Transfer Reactions of Excited-States Involved in UV Laser Desorption Ionization." Organic Mass Spectrometry **27**(7): 827-830.

Godovac-Zimmermann, J. and L. R. Brown (2001). "Perspectives for mass spectrometry and functional proteomics." Mass Spectrometry Reviews **20**(1): 1-57.

Goldsmith, S. C., J. Q. Guan, et al. (2001). "Synchrotron protein footprinting: a technique to investigate protein-protein interactions." Journal of Biomolecular Structure and Dynamics **19**(3): 405-418.

Guan, J.-Q. and M. R. Chance (2006). Footprinting Methods to Examine the Structure and Dynamics of Proteins. Encyclopedia of Molecular Cell Biology and Molecular Medicine, Wiley-VCH Verlag GmbH & Co. KGaA.

Guan, J.-Q., S. Vorobiev, et al. (2002). "Mapping the G-Actin Binding Surface of Cofilin Using Synchrotron Protein Footprinting†." Biochemistry **41**(18): 5765-5775.

Guan, J. Q. and M. R. Chance (2005). "Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectrometry." Trends in Biochemical Sciences **30**(10): 583-592.

Gukovskaya, A. S., M. Mouria, et al. (2002). "Food-derived polyphenols inhibit pancreatic cancer growth through mitochondrial cytochrome c release and apoptosis." International Journal of Cancer **98**(5): 761-769.

Hannink, M. and D. D. Zhang (2003). "Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress." Molecular and Cellular Biology **23**(22): 8137-8151.

Heidelbe (1970). "Chemical Carcinogenesis, Chemotherapy - Cancers Continuing Core Challenges - Clowes, Gha Memorial Lecture." Cancer Research **30**(6): 1549-&.

Henzel, W. J., T. M. Billeci, et al. (1993). "Identifying Proteins from 2-Dimensional Gels by Molecular Mass Searching of Peptide-Fragments in Protein-Sequence Databases." Proceedings of the National Academy of Sciences of the United States of America **90**(11): 5011-5015.

Heyduk, E. and T. Heyduk (1994). "Mapping Protein Domains Involved in Macromolecular Interactions: A Novel Protein Footprinting Approach." Biochemistry **33**(32): 9643-9650.

Heyduk, E., T. Heyduk, et al. (1997). "Conformational Changes of DNA Induced by Binding ofChironomus High Mobility Group Protein 1a (cHMG1a)." Journal of Biological Chemistry **272**(32): 19763-19770.

Hofstadler, S. A., J. H. Wahl, et al. (1994). "Capillary Electrophoresis Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry with Sustained Off-Resonance Irradiation for the Characterization of Protein and Peptide Mixtures." Journal of the American Society for Mass Spectrometry **5**(10): 894-899.

Hunt, D. F., J. E. P. Syka, et al. (2004). "Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry." Proceedings of the National Academy of Sciences of the United States of America **101**(26): 9528-9533.

Hvidt, A. and K. Linderstrom-Lang (1954). "Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions." Biochimica et Biophysica Acta **14**(4): 574-575.

Hvidt, A. and S. O. Nielsen (1966). "Hydrogen exchange in proteins." Advances in Protein Chemistry **21**: 287-386.

Ijames, C. F. and C. L. Wilkins (1990). "Surface-induced dissociation by Fourier transform mass spectrometry." Analytical Chemistry **62**(13): 1295-1299.

Itoh, K., M. Mochizuki, et al. (2004). "Transcription factor Nrf2 regulates inflammation by mediating the effect of 15-deoxy-D12,14-prostagiandin J2." Faseb Journal **18**(8): C94-C94.

Itoh, K., N. Wakabayashi, et al. (1999). "Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain." Genes & Development **13**(1): 76-86.

Janle, E. M. and J. E. Sojka (2000). "Use of ultrafiltration probes in sheep to collect interstitial fluid for measurement of calcium and magnesium." Contemporary Topics in Laboratory Animal Science **39**(6): 47-50.

Jemal, A., T. Murray, et al. (2005). "Cancer statistics, 2005." Ca-a Cancer Journal for Clinicians **55**(1): 10-30.

Jensen, O. N., A. Stensballe, et al. (2000). "Electron capture dissociation of singly and multiply phosphorylated peptides." Rapid Communications in Mass Spectrometry **14**(19): 1793-1800.

Kafka, A. P., T. Kleffmann, et al. (2011). "The application of MALDI TOF MS in biopharmaceutical research." International Journal of Pharmaceutics **417**(1-2): 70-82.

Karas, M. and F. Hillenkamp (1988). "Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons." Analytical Chemistry **60**(20): 2299-2301.

Kicman, A. T., M. C. Parkin, et al. (2007). "An introduction to mass spectrometry based proteomics - Detection and characterization of gonadotropins and related molecules." Molecular and Cellular Endocrinology **260**: 212-227.

Kobayashi, M. and M. Yamamoto (2005). "Molecular mechanisms activating the Nrf2-Keap1 pathway of antioxidant gene regulation." Antioxidants & Redox Signaling **7**(3-4): 385-394.

Kong, A. N. and W. G. Li (2009). "Molecular Mechanisms of Nrf2-Mediated Antioxidant Response." Molecular Carcinogenesis **48**(2): 91-104.

Lanman, J. and P. E. Prevelige, Jr. (2004). "High-sensitivity mass spectrometry for imaging subunit interactions: hydrogen/deuterium exchange." Current Opinion in Structural Biology **14**(2): 181-188.

Leegsma-Vogt, G., E. Janle, et al. (2003). "Utilization of *in vivo* ultrafiltration in biomedical research and clinical applications." Life Sciences **73**(16): 2005-2018.

Linhares, M. C. and P. T. Kissinger (1993). "Determination of Endogenous Ions in Intercellular Fluid Using Capillary Ultrafiltration and Microdialysis Probes." Journal of Pharmaceutical and Biomedical Analysis **11**(11-12): 1121-1127.

Linhares, M. C. and P. T. Kissinger (1993). "Pharmacokinetic Monitoring in Subcutaneous Tissue Using Invivo Capillary Ultrafiltration Probes." Pharmaceutical Research **10**(4): 598-602.

Liu, D., J. Guo, et al. (2007). "Screening for Ligands of Human Retinoid X Receptor-α Using Ultrafiltration Mass Spectrometry." Analytical Chemistry **79**(24): 9398-9402.

Maleknia, S. D., C. Y. Ralston, et al. (2001). "Determination of macromolecular folding and structure by synchrotron x-ray radiolysis techniques." Analytical Biochemistry **289**(2): 103-115.

Mária L. Simon, K. L., Márta Kotormán, Béla Szaján̈i (2001). "A comparative study of the conformational stabilities of trypsin and a-chymotrypsin." Acta Biologica Szegediensis **45**(1-4): 43-49.

McLafferty, F. W., R. A. Zubarev, et al. (2000). "Electron capture dissociation for structural characterization of multiply charged protein cations." Analytical Chemistry **72**(3): 563-573.

Medzihradszky, K. F. (2005). "Peptide sequence analysis." Biological Mass Spectrometry **402**: 209-244.

Medzihradszky, K. F., J. M. Campbell, et al. (2000). "The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer." Analytical Chemistry **72**(3): 552-558.

Mehta, R., G. Murillo, et al. (2010). "Cancer Chemoprevention by Natural Products: How Far Have We Come?" Pharmaceutical Research **27**(6): 950-961.

Mesecar, A. D., A. L. Eggler, et al. (2005). "Modifying specific cysteines of the electrophile-sensing human Keap1 disrupt binding to the protein is insufficient to Nrf2 domain Neh2." Proceedings of the National Academy of Sciences of the United States of America **102**(29): 10070-10075.

Mirsaleh-Kohan, N., W. D. Robertson, et al. (2008). "Electron ionization time-of-flight mass spectrometry: historical review and current applications." Mass Spectrometry Reviews **27**(3): 237-285.

Moran, M. F., D. Figeys, et al. (2001). "Mass spectrometry for the study of protein-protein interactions." Methods **24**(3): 230-239.

Motohashi, H. and M. Yamamoto (2004). "Nrf2–Keap1 defines a physiologically important stress response mechanism." Trends in Molecular Medicine **10**(11): 549-557.

Niture, S. K., A. K. Jain, et al. (2010). "Antioxidant-induced modification of INrf2 cysteine 151 and PKC-delta-mediated phosphorylation of Nrf2 serine 40 are both required for stabilization and nuclear translocation of Nrf2 and increased drug resistance (vol 122, pg 4452, 2009)." Journal of Cell Science **123**(11): 1969-1969.

O'Farrell, P. Z., H. M. Goodman, et al. (1977). "High resolution two-dimensional electrophoresis of basic as well as acidic proteins." Cell **12**(4): 1133-1141.

Osada, S., K. Ohta, et al. (2007). "Histone acetyltransferase MOZ acts as a co-activator of Nrf2-MafK and induces tumour marker gene expression during hepatocarcinogenesis." Biochemical Journal **402**: 559-566.

Pickett, C. B., P. Nioi, et al. (2005). "The carboxy-terminal Neh3 domain of Nrf2 is required for transcriptional activation." Molecular and Cellular Biology **25**(24): 10895-10906.

Plummer, S. M., K. A. Holloway, et al. (1999). "Inhibition of cyclo-oxygenase 2 expression in colon cells by the chemopreventive agent curcumin involves inhibition of NF-kappa B activation via the NIK/IKK signalling complex." Oncogene **18**(44): 6013-6020.

Sarkar, F. H., J. N. Davis, et al. (1999). "Genistein inhibits NF-kappa B activation in prostate cancer cells." Nutrition and Cancer-an International Journal **35**(2): 167-174.

Schmitz, A. and D. J. Galas (1979). "The interaction of RNA polymerase and lac repressor with the lac control region." Nucleic Acids Research **6**(1): 111-137.

Schneiderheinze, J. M. and B. L. Hogan (1996). "Selective *in vivo* and *in vitro* sampling of proteins using miniature ultrafiltration sampling probes." Analytical Chemistry **68**(21): 3758-3762.

Shea, M. A., B. R. Sorensen, et al. (2000). Proteolytic footprinting titrations for estimating ligand-binding constants and detecting pathways of conformational switching of calmodulin. Methods in Enzymology. G. K. A. Michael L. Johnson, Academic Press. **Volume 323:** 254-301.

Sims, P., P. L. Grover, et al. (1974). "Metabolic Activation of Benzo(a)Pyrene Proceeds by a Diol-Epoxide." Nature **252**(5481): 326-328.

Smith, R. D. and B. Bogdanov (2005). "Proteomics by FTICR mass spectrometry: Top down and bottom up." Mass Spectrometry Reviews **24**(2): 168-200.

Smith, R. D., J. S. Page, et al. (2004). "FTICR mass spectrometry for qualitative and quantitative bioanalyses." Current Opinion in Biotechnology **15**(1): 3-11.

Sporn, M. B. (1976). "Approaches to prevention of epithelial cancer during the preneoplastic period." Cancer Research **36**(7 PT 2): 2699-2702.

Steen, H. and M. Mann (2004). "The ABC's (and XYZ's) of peptide sequencing." Nature Reviews Molecular Cell Biology **5**(9): 699-711.

Surh, Y. J. (2002). "Anti-tumor promoting potential of selected spice ingredients with antioxidative and anti-inflammatory activities: a short review." Food and Chemical Toxicology **40**(8): 1091-1097.

Surh, Y. J. (2003). "Cancer chemoprevention with dietary phytochemicals." Nature Reviews Cancer **3**(10): 768-780.

Takamoto, K. and M. R. Chance (2006). Footprinting Methods to Examine the Structure and Dynamics of Nucleic Acids. Encyclopedia of Molecular Cell Biology and Molecular Medicine, Wiley-VCH Verlag GmbH & Co. KGaA.

Talalay, P., A. T. Dinkova-Kostova, et al. (2002). "Direct evidence that sulfhydryl groups of Keap1 are the sensors regulating induction of phase 2 enzymes that protect against carcinogens and oxidants." Proceedings of the National Academy of Sciences of the United States of America **99**(18): 11908-11913.

Talalay, P., J. W. Fahey, et al. (1995). "Chemoprotection against cancer by Phase 2 enzyme induction." Toxicology Letters **82-3**: 173-179.

Talbo, G. H., D. Suckau, et al. (2001). "MALDI-PSD-MS analysis of the phosphorylation sites of caseinomacropeptide." Peptides **22**(7): 1093-1098.

Tanaka, K., H. Waki, et al. (1988). "Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry." Rapid Communications in Mass Spectrometry **2**(8): 151-153.

Tanaka, T., K. I. Tong, et al. (2006). "Keap1 recruits Neh2 through binding to ETGE and DLG motifs: Characterization of the two-site molecular recognition model." Molecular and Cellular Biology **26**(8): 2887-2900.

van Breemen, R. B., C. G. Gu, et al. (1999). "Assays of ligand-human serum albumin binding using pulsed ultrafiltration and liquid chromatography-mass spectrometry." Combinatorial Chemistry & High Throughput Screening **2**(6): 353-359.

van Breemen, R. B., C. R. Huang, et al. (1997). "Pulsed ultrafiltration mass spectrometry: A new method for screening combinatorial libraries." Analytical Chemistry **69**(11): 2159-2164.

van Breemen, R. B. and D. Nikolic (1998). "Screening for inhibitors of dihydrofolate reductase using pulsed ultrafiltration mass spectrometry." Combinatorial Chemistry & High Throughput Screening **1**(1): 47-55.

van Breemen, R. B., D. Nikolic, et al. (2000). "Evaluation of cyclooxygenase-2 inhibitors using pulsed ultrafiltration mass spectrometry." Analytical Chemistry **72**(16): 3853-3859.

van Breemen, R. B., Y. K. Sun, et al. (2005). "Ultrafiltration tandem mass spectrometry of estrogens for characterization of structure and affinity for human estrogen receptors." Journal of the American Society for Mass Spectrometry **16**(2): 271-279.

Volmer, D. A. and L. Sleno (2004). "Ion activation methods for tandem mass spectrometry." Journal of Mass Spectrometry **39**(10): 1091-1112.

Wattenberg, L. W. (1990). "Inhibition of Carcinogenesis by Minor Anutrient Constituents of the Diet." Proceedings of the Nutrition Society **49**(2): 173-183.

Wells, J. M. and S. A. McLuckey (2005). "Collision-induced dissociation (CID) of peptides and proteins." Biological Mass Spectrometry **402**: 148-185.

WHO (2007). "WHO Cancer Brochure 2007."

Wilm, M., A. Shevchenko, et al. (1996). "Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry." Nature **379**(6564): 466-469.

Yamamoto, M., K. Itoh, et al. (1999). "Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain." Genes & Development **13**(1): 76-86.

Yamamoto, M., K. Itoh, et al. (2003). "Keap1 regulates both cytoplasmic-nuclear shuttling and degradation of Nrf2 in response to electrophiles." Genes to Cells **8**(4): 379-391.

Yamamoto, M., A. Kobayashi, et al. (2004). "Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate for proteasomal degradation of Nrf2." Molecular and Cellular Biology **24**(16): 7130-7139.

Yang, C. S., J. Y. Chung, et al. (1999). "Inhibition of activator protein 1 activity and cell growth by purified green tea and black tea polyphenols in H-ras-transformed cells: Structure-activity relationship and mechanisms involved." Cancer Research **59**(18): 4610-4617.

Yates, J. R. (1998). "Database searching using mass spectrometry data." Electrophoresis **19**(6): 893-900.

Zhang, C. J., H. X. Zhang, et al. (2010). "CHCA or DHB? Systematic Comparison of the Two Most Commonly Used Matrices for Peptide Mass Fingerprint Analysis with MALDI-MS." Spectroscopy **25**(2): 48.

Zhang, D. D. (2006). "Mechanistic studies of the Nrf2-Keap1 signaling pathway." Drug Metabolism Reviews **38**(4): 769-789.

Zhang, J., T. Hosoya, et al. (2007). "Nrf2 Neh5 domain is differentially utilized in the transactivation of cytoprotective genes." Biochemical Journal **404**(3): 459-466.

Zhang, W., A. N. Krutchinsky, et al. (2003). ""De novo" peptide sequencing by MALDI-quadrupole-ion trap mass spectrometry: a preliminary study." Journal of the American Society for Mass Spectrometry **14**(9): 1012-1021.

Zhao, Y. Z., R. B. vanBreemen, et al. (1997). "Screening solution-phase combinatorial libraries using pulsed ultrafiltration electrospray mass spectrometry." Journal of Medicinal Chemistry **40**(25): 4006-4012.

Zhong, M., L. Lin, et al. (1995). "A method for probing the topography and interactions of proteins: footprinting of myoglobin." Proceedings of the National Academy of Sciences **92**(6): 2111-2115.

Zhu, H., M. Bilgin, et al. (2003). "Proteomics." Annual Review of Biochemistry **72**: 783-812.

Zipper, L. M. and R. T. Mulcahy (2002). "The Keap1 BTB/POZ dimerization function is required to sequester Nrf2 in cytoplasm." Journal of Biological Chemistry **277**(39): 36544-36552.

Zubarev, R. A., M. Witt, et al. (2005). "Two-fold efficiency increase by selective excitation of ions for consecutive activation by ion-electron reactions and vibrational excitation in tandem Fourier transform ion cyclotron resonance mass spectrometry." Analytical Chemistry **77**(9): 2992-2996.

VITA

NAME:

ZHIYUAN SUN

EDUCATION

M.S., Medicinal Chemistry, University of Illinois at Chicago, Chicago, Illinois, 2011

M.S., Entomology, University of Kentucky, Lexington, Kentucky, 2009

B.S., Biotechnology, China Pharmaceutical University, Nanjing, China, 2006

RESEARCH EXPERIENCE

Mass spectrometric studies of Keap1-Nrf2 binding interactions 2009-2011

WORK EXPERIENCE

Summer Internship, Analytical Biochemistry group, Biogen Idec, Cambridge, MA, 2011

TEACHING EXPERIENCE

Teaching Assistant, College of Pharmacy, University of Illinois at Chicago, Chicago, Illinois, Principles of Drug Action and Therapeutics IV, 2009

Teaching Assistant, College of Pharmacy, University of Illinois at Chicago, Chicago, Illinois, Principles of Drug Action and Therapeutics V, 2010

PROFESSIONAL MEMBERSHIPS

American Society of Mass Spectrometry

PUBLICATIONS

R.Parthasarathy, Zhentao Sheng, **Zhiyuan Sun**, Subba R. Palli., 2010 Ecdysteroid regulation of ovarian growth and oocyte maturation in the red flour beetle, *Tribolium castaneum* Insect Biochem. Mol. Biol. 40 (2010) 429-439

R. Parthasarathy, **Zhiyuan Sun**, Hua Bai, Subba R. Palli., 2010. Juvenile hormone regulation of vitellogenin synthesis in the red flour beetle, *Tribolium castaneum*. Insect Biochem. Mol. Biol. 40 (2010) 405-414.

R. Parthasarathy, A.Tan, **Z. Sun**, Z. Chen, M. Rankin, S.R. Palli., 2009 Juvenile hormone regulation of male accessory gland activity in the red flour beetle, *Tribolium castaneum*. Mech. Dev. 126 (2009) 563 – 579