NOVEL METHODS FOR IDENTIFICATION OF POSTTRANSLATIONAL LIPID MODIFIED PROTEINS BY MASS SPECTROMETRY

by

RUCHIKA BHAWAL

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2015

Acknowledgements

This work would not have been possible without the help and support of many people. First, I would like to express my gratitude to my mentor Dr. Saiful Chowdhury for his guidance, patience and providing me a good atmosphere for doing my research in the lab and think independently for any challenges during my graduate research. This has helped me to stay focused and motivated in research. I would also like to thank for helping me advancing my understanding of proteomics and the related biological techniques through his lab meetings and one to one meetings every week.

I would express my sincere thanks to my committee member Dr. Purnendu Dasgupta for his invaluable support, encouragement and his trust on me. His encouragement motivated me to work harder and accomplish the best during my hard times in graduate studies. I always find him very creative, understanding and a true scientist in his thoughts. I would also like to thank my committee member Dr. Peter Kroll for helping and advising me whenever we meet during the research meeting. He is one of the best graduate advisor for us. I would also like to thank Dr. Alejandro Bugarin for his help in synthesizing some important compounds for my research.

I would like to acknowledge the Shimadzu Center for Advanced Analytical Chemistry (SCAAC) for the training to use various mass spectrometers and also always allowed us to use the mass spectrometers as and when needed. I like to express my thanks to Maciej Kukula and Liangqiao Bian for their help throughout. Thanks to my best friend Maheshika Wanigasekara for her trust and good wishes throughout. Thanks to all my fellow lab members Shahin, Jayanta, Kathy, Jonathan, Hamilton, Daria and many others for being helpful and supportive. I highly appreciate the assistance and support from all the staff of Chemistry and Biochemistry department and would like to thank Debbie, Jill, Jim, Jason and Brian.

My heartfelt gratitude to my parents and two sisters for their endless love, support and encouragement given to me through all those hard times. They taught me the value of hard work and gave me everything I needed to go through this journey. I wish to thank my husband's family for their support as well.

A big thanks to my friend and my husband Sumit Bhawal for being by me at all times, guiding and supporting me throughout my research career. Finally, I want to thank my sweetest daughter Alice for being so good and patient and always making me happy and letting me forget all the hard times and giving each new day for a fresh start.

November 23, 2015.

Abstract

NOVEL METHODS FOR IDENTIFICATION OF POSTTRANSLATIONAL LIPID MODIFIED PROTEINS BY MASS SPECTROMETRY

Ruchika Bhawal, PhD

The University of Texas at Arlington, 2015

Supervising Professor: Saiful M Chowdhury

Study of posttranslational modifications (PTMs) is very important for developing potential new drugs for the treatment of several diseases. Posttranslational lipid modifications are major PTMs which help proteins to anchor to cell membranes and facilitate protein-protein interactions. There are several kinds of lipid modifications of proteins. The major lipid modifications are palmitoylation, myristoylation and prenylation. Prenylation are of two types - farnesylation and geranylgeranylation. These modifications are involved with several human cancers and genetic diseases. Hutchinson-Gilford progeria syndrome (HGPS), a disease which causes premature aging for children is the product of prenylation. Understanding the mechanism of prenylation and their substrates is extremely needed to find the new drug target and develop new therapeutics for cancer and genetic diseases.

Current identification methods for prenyl protein are laborious and time consuming and in some cases tailored to specific proteins. Most of the studies on protein prenylation use radioactive labeling, chemical reporters, LC-MS analysis and tagging via substrate method. None of the method can identify prenyl proteins globally, their sites and types in a single experiment. A global or large-scale method is needed to identify all the proteins which go through prenylation during cellular events.

My dissertation work focused on the large scale identification of prenylated peptides/proteins by novel liquid-chromatography and mass spectrometric methods. We have developed mass spectrometry cleavable approaches to differentiate and locate the prenylation types. Prenyl proteins are hydrophobic and very difficult to analyze by LC-MS/MS. They have inconsistent gas phase fragmentation behavior during tandem mass spectrometry. We have studied in detail the CID and ETD mass spectrometry behavior of prenyl peptides. We have also converted prenyl peptides to more hydrophobic by oxidation chemistry. The fragmentation behavior of the synthesized prenylated peptides as well as the modified ones were studied in both collision induced dissociation (CID) and electron transfer dissociation (ETD) methods. On fragmentation, the loss of RSOH group (R = farnesyl/geranylgeranyl) from oxyprenylated peptides was found that result in the formation of a signature peak in the mass spectra. The mass loss distinguishes the type of prenylation. Oxidation chemistry was also incorporated epoxy groups in the prenyl side chain and increased the hydrophilicity and made them enrichable by chemical labeling method. In this dissertation, we developed a novel method to distinguish and locate a major lipid modification and we believe the study will significantly contribute in advancing cancer treatment and diagnosis.

Table of Contents

List of Illustrations

Chapter 1

Introduction to posttranslational modifications

1.1 Posttranslational lipid modification

Posttranslational modifications (PTMs) are mostly the chemical modifications that arise after the protein biosynthesis step, i.e., translation. They also arises due to proteolytic cleavage of protein N-terminal [1]. Many intra and extracellular events that causes some specific chemical change in biological system, which finally results in the chemical modifications of proteins [2]. PTMs involved in mediating their interaction with the target protein, add additional functional properties and diversities in the proteome. They are the key to a number of important biological processes [3]. Due to the chemical reactivity of certain amino acid residues, the chemical incorporation of various PTMs in the protein is possible. There are several types of posttranslational modifications found in biological system such as acetylation, formylation, phosphorylation, glycosylation, ubiquitination, amidation, and some lipid modifications such as prenylation, palmitoylation and Myristoylation [4]. Acetylation is the addition of acetyl group to the side chain amino group of lysine residues similarly formylation is the addition of formyl group to the N-terminal of the protein. Amidation is the attachment of amine group and its form amide in the c-terminal of protein. Phosphorylation generally add a phosphate group to the hydroxyl side chain of serine, threonine and tyrosine residues and glycosylation is the addition of a sugar moiety. Ubiquitination is the attachment of a small protein named ubiquitin to the lysine side chain and these proteins after ubiquitination undergoes degradation. Many of the PTMs acts as switches and (de)activate the signaling pathway by changing the protein-protein interactions [5]. Posttranslational lipid modifications make the target protein hydrophobic and they get translocated to the membrane and interaction

with the membrane protein increases. Since PTMS are not encoded by the genome directly, hence it is not accessible through the traditional genetic tools and it increases the complexity to the proteome.

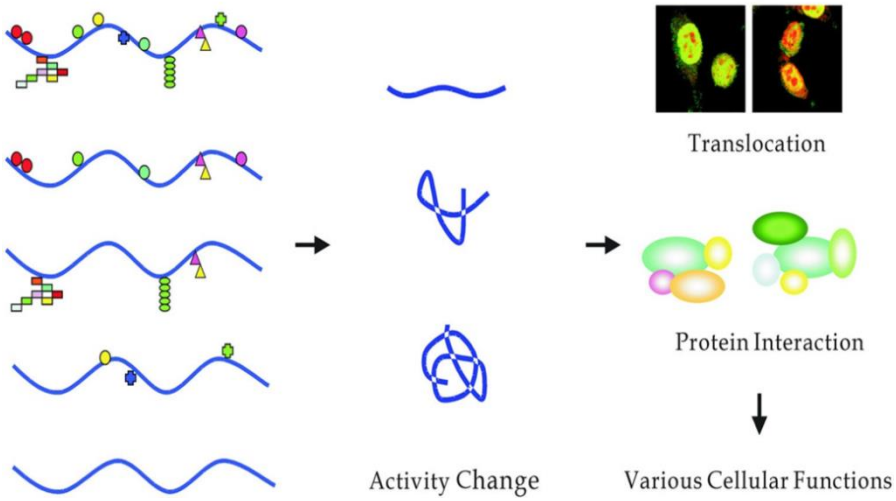


Figure 1-1 Posttranslational modifications (PTMs) of proteins

(The figure is adapted from the previous publication [6]).

Protein lipid-modifications including myristoylation, palmitoylation, and prenylation causes changes in protein-membrane interactions, protein interaction and enzyme activity [7]. It was found that they play a significant role in various types of cancers and tumors. In biological systems, N-myristoylation and S-palmitoylation comprise two major classes of protein lipid modifications that control enzyme activities of proteins [8]. N-myristoylation is a covalent modification of myristic acid to the N-terminal of the Gly residue of the protein through the enzyme N-myristoylase. S-palmitoylation is the incorporation of palmitic acid in cysteine residues of proteins catalyzed by protein S-acyltransferases. S-palmitoylated is uniquely reversible among fatty- acyl modifications of proteins and is involved in the translocation of the proteins to membranes. Prenylation is

the addition of isoprenoid, either farnesyl (15 carbon atoms) or geranylgeranyl (20 carbon atoms), to the cysteine residue of the carboxyl terminal of the protein through a thioether linkage. These modifications are enzymatically catalyzed by farnesyltransferase (FTase) or geranylgeranyl transferase (GGTase-I/II)[9]. Due to the involvement of farnesylation in oncogenesis various farnesyl inhibitors were synthesized for chemotherapy [10]. The substrates for prenylation, such as farnesyl/geranylgeranyl pyrophosphate are two intermediate products of cholesterol biosynthetic pathways. Unfortunately in cancer, to prevent tumor progression, inhibiting the activity of farnesyl transferase (FT) in well-known oncogene Ras proteins by farnesyl transferase inhibitors (FTI) has failed in clinical trials due to alternative geranylgeranyl transferase mediated (GGT) prenylation. Recent failed clinical outcomes of FTI emphasized the importance of the identification of the target of new prenyl substrates and their type of prenylation in the target proteins to know the roles of these modifications in cancers[11, 12].

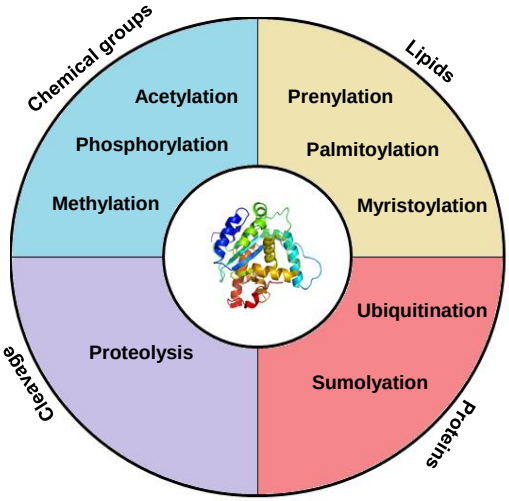
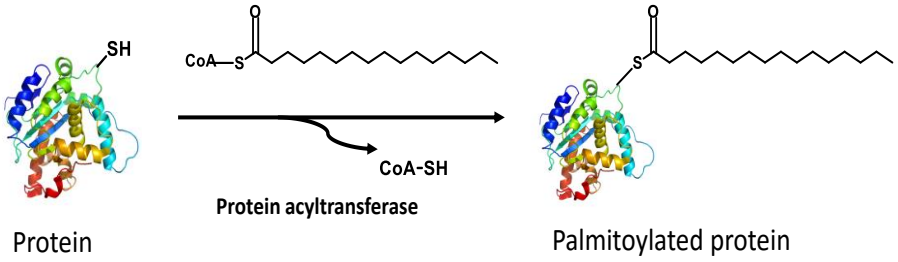


Figure 1-2 several types of posttranslational modifications (PTMs) of proteins. The figure is modified from the previous publication [13].

1.1.1 Palmitoylation

S-Palmitoylation is a one of the major lipid post-translational modification where the addition of a C16-carbon saturated fatty acyl chain takes place to cysteine amino acid residue of the proteins [31]. This modification is carried out by the enzyme palmitoyl-acyl Transferases and the opposite reaction are done by protein palmitoyl thioesterases. It occurs on the cytoplasmic face of membrane and it modifies both the soluble and integral membrane proteins. Palmitoyl acyl transferases catalyze the transfer of palmitoyl group to the cysteine residue of the protein via the formation of thio ester linkage. They are generally termed as S-palmitoylation [32]. Sometimes, the palmitoylation also occur via the formation of amide bond and they are named as N- palmitoylation. This is mostly happen when the migration of palmitoyl group takes place from N-terminal cysteine



residue to the free N-terminal of the protein.

Figure 1-4 S-palmitoylation of protein in biological system

Palmitoylation acts as a reversible lipid anchor bind the protein to membrane and the saturated hydrocarbon chain makes it more hydrophobic. Palmitoylation and depalmitoylation of proteins resulted in changes in modified protein locations and its interactions with other cellular components.[33]. Palmitoylation of the neuronal proteins is important for regulating proper development and function of the neurotransmitter.

Palmitoylation causes several diseases such as osteoporosis, neurodegenerative diseases like Alzheimer's disease, Huntington disease, and a variety of cancers [34, 35]. Not only that, it serve as a biomarkers for cardiovascular disease. There are several methods available for the global identification of palmitoylation sites within a complex biological system such as acyl-biotin exchange, biorthogonal palmitoylation reporters and mass spectrometric proteomic strategies [36].

1.1.2 Myristoylation

N-myristoylation of protein is the irreversible attachment of a saturated 14-carbon fatty acid chain (myristoyl group) to an N-terminal glycine residue of the protein via an amide linkage. The reaction is catalyzed by N-myristoyl transferase, which recognizes a MGXXXS/T signature sequence at the N-terminal of the protein [37-39]. In many cases, it was found that the N-terminal methionine residue was cleaved by an enzyme followed by myristoylation of the glycine takes place.

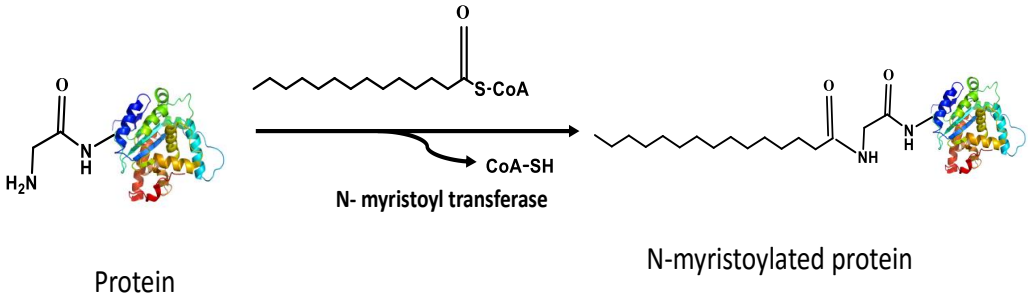


Figure 1-5 N-myristoylation of protein in biological system

One of the major functions of myristoylated group is assist in directing the protein towards the membrane. Mostly, for the membrane association, myristoylation of the N-terminal is the first signal followed by palmitoylation of the cysteine residues. Not all the myristoylated proteins are membrane bound, some of the kinases are found in cytosol. It

plays an important role in regulating intrinsic thermal stability, enzymic activity and protein stability. It was found that intramolecular myristoyl switch allows protein to undergo reversible membrane binding. Elevated levels of myristoylated proteins are found in several types of cancers and the enzyme[40], N-myristoyl transferase is a potential target for various infectious and neurological diseases.

1.1.3 Prenylation

Prenylation of proteins is the covalent addition of either 15 carbon (farnesyl) or 20 carbon (geranylgeranyl) groups in the c-terminal cysteine residues of protein having few specific motif. The most important aspects for protein farnesylation is the presence of CAAX (where C is the cysteine, A is any aliphatic and X is usually serine, threonine, glutamine, alanine or methionine) box [14, 15]. The most common farnesylated proteins are the members of Ras family [16] and nuclear lamins, such as type A and B. Farnesylation plays a significant role in several diseases like cancers, tumors, parasitic and aging [17-19]. Geranylgeranylation recognizes another type of motifs such as CXC and CC where both the cysteine residues undergo modifications. These modifications are generally found in heterotrimeric G proteins and in Ras, Rho and Rab [20]. Some of the most studied Ras and Rab proteins are critical for cell signaling during cell growth and differentiation. The involvement of mutated and over expressed Ras genes has been identified in 90% of pancreatic cancers, 50% of thyroid and colon cancers and 30% of lung cancers and myeloid leukemia diagnosis [21, 22]. Mutated and farnesylated Ras involved in several signal transduction pathways [23].

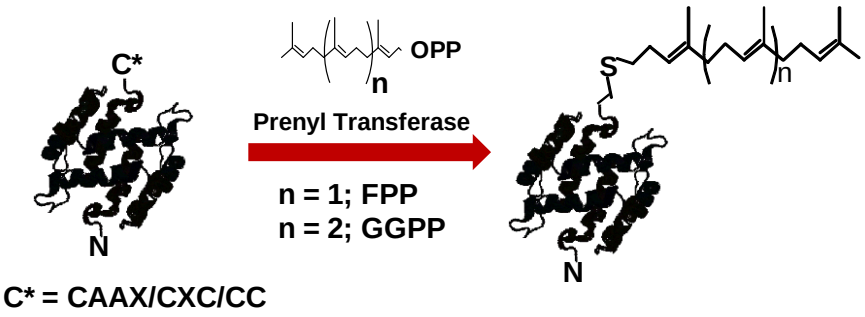


Figure 1-3 Scheme of prenylation of protein in biological system.

Prenylation involvement in cancer causing activities of proteins has led to the development of various anticancer drugs targeting enzyme prenyl transferases. Farnesyl transferase inhibitors showed high efficacy in killing the cancerous cells in animal model, but clinical trials were not so successful [17, 24]. But it showed tumor reductions in some cancer patients. Recently, it is a very important research area for understanding the role of these inhibitors for the treatment of certain tumors and cancer patients [25, 26] . In addition to oncogenesis, prenylation of proteins is necessary in normal cell functions. This hypothesis was proved by experimental data that genetic disruption of farnesyl transferase subunit in mice caused mouse embryo to die early.

Prenylation increases the hydrophobicity of the protein; hence the interaction of the prenylated protein with membrane increases. In RAS, the prenylation takes place either with FPP and farnesyltransferase or GGPP and geranylgeranyltransferase [27]. After, prenylation, the isoprenylcysteine carboxymethyltransferase formed OMe in place of AAX. It was found in the literature that if the protein has double CAAX box, it undergoes palmitoylation and prenylation at the two different cysteine residues. The X residue of CAAX is the deciding factor for the protein farnesylation or geranylgeranylation, but the

AA portion of CAAX is also important for its action [28]. It was found that Ras localization was markedly altered in mouse embryonic fibroblast having with and without the post prenylation processing of AAX provided the evidence for its importance in oncogenesis. So, these postprenylation enzymes are also found to be possible and attractive anticancer targets [16, 29, 30].

1.2 Methods for monitoring prenylated proteins

Several biochemical and molecular biological methods are available to study protein prenylation. Radioactive labeling is the common method for identification of prenylation sites where a radioactive precursor [^{3}H]-mevalonic acid was incorporated into the cell or animal tissues either by injecting it or by feeding the animal [41]. This [^{3}H]-mevalonic acid takes active part in the biosynthesis of cholesterol [42]. During cholesterol biosynthesis, radioactive farnesylpyrophosphate and geranylgeranylpyrophosphates were formed as the byproduct, which are the substrates for the synthesis of prenylated proteins. Generally, the proteins are separated from the whole cell lysate and then SDS-page gel helps the proteins to separate from each other and detected by autoradiography. This method does not interfere with any other post translational modification, but it is a time consuming method, needs expensive precursors and it is hazardous to health.

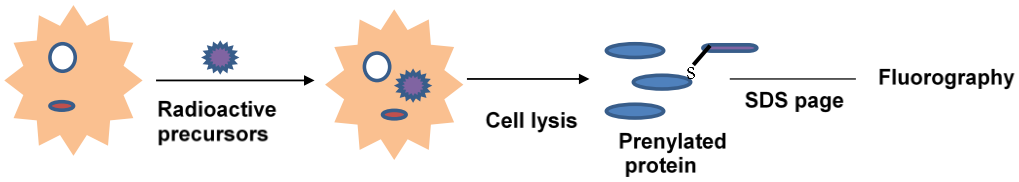


Figure 1-6 Metabolic radioactive labeling method for detection of prenylated proteins. The figure is modified from the previous publication. [43]

Different methods were developed for more better and quick detection of prenylated proteins such as using chemical probes which gave better results than the radioactive labeling method [44]. Hang et. al. found few isoprenoid proteins such as small GTPases in macrophages using an alkyne-farnesol chemical reporter [45] and improved bio-orthogonal proteomic methods. The same group identified unannotated S-prenylated substrates such as the zinc finger antiviral proteins [46] using the same chemical reporters also. A Tagging-via-substrate (TAS) approach for the identification and proteomic analysis of farnesylated proteins was done by Kho et. al. [47, 48]. TAS technology involves metabolic incorporation of a synthetic azido-farnesyl analog into the protein during prenylation and using a biotinylated phosphine capture reagent to capture the prenylated protein. Affinity purification of the prenylated protein was done either by streptavidin-linked horseradish peroxidase or agarose beads. They identified 18 farnesylated proteins by this method. However, this method was not used to identify the exact location of the modification and also the experiment will not be carried out in native state as this method needs pure Prenyltransferase as well as the substrates for prenylation.

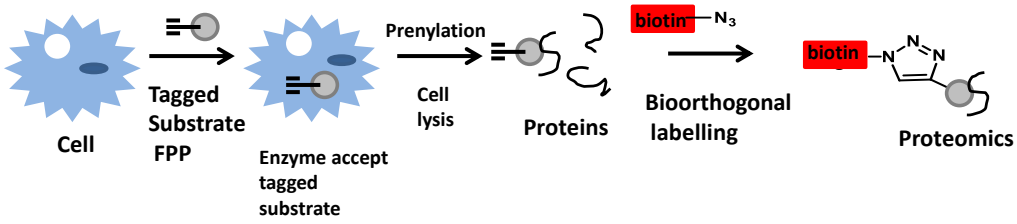


Figure 1-7 Tagging- via- substrate method for detection of prenylated proteins. The figure is modified from the previous publication [49].

Another method for identification of prenylated proteins used matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) where the sample requirement is very less and it provides important information on modifications as well as the molecular masses of modified proteins [50]. However, in this method, proteins are separated and purified by gel electrophoresis before the mass analysis. The separated proteins are then digested within the gel by the enzyme trypsin. This lipid modified proteins are highly hydrophobic and hence it sticks to the gel. They are not eluted efficiently from SDS page gel and long hydrocarbon chains decreases the ionization efficiency in mass spectrometer. Different mass spectrometric based methods for identification of farnesylated peptides were developed. In the literature, it was found that discovered that neutral loss of 204 Da farnesyl group was produced in the fragmentation of the farnesylated peptides using mass spectrometry [51, 52]. But, the peak which is formed after the loss has very low intensity in the mass spectrum. Most of the work was done for peptides only, it is difficult to analyze the proteins in complex samples because of its low abundance [53]. This fragmentation of the peptides was also found to be not consistent and had low intensity. Hence it was not used as marker ions for the detection of prenylated proteins in a complex samples.

Currently there are no methods available which can identify and distinguish two types of prenylation in a single experiment set up. Since these types of modifications are very common in certain types of cancers and diseases, so the identification and differentiation studies of prenylated peptides/proteins will be of prime importance in cancer research.

A proteome –wide approach or a global prenylome method is needed for identification of all the types of prenylated proteins in a complex matrix, which will help in cancer diagnosis research in future.

1.3 Introduction to mass spectrometry based proteomics

Mass spectrometry (MS) is the most common and sensitive analytical methods used in various posttranslational modifications discovery in recent times [54]. Briefly, the mass spectrometric workflows for proteomics generally require preparation of the sample before analysis steps. In this sample preparation step, the proteins are precipitated out and isolated from the cell or tissue. Then the protein mixture or proteins are separated from each other by their molecular weight in SDS page gel electrophoresis method. Then the gels are cut into various segments and in-gel digestion of the proteins takes place by the enzyme trypsin.

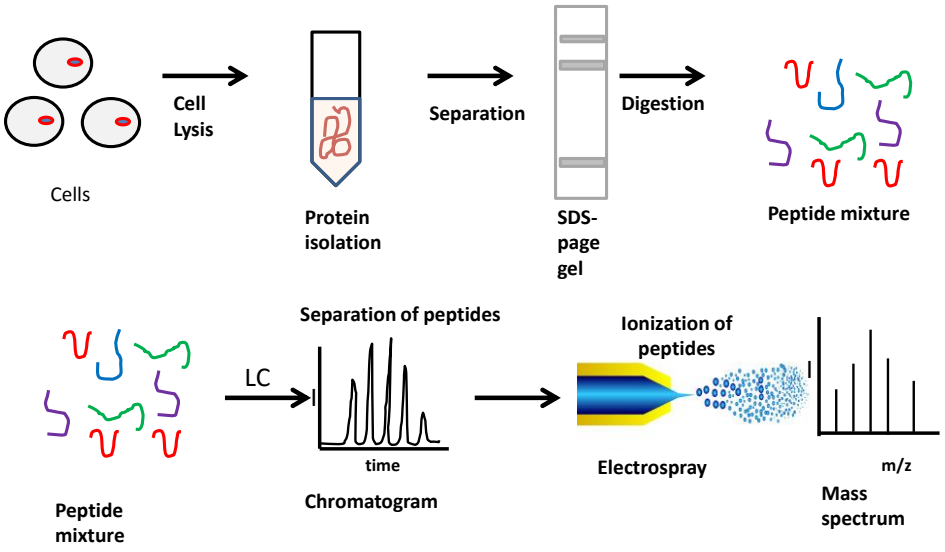


Figure 1-8 Generic mass spectrometric based proteomics workflow.

The figure is modified from the previous publication [55].

In the mass spectrometer, the digested protein or peptide mixture was injected into the ion source after liquid chromatography (LC). In LC, the peptides are separated

based on their retention time in the column and then directly injected into ion source. The ionization of the peptide in their gas phases are done either with electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI). Both of these methods are soft ionization techniques and they ionized the peptides with high sensitivity.

After the ionized molecules are formed, they are subjected towards mass analyzers like time of flight (TOF) and linear ion trap (IT) where the separation of the ions are based on their mass to charge ratio. The separated ions are then detected and forms peaks in the mass spectrum. Further fragmentation of the ions of interest are done using different methods like collision induced dissociation (CID) and electron transfer dissociation (ETD), form fragments of the ion. The fragmentation studies give the confirmation for the site of fragmentation as well as the type of prenylation in the protein. Protein identification based on peptide CID is clearer than those achieved by other methods because in addition to the peptide mass, it provides information about the protein sequence.

1.3.1 Ionization techniques

There are several ionization methods used for analyzing small molecules, but for proteomic studies two soft ionization techniques capable of ionizing nonvolatile and thermos unstable biological samples such as ESI and MALDI are mostly used. In MALDI, the Arg-containing peptides show a very high signal while ESI is less affected by the type of the amino acid residues. In ESI, firstly the peptide mixtures are separated through some reverse phase columns where they are separated based on their retention time [56]. The ionization of the peptide is done in gas phase and it works at atmospheric pressure. Here, an analyte is sprayed into the transfer tube into a strong electric field in the presence of nitrogen gas, which helps in nebulization and desolvation. The charged

ions cluster formed in this region breaks down into small ions. The ions are allowed to pass through the vacuum to the analyzer, where multiple charged ions are formed. In the analyzer, the ions are separated from each other based on their mass to charge ratio and then detected by the detector. The peaks are formed in the mass spectrum and each peak corresponds to the single mass ions.

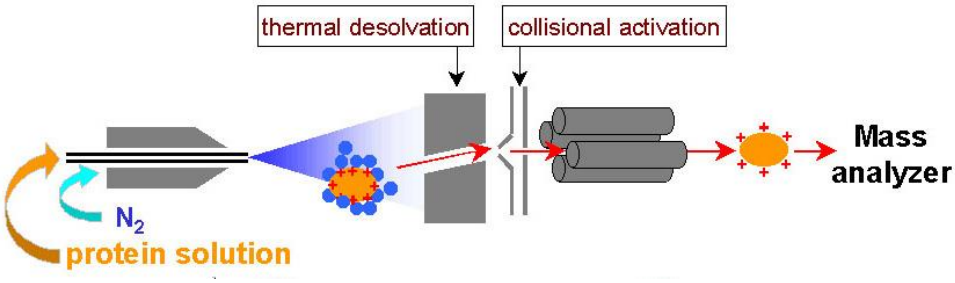


Figure 1-9 Electrospray ionization method in mass spectrometry.

The figure is adapted from website. (http://www.umass.edu).

Matrix assisted laser desorption ionization (MALDI) is another soft ionization technique, which is mostly frequently used for peptides and protein identifications [57]. Here the analyte or the sample is mixed with relatively lots of matrix and the mixture is put on a metal plate, which results in crystallization of the analyte and matrix with air drying. Then, a pulsed laser irradiates the sample that causes desorption of both analyte and matrix. The matrix absorbs the laser heat and results in the formation of more molecules and ions in the gaseous state. The analyte molecules are then undergoes ionization with the matrix which acts as proton acceptor or donor and can then be analyzed by mass analyzers mostly with time of flight (TOF) [58]. The arrival time of the

ions are depending upon the kinetic energy of the ions. Different ions have different kinetic energies; hence they are separated from each other and analyzed. The matrix can be 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid), α-cyano-4-hydroxycinnamic acid (CHCA) and 2,5-dihydroxybenzoic acid (DHB) depends upon the type of analytes.

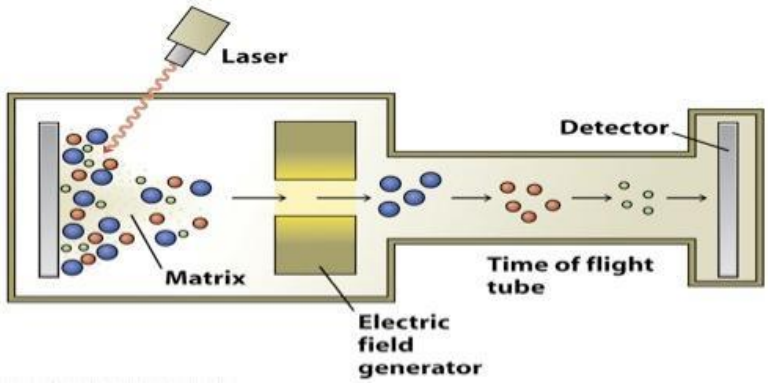


Figure 1-10 Matrix assisted laser desorption ionization method.

The figure is modified from the website (http://www.sigmaaldrich.com).

1.3.2 Tandem mass spectrometric methods

Tandem mass spectrometry is the technique in mass spectrometry that tells about the structure of the molecules. It is used for getting the structural information of peptides/proteins[59]. Tandem mass spectrometry (MS/MS) is used to first isolate the desired ions from the mass spectrum and then fragment it by collision with Argon gas. After that fragment ion spectrum determines the m/z of the fragment ions produced after collision. The spectra, where the desired ions are selected are named as MS spectra, whereas the fragment ions of the desired ion are found in MS/MS spectrum. There are two tandem mass spectrometry methods one is tandem in space and other is tandem in time. In tandem in space, the ions flow through the instrument from one mass analyzer to

another. The second analyzer acts as a collision cell, where collision of the ions with the neutral gases and among them takes place that results in the formation of fragment ions. These fragments ions are analyzed in the last mass analyzer before going for detection with the detector. The ions are separated by their *m/z* values and then the ions of specific m/z are selected by the first quadrupole (Q1). The selected ions are then excited to undergo fragmentation by collision with the neutral gas like Argon gas within the collision cell (Q2). The fragments of the ions that are generated continuously enter the third quadrupole (Q3) which scanned through the entire *m/z* range and the ions are then detected by the detector to generate a mass spectrum of the fragments (MS/MS).

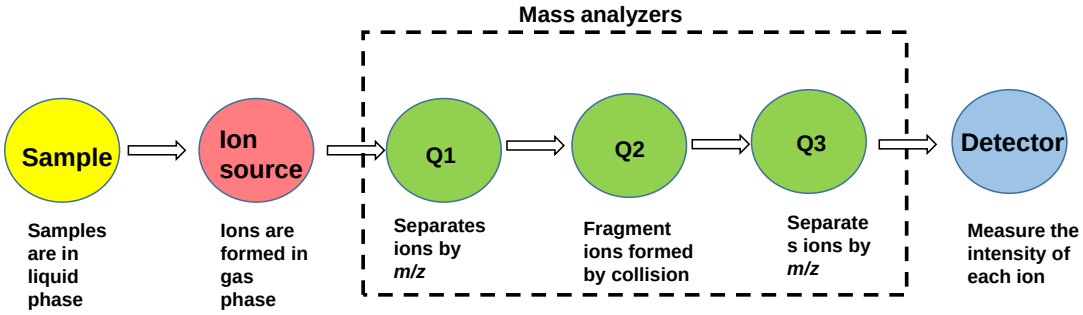


Figure 1-11 General scheme of tandem in space mass spectrometry method.

In tandem in time mass spectrometry method, the ions are trapped in ion traps where all ions of all m/z are present, then all the other ions except the specific selected ion of particular m/z which is to be fragmented is ejected from the trap [60]. The trapped ion is then excited under low vacuum and high voltage to undergo fragmentation by collision with the neutral gas and then fragments ions formed in the trap where sequentially ejected onto a detector to generate a mass spectrum of the fragments ions. This process is called MS2. The trapping and fragmentation process may be repeated as desired to

get the more information about the structure by several times in order to fragment the fragments and they are called MS^n where n is the number of times of fragmentation.

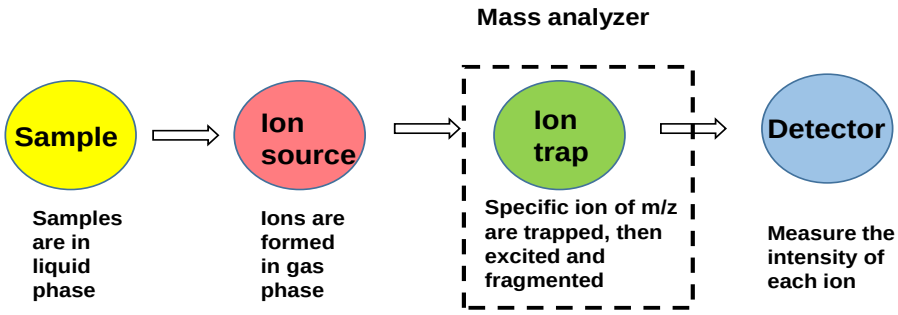


Figure 1-12 General scheme of tandem in time mass spectrometry method.

There are two common MS/MS methods that are typically used for fragmentation of the peptide to get the information about the sequence as well as any posttranslational modifications of the peptides. Collision induced dissociation (CID) is most commonly used fragmentation method where the isolation of the ions of interest takes place either in the trap or using the quadrupole and then they are excited in high vacuum and low pressure[61]. The excited ions are then colliding with the inert target gas such as helium or Argon. By this process, the weaker bond of the peptide backbone undergoes cleavage results in the formation of various fragments of the precursor ions. The structural information of the peptide is obtained from the fragment ions of the tandem MS spectra. Not only that, it also shows the position of posttranslational modification sites of the peptides. CID cleavage a major portion of the peptide backbone and sometime the labile ptms are cleaved. So, it makes the MS/MS spectra more complex and difficult to predict all the fragment ions. CID generates mostly y- and b-type fragment ions by the cleavage of peptide bonds. It was found in literature that cleavage near the amino acids E, D, and

P are generally occur in CID modes and lots of internal fragmentation along with loss of PTMs also seen in many MS/MS spectra [61].

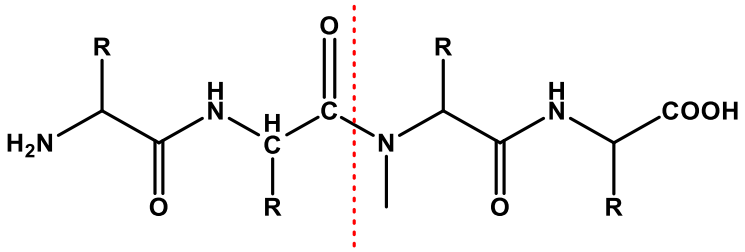


Figure 1-13 Fragmentation pattern of peptide in CID.

Electron-transfer dissociation (ETD) is another fragmentation method most widely used in mass spectrometer based proteomics for identification of ptms sites in peptides [62]. Here, selected large peptides are isolated from the rest of the peptide mixture in a trap. Then the fragmentation was done by reacting with the fluoranthene radical anion, generated by chemical ionization method. These anions resulted in the formation of fragments ions from the peptide by cleaving C-N bond and generate c and z fragments. The technique only works well for higher charge state ions and those peptides have labile ptms than the collision-induced dissociation (CID), ETD is advantageous for the fragmentation of longer peptides with higher charged states and proteins. ETD is a found to be effective fragmentation technique for the identification of modification sites of labile post-translational modifications for the proteomic studies [63].

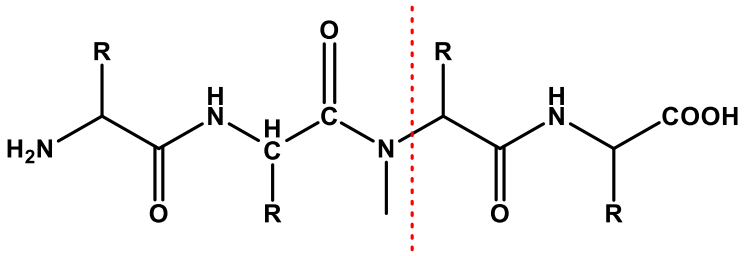


Figure 1-14 Fragmentation pattern of peptide in ETD.

1.4 Overview of the thesis

This dissertation is centered on the development of mass spectrometric based method for the large scale identification of posttranslational lipid modifications of proteins. Our method is mainly focused on the particular type of lipid modification called prenylation. They are basically two types of prenyl modifications such as farnesylation and geranylgeranylation. These modifications are observed in the carboxyl terminal cysteine residue of the proteins. The first portion (Chapter 1) of the dissertation focuses on the introduction of posttranslational modifications. Here, the three different types of posttranslational lipid modifications are discussed-prenylation, palmitoylation and N-myristoylation. Prenylated proteins are identified in various cancers and genetic diseases. Several articles were published recently and inhibitors were synthesized to block the activities of prenyl transferases. Other two lipid modifications, palmitoylation and N-myristoylation were described very briefly in the introductions. Different methods for identification of prenylated proteins were discussed elaborately. Then a short description of the introduction to general design and the workflow for mass spectrometric based proteomics was described. The different ionization techniques such as ESI and MALDI as well as the fragmentation of the molecules by CID and ETD method were discussed.

Chapter 2 describes the development of new mass spectrometry- cleavable strategy for the identification of prenylation of protein. Here, we first synthesized the farnesylated and geranylgeranylated peptides using the reagents. The products and their fragmentations were studied in ESI-IT-TOF and MALDI-QIT-TOF under CID mode. The oxidized and epoxidized prenylated peptides were also synthesized using oxidizing reagents. During MS/MS fragmentation of oxidized modified prenyl peptides, the thio ether bonds cleaved selectively in the gas phase and generate a signature fragment ion with the loss of RSOH and RSOH + nO group respectively (R= farnesyl/geranylgeranyl group and n is the number of oxygen incorporated in the prenyl chain during epoxidation). The mass loss resulted in the differentiation of the type of prenylation in a single experimental set up. It is the first time in the literature; the identification and differentiation of farnesylation and geranylgeranylation were studied by mass spectrometry. Chapter 3 highlights only the fragmentation methods of the prenylated and modified prenylated peptides by CID and ETD. The comparison as well as the advantages of each method of fragmentation is the first time discussed in the literature. This fragmentation studies creates a protein MS/MS database and results in the studies of prenylated proteins in cancer research. Chapter 4 throws a light on the future direction of the research project. Enrichment strategies were designed using epoxy modified prenyl peptides and we have demonstrated few preliminary data here. The thesis completes with a conclusion paragraph.

Chapter 2

Mass spectrometric cleavable strategy for the identification and differentiation of prenylated peptides

2.1 Abstract

Farnesylation (15 carbon prenyl group) and geranylgeranylation (20 carbon prenyl group) are two types of post-translational lipid modification of proteins. They modify the cysteine residues of the carboxyl terminal proteins having a specific motif like CAAX or CC or CXC. They are involved in several human cancers, such as pancreatic, colon and acute myeloid leukemia as well as Hutchinson-Gilford progeria syndrome (HGPS), a genetic disease that is associated with premature aging for children. Available biochemical methods in literature are not efficient in identifying and differentiating the type of prenylation. Most of the methods are tailored to specific proteins, some are expensive and some have low throughput. Studying the identification and differentiation of two types of prenylated protein/peptide using the mass spectrometric technique is also very challenging as they are highly hydrophobic and their long hydrophobic tail makes their elusion from the SDS-PAGE gel band and reverse phase C18 column is more difficult. Hence there are limited methods available for large-scale detection of prenylated proteins using mass spectrometry and almost no method currently available which can distinguish farnesylation and geranylgeranylation in a single experiment set up. A proteome wide approach to detect and differentiate the types of prenylation is needed. A simple and novel method for detection of large scale prenylated protein using mass spectrometry cleavable strategy was developed. The method utilizes the oxidation chemistry using hydrogen peroxide and m-chloroperoxybenzoic acid for the oxidation and

epoxidation of prenylated peptides. The formation of signature mass fragment during tandem mass spectrometry is observed, after the cleavage of RSOH group where R is farnesyl/geranylgeranyl group from the modified prenylated peptides. This method effectively distinguishes the types of prenylation in a single experiment set up. Proof of concept of this method was established and effective enrichment strategies were proposed using alkyne azido click chemistry reagents. Our belief is this method will improve the understanding of the role of these prenylated proteins in cancer research.

2.2 Introduction

Posttranslational modifications of proteins arises either due to proteolytic cleavage or a chemical change that takes place in the side chain of amino acid residue in a protein after the biosynthesis of protein or the translational step. They occur due to the chemical reactivity of the side chains of amino acids that changes the functional property of the protein. There are basically three types of posttranslational lipid modification, such as myristoylation, palmitoylation and prenylation [31, 37]. One of the important posttranslational lipid modifications is protein prenylation, which occurs on proteins in all eukaryotic systems. All monomeric and heterotrimeric G proteins are prenylated [64, 65]. In order for them to propagate impulses through various signaling transduction pathways, these proteins must attach to the cell membrane through their prenylcysteine moieties. There are two types of prenylation: 1) farnesylation in which the attachment of 15-carbon isoprenoid (204 Da) to the cysteine at the C terminal protein is catalyzed by farnesyltransferase (FTase), and 2) geranylgeranylation which is the attachment of 20 carbon isoprenoid (272 Da) catalyzed by geranylgeranyltransferase I (GGTase I) and bisgeranylgeranylation, catalyzed by geranylgeranyltransferase II (GGTase II), which transfers two 20 carbon isoprenoid to the C- terminus of proteins [50, 66]. The

farnesylation and later two forms of geranylgeranylation use farnesyl pyrophosphate (FPP) and geranylgeranyl pyrophosphate (GGPP) as their substrates. The most important for protein prenylation is the presence of CaaX box residue at the carboxyl end of the protein. The major determining factor whether farnesylation or geranylgeranylation takes place depends upon the amino acid 'X' [67]. The CaaX box (where "C" is the cysteine, "a" is any aliphatic amino acid and "X" is usually serine, threonine, glutamine, alanine or methionine) allows the farnesylation of protein by the help of farnesyl transferase whereas geranylgeranylation occurs with the help of geranylgeranyltransferase when the CaaX sequence ends in leucine. The most common farnesylated proteins are the members of the Ras family, but there are some other proteins are now known to be farnesylated. These include two of the nuclear lamins [16, 68, 69]. Proteins that have been shown to be prenylated by GGT I include the y subunits of heterotrimeric G proteins and other members of the small G protein family such as Ras, Rho and Rab [70]. The second factor for prenylation is CXC which commonly results in geranylgeranylation on both cysteine residues and carboxyl methylation. The last factor is CC, which is mostly found in Rab family also undergoes geranylgeranylation without further methylation [71].

Protein prenylation increases protein hydrophobicity. This increased hydrophobicity of the protein increases the affinity towards the cell membranes. It plays an important role for regulation of signal transduction pathway by increasing the protein interaction with the membrane. Although mutation in the Ras genes and involvement of prenyl modification in cancer was identified more than 20 years ago, we still don't have enough understanding about the global substrates of prenyl transferase (PT), the enzyme that catalyzes the prenyl modification in proteins. Unfortunately in cancer, inhibiting farnesyl transferase (FT) in Ras proteins by farnesyl transferase inhibitors (FTI) has failed

to prevent tumor progression due to alternative geranylgeranyl transferase mediated (GGT) prenylation [72, 73]. Recent failed clinical outcomes of FT warrants the identification of new prenyl substrates. Due to the lack of the high-throughput needed to facilitate effective identification methods, the overall status of prenyl modification in large-scale studies has never been studied. A robust method is needed which will be simple and effective and can distinguish the types of prenylation in a single experiment, and allow large-scale prenylation (prenylome) studies *in vivo*.

Several biochemical and molecular biological methods are available in the literature to study protein prenylation. Most of these studies are tailored to individual proteins. Radioactive labeling is a common method for identification of prenylated sites. In this technique, a radioactive precursor [^{3}H]-mevalonic acid, is incorporated into the cell either by injecting in the cell or by feeding the animal, which in turn participated in cholesterol biosynthesis [74-76]. The radioactive prenylated proteins thus obtained are separated using SDS-PAGE and then further detected by autoradiography. The method is very tedious with low throughput; it is also hazardous in some respect. The major drawback of this method is the length of time required for autoradiography (up to 3 months) [25]. Recently, a variation of this method was reported, where the incorporation of a thin layer chromatography analysis was used to avoid long exposure time of [^{3}H]-autoradiography. Further research led to a better method using chemical probes to identify prenylated proteins; hence, better results than those obtained using traditional radioactive labeling were developed. The application of these lipid chemical reporters has enabled the enrichment and identification of farnesylated or geranylgeranylated protein subsets, but the general proteomic coverage of prenylated proteins has been limited. Using an alkyne–farnesol chemical reporter and improved bio-orthogonal proteomic methods, Hang et al. identified isoprenoid proteins in macrophages, including

small GTPases as well as unannotated S-prenylated substrates such as the zinc finger antiviral protein (ZAP). Various tagging-via-substrate approaches were recently developed for the detection of prenylated protein using SDS-PAGE and western blotting. In these methods, the isoprenoid groups of the prenylated peptides were modified by chemical reactive groups such as azides or alkynes groups. These chemically modified prenylated groups were selectively conjugated to modified biotin moiety or fluorophore that reacts selectively with azido or alkynyl group via a Staudinger ligation or a Cu(I)-catalyzed [3+2] Huisgen cycloaddition reaction (click chemistry). Tagging has a great advantage for determining whether modifications have occurred. However, this identification was not used to identify the exact location of the modification and also the experiment did not carried out in the native settings due to the use of purified FPP and GGPP and prenyl transferase (PT). Another method of identification of prenylated proteins used matrix- assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) [43]. This method of analysis requires fewer samples than any other technique, and it provides important information on modifications as well as the molecular masses of modified proteins. In this method, proteins often are subjected to gel electrophoresis (SDS-PAGE) before the mass analysis. The separated proteins usually are digested within the gel using trypsin. However, lipid-modified peptide fragments are not efficiently eluted from the gel and long hydrocarbon chains in the lipid decrease the efficiency of ionization of the fragments. LC-MS analysis of farnesylated peptides was also carried out by Wotske et al [77]. Due to their hydrophobicity and low abundances it is difficult to analyze them in complex samples. Hoffmann and Kast discovered that the neutral loss of 204 Da (farnesyl group) was produced in CID fragmentation of farnesylated peptides using mass spectrometry. This fragmentation was not very consistent; the neutral fragment had low intensity, and to this date, the Hoffman-Kast

method has not been used as a marker to identify prenylated proteins in complex samples[51].

Current large-scale methods utilize purified farnesyl transferase and tagged farnesyl pyrophosphate to identify farnesylated proteins. In addition, the methods identify the proteins by looking at the unmodified peptides due to the inherent problem in identification of prenylated peptides by mass spectrometry. Currently, there are no methods available, which can identify and distinguish two types of prenylation in a single experimental set up. A proteome-wide approach or a global prenylome method is needed for identification of all the types of prenylated proteins that are affected by FTIs and /or GGTIs, which in turn will link the physiological effects of the inhibitors to their molecular targets. This will immensely advance the field of cancer and its diagnostics.

The mass spectrometry-cleavable reporter ion-based labeling approach is very efficient in identifying specific protein modifications. A novel strategy for the study of global protein prenylation as well as changes in the prenylation status of proteins was developed. Our new method employs simple oxidation chemistry on the prenyl groups and the cleavable properties of the sulfoxide group in a gas phase to produce a signature mass spectrum during tandem mass spectrometric events. We have also demonstrated introduction of epoxy groups in the prenylation sites of the proteins to make the prenyl peptide more hydrophilic. In addition, a strategy to introduce azide functionality was proposed to react with the epoxides in order to enrich farnesylated and geranylgeranylated peptides/proteins from complex samples. In this study we confirmed the proof-of-concept utilizing the peptides which mimics farnesyl and geranylgeranylation at native biological conditions. We have successfully demonstrated that farnesylated and geranylgeranylated peptides can be distinguished by the signature mass loss from the oxidized modified peptides. To the best of our knowledge, it is the only method currently

available which can identify and distinguish types of prenylation in a single experimental set up. We believe this method will advance large-scale identification of protein prenylation and will initiate the way to study global prenylation (prenylome) under the exposure of environmental and external factors.

2.3 Experimental

2.3.1 Materials

Peptide RGDC was obtained from Genscript (New Jersey, USA). REKKFFCAIL and REKKFFCAIM peptides were custom synthesized by Biosynthesis (Texas, USA) and Genscript (New Jersey, USA). Protein BSA was purchased from Sigma Aldrich (USA). *Trans trans*-farnesyl bromide was obtained from Sigma Aldrich (USA), geranylgeranyl bromide was synthesized from geranyl geraniol (Sigma Aldrich, USA) and phosphorus tribromide (Sigma Aldrich) and the details are in the supporting information [78, 79]. All other chemical and reagents m-chloroperoxybenzoic acid (mCPBA), sodium carbonate, hydrogen peroxide (30%v/v), dichloromethane, acetic anhydride (>95%) were obtained from Sigma Aldrich, USA. HPLC grade solvents were purchased from also from Sigma Aldrich.

2.3.2 Synthesis of geranylgeranyl bromide.

To an oven-dried 5 mL round bottom flask equipped with a magnetic stirring bar and a rubber septum, under an argon atmosphere, a solution of phosphorus tribromide (4.30 µL, 0.045 mmol, 0.65 equiv) in 1 mL of THF was added to an ice-cold solution of geranylgeraniol (20 mg, 0.069 mmol, 1.0 equiv) in 1 mL of THF and stirred at 0 °C. The

reaction mixture was monitored by TLC, until the starting material was completely consumed (around 30 min). Saturated NaHCO3 (3 mL) was added and the mixture was extracted with n-hexane (3×10 mL). The combined organic layers were dried over Na2SO4 and concentrated under reduce pressure to afford the pure product as a pale yellow oil in 74% yield (18 mg, 0.051 mmol) and use immediately. 1H NMR (CDCl3, 300 MHz) 1.58 (s, 6H, CH3), 1.59 (s, 3H, CH3), 1.68 (s, 3H, CH3), 1.73 (s, 3H, CH3), 2.07 (m, 12 H, CH2), 4.06 (d, J = 8.58 Hz, 2H, CH2Br), 5.10 (m, 3H, CH=C), 5.53 (t, J = 8.25 Hz, 1H, CH=C).

2.3.3 Synthesis of Prenylated peptides

1:1 molar ratio of peptide and *trans trans*-farnesyl bromide was dissolved in 1 mL of 7M ammonia in methanol and kept at 4 °C for about 3 hours and then the solvent was removed or dried under vacuum. The method provided a yield > 90%. The synthesized peptide was further reconstituted in 0.1 % formic acid (FA) for mass analysis studies without purification. This farnesyl peptide product was then used for epoxidation. Similarly, the geranylgeranylation of peptide was prepared in the same medium using *trans trans* geranylgeranylbromide [80].

2.3.4 Synthesis of epoxy prenylated peptides

About 1:2 molar ratio of prenylated peptides dissolved in dichloromethane was made to react with mCPBA solution in dichloromethane and then the reaction was kept at ice cold water bath for 15 minutes with constant shaking [81]. The reaction was stopped using few drops of sodium bicarbonate and then the solvent was dried under vacuum. The impure epoxy product was then reconstituted in water and the product was desalted by reverse phase HPLC C-18 column with the eluting mobile phase of 80:20:0.1 acetonitrile, water and formic acid for further mass spectrometric studies. The stability of

the epoxidized products are also studied in water as well as in 0.1% formic acid. Samples were kept at -20° C for later use.

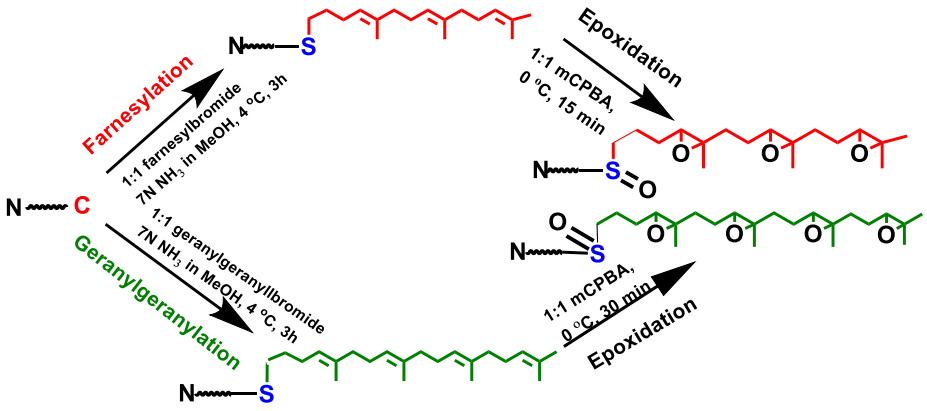


Figure 2-1 Chemical synthesis of prenylated and epoxy prenylated peptides.

2.3.5 Stability test for epoxy prenylated peptides

Aliquots of epoxy farnesylated peptides and epoxy geranylgeranylated peptides were studied in different pH conditions at 0.1% formic acid as well as pure water at room temperature. At several time points, a 2 µl of the aliquot was taken and was mixed with equal amount of 2,5-dihydroxy benzoic acid (DHB) (25 mg/ml in 50% ACN/0.1% FA). Then the mixture was co-crystallized on a steel MALDI target plate and later analyzed on MALDI-QIT-TOF mass spectrometer (Shimadzu Scientific Instruments) with 25-50% laser power. For each spectrum, 200-2000 laser shots were accumulated and the average spectrum was recorded. The MALDI –QIT-TOF mass spectra were analyzed using the Shimadzu Biotech Launch pad software.

2.3.6 Protein digestion

0.5mM of BSA(bovine serum albumin) protein was first treated with 10 mM dithiothreitol (DTT) at 56° C to reduce the disulfide bonds of the protein and then with

55mM iodoacetamide to further modify the newly formed –SH and prevent the reformation of disulfide bonds. The digestion of the protein was done with the enzyme trypsin in 1:100 w/w in presence of the digestion buffer, 50 mM ammonium bicarbonate and the digest was incubated with constant shaking at 37° C for overnight and subsequently acidified with 0.1% FA to stop digestion. The peptides thus obtained after digestion were desalted using Pierce C18 tips of bed volume 100 µl and then analyzed in mass spectrometer. Aliquots were kept at -20° C for later use.

2.3.7 Mass spectrometric parameters

All the mass spectrometric studies were performed by positive ion mode using matrix-assisted laser desorption ionization-quadrupole ion trap-time-of-flight (MALDI-QIT-TOF) (Axima resonance) and nano-electrospray ionization-ion trap-time-of-flight (ESI-IT-TOF) from Shimadzu Scientific Instruments. For the ESI-IT-TOF instrument, data dependent acquisition was performed by switching between one MS scan (r = 60 000 to 18 000 at m/z 400 to 1000 m/z) and seven MS/MS events (r = 7500). The seven most abundant ions with charge state ≥2 were isolated with a window of ±3 *m/z*, and the normalized collision energy was set at 35% for CID. For the MALDI-QIT-TOF instrument, the ion gate width was set at 2 Da for precursor ion mass of 1000 *m/z* and 4 Da for 2000 *m/z*. The synthesized farnesylated and geranylgeranylated peptides were introduced in MALDI by spotting a mixture of 1:1 mixture of peptide/protein and the matrix either 2,5 dihydroxy benzoic acid (DHB) or α-cyano-4-hydroxy –cinnamic acid (CHCA) in the MALDI plate and the spot was dried under room temperature. Intense peaks were observed corresponding to the mass of the respective prenylated peptides. The desired peaks were selected from the mass spectrum and then further collision induced dissociation (CID) fragmentation was done to obtain MS/MS spectra. The mass spectra

were collected by adding individual spectra obtained from 200 laser shots to improve the statistics of the measurement and increase signal to noise ratio. The spectrum was initially calibrated in the low and mid mass range using the peptide calibration standard (Proteo Chem). The mass spectra of the epoxidized farnesylated and geranylgeranylated were collected similarly. The fragmentation of the peptides was done with CID power of about 300 in the range of 100-1000.

2.4 Results and Discussion

2.4.1 Scheme to identify and differentiate prenylated peptides

In order to develop an effective strategy for identification and differentiation of prenyl modification of proteins in a complex sample, we developed a novel and effective strategy. Figure 2-2 depicts the schematic workflow for identification and enrichment of prenylated peptides from proteins. The targeted protein is first digested with trypsin followed by the farnesylation and geranylgeranylation using the chemical methods as well as by the use of respective enzymes (step 1). These chemical and enzymatic steps are not necessary in cell studies as real biological systems will have prenylation due to the cellular processes. Next, both the farnesylated and geranylgeranylated peptides were subjected to oxidation either with 3% Hydrogen peroxide (3% H_2O_2) (step 2a) or meta-chloroperoxybenzoic acid (mCPBA) (step 2b). Oxidation with H_2O_2 produced mono-oxidized thio-ether bond in the prenyl modification sites. To make prenyl modified peptide hydrophilic and to incorporate enrichment functionality we utilized mCPBA as an oxidizing agent. Oxidation with mCPBA served two purposes, 1) To generate a mono-oxidized thio-ether bond at the prenyl modification sites, and 2) To convert the isoprenoid groups of prenyl modification into epoxy groups. Chowdhury *et. al* demonstrated that mono-oxidized thio-ether bond is very labile (easily breakable) and specific cleavage of

this thio-ether bond of the cysteine residue take place with a neutral fragment loss in the gas phase of the mass spectrometer [82, 83]. These thio-ether bonds were cleaved selectively at the gas-phase in the mass spectrometer generating distinct mass loss which distinguished farnesylation and geranylgeranylation modification in the proteins. Incorporation of epoxy group was helped to make peptide more hydrophilic and to introduce azide functionally in the epoxy sites. The type of prenylation will be distinguished from the signature mass loss from the precursor peptide, which will lose the –RSOH (R= epoxidized farnesyl or geranylgeranyl side chain) fragment from the precursor peptides. Further MS^3 of the newly generated precursor will identify the sequence of peptide efficiently because it already lost it prenyl side chain. In these studies we have demonstrated the feasibility of large-scale studies utilizing the signature mass ions and also demonstrated the proof of concept of incorporating an enrichment tag in the epoxy sites for further enrichment of prenylated peptides. A computer search algorithm was also developed to identify the prenylation peptides using the pattern of signature mass lost during MS^2 of oxidized-farnesylated peptides.

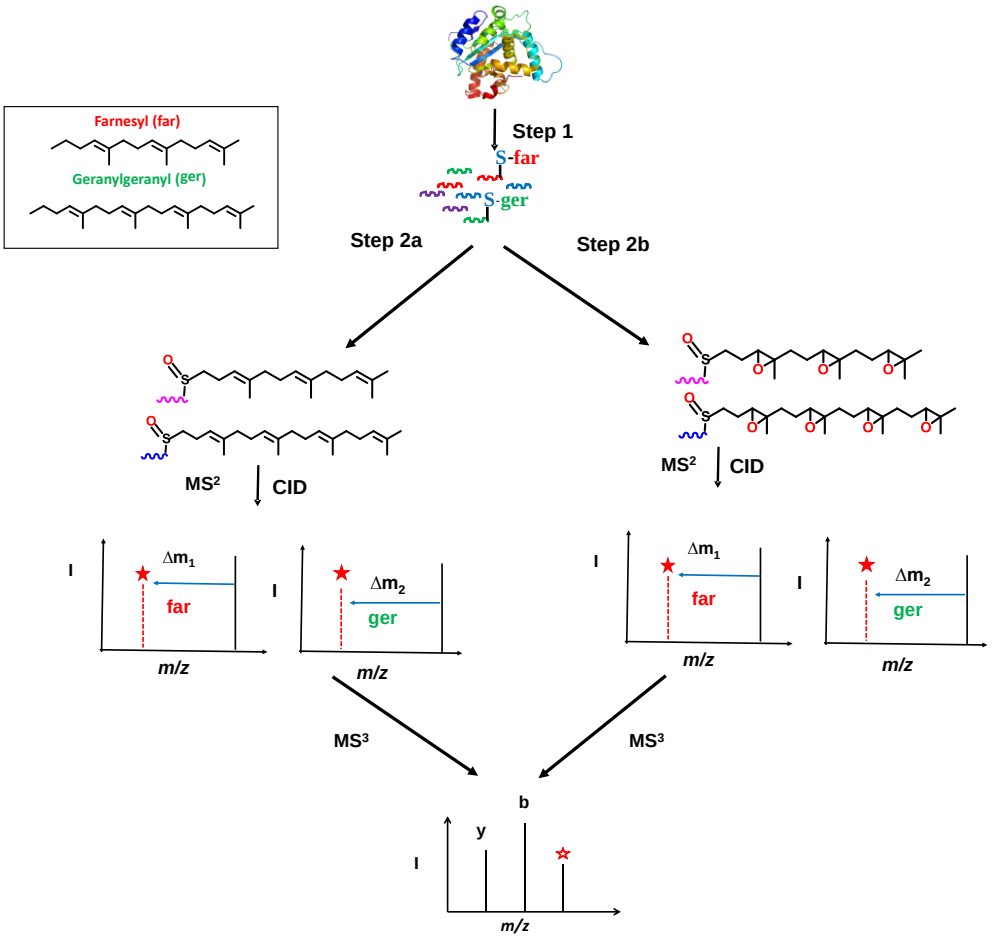


Figure 2-2 Schematic workflow for identification and enrichment of prenylated peptides.

2.4.2 MS/MS fragmentation study of prenylated peptides

In order to demonstrate our strategy, we first synthesized prenylated peptides which contain c-terminal prenyl motifs (RGDC, REKKFFCAIM and REKKFFCAIL) [84]. To produce a farnesylated peptide, a small peptide RGDC with a terminal cysteine residue (monoisotopic mass 450.1765) was chemically farnesylated. The expected monoisotopic mass of the synthesized farnesylated peptide RGDC was 654.19, with the mass addition of 204.02 Da which is due to the addition of farnesyl group. In Figure 2-3,

A (inset), intense peak at *m/z* 654.19 was detected that corresponds to the farnesylated RGDC. Small sodiated peak with the added mass of Na atom was found in the mass spectrum, which is obvious as these chemically synthesized peptide was not purified by HPLC. The peptide peak at *m/z* 654.19 was isolated and subjected to fragmentation and the MS/MS spectrum as in Figure 2-3, A, was analyzed and found that a small peak with the loss of farnesyl group at *m/z* 450.40 and the peptide fragments of *m/z* 198.09 and 330.14 were formed. The fragmentation of the farnesylated peptide showed a loss of 203.79 Da upon low energy CID [85]. In case of a C terminal CaaX motif peptide REKKFFCAIM (theoretical monoisotopic mass value =1272.65 Da), which was predicted to be farnesylated in biological medium using the enzyme farnesyltransferase, a peak corresponds to the farnesylated peptide at *m/z* 1476.22 was observed (Figure 2-3, B, inset). The MS/MS analysis of the peak at *m/z* 1476.22, as shown in Figure 2-3, B, revealed the fragments of the peptides. The peak correspond to the loss of farnesyl group is not prominent in the spectrum. Hence it is clear that the detection of low abundance modified peptide using this loss in a complex sample is quite challenging. Both RGDC and the CaaX motif peptide REKKFFCAIM were geranylgeranylated chemically with geranylgeranyl bromide in the similar reaction condition as mentioned earlier. The subsequent addition of geranylgeranyl group having the mass of about 272 Da took place at the cysteine residue of the peptides. The geranylgeranylated peptide (theoretical monoisotopic mass = 722.17 Da) was observed at 721.86 as shown in Figure 2-3, C**.** Further fragmentation of the peptide confirmed the attachment of the geranylgeranyl group to the cysteine residue. The CaaX motif geranylgeranylated peptide (theoretical monoisotopic mass = 1544.65) peak was found at *m/z* 1544.41 (Figure 2-3, D).

Another CaaX motif peptide REKKFFCAIL (monoisotopic masses= 1254.7026) was also farnesylated and geranylgeranylated chemically and the farnesylated peak with the addition of 204.38 Da and geranylgeranylated peak with the attachment of 272.46 Da was observed at *m/z* 1459.09 and m/z 1527.17 respectively. It was observed from Figure 2-3, A, the MS/MS of the farnesylated RGDC results in a small peak of *m/z* 450.40 with a loss of farnesyl group and the fragments of the peptides. This is also according to the fragmentation as depicted in the literature by Hoffmann and Kast. The peak with a loss of geranylgeranyl group was not at all seen in the MS/MS spectra because the geranylgeranyl group was labile and it undergoes fragmentation easily. Thus, it is very clear that locating these two modifications in unknown prenylated peptides using MS/MS pattern will be very difficult to predict in a complex mixture. The Figure 2-3 data were shown using a MALDI-QIT-TOF mass spectrometer but our ESI-MS/MS data also delivered similar MS/MS pattern for these peptides. We would also like to mention that there is not many MS/MS study on geranylgeranyl peptide having the prenyl motifs, and these MS/MS studies could serve as a reference for that purpose. The detailed synthesis of farnesylated and geranylgeranylated peptides was mentioned in the material and method section.

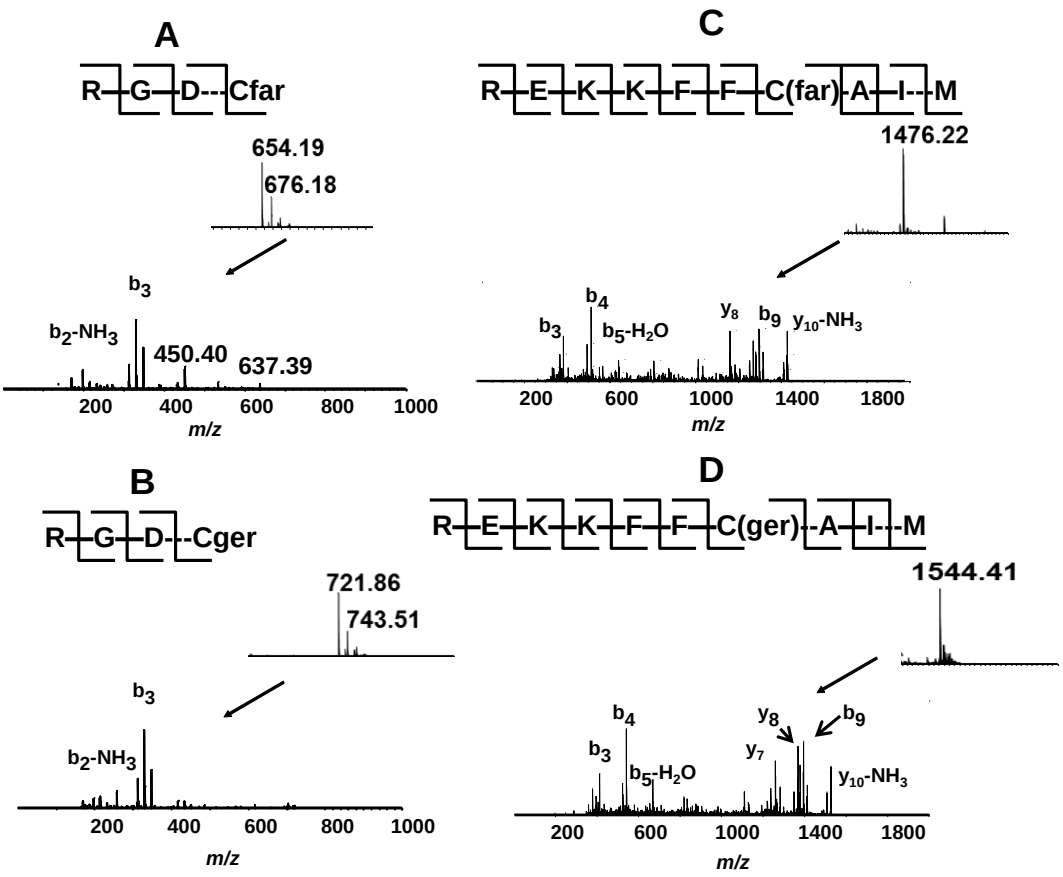


Figure 2-3 Mass spectra of the synthesized farnesylated and geranylgeranylated peptides: (A) RGDC(far) at *m/z* 654.19 (inset) and MS/MS of 654.19, (B) RGDC(ger) at *m/z* 721.86 (inset) and MS/MS of 721.86, (C) REKKFFC(far)AIM at *m/z* 1476.22 (inset) and MS/MS of 1476.22, and (D) REKKFFC(ger)AIM *m/z* 1544.41 (inset) and MS/MS of 1544.41.

2.4.3 Oxidation and epoxidation of prenylated peptides

The oxidations of the prenylated peptides (RGDC(far/ger) and REKKFFC(far/ger)AIM) were done initially using 3% hydrogen peroxide solutions to oxidize the sulfur atom of the cysteine. It was already reported that fragmentation of mono-oxidized thio-ether containing peptides results in loss of specific fragment (sulfenic group, RSOH) from the peptides. The idea about the oxidation and fragmentation of the oxidized prenylated peptides leads us to think about the strategy of identification and differentiation of prenylated peptide from the mass loss from the precursor peptides. Although 3% H2O2 treatment generated thio-ether bond which is labile, we found same thio-ether bond can be incorporated along with the epoxy groups utilizing m-chloroperoxybenzoic acid, a common epoxy reagent.

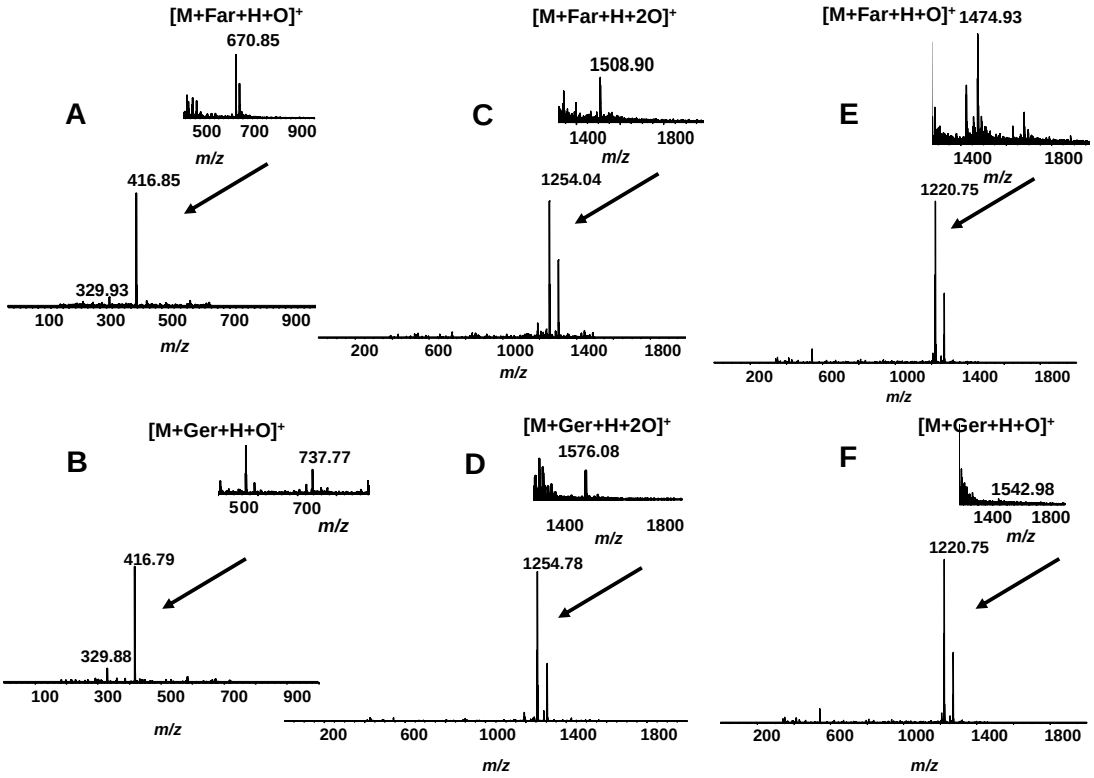


Figure 2-4 Mass spectra of the oxidized prenylated peptides (formed by the oxidation with 3% H_2O_2). (A) Oxidized RGDC(far) with m/z 670.85 (inset) and MS/MS of 670.85. (B) Oxidized RGDC(ger) with m/z 737.77 (inset) and MS/MS of 737.77. (C) Oxidized REKKFFC(far)AIM with m/z 1508.90 (inset) and MS/MS of 1508.90. (D) Oxidized REKKFFC(ger)AIM with m/z 1576.08 (inset) and MS/MS of 1576.08. (E) Oxidized REKKFFC(far)AIL with m/z 1474.93 (inset) and MS/MS of 1474.93. (F) Oxidized REKKFFC(ger)AIL with m/z 1542.98 (inset) and MS/MS of 1542.98.

Epoxidation was very similar to the oxidation step where along with the oxidation of the prenylated peptides, the incorporation of the epoxy group into the prenyl group took place [86]. Since there are three conjugative double bonds are present in the farnesyl group, incorporation of maximum of three epoxy groups are expected (addition of three oxygens) along with the oxidation of sulfur atom (S=O). Additionally, if the peptide has methionine residue, the S atom of the methionine should also be oxidized. The confirmation of the oxidation of the S atoms of the cysteine and that of methionine was done by the treatment of the prenylated peptide with 3% hydrogen peroxide for about 20 minutes to get only the oxidized peptides. After we observed the desired thio-ether cleavage during MS/MS, we moved forward with our oxidation reaction with mCPBA. Our treatment of farnesylated and geranylgeranylated RGDC and c-terminal CaaX motif peptide REKKFFCAIM with mCPBA both generated several epoxy groups in the peptides. The several epoxidized peptides are obtained with the incorporation of different number of atoms of oxygen, which was controlled by optimizing the molar ratio of the peptide and mCPBA. At the beginning we were concerned about several epoxidized products but these products found very useful in locating prenyl peptide by MS/MS by generating several targets for tandem mass spectrometry. In (Figure 2-4, panel A), mass spectra of the farnesylated and epoxidized RGDC peptide was shown. For the RGDC epoxidized farnesylated peptide, the variable oxygen products were obtained at m/z 670.17 (M+H+Far+O), 685.67 (M+H+Far+2O), 701.67 (M+H+Far+3O) and a very small peak at 717.67 (M+H+Far+4O). Each epoxidized products have a mass difference of nearly 15.99 Da, which corresponds to the mass of an oxygen atom (Figure 2-4, panel A). After MS/MS of each precursor peptide, a very intense peak at about m/z 416.85 was formed. The loss corresponds to the calculated mass of farnesyl group with

added oxygen from epoxy group and loss of sulfenic group (RSOH, R = farnesyl/geranylgeranyl) (Figure 2-4). It is clear that mono-oxidized thio-ether bond in farnesyl side chain cleaved effectively along with the added mass due to the epoxidation. Interestingly all the variable epoxidized products generated same precursor peptide following a signature mass loss from the original modified peptide (Figure 2-4, panel A and B). Further MS/MS of m/z 416.85 clearly identified the original peptide sequence. Similar experiment with geranylgeranylated RGDC clearly showed the same pattern and signature mass loss with added masses, with a distinguished labile peak at the mass spectra. The epoxidation of geranylgeranylated peptide RGDC of m/z 721.83 resulted in the formation of variable oxidation products of m/z 737.91 (M+H+Ger+O), 753.78 (M+H+Ger+2O) and 769.57 (M+H+Ger+3O) respectively as shown in Figure 2-4, panel B. Here, we also observed the mass difference between the epoxidized products which is the mass of an oxygen atom i.e. 15.99 Da. The similar signature peak at m/z 416.79 was obtained in every case with the loss of prenyl group with added oxygen atoms and SOH group (Figure 2-4, panel B). The mass difference between the signature peak and that of precursor defined the type of prenylation. It is clear from the figure 3 that using these signature mass loss both farnesyl and geranylgeranylated peptide can be distinguished unambiguously. The same loss pattern of SOH group from the oxidized peptides were reported previously in the literature by Chowdhury et al.[82], if the cysteine or the methionine S atoms are oxidized to form S=O group. The identical or the signature peak obtained in all the fragmentation of the epoxidized farnesylated peptides confirmed the theory of loss of SOH group along with the epoxy prenyl group. The mass difference of the signature peak and the precursor ion would define the type of prenylation in the peptides. The same signature peak was obtained from each variable oxidation products

during MS/MS, which increased the count of number of targets for same prenylated peptide to be identified by this strategy.

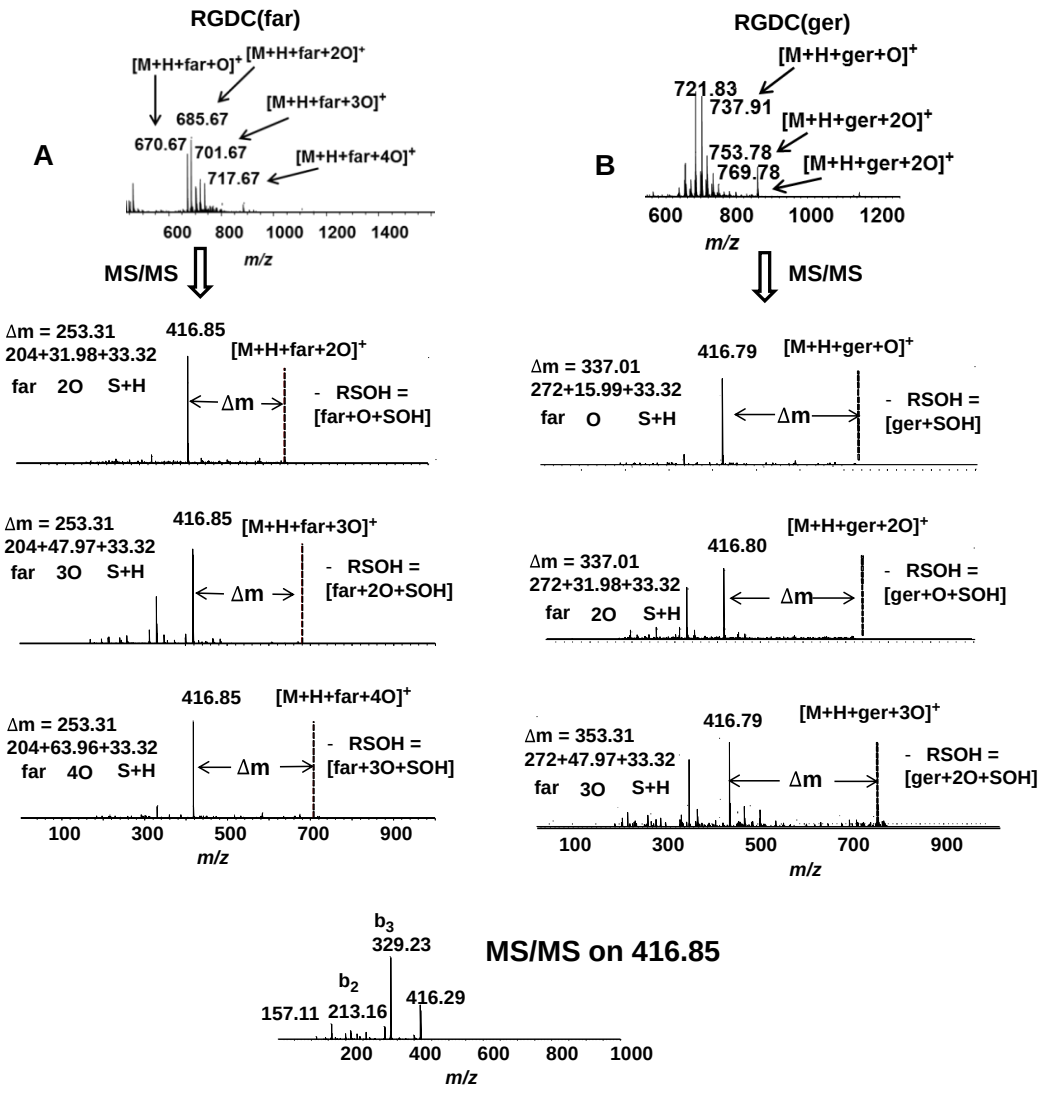


Figure 2-5 MALDI-QIT-TOF MS/MS spectra of the epoxidized farnesylated and geranylgeranylated RGDC peptides. Panel A: Top - RGDC(far) epoxides at *m/z* 670.67, 685.67, 701.67 and 717.67. Bottom - MS/MS of peptides having *m/z* 685.67, 701.67 and 717.67 (from up to down). Panel B: Top RGDC(ger) epoxides at *m/z* 737.91, 753.78 and

769.78. Bottom- MS/MS of peptides having *m/z* 737.91, 753.78 and 769.78 (from up to down).

The formation of various epoxidized products of CaaX peptide REKKFFCAIM with *m/z* of 1508.27 (M+H+Far+2O), 1524.24 (M+H+Far+3O), 1540.27 (M+H+Far+4O) and 1556.23 (M+H+Far+5O) were confirmed in Figure 2-5, panel A. The peak at *m/z* 1508.27 was formed with the mass addition of 2 x15.99 Da, i.e. the addition of two oxygen atoms (one S=O for cysteine and another S=O for methionine. This also confirmed that these two oxygens are not situated in methionine. The selectivity of sulfoxide formation was also confirmed by the reaction of this peptide with hydrogen peroxide that resulted in a peak at *m/z* 1508.90 (M+H+Far+2O). MS/MS of all epoxides products, at *m/z* 1508.27 (M+H+Far+2O), 1524.24 (M+H+Far+3O) and 1540.27 (M+H+Far+4O) generated same signature peak at *m/z* 1254.45, which was formed with the loss of the prenyl group with the added oxygen atoms and SOH group (Figure 2-5 , panels A).

The MS/MS spectrum of all the epoxidized farnesylated peptides were less noisy as compared to that of pure farnesylated peptide, that also confirms that the ionization of modified hydrophilic peptides are better in gas phase in the mass spectrometer.

The epoxidation of CaaX motif geranylgeranylated peptide results in the formation of the corresponding epoxides products at *m/z* of 1576.09 (M+H+Ger+2O), 1592.09 (M+H+Ger+3O), 1609.09 (M+H+Ger+4O), 1625.08 (M+H+Ger+4O) and 1641.08 (M+H+Ger+4O). In Figure 2-5, panel B, in the top we pointed the arrow on the higher oxidized products. MS/MS of 3, 4, 5 oxygen containing products generated the same signature peaks at *m/z* 1254.76 after the loss of RSOH (R =- epoxidized ger) (Figure 2-5, panel B). MS/MS 1254.49 was also shown for identification of the sequence of the peptide.

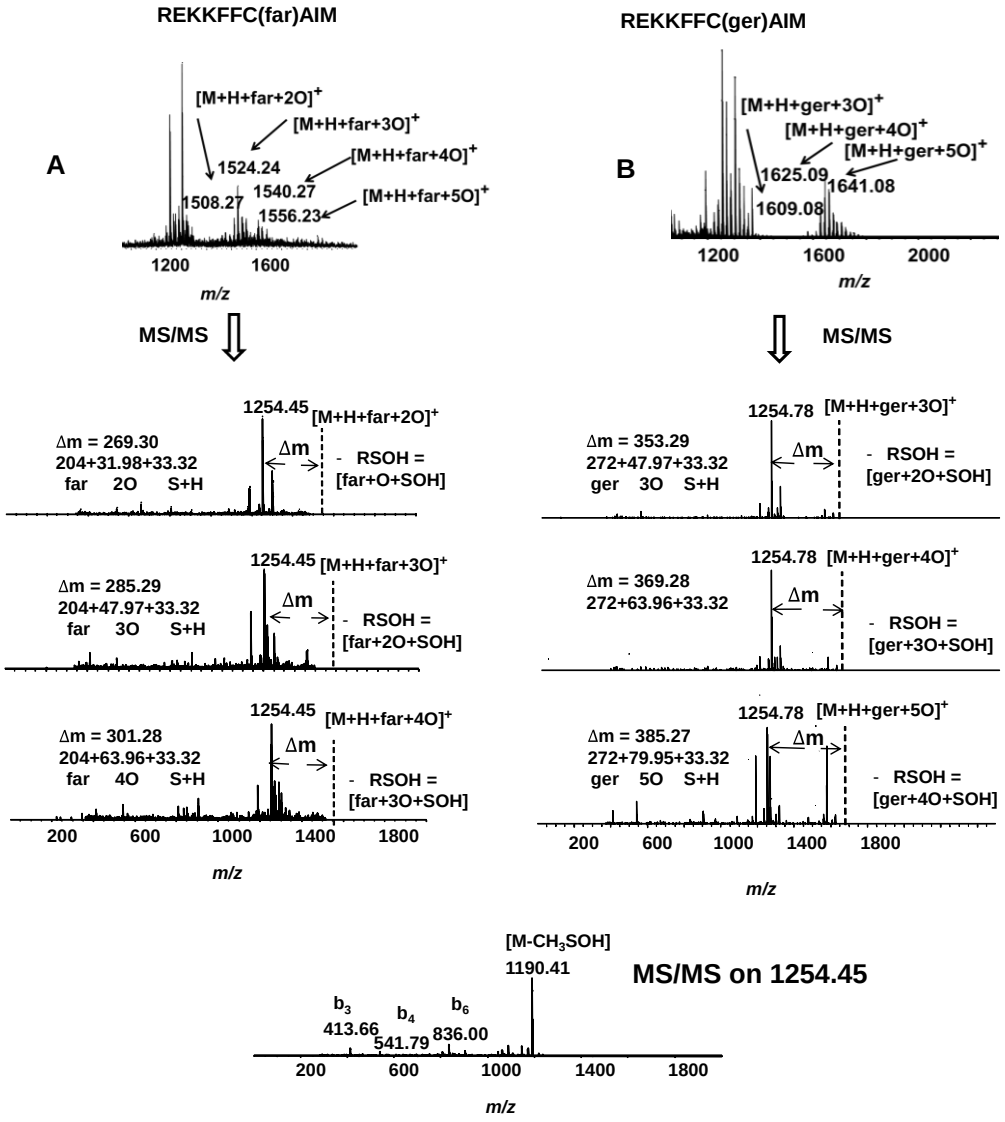


Figure 2-6 MALDI-QIT-TOF MS/MS spectra of the epoxidized farnesylated and geranylgeranylated REKKFFCAIM peptides. (A) Top, several epoxidized products of REKKFFC(far)AIM; bottom, MS/MS of peptides having m/z 1524.24, 1540.27, and 1556.23 (from up to down). (B) Top, several epoxidized products of REKKFFC(ger)AIM;

bottom, MS/MS of peptides having m/z 1609.09, 1625.08, and 1641.08 (from up to down).

2.4.4 Epoxidation of the prenylated peptide in a complex matrix

To validate our method for identification of prenylated peptides in a complex sample, we did the spiking study of the prenylated peptides in a matrix of non prenylated peptides of the protein BSA. Three different prenylated peptides RGDC(far) (*m/z* 654.19), REKKFFC(far)AIM (CaaX peptides having *m/z* 1476.22) and another CAAX peptide REKKFFC(far)AIL having (*m/z* 1458.70) were added in the tryptic digest of 10 µl of 0.5 mM BSA protein. Then the epoxidation reaction was carried out in aliquot of mCPBA. That resulted in the formation of the epoxidized peaks of the prenylated peptides with the mass difference of 15.99 Da (mass of an oxygen atom). The epoxidized peaks were isolated for further fragmentation in MALDI-QIT-TOF. The fragmentation of the epoxidized prenylated peptides showed the similar fragmentation pattern and the formation of fragments with the loss of the prenyl group and the sulfenic acid groups were observed (Figure 2-6). These data provides the proof that selective oxidation and epoxidation is possible in complex background utilizing mCPBA. Our subsequent MS/MS of these peptides also showed that these peptides can also be selectively fragmented for further identification.

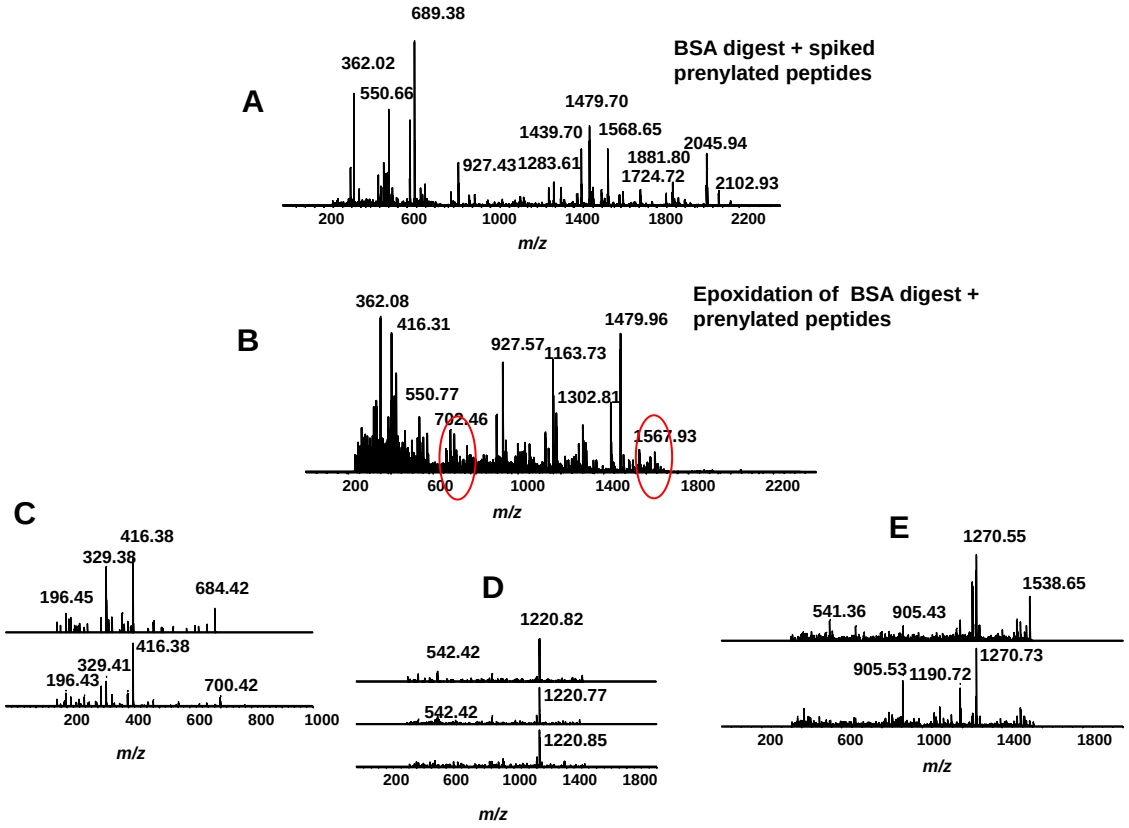


Figure 2-7 Feasibility of selective epoxidation in complex peptide mixtures. (A) MALDI-QIT-TOF mass spectrum of the tryptic digests of BSA and spiked prenylated peptides. (B) Mass spectrum of the BSA digest with spiked prenyl peptides after epoxidation. (C) MS/MS spectra of the epoxidized RGDC(far) at m/z 702.46 (M + Far + H + 3O) and m/z 718.46 (M + Far + H + 4O). (D) MS/MS spectra of the epoxidized REKKFFC(far)AIL at m/z 1490.96 (M + Far + H + 2O), 1506.96 (M + Far + H + 3O), and 1522.95 (M + Far + H + 4O). (E) MS/MS spectra of the epoxidized REKKFFC(far)AIM at m/z 1556.89 (M + Far + H + 5O) and 1572.92 (M + Far + H + 6O). Precursor fragments correspond to the mass of dioxidized methionine.

2.4.5 Stability of epoxy prenylated peptides and analysis by RP-HPLC- ESI-MS

We analyzed all the epoxy modified peptides utilizing MALDI-QIT-TOF MS. In order to see if the same signature mass was retained in ESI-MS, we studied the same peptide using LS-MS/MS condition. Although, the main purpose of adding an epoxy group is to develop an enrichment strategy and to make the peptide more hydrophilic, the method is quite efficient in identifying the prenyl modification using the signature peaks without enrichments due to the formation of various epoxy products to be identified by tandem mass spectrometry. Several HPLC methods have been developed to analyze very hydrophobic peptides and peptides with lipid modifications. Due to their dramatic difference in hydrophobicity, it is challenging to analyze these hydrophobic peptides and their unmodified counterparts in a single HPLC run. While it was possible to analyze lipid modified peptides on a C18 column, it requires the use of strong organic solvent. The retention time for RGDC farnesylated and geranylgeranylated ones were 46.690 and 66.338 min in 80 min 5-80% ACN gradient run through the monolithic C-18 column. It was quite obvious that hydrophobic peptides eluted at higher organic solvent composition, but sometimes, it still retained in the column for longer time.

The epoxidized farnesylated RGDC products were subjected to separation in reverse phase column in nano LC system. The epoxidized products are much more hydrophilic as compared to their farnesylated ones and all the epoxidized products have the similar hydrophilicity and a little difference in it is due to the number of oxygen atoms incorporation. It is quite challenging to separate them in 10 cm monolithic C-18 column by using the gradient method. Hence, we separated them using a long run gradient of about 80 min with 5-80% ACN. The oxidized peptides were separated out with a time difference of about 1 min. The retention times for the mono-epoxidized, di-epoxidized

and tri-epoxidized peptides were 37.537, 36.017 and 34.497 min respectively (Figure 2-7). The separated peaks were analyzed online by ESI-IT-TOF.

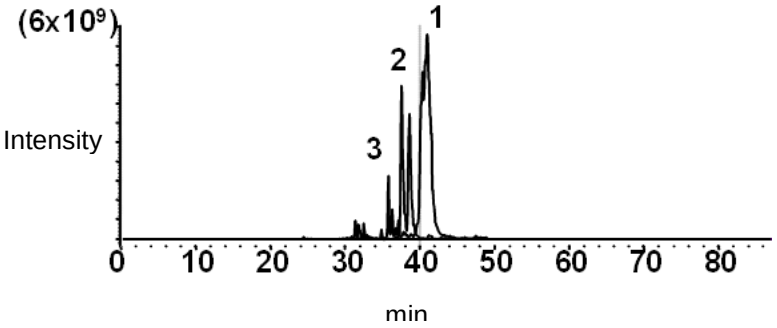


Figure 2-8 Chromatogram of the mixture of the epoxidized farnesylated RGDC having retention time 37.537(no. 3), 36.017(no. 2) and 34.497(no. 1) min.

The CID fragmentation was done so as to get the fragments at a very low energy. The CID spectra of all the epoxidized products are also showing the loss of the marker RSOH ions where R stands for the farnesyl group with the number of added oxygen atoms in that group.

After we found that epoxides groups are stable in long gradient in LC-MS, we performed the LC-MS/MS study with our CaaX motif containing peptide, REKKFFC(far)AIM.

In figure 5, we show our LC-ESI-MS/MS experiment using an ESI-IT-TOF mass spectrometer using the CaaX box farnesylated peptide. All the variable oxidized products in Figure 2-8 are shown along with their MS/MS. It shows the similar pattern which we observed in MALDI-QIT-TOF mass spectrometer. It is very important to note that after 2h gradients the epoxy groups were very stable at LC-MS/MS conditions. It is clear that targeted MS/MS of these oxidized peptides unambiguously identified the prenyl modifications and their types.

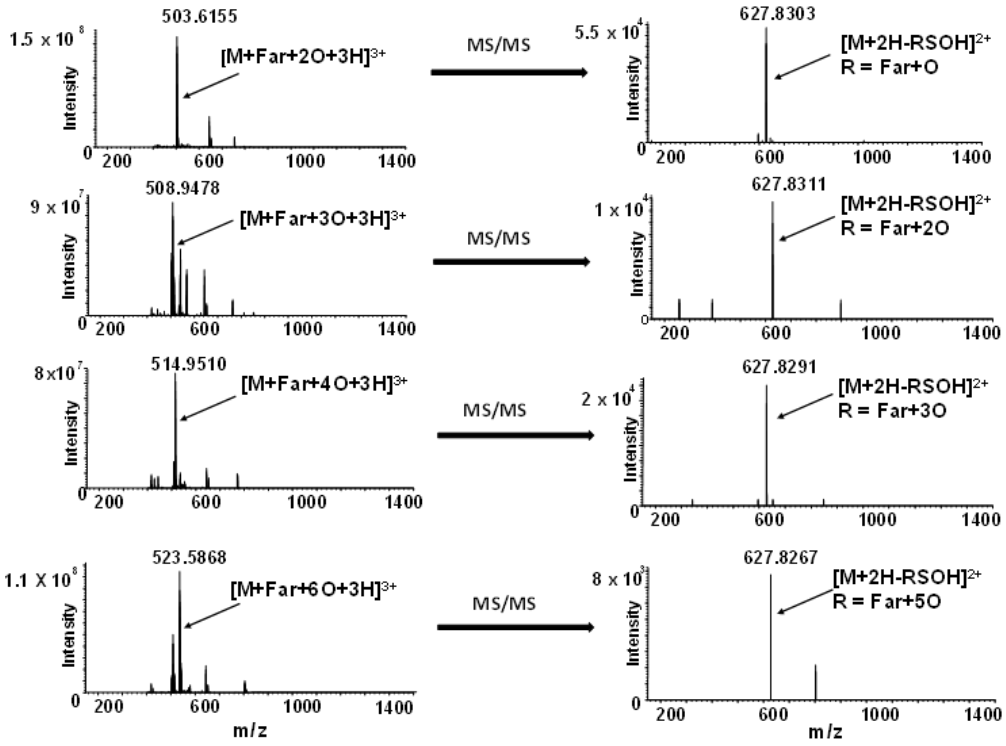


Figure 2-9 LC-MS/MS spectra of the epoxidized farnesylated REKKFFC(far)AIM peptides in nano ESI-IT-TOF and MS/MS of the epoxidized peaks of triple charged species of m/z - A) 1507.8465 $(M+Far+3H+2O)^{3+}$, B) 1523.8434 $(M+Far+3H+3O)^{3+}$, C) 1541.853 $(M+Far+3H+4O)^{3+}$ and D) 1569.6556 $(M+Far+3H+6O)^{3+}$(from top to bottom).

2.4.6 Enrichment studies: proof of concept

It is quite evident that addition of epoxy groups makes these peptides more hydrophilic and we found epoxy group was also quite stable during LC-MS/MS. We are very confident that using this fragmentation signature we can unambiguously detect the

location and types of prenylation without further enrichment. Since we have successfully incorporated epoxy groups in the prenyl modification sites, we designed an enrichment strategy of these epoxidized peptides. In order to demonstrate that this peptide can be enriched using click chemistry, we tested the incorporation of azide moiety in the epoxy farnesylated peptide. The peaks at m/z 745.3570 (M+H+Far+3O+azide) and 761.3491 (M+H+Far+4O+azide) were adducts of the epoxidized farnesylated RGDC with one azide moiety are shown in Figure 2-9. These peaks were separated from the rest of the epoxidized counterparts as they have slightly higher retention time as predicted. The CID fragmentation of these adducts also showed same fragment at m/z 416.1839 with the loss of RSOH group (where the R = epoxy farnesyl + one azide group). It is very clear that even after addition of azide moiety the same signature mass pattern was retained which showed the strength of this approach to identify prenylated peptide before and after enrichment using the same signature mass loss.

The similar experiment was performed on the farnesylated epoxidized CaaX peptide REKKFFCAIM with the sodium azide for specific reaction between the epoxy and azide group and the similar result with only one azide group incorporation in the epoxy prenyl group were observed in case of two of the epoxides with m/z 1568.85 (M+H+Far+3O+azide) and 1584.85 (M+H+Far+4O+azide). Further fragmentation of these adducts showed same fragment at m/z 1254.78 with the loss of RSOH group.

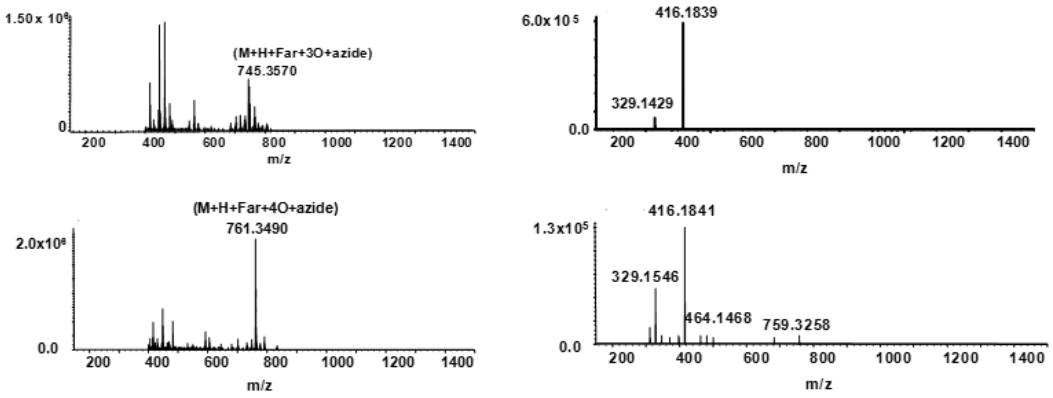


Figure 2-10 LC-MS spectrum of the azide incorporated epoxidized farnesylated RGDC peptides in nano ESI-IT-TOF and MS/MS of the peaks with m/z 745.3570 and 761.3490.

The incorporation of azide clearly showed that this peptide can be enriched using click chemistry. One puzzling thing we see that instead of incorporation of two or three azide moiety we see incorporation of one azide. We are currently working on to understand this phenomenon and click chemistry approach to enrich this peptide selectively from the complex mixtures. We are also working on to introduce charged species into the epoxides so prenyl peptides can be enriched utilizing ion-exchange chromatography. These data clearly showed that introduction of azide moiety is feasible in the in the epoxy farnesylated/geranylgeranylated peptides and same signature fragmentation is retained during MS/MS.

2.5 Conclusions

Prenylated peptides are difficult to detect in mass spectrometer due to its low abundance and hydrophobicity of the prenyl group. However, after oxidation and epoxidation of prenylated peptides/proteins, the site of prenylation can be detected by fragmenting the sample in CID in gas phase in a mass spectrometer. In this study we demonstrated that prenyl peptide can be identified and distinguished using a signature mass loss peak in the mass spectra. This peak was generated due to a selective oxidation reaction which introduces a labile sulfoxide group and hydrophilic epoxy functionality in the prenyl modification sites. In order to validate our method, we studied three types of prenylated peptides. One of the peptide has C-terminal cysteine and the other two has CaaX motif at the carboxyl terminal, which undergoes farnesylation and geranylgeranylation in biological medium depending upon “X”. The drastic difference in hydrophobicity between the epoxidized prenylated and their prenylated forms enables the effective separation by LC-MS. The stability of the epoxidized peptides are also seen in two hour long gradient run in LC-MS. The signature mass loss on CID fragmentation of epoxidized peptides results in identification and differentiation of farnesylation and geranylgeranylation of prenylated peptides. The epoxidation study was also done by spiking the prenylated peptide in a tryptic digest of BSA. The method is quite efficient to identify and distinguish type of prenylation in complex mixtures using the signature mass loss and due to the generation of several epoxidized products. Epoxidation made the prenyl peptides more hydrophilic and gave an opportunity to introduce enrichment functionality such as azide groups. Future enrichment strategy can be developed using the biotinylated click reagent that selectively reacts with azide. We are currently generating our farnesylated protein and farnesyl transferase to demonstrate this in purified proteins, thus we can transfer this technique in complex cell lysate. A computer

search algorithm was also developed to search the prenylation in large-scale studies. Currently we are developing the software so this method can be available to cancer researchers for general use.

It is almost 20 years the modification of Ras, an oncogenic protein and its prenyl modifications was identified, we still don't have enough understanding about the effect of this modification in cancer. Due to the lack of high-through put method, discovery of new prenyl substrate are hindered. There are few available methods to study farnesylation *in vit*ro but no method is capable of distinguishing types of prenylation *in vivo*. This method will open the possibility of studying these two types of prenylation in systems-level studies under exposure to environmental and external agents. We believe this method will initiate the way to study this modification in large-scale in a clinical sample and thus will identify the potential new target for cancer drugs.

Chapter 3

Gas-phase fragmentation behavior of prenyl peptides by differential tandem mass spectrometry

3.1 Abstract

Farnesylation and geranylgeranylation are the two types of prenyl modification of proteins. Identification and differentiation between farnesylated and geranylgeranylated peptides with mass spectrometry is quite challenging. Prenylated peptides are highly hydrophobic and also their abundances in the biological samples are low. Detailed mass spectrometric fragmentation behavior of two types of prenylated peptides was not reported due to the fact that most of the studies are identifying prenyl proteins by looking at the unmodified peptides. At present, there is no study available where characteristics mass- spectrometric fragmentation of geranylgeranyl peptides was shown. In this report, we studied the prenylated peptides by mass spectrometry and identified them by Collision Induced Dissociation (CID) and Electron Transfer Dissociation (ETD) tandem mass spectrometry. Modified prenyl peptides were generated utilizing strong and low strength oxidizing agents to selectively oxidized cysteine sulfur and also selectively epoxidized prenyl side chain. We oxidized and epoxidized them to make them hydrophilic in nature and to make prenyl side chain labile in the gas-phase. We selected three peptides with prenyl motifs and synthesized their prenylated versions. The detailed characteristics fragmentations of farnesylated and geranylgeranylated peptides along with their modified version were studied. CID and ETD mass spectrometry clearly distinguished the unmodified and modified version of these peptides. ETD mass spectrometry clearly sequenced the highly labile modified prenyl peptides and showed

different characteristics fragmentations compared to CID. No comprehensive studies on the tandem mass spectrometry behavior of two types of prenylated peptides ever studied in differential tandem mass spectrometry and this report will significantly improve our knowledge of the behavior of these peptides in the gas-phase.

3.2 Introduction

Prenylation is a type of posttranslational lipid modifications which generally occurs at the cysteine residues situated at the carboxyl terminal of the protein. Farnesylation and geranylgeranylation are the two types of prenylation [87]. Farnesylation is a type of post translational lipid modification where a 15 carbon prenyl group (~204 Da) is attached to cysteine residue of the carboxyl terminal whereas geranylgeranylation is 20 carbon prenyl group (~272 Da) attached to the cysteine residue [88]. These modifications of proteins are connected with several human cancers, such as pancreatic, colon and acute myeloid leukemia and other diseases like progeria, aging, parasitic diseases and bacterial and viral infections [28]. Potential inhibitors of prenyl transferase, the enzyme which causes the prenyl modifications, were developed and investigated to suppress the cancer in oncogenic cells and tissues, but clinical trials using these inhibitors were not found to be very effective. This challenge created an intense area of investigation to find new prenylation sites and distinguish their types in proteins and large-scale samples and define their roles in several diseases so appropriate inhibitors can be synthesized. There are limited methods available for large-scale detection of prenylated proteins using mass spectrometry. Currently available methods utilize radiolabeled or biotinylated isoprene and biorthogonal isoprene reporters and click chemistry based enrichment strategy. Several bottlenecks hampers these methods, such as difficulty of synthesis of probes, toxic nature of the probes, and inefficient radioactive

labeling. Most methods identify the proteins by sequencing un-modified peptides after enrichments.

The matrix- assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS) was used for the determination of prenylated proteins [89]. However, in this method, proteins often are subjected to gel electrophoresis (SDS-PAGE) for separation and then the gel are cut into various segments for further digestion with the enzyme trypsin. After digestion, the peptide mixture was analyzed by mass spectrometry and further data analysis helps in protein identification. It was found that mostly all prenylated proteins are hydrophobic and hence they are difficult to elute from SDS page gel after in-gel digestion. As a result, the abundance of the sample is quite less in the sample for analysis. Few mass spectrometric studies was performed on farnesylated peptides by researchers, however, a very low abundance of neutral loss fragment of mass 204 Da (farnesyl group) and 272 Da (geranylgeranyl group) are formed with very low intensity in the CID fragmentation of these prenylated peptides. In the analysis of farnesylation of proteins, Anderegg *et al*. found a mass shift of 204 in their FAB MS/MS analysis of mating hormone factor. Same type of fragmentation was found by Hoffmann and Kast but the fragment was not that very distinct in the spectrum.

Overall, this fragmentation signature was not very consistent and this method has never been used for the identification of farnesylated peptides/proteins using mass spectrometry. Electrospray ionization and Matrix assisted laser desorption ionization are two most commonly used soft ionization techniques which are widely being used in mass spectrometry-based proteomics studies. But there is several challenges to detect the prenylated peptides in a large pool of non prenylated ones, which are generally obtained by enzymatic digestion of the protein. The ionization efficiency of prenylated peptides in positive ion mode MS typically is lower compared to non prenylated peptides due to the

attachment of a long hydrophobic lipid side-chain group. Thus, mass spectrometric fragmentation studies on prenylated peptides are less studied as compared to other type of posttranslational modifications like phosphorylation and glycosylation. It is also to note that, although few studies can be found for farnesylated peptides but almost no studies in the literature can be found on the fragmentation behavior of geranylgeranyl peptides, a major type of prenylation in proteins. We recently developed a method and published where we modified the prenyl peptides using strong and weak oxidizing agents. The modified peptide improved their hydrophobicity and showed labile nature in CID mass spectrometry. Using these labile characteristics, we distinguished the types of prenylation.

Collision induced dissociation (CID) and electron transfer dissociation (ETD) are two major fragmentation techniques in mass spectrometry. They are being widely used to sequence peptides with modifications. Due to the fragmentation in amide bond, ETD has become popular for sequencing peptides with labile PTMs. Major drawback of ETD is that it requires peptides to be highly charged, therefore peptide with charge state over 3 shows better fragmentation in ETD. Combination of these two fragmentations provides more information and currently widely being used for sequencing peptides.

In this report we are showing the comprehensive fragmentation studies with CID and ETD mass spectrometry of modified and unmodified prenyl peptides. We synthesized several prenyl peptides and studied their fragmentation behavior in gas phase utilizing CID and ETD mass spectrometry. We modified these peptides and in addition to CID, for the first time in the literature we are providing the concurrent ETD mass spectrometry of these peptides along with their modified counter parts. These fragmentation characteristics will significantly improve our understanding of the

fragmentation of different type of prenyl peptides by mass spectrometry and provide confidence in identifying these peptides in large-scale samples.

3.3 Materials and methods

Peptides with sequence REKKFFCAIL was custom synthesized by Genscript (New Jersey, USA) and peptides KHSSGCAFL and DAEFRHDSGYEVC were from AnaSpec, CA. Protein BSA and *Trans trans*-farnesylbromide was purchased from Sigma Aldrich (USA). Geranylgeranyl bromide was synthesized from geranylgeraniol (Sigma Aldrich, USA) and phosphorus tribromide (Sigma Aldrich).[90] All other chemical and reagents m-chloroperoxybenzoic acid (mCPBA), sodium carbonate, hydrogen peroxide (30%v/v), dichloromethane and acetic anhydride (>95%) were obtained from Sigma Aldrich, USA. HPLC grade solvents were also purchased from Sigma Aldrich. All aqueous solutions were prepared using water from a Milli-Q water system. (Aries Filterworks).

3.3.1 Prenylated and oxidized modified peptides syntheses

Chemical synthesis method was utilized for synthesizing prenylated peptides. Cysteinyl peptides with prenyl motifs were reacted with trans trans-farnesyl and geranylgeranyl bromide. Oxidized peptides were generated with 3% H2O2 and epoxidized peptides were obtained using strong oxidizing agent mCPBA. Detail synthesis protocol of prenyl peptides and their oxidized and epoxidized version was reported before.

3.3.2 LC-ESI-IT analysis parameters

The farnesylated/geranylgeranylated peptides as well as the oxidized and epoxidized products were separated in nano LC C-18 column (100 um X 15 µm, C18, 2.1 µm, 100 A) and analyzed by a LC-ESI-IT-TOF (Shimadzu) and LC-ESI-LIT (Thermo

Velos Pro). HPLC was performed with a gradient of about 60 min from 5-90% organic phase. For Shimadzu nano LC- ESI-IT-TOF, the parameters were set as interface voltage 2.6 kV, nebulizing gas (N2, 1.5 L/min), detector voltage 1.7 kV, collision gas (Ar, 35% 98 kPa) and frequency parameter (q = 0.329). For Thermos Velos Pro experiment, peptides were injected into Ultimate 3000 nano UPLC system (Dionex, Sunnyvale Ca, USA) and it is connected to LTQ Thermo Velos Pro. A nano Viper Acclaim PepMap 100 column (C18, 3 um, 100 Å) with a capillary of 15 cm bed length was used for separation by nano LC. A flow rate of 300 nl/min and a mobile phase gradient from 4% B – 90% B, where B is the organic phase having the composition 95: 5: 0.1 acetonitrile: water: formic acid. in 60 min was used for separation. Direct infusion experiment was also done in Thermo Velos Pro to observe the fragmentation behavior for farnesylated/ geranylgeranylated peptides in both CID and

ETD experiments: The method used for the isolation of the top five intense ions with 90s exclusion. It was found that 0.35 collision energy isolation widths of 2 Da and the activation Q to 0.25 is the best for these types of modifications. The activation time for CID was 10 ms and 55 ms was set for ETD. Here, the peptide samples were diluted with (1:1 MeOH: H_2O with 2% acetic acid) and using the syringe pump, the sample is injected in ESI-IT mass spectrometer, where they are ionized and analyzed. For ETD experiments, the reagent ion source emission current, reagent ion electron energy and reagent ion source CI pressure were set to 50 µA, −70 V and 20 psi, respectively. For alternating CID and ETD experiments, the same precursor ion was fragmented successively by CID and ETD.

3.4 Results and discussions

3.4.1 Farnesylated and geranylgeranylated peptides CID fragmentation studies

The syntheses of the farnesylated and geranylgeranylated peptides were mentioned in the material and methods. For fragmentation studies we have taken three peptides having the sequence REKKFFCAIL (N-terminal basic group and having prenyl motif, CAAX at carboxyl end), KHSSGCAFL (N-terminal basic group and CAAX at carboxyl end) and DAEFRHDSGYEVC (N-terminal acidic group and C at the carboxyl end). All three peptides mimic the farnesyl and geranylgeranylation motifs in biological systems. REKKFFCAIL (*m/z* 1254.70) was farnesylated and the mass addition of 204.20 Da was confirmed by mass spectrometry (*m/z* =1476.90). To observe the fragmentation in ESI-IT-TOF under CID with 35 % collision energy, a direct infusion method was used where the farnesylated peptide was injected by a syringe pump and the peptide was dissolved in 48:48:02 of MeOH:AcOH:FA. The molecular ion of doubly charged peak at *m/z* 729.85 $(M+2H+Far)^{2+}$ was found and it was isolated with for further fragmentation under CID. The tandem mass spectrometry experiment performed on *m/z* 729.85 results in the formation of ions at *m/z* 414.25(b_3^+), 542.34(b_4^+), 836.42(b_6^+), 562.44 (b_7^{2+}-H_2O), 720.70(y_{10}^{2+}-H_2O), 1123.57(b_7^+-H_2O), 1214.75(b_8^+), 664.53 (b_9^{2+}) and a peak with loss of farnesyl group appeared at low intensity at *m/z* 1254.65 (Figure 3-1, A) (detailed information's are provided in the supplementary excel file) . The geranylgeranylated peptide, REKKFFC(ger)AIL undergoes fragmentation to yield same fragment ions at *m/z* 542.34(b_4^+),664.53 (b_9^{2+}), 836.42(b_6^+),1113.46(y_7^+), and a peak with a loss of geranylgeranylated group at *m/z* 1254.65 (Figure 3-1, B) But the peaks with mass loss of farnesyl group (204.20 Da) and geranylgeranyl group (272.20 Da) were not so significant in the spectra. For the peptide, KHSSGC(far)AFL on tandem mass fragmentation yielded various b and y fragments due to backbone cleavage of the peptide at *m/z* 279.09(y_2^+),

571.76 (b_8^{2+}), 818.35 ($y_6^{+}+H_2O$), 875.43(b_7^{+}) and 1022.60 (b_8^{+}).Here, in the CID fragmentation of KHSSGC(far)AFL, the two peaks obtained at 475.17$(M+2H-Far)^{2+}$ and at 949.43$(M+H-Far)^{2+}$ were significant (Figure 3-1, C). To confirm the loss of farnesyl/geranylgeranyl group from the precursor ions, the CID fragmentation of KHSSGC(ger)AFL was studied and the pattern was found to be similar to that (Figure 3-1, D). It is clear that loss of farnesyl or geranylgeranyl was not very consistent . To confirm that, N-terminal acidic amino acid peptide, DAEFRHDSGYVEC(far/ger) was synthesized and CID fragmentation behavior was studied (Figure3-1, C). This peptide fragmented very quickly at low collision energy of about 35%. According to some of our previous studies, the same fragmentation behavior was observed for all the farnesylated peptides in low energy CID fragmentation in MALDI-QIT-TOF mass spectrometer.

It is clear from our studies with these three peptides that CID fragmentation behavior of these three peptides with different prenyl motifs is different. There is no clear evidence of consistent loss of lipid side-chain from these peptides. Identifying these peptides with a prenyl marker ion is not feasible. It is also worth to mention that, there is no fragmentation study is available from geranylgeranylated peptides and this is the first example of fragmentation behavior of these peptides.

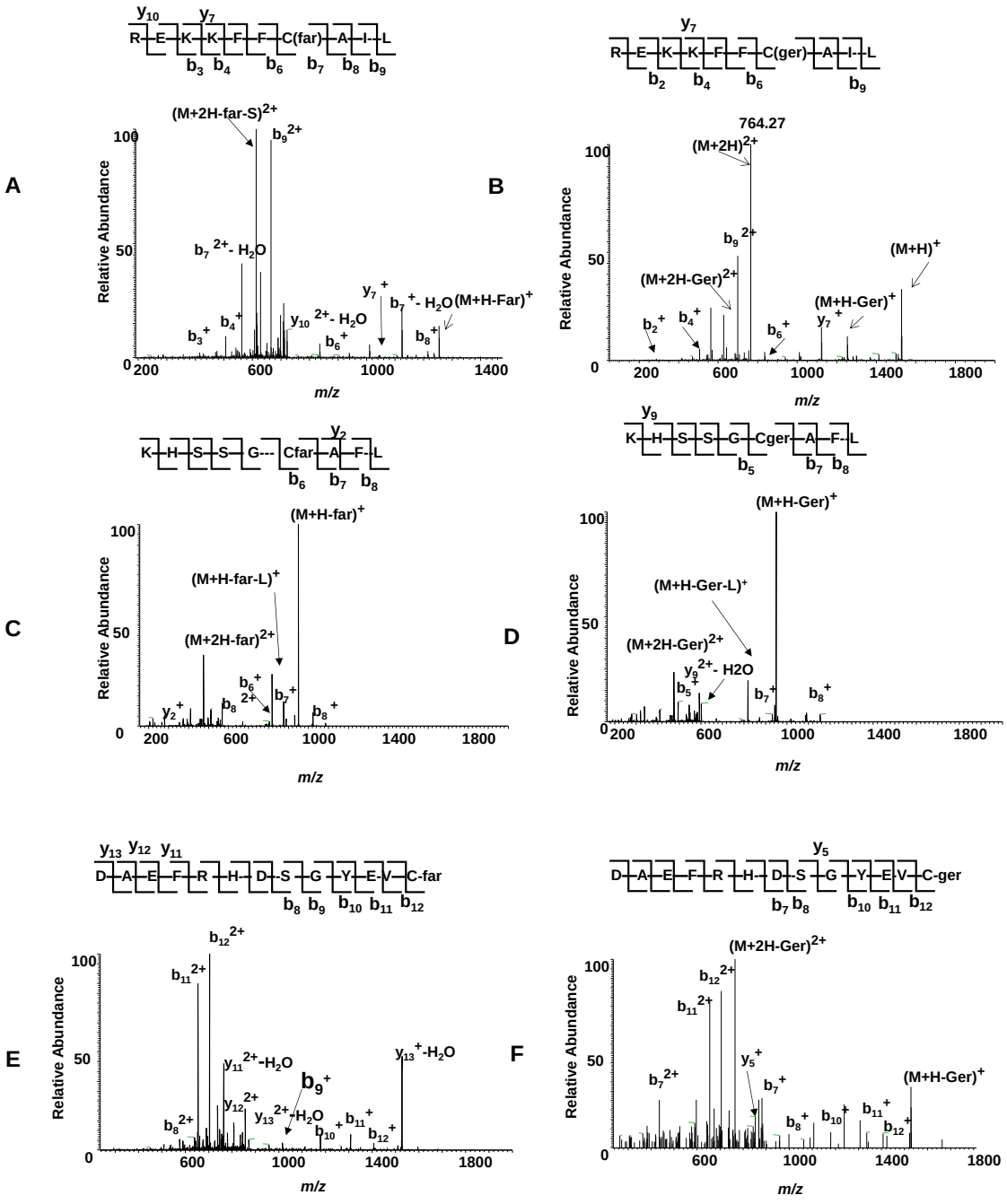


Figure 3-1 CID-MS/MS Fragmentation of farnesylated peptides (left panel) and geranylgeranylated peptides (right panel).

3.4.2 Farnesylated and geranylgeranylated peptides ETD fragmentation studies

The masses of synthesized farnesylated and geranylgeranylated peptides were analyzed in ESI (Thermo Velos Pro Dual Pressure Linear Ion trap) in direct injection mode and prenyl peptide masses were confirmed. Fragmentations of highly charged species of prenyl peptides were done using ETD tandem mass spectrometry. The triply charged farnesylated peptide REKKFFC(far)AIL of *m/z* 487.23 $(M+3H)^{3+}$ was fragmented in ETD mode and fragments of *m/z* 303.27 (c_2), 431.32 (c_3), doubly charged 673.15 (c_9), 853.33(c_6), 1029.69(z_7) and 1157.75 (z_9)were observed. We also found the doubly charged and singly charged molecular peaks at 731.09 and 1458.88 respectively (Figure 3-2, A**)**. In the fragmentation of farnesylated peptide, the loss of farnesyl group is not observed. Similar ETD fragmentation pattern was observed in case of triply charged geranylgeranylated peptide REKKFFC(ger)AIL of *m/z* 508.90. In the case of geranylgeranylated fragmentation, corresponding c and z ions were observed with no loss of geranylgeranyl group (Figure 3-2, B). For the N-terminal K peptide, triply charged KHSSGC(far)AFL of *m/z* 385.48 on ETD fragmentation yielded ions at *m/z* 283.17(c_2) , 458.76 (c_3) and 872.43 (z_7) and 1205.44(z_8) respectively with the formation of doubly charged molecular ion peak (Figure 3-2, C). The fragmentation of triple charged KHSSGC(ger)AFL at *m/z* 413.31 was also studied and it was found that several fragments at *m/z* 283.17(c_2) , 458.76 (c_3), doubly charged 554.2 (c_8), 889.28 (c_6) and 1077.25 (z_8) along with doubly and singly charged molecular ions are formed (Figure 3-2, D). In case of N-terminal acidic amino acid peptide, DAEFRHDSGYEVC(far) was synthesized and CID fragmentation study is shown (Figure 3-1, E and F). It was found that, this peptide is highly labile and broken down into fragment ions even at very low collision energy in CID mode. We saw complete fragmentation of this farnesylated peptide in ETD with the formation of various c and z ions, but no loss of farnesyl group

was seen (Figure 3-2, E). The fragmentation pattern was also studied for DAEFRHDSGYEVC(ger) and it was found that many z fragment ions were formed along with a small peak where loss of geranylgeranyl group of about 272 Da was observed.(Figure 3-2, F).

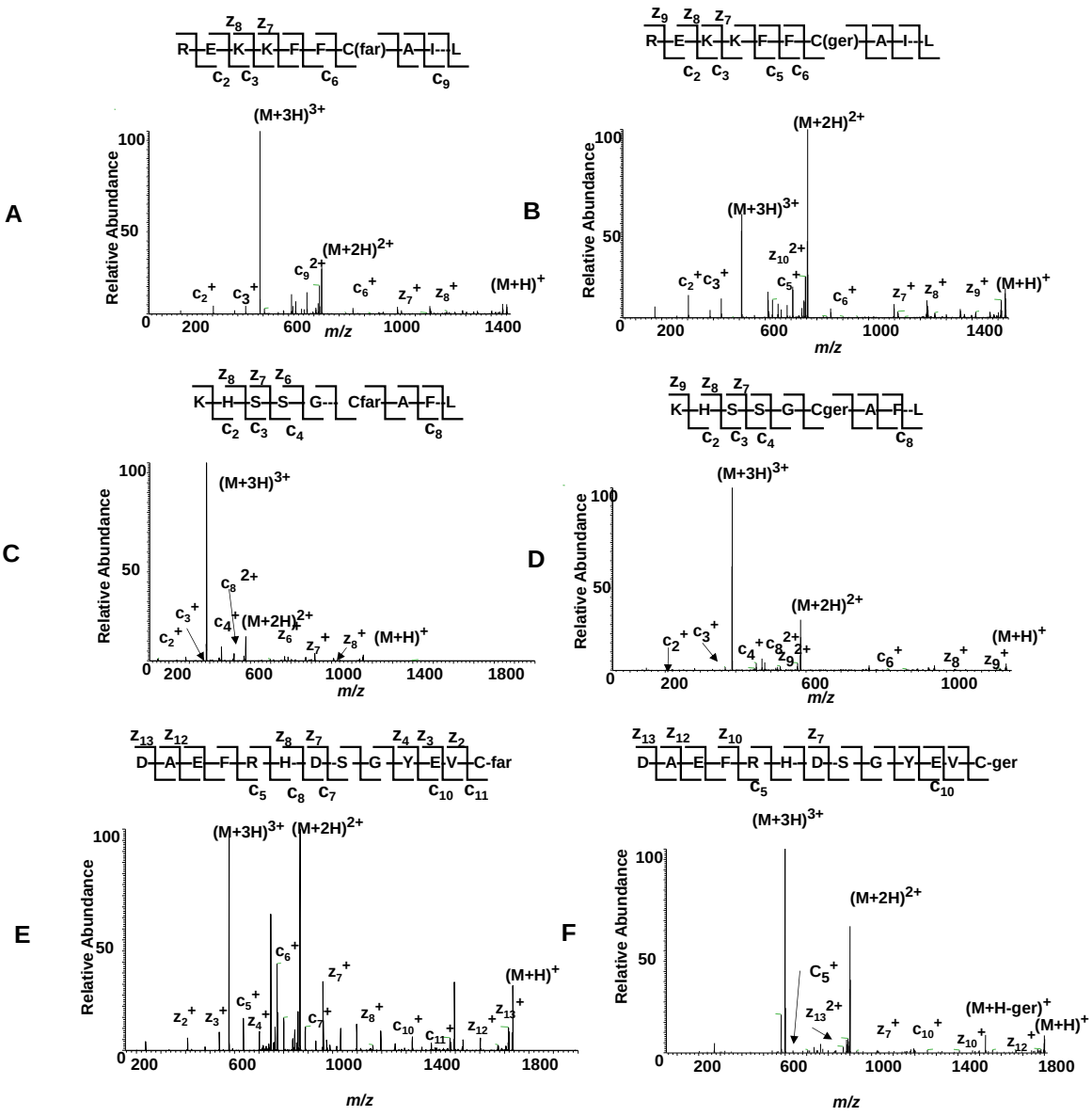


Figure 3-2 ETD-MS/MS fragmentation of farnesylated peptides (left panel) and geranylgeranylated peptides (right panel).

3.4.3. Oxidized farnesylated and geranylgeranylated peptides CID fragmentation studies

We recently published a paper where we oxidized and epoxidized the prenyl peptides to make prenyl side-chain labile and hydrophilic. Mono-oxidation of sulfur of cysteine made the prenyl peptides labile and epoxidation helped them to enrich from the complex samples. Mono-oxidation of sulfur helped the peptide selectively break in the prenyl side chain in the gas phase (loss of RSOH, where R = prenyl side chain). In this report we are also providing the fragmentation behavior of these modified peptides in CID and ETD in order to identify them in LC-MS/MS studies. As we mentioned before that prenyl peptides are hydrophobic and these modification will improve their hydrophilicity to efficiently retain and elute them from C18 column during LC-MS/MS based experiments.

Fragmentation studies on mono-oxidized thio-ether containing peptides were performed and it showed the loss of specific fragment (sulfenic group, RSOH) from the peptides in tandem mass spectrometry. This idea about the oxidation and fragmentation of the oxidized prenylated peptides lead us to make a mass spectrometry fragmentation strategy for identification and differentiation of prenylated peptide from the signature mass loss from the precursor peptides and the strategy was published recently (REF again). To demonstrate the CID and ETD mass spectrometry behavior together here we showed the fragmentation behavior of these three peptides in different instrumental platforms with different ionization and tandem mass spectrometric techniques. Mono-oxidized farnesylated and geranylgeranylated peptides were prepared with 3% hydrogen peroxide and the details are in the experimental section. On oxidation of the peptide, the thio-ether bond was selectively oxidized to sulfoxide group. The CID fragmentation of the oxidized farnesylated REKKFFCAIL peptide was done in ESI-IT-TOF-MS.(Figure 3-3, A) Since, proteolytic peptides were separated using C-18 reverse phase columns, the farnesylated peptides were analyzed on the same material. The modified farnesylated

peptides were eluted with 50-60% acetonitrile in organic phase and were still observed by the mass spectrometer ESI-IT-TOF. The retention time was 53.63 min in 70 min long run gradient from 0-80 % organic phase compared to 66 min for the farnesylated one. In ESI-IT-TOF, a doubly charged ion of REKKFFC(oxyfar)AIL was found at *m/z* 738.34.(Figure 3-3, A) On CID fragmentation with the collision energy at 35 %, the two doubly charged fragment ions at comparable high intensity were formed at *m/z* 610.86 $(M+2H-RSOH)^{2+}$ and *m/z* 763.01$(M+2H-H_2O)^{2+}$ along with the formation of single charged peaks at *m/z* 706.43 (y_4), 836.51(b_6),1113.68 (y_7-NH_3) and a peak with loss of RSOH group is obtained. The fragment at *m/z* 610.86 was obtained with the loss of RSOH group where R= farnesyl group. Same fragments at high intensity were observed during the CID fragmentation of the REKKFFC(oxyger)AFL (Figure 3-3, B) The doubly charged highest intensity peak is the signature peak with the expected mass loss of RSOH group where R is the farnesyl/geranylgeranyl group. The retention time for this product is about 53.60 min, similar to oxyfarnesylated one, whereas the non oxygeranylgeranylated peaks were obtained between 65-70 min in the same gradient run. The fragmentation study of DAEFRHDSGYEVC (oxyfar) was also done using CID in ESI-MS. It was found that doubly charged 589.75(b_{10}), 639.18(b_{11}) and 703.8 (b_{12}) and 874.31(y_{13}) along with their single charged ions were formed at very low intensity as compared to the highest doubly charged peak at 747.84 with the loss of RSOH group (R=far) **(**Figure 3-3, C**)** . It is confirmed from the MS/MS spectrum obtained from CID, the formation of signature fragment from each of the oxyfarnesylated and oxygeranylgeranylated peptides. (Figure 3-3, D)

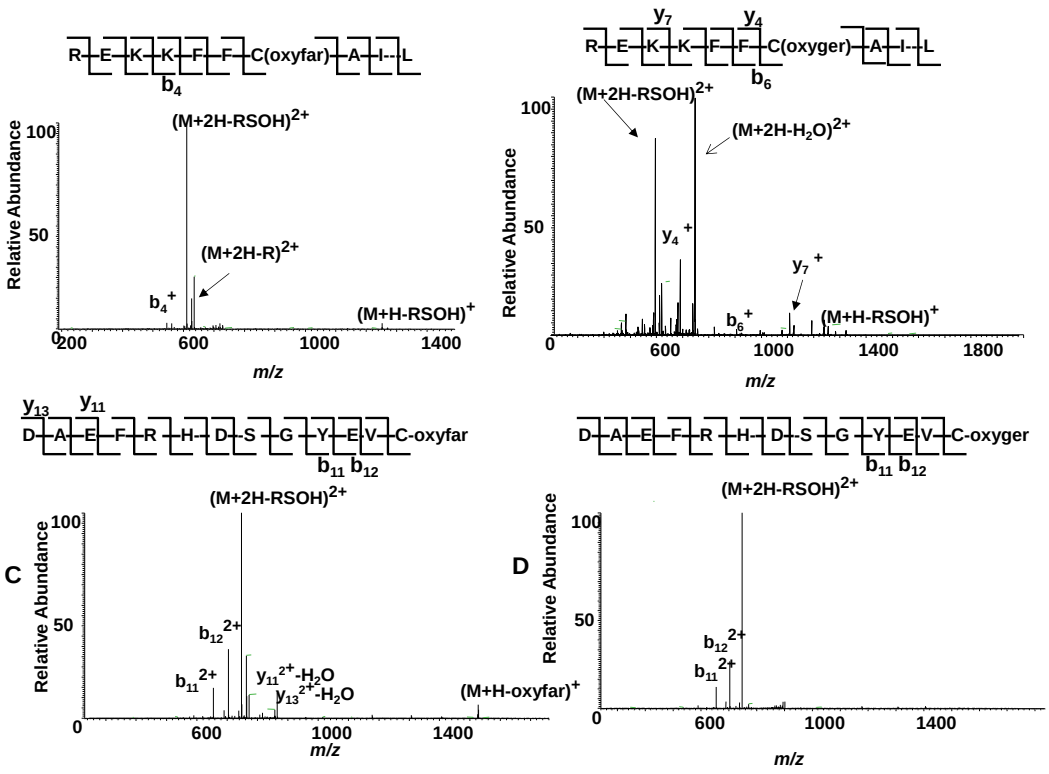


Figure 3-3 CID-MS/MS fragmentation of oxyfarnesylated peptides (left panel) and oxygeranylgeranylated peptides (right panel).

The same expected peak with the same loss was also formed at very high intensity as compared to other peaks in case of CID fragmentation of KHSSGC(oxyfar)AFL and KHSSGC(oxyger)AFL also and that was shown in Figure 3-4. Here, the doubly charged signature peak is observed at *m/z* 458.23 with the loss of RSOH group from *m/z* 585.32 and 619.35 respectively.

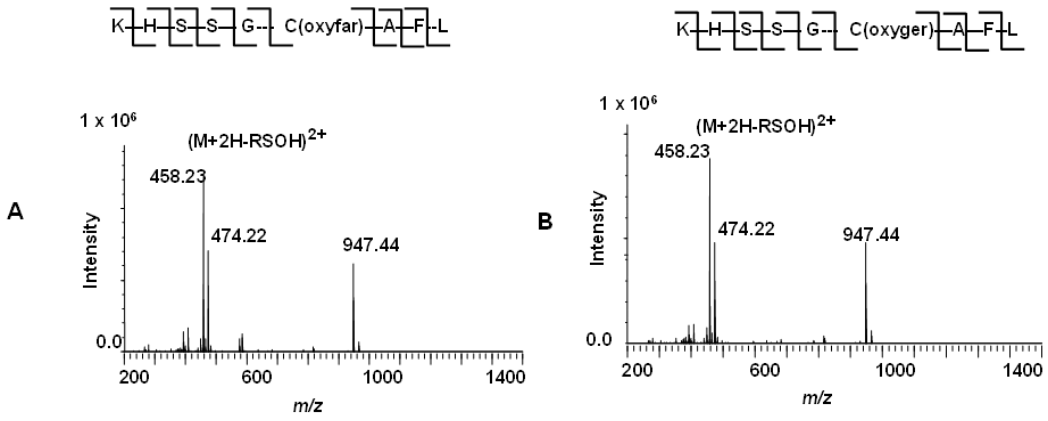


Figure 3-4 CID-MS/MS fragmentation of oxyfarnesylated peptide KHSSGCAFL (left panel) and oxygeranylgeranylated peptide KHSSGCAFL (right panel).

3.4.4. Oxidized farnesylated and geranylgeranylated peptides ETD fragmentation studies

The fragmentation of this oxidized prenylated products were done in ETD method to see the loss of distinct mass loss of RSOH group, where R= farnesyl/geranylgeranyl group.

ETD fragmentation of triply charged oxyfarnesylated REKKFFC(oxyfar)AIL peptide of *m/z* 492.56 $(M+3H+far+O)^{3+}$ was studied in thermos Velos ESI-Iontrap using direct infusion method (Figure 3-5, A). Unlike the fragmentation of these oxidized farnesylated products in CID, the loss of RSOH from the precursor ions were not observed, instead various c and z ions with doubly charged and precursor ion peak were formed. The ions at *m/z* 303.18(c_2), 431.37(c_3), 853.53(c_6) and 1045.62 (z_7) and 1173.68(z_8) were shown in Figure 3-5, A. We published recently that loss of RSOH group during the low-energy CID fragmentation was the key for identification and differentiation between the type and site of prenylation of peptides. In ETD spectra, the backbone cleavage was significant and various c and z ions were formed. The similar fragmentation

was observed for triply charged REKKFFC(oxyger)AIL peptide of *m/z* 515.23 $(M+3H+ger+O)^{3+}$ (Figure 3-5, B).

ETD fragmentation of triply charged DAEFRHDSGYEVC(oxyfar) at *m/z* 578.00 resulted in the cleavage of backbone of peptide with no formation of the signature fragment and formed various c and z ions like *m/z* 480.26(c_4), 636.31(c_5), 773.37(c_6), 888.43 (c_7) and doubly charged 701.43 (z_{11}) and 1097.52 (z_7), 1254.60(z_8), 1195.52(c_{10}) and 1601.78(z_{12}). (Figure 3-5, C) Similar type of c and z ions were observed for DAEFRHDSGYEVC(oxyger) of *m/z* 600.27 on ETD fragmentation. (Figure 3-5, D).

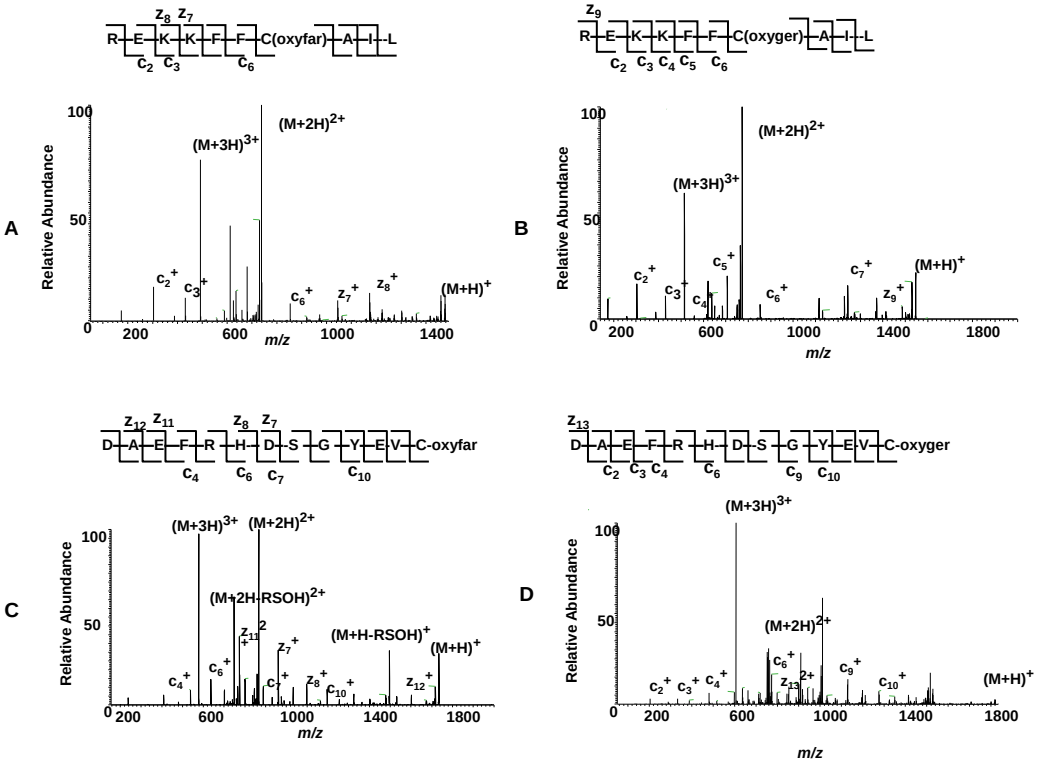


Figure 3-5 ETD-MS/MS fragmentation of oxyfarnesylated peptides (left panel) and oxygeranylgeranylated peptides (right panel).

For the oxyfarnesylated and oxygeranylgeranylated KHSSGCAFL of *m/z* 390.26 and 408.15 respectively, the ETD fragmentation studies are also done and it was found that the neutralization of triply charged precursor ions to its double and single charged ions are more prevalent than the backbone fragmentation peaks at *m/z* 283.25 (c_3), 370.24 (c_4) and 698.42 (z_5). (Figure 3-6). It is clear from these CID and ETD-MS/MS studies of the modified peptides that this combination will un-ambiguously detect the prenylation sites and types in the peptides.

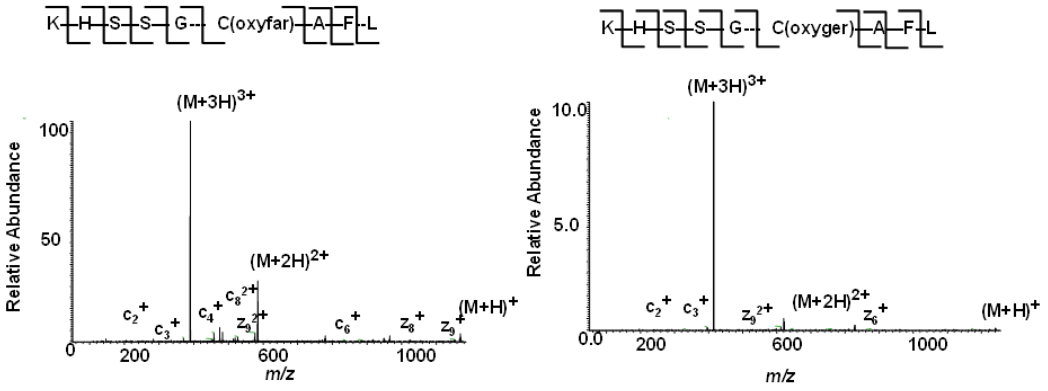


Figure 3-6 ETD-MS/MS fragmentation of oxyfarnesylated peptide KHSSGCAFL and oxygeranylgeranylated peptide KHSSGCAFL.

3.4.5. Epoxidized farnesylated and geranylgeranylated peptides CID fragmentation studies

After we studied the CID and ETD fragmentation studies of the mono-oxidized prenyl peptide in CID and ETD we extended our studies on the epoxidized prenyl peptides. Detail of epoxidized prenyl peptides behavior was studied in CID and published recently (REF) with MALDI ionization. The goal of epoxidation is to incorporate

enrichment functionality as well as making them more hydrophilic hence LC-ESI-MS studies will be efficiently performed. Here in the following section we demonstrate the CID as well as ETD behavior for the first time of these epoxy prenyl peptides.

Epoxidation reaction of prenyl peptides converts isoprenoid groups to epoxides along with the oxidation of thio-ether bond (S=O) in the prenylated peptides. The number of oxygen incorporated is directly related to the reaction time. This resulted in increase of hydrophilicity of the epoxidized prenylated peptides and also made it easier to design an enrichment strategy using the reactivity of epoxy group for these low abundance modification. Incorporation of epoxy group in the prenyl chain depends upon the number of unsaturated bonds. It was found that several epoxidized products were obtained for both the farnesylated and geranylgeranylated peptides due to the number of epoxy group formation. These variable epoxidized products are found very useful as targets to generate the signature mass fragment in tandem mass spectrometry as all undergoes similar fragmentation with the loss of RSOH group where R= farnesyl/geranylgeranyl group + nO (n is the number of oxygen atoms in the prenyl chain). Each epoxidized product acts as precursor for the generation of marker ions at low energy fragmentation in a mass spectrometer. Loss of signature fragment from the epoxidized peptides were demonstrated by our group before, here we demonstrated the ESI-CID-MS/MS of different prenylated peptide with specific prenyl motifs.

The epoxidized farnesylated peptides KHSSGC(epoxyfar)AFL were eluted by C-18 column and the retention time was in between 42-47 min in 70 min long run gradient from 0-80 % organic phase in nanoLC-ESI-IT-TOF-MS. The doubly charged epoxidized peaks were obtained with *m/z* 593.32 $(M+2H+Far+2O)^{2+}$, 602.32 $(M+2H+Far+3O)^{2+}$ and 610.32 $(M+2H+Far+4O)^{2+}$ and the CID fragmentations of collision energy 35% of all the peaks resulted in the formation of doubly charged marker ion at *m/z* 458.23 (loss of

RSOH). Doubly charged epoxidized products of KHSSGC(epoxyfar)AFL at *m/z* 593.32 $(M+2H+far+2O)^{2+}$ undergoes a loss of expected distinct mass same as the mass of RSOH group (where R= farnesyl + nO, n= number of epoxy group) and formed a signature marker ion at *m/z* 458.23 $(M+2H-RSOH)^{2+}$(Figure 3-7, A). Similar fragmentation was observed for higher epoxidized farnesylated KHSSGC(epoxyfar)AFL peptides of *m/z* 602.32 $(M+2H+Far+3O)^{2+}$ and 610.32 $(M+2H+Far+4O)^{2+}$. (Figure 3-7, B and C)

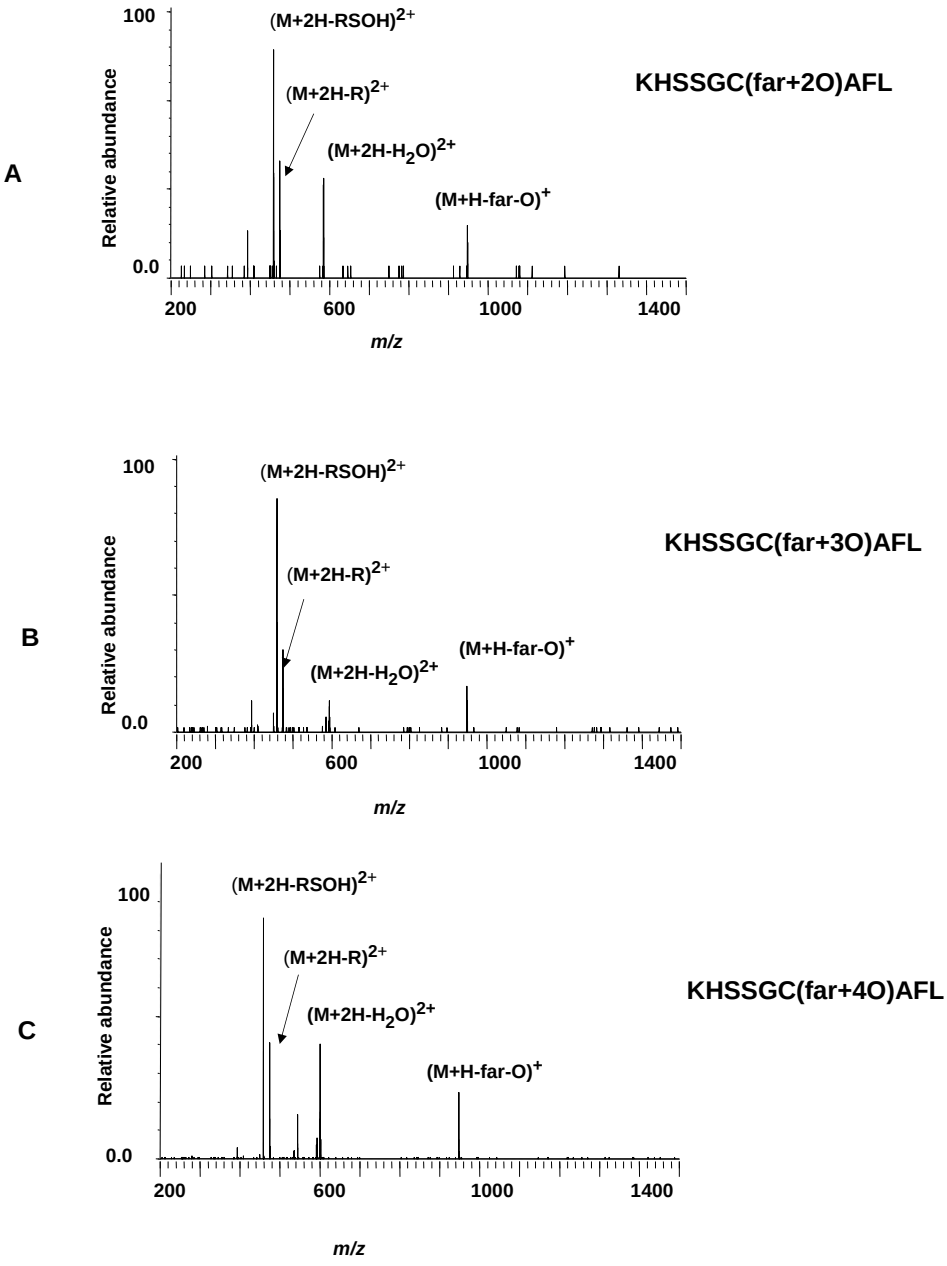


Figure 3-7 CID-MS/MS fragmentation of KHSSGC(epoxyfar)AFL (A) *m/z* 593.32 $(M+2H+Far+2O)^{2+}$ (B) 602.32 $(M+2H+Far+3O)^{2+}$ and (C) 610.32 $(M+2H+Far+4O)^{2+}$ peptides with ESI-IT-TOF-MS.

The doubly charged epoxidized geranylgeranylated peptides DAEFRHDSGYVEC with m/z 916.30 $(M+2H+Ger+2O)^{2+}$, 924.29 $(M+2H+Ger+3O)^{2+}$ and 932.28 $(M+2H+Ger+4O)^{2+}$ are observed in the mass spectrum. The CID fragmentation of the peaks were also done at 35% collision energy and it was found that in case of m/z 916.30 $(M+2H+Ger+2O)^{2+}$, a very high intensity peak is obtained at m/z 747.48. This is the doubly charged peak with the predicted mass loss of RSOH group where R= geranylgeranyl + O (Figure 3-8, A). Similar fragmentations were observed in case of other epoxy geranylgeranylated peaks at m/z 924.29 $(M+2H+Ger+3O)^{2+}$ and 932.28 $(M+2H+Ger+4O)^{2+}$ (Figure 3-8, B and C).

The epoxidation of the prenylated peptide, REKKFFC(far)AIL (monoisotopic mass= 1458.70) was done with epoxydising agent and several epoxidized products at m/z 1490.67 (M+H+far+2O), 1506.71(M+H+far+3O) and 1522.72 (M+H+far+4O) were formed, which we analyzed and confirmed by mass spectrometry. These epoxidized products formations were confirmed by mass analysis in ESI-IT-TOF and the double and triple charged species are found in the mass spectrum for the REKKFFC(epoxyfar)AIL peptide . Doubly charged epoxidized products of REKKFFC(epoxyfar)AIL at m/z 746.34 (M+2H+far+2O)2+ undergoes a loss of expected distinct mass same as the mass of RSOH group (where R= farnesyl + nO, n= number of epoxy group) and formed a signature marker ion at m/z 610.95 $(M+2H-RSOH)^{2+}$ (Figure 3-9, A). The other doubly charged products of REKKFFC(epoxyfar)AIL at m/z 754.33 $(M+2H+far+3O)^{2+}$ and 762.21$(M+2H+far+4O)^{2+}$ were also fragmented under CID condition in a similar pattern and in every spectrum, the loss of RSOH group was found as a signature fragment to identify the prenylation type and their sites (Figure 3-9, B, C).

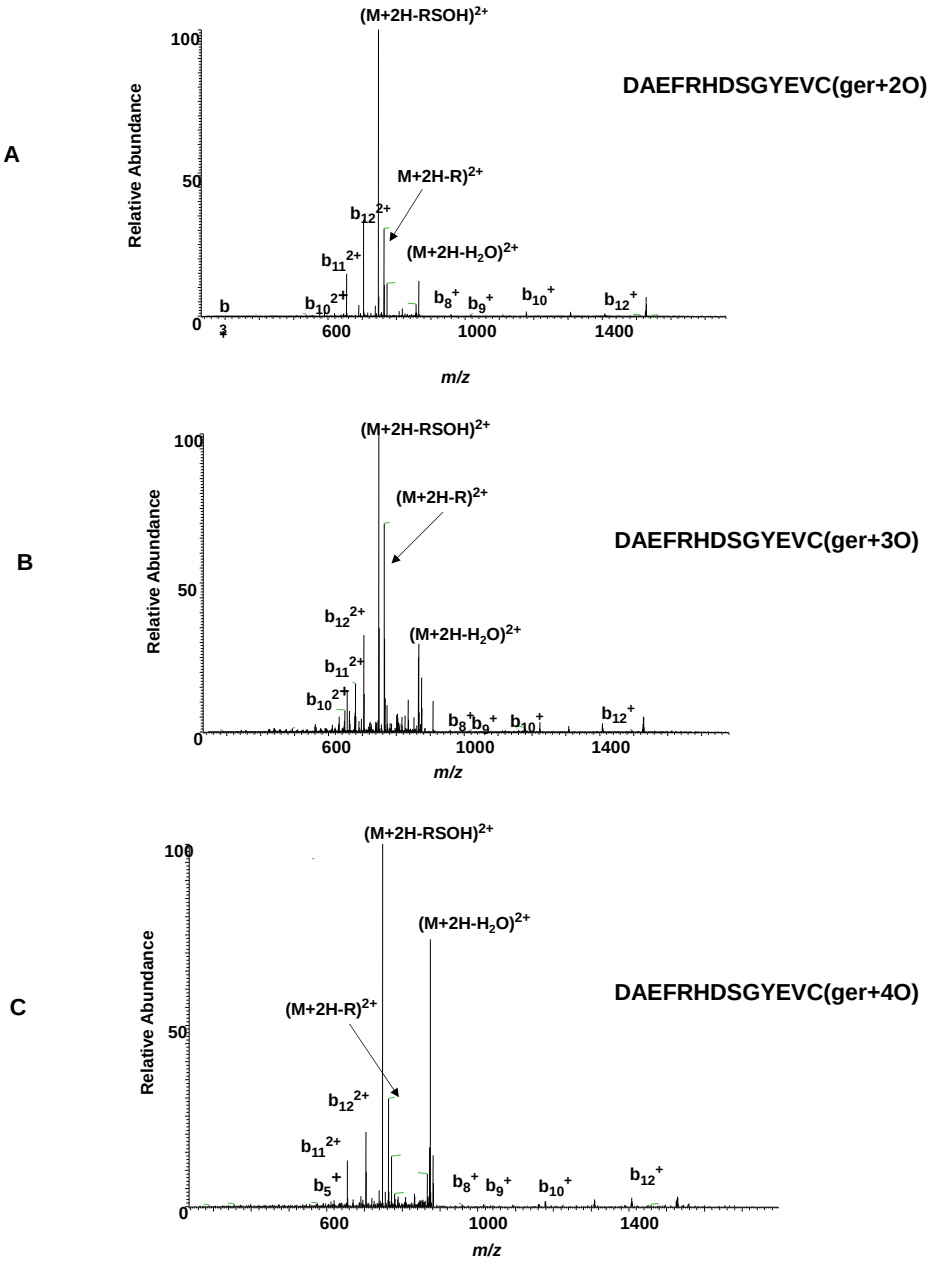


Figure 3-8. CID-MS/MS fragmentation of DAEFRHDSGYVEC(epoxyger) with ESI-IT--MS. (A) *m/z* 916.30 $(M+2H+Ger+2O)^{2+}$,(B) 924.29 $(M+2H+Ger+3O)^{2+}$ and (C) 932.28 $(M+2H+Ger+4O)^{2+}$.

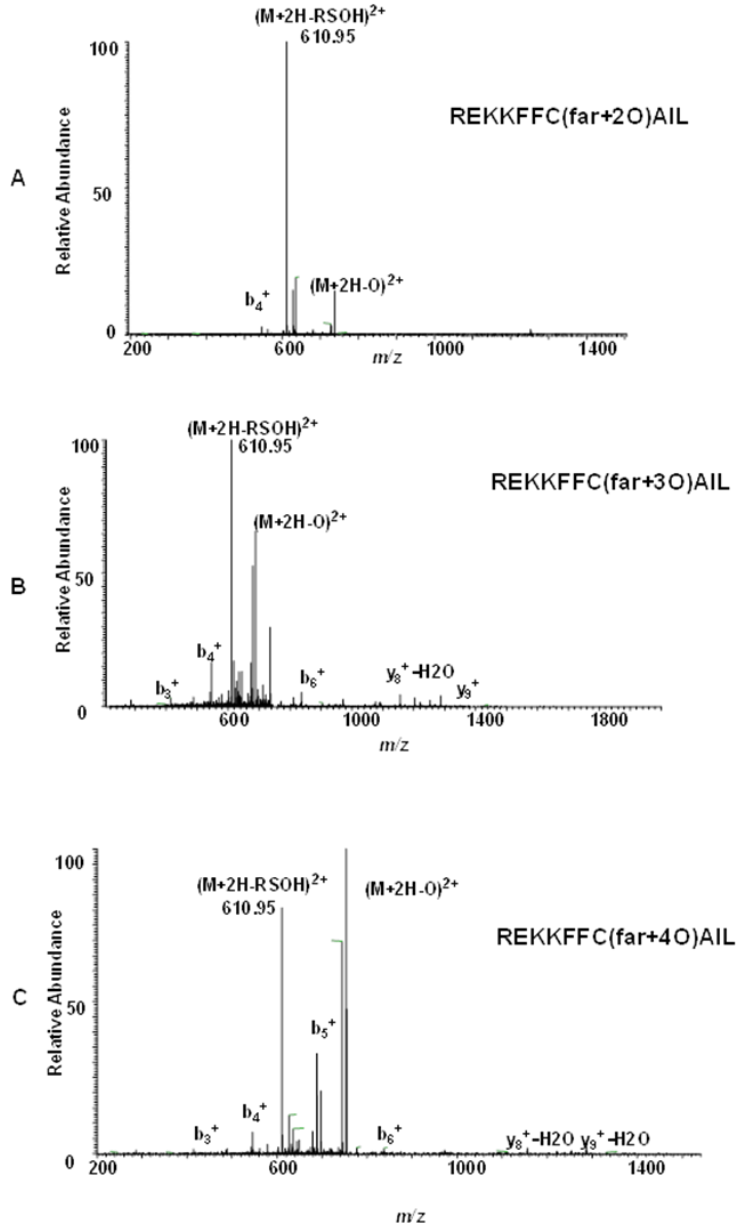


Figure 3-9. CID-MS/MS fragmentation of REKKFFC(epoxyfar)AIL with ESI-IT-TOF-MS (A) *m/z* 746.34 $(M+2H+Far+2O)^{2+}$, (B) 754.33$(M+2H+Far+3O)^{2+}$ and (C) 762.21$(M+2H+Far+4O)^{2+}$.

3.4.6. Epoxidized farnesylated and geranylgeranylated peptides ETD fragmentation studies

To determine the ETD fragmentation pattern of epoxyfarnesylated peptide, epoxyfarnesylated REKKFFCAIL was chosen and the fragmentation of the triply charged epoxy products were studied with 55 ms acquisition time for ETD. The fragments obtained from m/z 497.87 $(M+3H+Far+2O)^{3+}$ showed fragment ions at m/z 303.17(c2), 431.34(c3), 507.34(c6) along with the doubly charged and singly charged precursor ion peaks. Similar to the ETD fragmentation pattern for the oxyfarnesylated peptides, the peaks were observed with backbone cleavage but no formation of signature ion peak with the distinct mass loss. To evaluate the fragmentation behavior in ETD, we studied the fragmentation pattern of the three epoxyfarnesylated products of peptide KHSSGCAFL. They were isolated in the MS spectrum at m/z 395.51 $(M+3H+Far+2O)^{3+}$ (Figure 3-10, A), , 400.59 $(M+3H+Far+3O)^{3+}$ (Figure 3-10, B), and 406.17 $(M+3H+Far+4O)^{3+}$ (Figure 3-10, C), and the ETD fragmentation of these three epoxidized products were obtained. It was found that several c and z ions (at m/z's 283.19(c2), 370.22(c3), 457.25(c4), 514.27(c5) along with c6, z8 and z9 ions, see the excel file at supplementary data) were formed in each case along with the neutralization of triply charged molecular ion species (Figure 3-10). However, we did not found any loss of signature mass in most of the epoxidized farnesylated peptide fragmentations, which we always found in case of CID fragmentation.

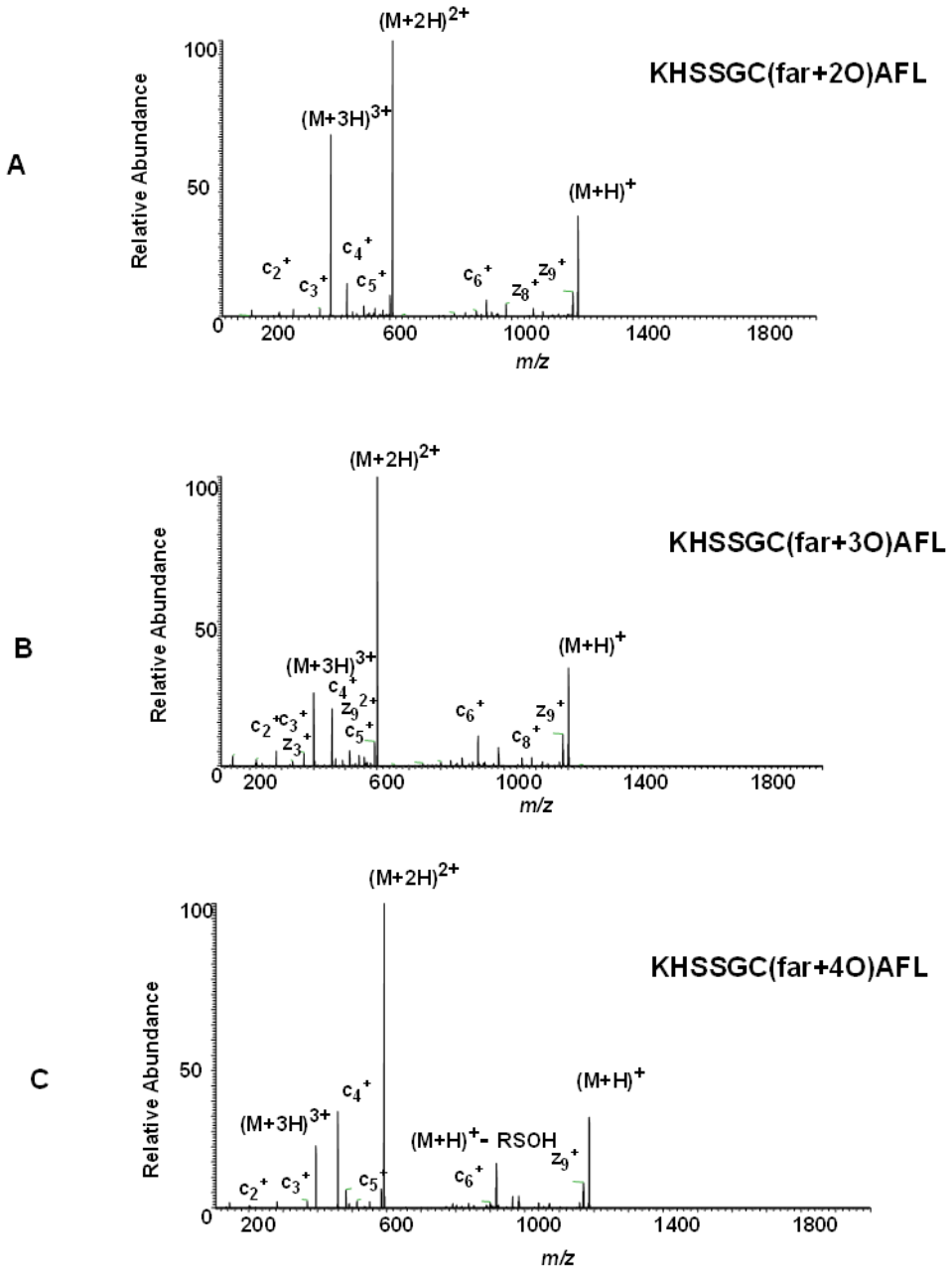


Figure 3-10. ETD-MS/MS fragmentation of KHSSGC(epoxyfar)AFL with ESI-IT-MS. (A) m/z 395.51 $(M+3H+Far+2O)^{3+}$,(B) 400.59 $(M+3H+Far+3O)^{2+}$ and (C) 406.17 $(M+3H+Far+4O)^{2+}$.

ETD fragmentation studies on epoxyfarnesylated and epoxygeranylgeranylated DAEFRHDSGYVEC was studied in detail. It was observed that the ETD fragmentation of epoxygeranylgeranylated products of m/z 606.83 $(M+3H+ger+2O)^{3+}$(Figure 3-11, A), 611.71 $(M+3H+ger+3O)^{3+}$ (Figure 3-11, B), and 617.73 $(M+3H+ger+4O)^{3+}$ (Figure 3-11, C), resulted in the formation of mostly the c fragments along with the formation of doubly charged and molecular ion peak.

In case of fragmentation of epoxyfarnesylated products at m/z 587.43 $(M+3H+far+2O)^{3+}$, 593.51 $(M+3H+ger+3O)^{3+}$ and 599.43 $(M+3H+ger+4O)^{3+}$, showed both the c and z ions but of much lesser intensity as compared to that of epoxy-geranylgeranylated ones. The ETD fragmentation studies were also done for other triply charged epoxyfarnesylated peptides of REKKFFCAIL [at *m/z's* 503.22 $(M+3H+Far+3O)^{3+}$ and 508.47 $(M+3H+Far+4O)^{3+}$] and it showed the similar fragmentation pattern that we observed for other epoxyfarnesylated and epoxygeranylgeranylated peptides (Figure 3-12). The ETD fragmentation of epoxy farnesylated/geranylgeranylated peptides showed efficient backbone cleavage in different epoxy prenylated products whereas CID showed efficient gas-phase cleavage on the mono-oxidized thio-ether bonds on the prenyl side-chains.

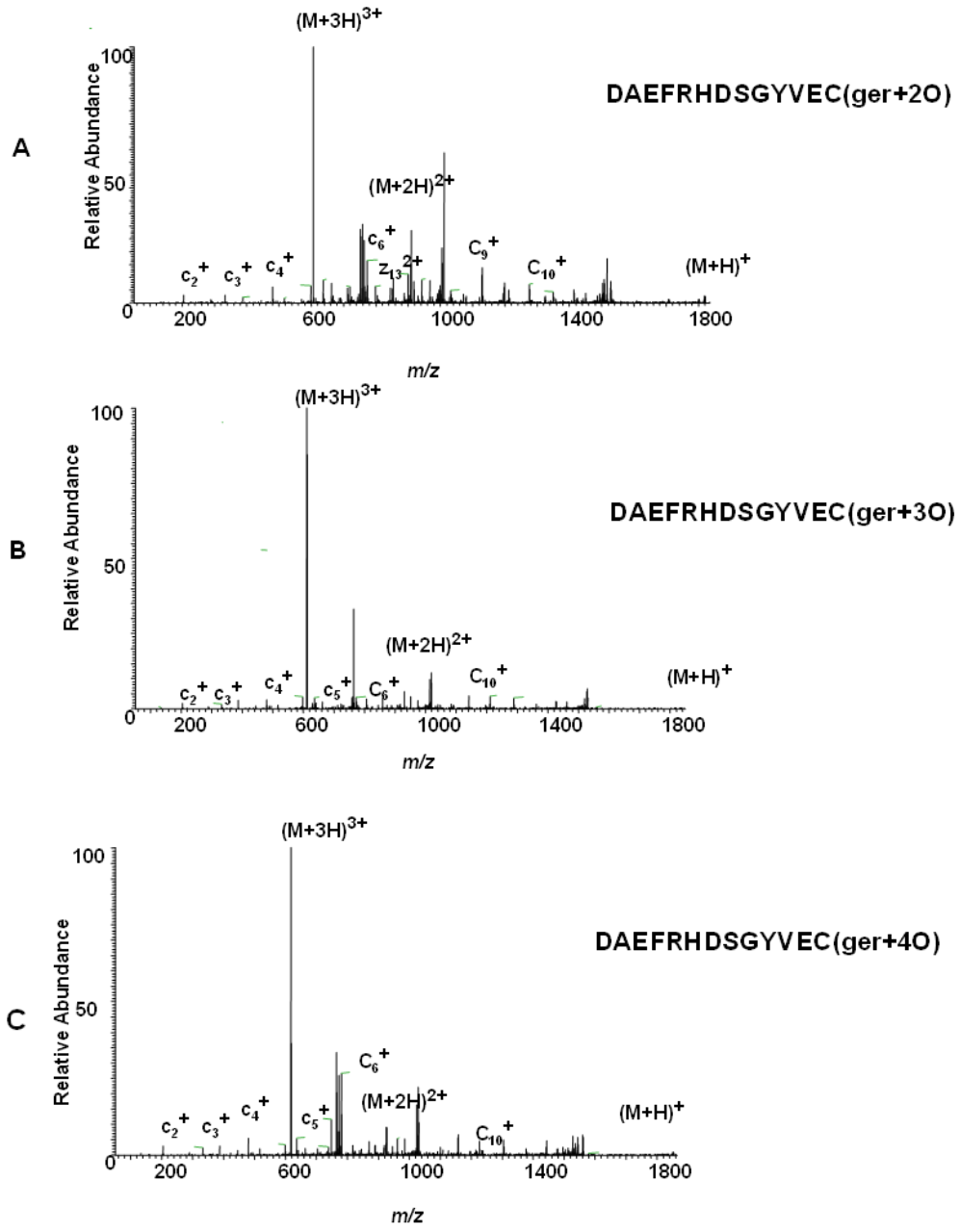


Figure 3-11. ETD-MS/MS Fragmentation of DAEFRHDSGYVEC(epoxyger) with ESI-IT-MS.(A) *m/z* 606.83 $(M+3H+Ger+2O)^{3+}$, (B) 611.71 $(M+3H+Ger+3O)^{3+}$ and (C) 617.73 $(M+3H+Ger+4O)^{3+}$.

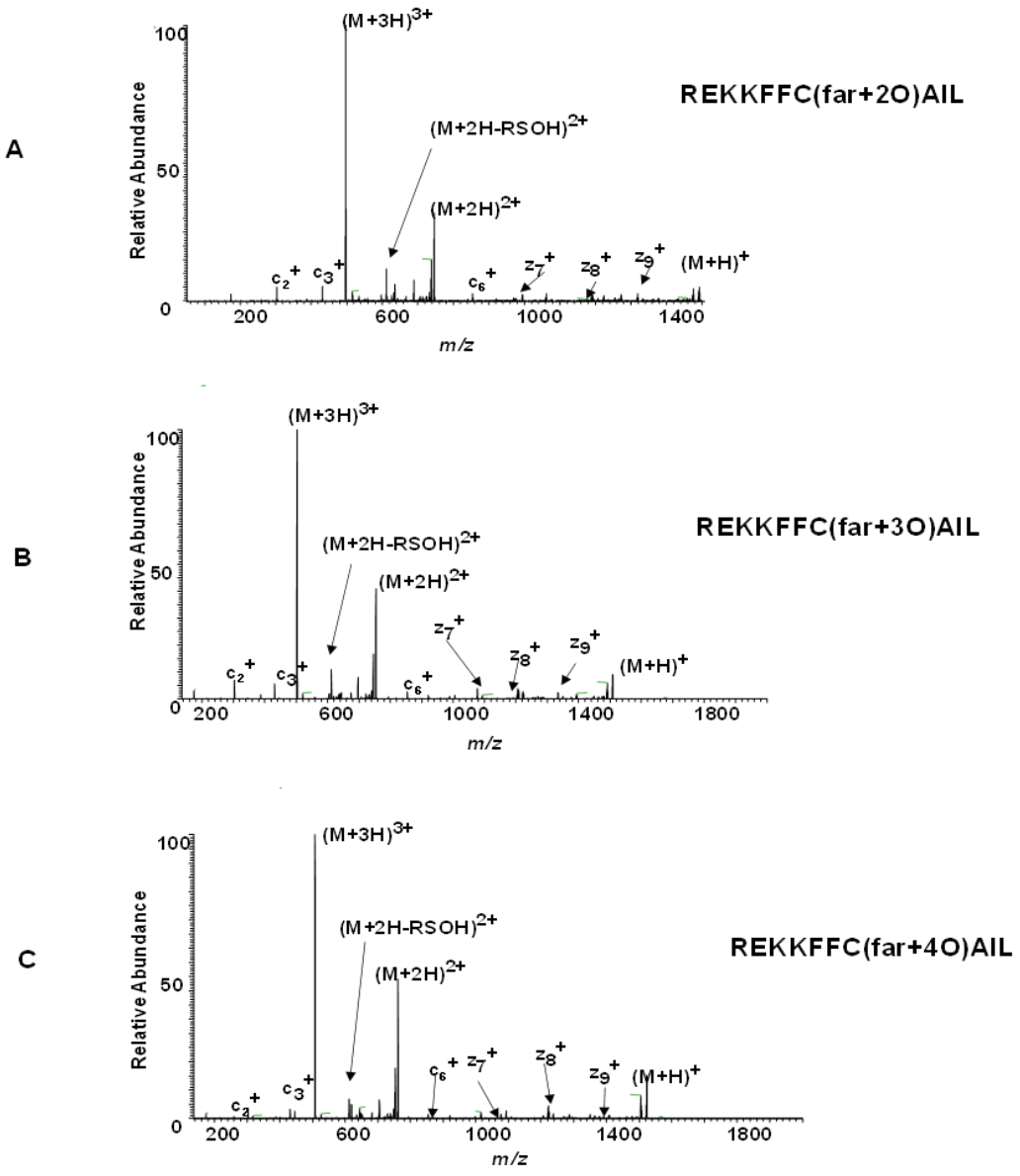


Figure 3-12. ETD-MS/MS fragmentation of REKKFFC(epoxyfar)AIL with ESI-IT-MS (A) ETD-MS/MS of m/z 497.87 $(M+3H+Far+2O)^{3+}$ (B) ETD-MS/MS of 503.22 $(M+3H+Far+3O)^{3+}$ and (C) at m/z 508.47 $(M+3H+Far+4O)^{3+}$.

3.5 Conclusions

We have demonstrated the fragmentation behavior of farnesylated and geranylgeranylated peptides and their oxidized products in both low energy CID and ETD fragmentation. We have selected three cysteinyl peptides with prenyl motifs and corresponding farnesylated and geranylgeranylated peptides were synthesized. Prenyl peptides are hydrophobic so their corresponding oxy-and epoxy products were also prepared. These modifications changed the un-labile character of prenyl side chain in peptides into labile nature in the gas phase tandem mass spectrometry experiments. We demonstrated the comprehensive CID and ETD fragmentation of these peptides in this study. Oxidized peptides generally release a signature ion due to a sulfoxide group formation in the prenyl side chain. ETD fragmentation clearly showed the different fragmentation characteristics in the gas phase. The labile nature of modified peptides was clearly demonstrated in ESI-CID-MS/MS whereas ETD showed mostly peptides backbone cleavage. These studies for the first time demonstrated the combined CID and ETD fragmentation of unmodified and modified geranylgeranylated peptides. We believe these studies will help understand fragmentation behavior of these peptides in gas phase and hence this knowledge can be implemented in the search algorithms for better data analysis of prenylated proteins after proteomic experiments. We believe this information is very important for proteomic research because prenylation is one of the most common lipid modifications that are involved in several types of cancer and no efficient mass spectrometric method is available to locate and differentiate the prenylation types. We hope this study will open a new avenue for analyzing these peptides by mass spectrometry very efficiently.

Chapter 4

Enrichment strategies developed for PTM peptides

The posttranslational lipid modified proteins are found to be in low abundance in cell and hence high sensitivity method is desirable to identify them. Mass spectrometry based proteomics is the fast and sensitive method of identification of various PTM sites. Here, the protein of interest is isolated from cells and digested into small peptides with a suitable enzyme such as trypsin. Then, the resulting proteolytic peptides are enriched using a suitable enrichment strategy to separate the PTM peptides of interest from the rest of the peptides. After that, the separated PTM-peptides are then analyzed by nano-HPLC/MS/MS for identification of PTM sites in peptides. Various proteomic software are developed which helps in more accurate identification of the peptide/protein. The detection sensitivity of the method depends on several factors like affinity enrichment, sensitivity of the nano-HPLC/MS/MS system, and complexity of the sample.

There are different methods used for enrichment of various PTMS. Major of them are Antibody-based affinity enrichment, Tagging PTMs and Ionic interaction-based enrichment.

4.1 Antibody-based affinity enrichment

Antibody based affinity purification strategy is based on charge properties or antibody recognition of the PTM peptides. This binding is highly specific and in peptide level. In this method, generally proteins are digested into peptides and then binding regions of peptides of interest binds specifically with the immobilized antibody in the chromatography column and forms immune complex. All other peptides are washes away and then they are analyzed by various other techniques. Here, less non-specific binding

is likely to occur for PTM peptides than proteins. Antibody based enrichment method is applicable for identification of protein lysine acetylation [91], arginine/lysine methylation [92], tyrosine nitration, and tyrosine phosphorylation [93, 94]. Antibodies are also used to identify kinase substrates and quantify the PTM changes in the peptides [95]. Pure antibodies are not always available for PTMs of interest because of their small sizes, difficult to generate and also have low binding affinity towards particular PTM peptides. Western blotting is the most common analysis method for immune complex and quality control is the major issue for that. Antibody are very specific and need to be pure too.

4.2 Enrichment by tagging PTMs

Different chemical methods are used to tag the PTMS either by chemical reactions or by metabolic labeling. Azide is a small reactive chemical group used for metabolic labeling of PTMs because of its biorthogonal nature. After that, the chemically labeled PTM proteins is conjugated to biotin [96]. Biorthogonal method using biotin-affinity is mostly used for the identification of various protein farnesylation, O-GlcNAc modifications [93], palmitoylation [97] and myristoylation [98]. Different PTMs are derivatized with specific chemically reactive biotinylated reagent for its detection and affinity isolation [99]. In- vitro chemical derivation method for S-palmitoyl group into a tractable tag for affinity-enrichment and analyzed by HPLC/MS/MS analysis were already developed [100]. However, multiple reactions for the chemical derivatization is not desirable as it may leads to various unwanted chemical side products.

Here, we developed also a chemical derivatization based affinity enrichment methods for both the farnesylated and geranylgeranylated peptides for complex protein samples. In our method, the protein enzymatic digestion was done to form peptides and prenylated peptides also. Then the epoxidation reaction was done on the prenyl chain

using the chemical reagents which resulted in the formation of various epoxidized prenylated peptides depending upon the stoichiometric ration of reagent to the peptides. After that, the specific tagged biotin was used to tag the epoxy group and enriched using affinity purification. The isolated peptides of interest were analyzed in nano HPLC-MS/MS system.

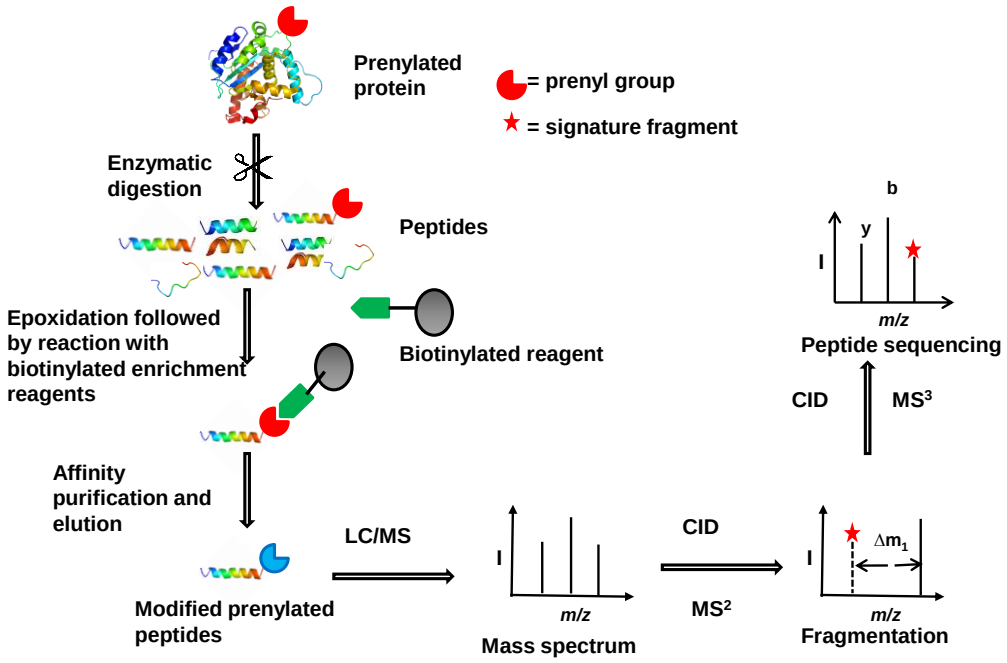


Figure 4-1 General scheme for enrichment of prenylated peptides using chemical derivatization method.

4.3 Ionic interaction-based enrichment

The most successful enrichment strategies for phosphopeptides is ionic interaction based enrichment method. This method is effective because of negative charge of the phosphate group and its tendency to form coordinate covalent bonding with immobilized metal ions. It results in the development of immobilized metal affinity

chromatography (IMAC) for isolating phosphopeptides in some phosphoproteomics studies [101]. Because the interaction with ion exchange beads, both strong cation exchange chromatography (SCX) have also been developed for phosphoproteomics studies for better efficiency[102]. Another solid matrix of titanium dioxide (TiO2) has been used for enriching phosphopeptides is widely recognized because of the unique selectivity [103, 104]. Zirconium and aluminum oxides were also tried for enriching phosphopeptides [100, 102, 105]

4.4 Our approach for enrichment of prenylated peptides

Enrichment of prenylated peptides are quite difficult as the prenyl group is less reactive. Our approach is the addition of epoxy groups by strong oxidizing agents so prenylated peptides will be more hydrophilic. The chemical reactivity of epoxy group can be used as a chemical tag for the development of enrichment strategy. Since we have successfully incorporated epoxy groups in the prenyl modification sites, we designed an enrichment strategy of these epoxidized peptides. In order to demonstrate that this peptide can be enriched using click chemistry, we tested the incorporation of azide moiety in the epoxy farnesylated peptide. On the other hand, we use Dithiothreitol reactivity towards the epoxy group of epoxyprenylated peptide is also studied. It has found that in case of DTT, the incorporation of a one and two DTT molecules are incorporated in epoxy chain.

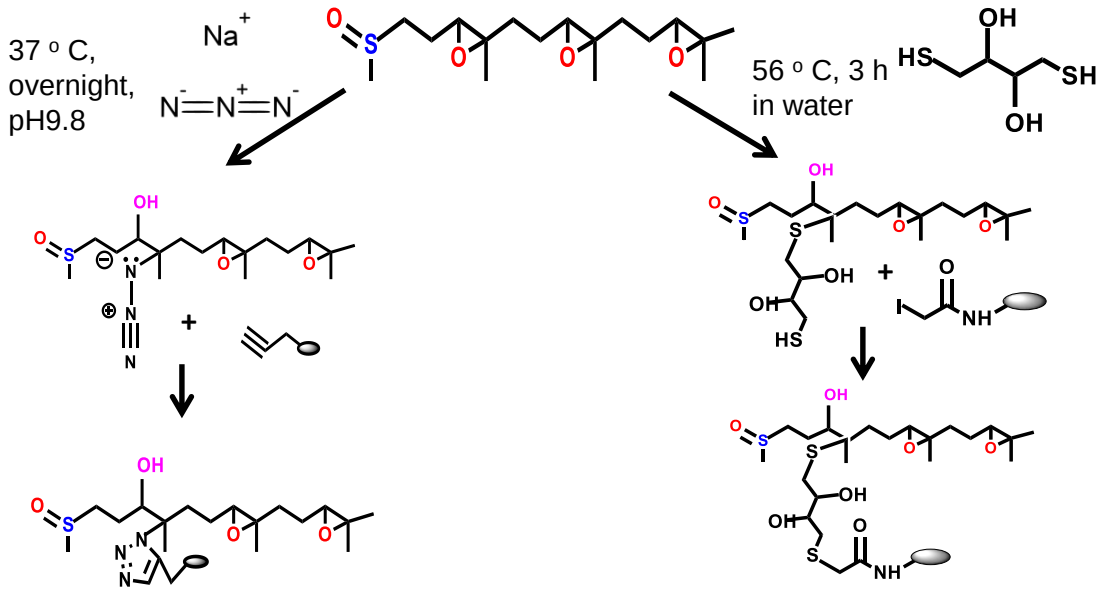


Figure 4.2 Two different approaches developed for enrichment of prenylated peptides. Enrichment with sodium azide and then enriched by alkyne biotinylated reagent (left panel) and enrichment with dithiothreitol and then enriched with iodoacetamide biotinylated reagent 9right panel).

In Figure 4.2, the two different enrichment approaches are shown, that are successfully carried out. The reaction of sodium azide with the epoxyprenylated peptide was done and it was found that the incorporation of the azide moiety takes place in the epoxy chain. We found mostly the single addition of azide moiety as because the reaction was carried out in high pH and it was found the epoxy group undergoes hydrolysis in base medium to form hydroxyl groups. Further, the alkyne biotinylated reagent is added that selectively reacts with azide via biorthogonal ligation and analyzed by mass spectrometry. Since, this reaction is not effective, we proceed the other way for enrichment. The reaction of

dithiothreitol with the epoxy group takes place in neutral pH and it was studied before the epoxy groups are stable in neutral medium. Hence, several adducts of dithiothreitol with the epoxyprenylated peptides were observed. The reaction is effective for larger peptides also and the enrichment of adduct can be done with iodoacetamide biotinylated reagents.

To demonstrate our enrichment approach, the enrichment reaction of a simple peptide RGDC was done with sodium azide. Then the products are analyzed using MALDI-QIT-TOF. The peaks at *m/z* 728.69 (M+H+Far+3O+azide), 742.49 (M+H+Far+3O+azide) and 760.49(M+H+Far+4O+azide) were the products formed by chemical derivatization of the epoxidized farnesylated RGDC with one azide moiety (Figure 4-3). These peaks were seen in the MALDI mass spectrum and isolated from the rest of the epoxidized counterparts. The CID fragmentation of these adducts was shown in Figure 4-4 with the formation of same fragment at m/z 415.51 with the loss of RSOH group (where the R = epoxy farnesyl + one azide group).

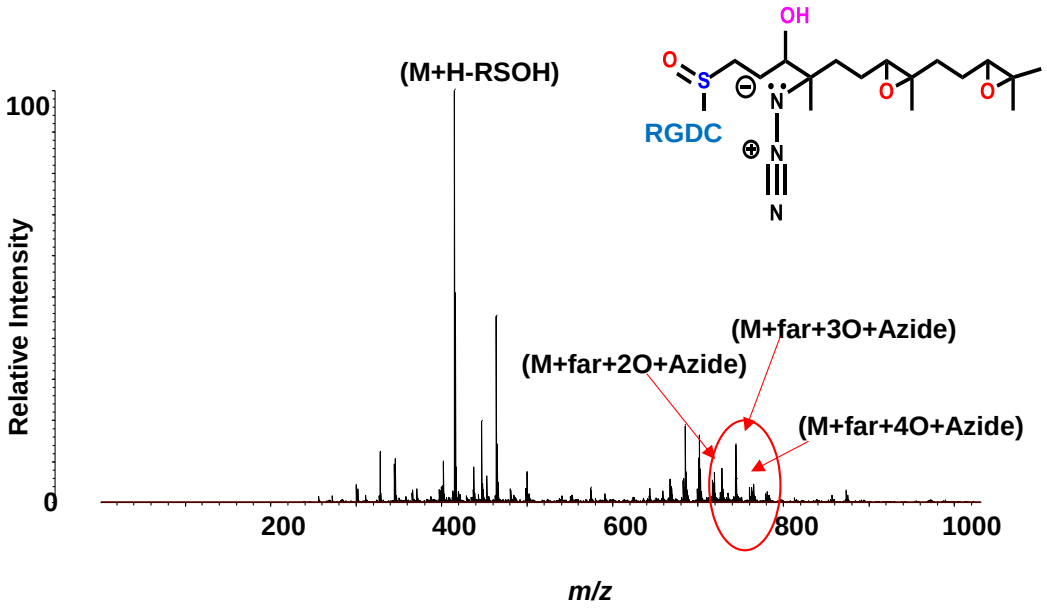


Figure 4-3 Chemical derivatization of epoxy prenylated peptide RGDC with azide.

It is very clear that even after addition of azide moiety the same signature mass pattern was retained which showed the strength of this approach to identify prenylated peptide before and after enrichment using the same signature mass loss.

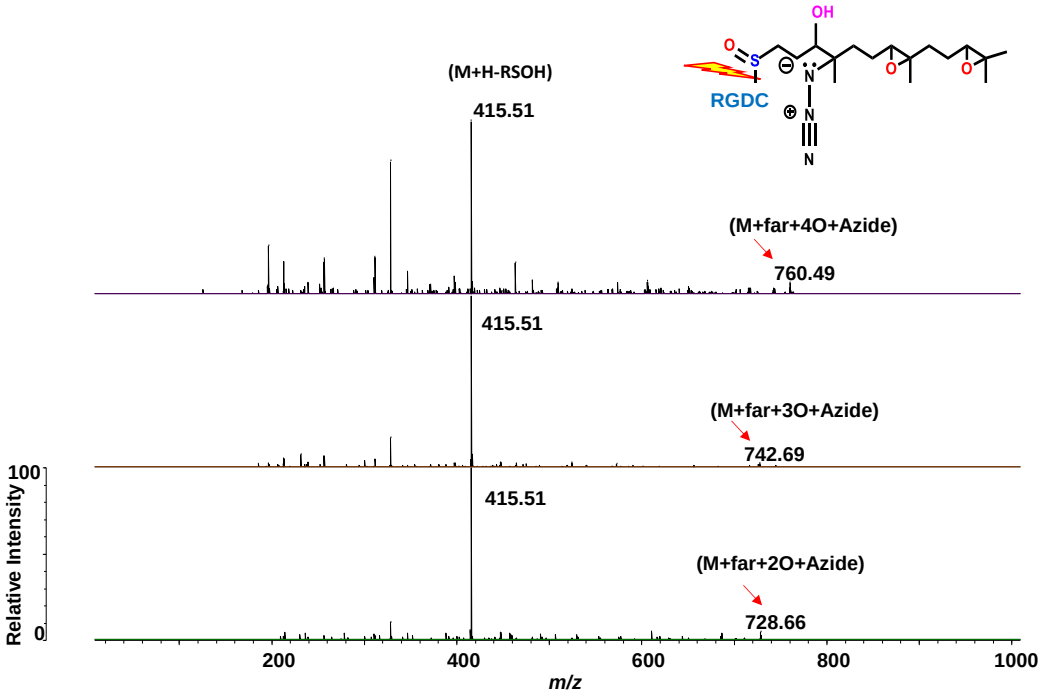


Figure 4-4 MALDI-MS/MS fragmentation of the products formed at *m/z* 728.69 (M+H+Far+3O+azide), 742.49 (M+H+Far+3O+azide) and 760.49(M+H+Far+4O+azide).

The similar experiment was performed on the farnesylated epoxidized CaaX peptide REKKFFCAIM with the sodium azide for specific reaction between the epoxy and azide group and the similar result with only one azide group incorporation in the epoxy prenyl group were observed in case of two of the epoxides with m/z 1568.85 (M+H+Far+3O+azide) and 1584.85 (M+H+Far+4O+azide). Further fragmentation of adducts also showed a signature peak at m/z 1254.78 with the loss of RSOH group. The incorporation of azide clearly showed that this peptide can be enriched using click chemistry. We also look for another method for enriching the epoxy group by the chemical reaction of epoxidized prenylated peptides with Dithiothreitol (DTT). The reaction of DTT with the epoxy group takes place at neutral pH. The DTT derivatization of

the farnesylated epoxidized CaaX peptide REKKFFCAIM was performed. The reaction between the epoxy and sulfhydryl group took place and it resulted in the formation of various adducts with either one DTT or two DTT molecules incorporated in the epoxy prenyl group.

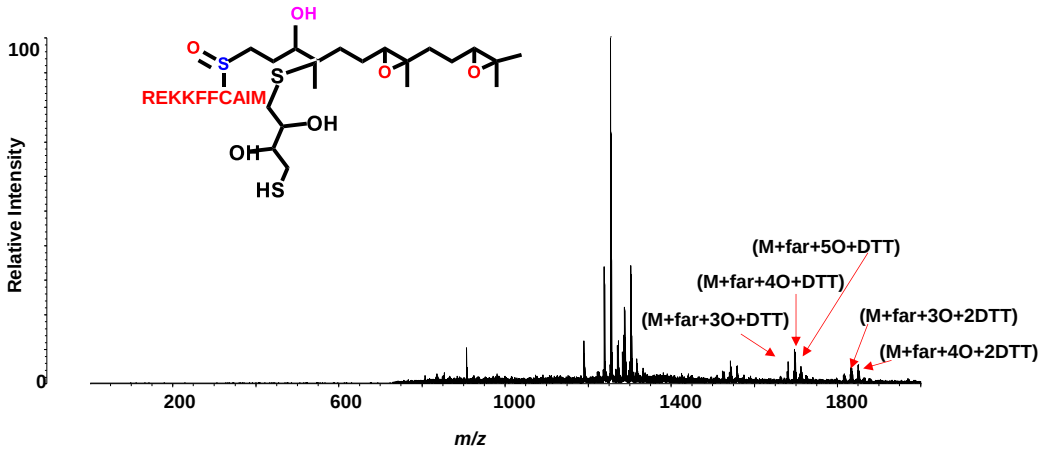


Figure 4-5 Chemical derivatization of epoxy prenylated peptide REKKFFCAIM with Dithiothreitol.

Several products at *m/z* 1680.84 (M+far+3O+DTT), 1696.82 (M+far+4O+DTT), 1713.80 (M+far+5O+DTT), and 1834.80 (M+far+3O+2DTT) and 1850.82 (M+far+4O+DTT) were observed. These products were isolated IN MALDI MS spectrum and CID fragmentation of these products resulted in the formation of signature ion peak at *m/z* 1254.58 with the loss of RSOH where R= far + nO + nDTT (n is the numbers) We anticipate that this complex can be enriched with maleimide–biotin, which is the most common for enriching the cysteinyl peptides.

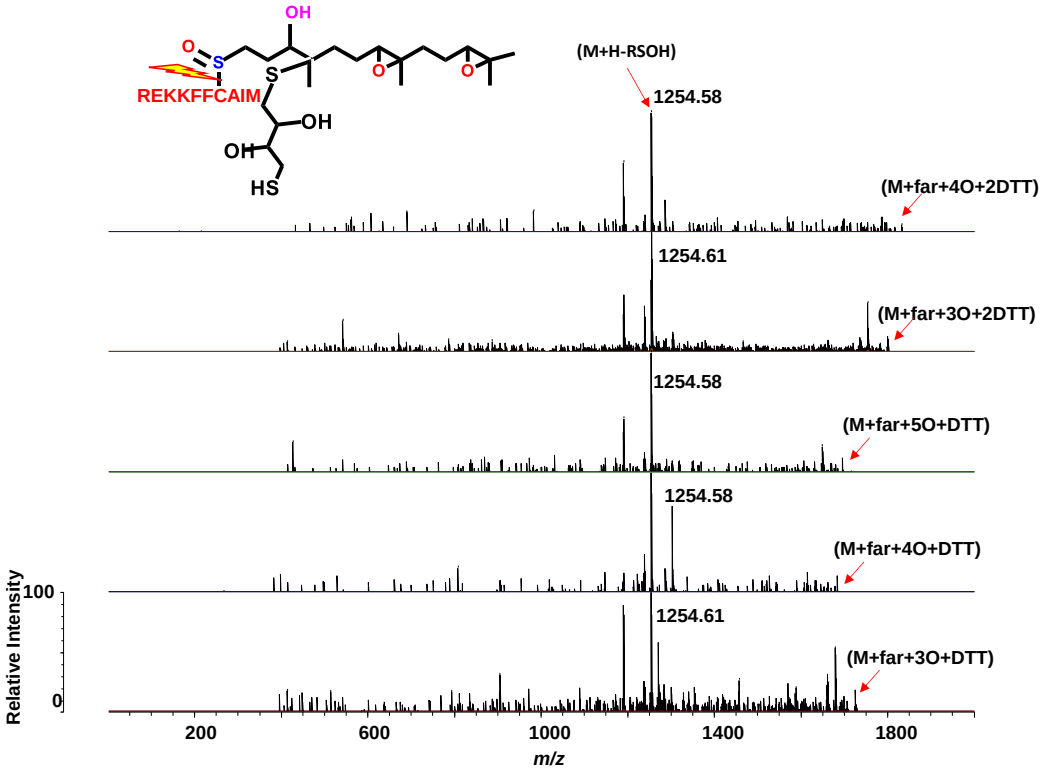


Figure 4-6 MALDI-MS/MS fragmentation of the products formed at *m/z* 1680.84 (M+far+3O+DTT), 1696.82 (M+far+4O+DTT), 1713.80 (M+far+5O+DTT), and 1834.80 (M+far+3O+2DTT) and 1850.82 (M+far+4O+DTT).

Chapter 5

Summary and future work

In this dissertation, a novel mass spectrometric cleavable strategy for the identification and differentiation of the types of prenylation was developed and applied to the complex sample like protein. Protein prenylation is a type of covalent posttranslational lipid modification found in all biological system, where either a farnesyl (15 carbon prenyl group) or geranylgeranyl (20 carbon prenyl group) is attached to cysteine amino acid residue at the carboxyl terminal of protein. These modifications are carried out with the help of an enzyme called Prenyltransferase in biological system. For any protein to be prenylated, it should have specific amino acid sequence in carboxyl terminal. They are connected with several human cancers, such as pancreatic, colon and acute myeloid leukemia and other diseases like progeria, aging, parasitic diseases and bacterial and viral infections. Potential inhibitors of prenylation were developed and investigated to suppress the cancer activity, but clinical trials are not so effective. This challenges creates an intense area of investigation of prenylation sites and the type of prenylation to define its roles in several diseases and synthesize appropriate inhibitors. Our method involved the chemical oxidation and epoxidation of prenyl group in the prenylated peptides using the common epoxydising agent m-chloroperoxybenzoic acid. On oxidation, the sulfoxide (S=O) was formed at cysteine sulfur in the prenylated peptides. In literature, it was studied that when the cysteine form sulfoxide, the C-S bond undergoes fragmentation in mass spectrometer at very low energy. Our method used the same principle to get the signature mass fragment with the loss of RSOH group (R = farnesyl/geranylgeranyl) during the fragmentation in mass spectrometry. The epoxy

groups increases the hydrophilicity of the prenylated proteins and make prenyl peptides enrichable. Their mass loss also distinguishes the type of prenylation.

We studied the fragmentation of the prenylated and the oxidized prenylated peptides in mass spectrometer using two different methods- Collision Induced dissociation (CID) and electron transfer dissociation (ETD). It was found the signature fragment was very distinct during the CID fragmentation, but not in ETD method. Oxidized peptides generally release a signature ion due to a sulfoxide formation in the cysteine sulfur. ETD fragmentation clearly showed the different fragmentation characteristics in the gas phase. The labile natures of modified peptides were clearly demonstrated for few peptides in the ESI-CID-MS/MS whereas ETD kept the labile modification intact. These studies for the first time demonstrated the CID and ETD fragmentation of unmodified and modified geranylgeranylated peptides.

Prenylated protein has low abundance as compared to the whole proteome. So for sensitive and effective identification of the prenylation site and prenylation type, the need of enrichment strategy is very important in near future. The epoxy group is introduced in the prenyl chain by epoxidation and the reactivity of epoxy group helps in the selective reaction with epoxy-reactive biotinylated reagents. Some preliminary studies of the reaction of epoxy with azide-biotin and maleimide biotin were done. In future, our method will be used for identifying the status of the prenylated protein in large-scale samples.

Appendix A

Copyright permission for Chapter 2

References

1. Resh, M.D. Covalent lipid modifications of proteins. *Curr. Biol.* **2013**, 23, R431-5.
2. Basle, E.; Joubert, N; Pucheault, M. Protein chemical modification on endogenous amino acids. *Chem. Biol.* **2010,** 17, 213-27.
3. Hannoush, R.N.; Sun, J.The chemical toolbox for monitoring protein fatty acylation and prenylation. *Nat. Chem. Biol.* **2010**, 6, 498-506.
4. Seo, J.; Lee, K.J. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *J. Biochem. Mol. Biol.* **2004**, 37, 35-44.
5. Deribe, Y.L.; Pawson, T.; Dikic, I. Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **2010**, 17, 666-72.
6. Leung, K.F.; Baron, R.; Seabra, M.C. Thematic review series: lipid posttranslational modifications. geranylgeranylation of Rab GTPases. *J. Lipid Res.* **2006**, 47, 467-75.
7. Basso, A.D.; Kirschmeier, P.; Bishop, W.R. Lipid posttranslational modifications. Farnesyl transferase inhibitors. *J. Lipid Res.* **2006**, 47,15-31.
8. Resh, M.D. Trafficking and signaling by fatty-acylated and prenylated proteins. *Nat. Chem. Biol.* **2006**, 2, 584-90.
9. Resh, M.D. Targeting protein lipidation in disease. *Trends Mol. Med.* **2012**, 18, 206-14.
10. Schafer, W.R.; Rine, J. Protein prenylation: genes, enzymes, targets, and functions. *Annu. Rev. Gene.*, **1999,** 26, 209-37.

11. Lane, K.T.; Beese, L.S. Thematic review series: lipid posttranslational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I. *J. Lipid Res.*, **2006**, 47, 681-99.

12. Du, W.; Lebowitz, P.F.; Prendergast, G.C. Cell growth inhibition by farnesyltransferase inhibitors is mediated by gain of geranylgeranylated RhoB. *Mol. Cell Biol.* **1999**, 19, 1831-40.

13. Wang, Y.C.; Peterson, S.E.; Loring, J.F. Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res.* **2014**, 24,143-60.

14. Ahearn, I.M. Regulating the regulator: post-translational modification of RAS. *Nat. Rev. Mol. Cell Biol.* **2012**, 13, 39-51.

15. Khosravi-Far, R. Ras (CXXX) and Rab (CC/CXC) prenylation signal sequences are unique and functionally distinct. *J. Biol. Chem.* **1992**, 267, 24363-8.

16. Wennerberg, K.; Rossman, K.L.; Der, C.J. The Ras superfamily at a glance. *J. Cell Sci.* **2005**, 118, 843-6.

17. Der, C.J.; Cox, A.D. Isoprenoid modification and plasma membrane association: critical factors for ras oncogenicity*. Cancer Cells.* **1991**, 3, 331-40.

18. Dominguez-Gerpe, L.; Araujo-Vilar, D. Prematurely aged children: molecular alterations leading to Hutchinson-Gilford progeria and Werner syndromes. *Curr. Aging Sci.* **2008**, 1, 202-12.

19. Young, S.G. Targeting protein prenylation in progeria*. Sci. Transl. Med.* **2013**, 5, 171ps3.

20. Chia, W.J.; Tang, B.L. Emerging roles for Rab family GTPases in human cancer. *Biochim. Biophys. Acta.* **2009**, 1795, 110-6.

21. Garcia-Ruiz, C.; Morales, A. ; Fernandez-Checa, J.C. Statins and protein prenylation in cancer cell biology and therapy. *J. Med. Chem*. **2012**, 12, 303-15.

22. Scharovsky, O.G. Inhibition of ras oncogene: a novel approach to antineoplastic therapy. *J. Biomed. Sci.* **2000**, 7, 292-8.

23. Marshall, C.J. Protein prenylation: a mediator of protein-protein interactions. *Science*. **1993**, 259, 1865-6.

24. Bos, J.L. Ras oncogenes in human cancer: a review. *Cancer Res.* **1989**, 49, 4682-9.

25. Berndt, N.; Hamilton, A.D.; Sebti, S.M. Targeting protein prenylation for cancer therapy. *Nat. Rev. Cancer*. **2011**, 11, 775-91.

26. Casey, P.J. p21ras is modified by a farnesyl isoprenoid. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, 86, 8323-7.

27. Fernandez-Medarde, A.; Santos, E. Ras in cancer and developmental diseases. *Genes Cancer*, **2011**, 2, 344-58.

28. Gamblin, D.P. Chemical site-selective prenylation of proteins. *Mol. Biosyst.* **2008**, 4, 558-61.

29. Gelb, M.H. Therapeutic intervention based on protein prenylation and associated modifications. *Nat. Chem. Biol.* **2006**, 2, 518-28.

30. Greenwood, J.; Steinman, L.; Zamvil, S.S. Statin therapy and autoimmune disease: from protein prenylation to immunomodulation. *Nat. Rev. Immunol.* **2006**, 6, 358-70.

31. Buglino, J.A.; Resh, M.D. Palmitoylation of Hedgehog proteins. *Vitam. Horm.* **2012**, 88, 229-52.

32. Linder, M.E.; Deschenes, R.J. Palmitoylation: policing protein stability and traffic. *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 74-84.

33. Resh, M.D. Palmitoylation of ligands, receptors, and intracellular signaling molecules. *Sci. STKE.* **2006**, 359, re14.

34. Frohlich, M. S-palmitoylation represents a novel mechanism regulating the mitochondrial targeting of BAX and initiation of apoptosis. *Cell Death Dis.* **2014**, 5,1057.

35. Young, F.B. Putting proteins in their place: palmitoylation in Huntington disease and other neuropsychiatric diseases. *Prog. Neurobiol.* **2012**, 97, 220-38.

36. Wan, J. Palmitoylated proteins: purification and identification. *Nat. Protoc.* **2007**, 2,1573-84.

37. Patwardhan, P.; Resh, M.D. Myristoylation and membrane binding regulate c-Src stability and kinase activity. *Mol. Cell. Biol.* **2010**, *30*, 4094-4107.

38. Gordon, J.I.; Duronio, R.J.; Rudnick, D.A.;Adams, S.A.;Gokel, G.W. Protein N-myristoylation. *J. Biol. Chem.* **1991**, 266, 8647-50.

39. Farazi, T.A.; Waksman, G.; Gordon, J. I.The biology and enzymology of protein N-myristoylation. *J. Biol. Chem.* **2001**, 276, 39501-4.

40. Wright, M.H.; Heal, W.P.;Mann, D.J.;Tate, E.W. Protein myristoylation in health and disease. *J. Chem. Biol.* **2010**, 3, 19-35.

41. Kelman, Z.; Naktinis, V.; O'Donnell, M. Radiolabeling of proteins for biochemical studies. *Methods Enzymol.* **1995**, 262, 430-42.

42. Chakrabarti, D.; Da Silva, T.; Barger,J.; Paquette,S.; Patel, H.;Patterson, S.; Allen, C.M. Protein farnesyltransferase and protein prenylation in Plasmodium falciparum. *J. Biol. Chem.* **2002**, 277, 42066-73.

43. Suzuki, T.; Ito, M.;Ezure, T.; Shikata, M.; Ando, E.; Utsumi, T.; Tsunasawa, S.; Nishimura, O. Protein prenylation in an insect cell-free protein synthesis system and identification of products by mass spectrometry. *Proteomics*. **2007**, 7, 1942-50.

44. Hang, H.C.; Geuties, E.J.;Grotenbreg, G.; Pollington, A.M.; Bijlmakers, M.J.; Ploegh, H.L. Chemical probes for the rapid detection of Fatty-acylated proteins in Mammalian cells. *J. Am. Chem. Soc.* **2007**, 129, 2744-5.

45. Charron, G.; Tsou, K.; Maguire, W.; Yount, J.S.; Hang, H.C. Alkynyl-farnesol reporters for detection of protein S-prenylation in cells. *Mol. Biosyst.* **2011**, 7, 67-73.

46. Hang, H.C.; Wilson, J.P.; Charron, G. Bioorthogonal chemical reporters for analyzing protein lipidation and lipid trafficking. *Acc. Chem. Res.* **2011**, 44, 699-708.

47. Kho, Y.; Kim, S.C.; Jiang, C.; Barma, D.; Kwon, S.W.; Cheng, J.; Jaunbergs, J.; Weinbaum, C.; Tamanoi, F.; Falck, J.; Zhao, Y. A tagging-via-substrate technology for detection and proteomics of farnesylated proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 12479-84.

48. Onono, F. O.; Morgan, M. A.; Spielmann, H. P.; Andres, D. A.; Subramanian, T.; Ganser, A.; Reuter, C. W. A tagging-via-substrate approach to detect the farnesylated proteome using two-dimensional electrophoresis coupled with Western blotting. *Mol. Cell Proteomics.* **2010**, 9, 742-51.

49. Charron, G.; Li, M. M.; MacDonald, M. R.; Hang, H. C. Prenylome profiling reveals S-farnesylation is crucial for membrane targeting and antiviral activity of ZAP long-isoform. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, 110, 11085-90.

50. Sinensky, M., Recent advances in the study of prenylated proteins. *Biochim. Biophys. Acta.* **2000**, 1484, 93-106.

51. Hoffman, M.D.; Kast, J. Mass spectrometric characterization of lipid-modified peptides for the analysis of acylated proteins. *J. Mass Spectrom.* **2006**, 41, 229-41.

52. Hoffman, M.D.; Sniatynski, M.J.; Kast, J. Current approaches for global post-translational modification discovery and mass spectrometric analysis. *Anal. Chim. Acta.* **2008**, 627, 50-61.

53. Anderegg, R.J.; Betz, R.; Carr, S.A.; Crabb, J. W.; Duntze, W. Structure of Saccharomyces cerevisiae mating hormone a-factor. Identification of S-farnesyl cysteine as a structural component. *J. Biol. Chem.* **1988**, 263, 18236-40.

54. Domon, B.; Aebersold, R. Mass spectrometry and protein analysis. *Science.* **2006**, 312, 212-7.

55. Walther, T.C.; Mann, M. Mass spectrometry-based proteomics in cell biology. *J. Cell Biol.* **2010**, 190, 491-500.

56. Ho, C.S.; Lam, C.W.; Chan, M.H.; Cheung, R.C.; Law, L.K.; Lit, L. C.; Ng, K. F.; Suen, M. W.; Tai, H. L. Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin. Biochem. Rev.* **2003**, 24, 3-12.

57. Webster, J.; Oxley, D. Protein identification by MALDI-TOF mass spectrometry. *Methods Mol. Biol.* **2012**, 800, 227-40.

58. Baldwin, M.A. Protein identification by mass spectrometry. *Mol. Cell Proteomics*. **2004**, 3, 1-9.

59. Pitt, J.J. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *Clin. Biochem. Rev.* **2009**, 30, 19-34.

60. Yost, R.A.; Boyd, R. K. Tandem mass spectrometry: quadrupole and hybrid instruments. *Methods Enzymol.* **1990**, 193,154-200.

61. Sobott, F.; Watt, S .J.; Smith, J.; Edelmann, M. J.;Kramer, H. B.; Kessler, B. M. Comparison of CID versus ETD based MS/MS fragmentation for the analysis of protein ubiquitination. *J. Am. Soc. Mass Spectrom.* **2009**, 20, 1652-9.

62. Mechref, Y. Use of CID/ETD mass spectrometry to analyze glycopeptides. *Cur. Protoc. Protein Sci.* **2012,** Chapter 12, Unit 12, 11, 1-11.

63. Frese, C.K.; Altelaar, A. F.; Hennrich, M. L.; Nolting, D.; Zeller, M.; Griep-Raming, J.; Heck, A. J.; Mohammed, S. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J. Proteome Res.* **2011**, 10, 2377-88.

64. Roberts, P. J.; Mitin, N.; Keller, P. J.; Chenette, E. J.; Madigan, J. P.; Currin, R. O.; Cox, A. D.; Wilson, O.; Kirschmeier, P.; Der, C. J. Rho Family GTPase modification and dependence on CAAX motif-signaled posttranslational modification. *J. Biol. Chem.* **2008**, 283, 25150-63.

65. Joberty, G.; Tavitian, A.; Zahraoui, A. Isoprenylation of Rab proteins possessing a C-terminal CaaX motif. *FEBS Lett.* **1993**, 330, 323-8.

66. Casey, P.J.; Seabra, M.C. Protein prenyltransferases. *J. Biol. Chem.* **1996**, 271, 5289-92.

67. Manolaridis, I.; Kulkarni, K.; Dodd, R. B.; Ogasawara, S.; Zhang, Z.; Bineva, G.; O'Reilly, N.; Hanrahan, S. J.; Thompson, A. J.; Cronin, N.; Iwata, S.; Barford, D. Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1. *Nature*. **2013**, 504, 301-5.

68. Lammerding, J.; Fong, L. G.; Ji, J. Y.; Reue, K.; Stewart, C. L.; Young, S. G.; Lee, R. T. Lamins A and C but not lamin B1 regulate nuclear mechanics. *J. Biol. Chem.* **2006**, 281, 25768-80.

69. Sinensky, M.; Fantle, K.; Trujillo, M.; McLain, T.; Kupfer, A.; Dalton, M. The processing pathway of prelamin A. *J. Cell Sci.* **1994**, 107 , 61-7.

70. Seabra, M. C.; Goldstein, J. L.; Sudhof, T. C.; Brown, M. S. Rab geranylgeranyl transferase. A multisubunit enzyme that prenylates GTP-binding proteins terminating in Cys-X-Cys or Cys-Cys. *J. Biol. Chem.* **1992**, 267, 14497-503.

71. Smeland, T. E.; Seabra, M. C.; Goldstein, J. L.; Brown, M. S. Geranylgeranylated Rab proteins terminating in Cys-Ala-Cys, but not Cys-Cys, are carboxyl-methylated by bovine brain membranes in vitro. *Proc. Natl. Acad. Sci. U. S. A.* .**1994**, 91, 10712-6.

72. Kohl, N. E.; Wilson, F. R.; Mosser, S. D.; Giuliani, E.; deSolms, S. J.; Conner, M. W.; Anthony, N. J.; Holtz, W. J.; Gomez, R. P.; Lee, T. J. Protein farnesyltransferase inhibitors block the growth of ras-dependent tumors in nude mice. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, 91, 9141-5.

73. Kazi, A.; Carie, A.; Blaskovich, M. A.; Bucher, C.; Thai, V.; Moulder, S.; Peng, H.; Carrico, D.; Pusateri, E.; Pledger, W. J.; Berndt, N.; Hamilton, A.; Sebti, S. M. Blockade of protein geranylgeranylation inhibits Cdk2-dependent p27Kip1 phosphorylation on Thr187 and accumulates p27Kip1 in the nucleus: implications for breast cancer therapy. *Mol. Cell Biol.* **2009**, 29, 2254-63.

74. Jackson, J. H.; Cochrane, C. G.; Bourne, J. R.; Solski, P. A.; Buss, J. E.; Der, C. J. Farnesol modification of Kirsten-ras exon 4B protein is essential for transformation. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87, 3042-6.

75. Sepp-Lorenzino, L.; Ma, Z.; Rands, E.; Kohl, N. E.; Gibbs, J. B.; Oliff, A.; Rosen, N. A peptidomimetic inhibitor of farnesyl:protein transferase blocks the anchorage-dependent and -independent growth of human tumor cell lines. *Cancer Res*. **1995**, 55, 5302-9.

76. Wang, C.; Gale, M., Jr.; Keller, B. C.; Huang, H.; Brown, M. S.; Goldstein, J. L.; Ye, J. *I*dentification of FBL2 as a geranylgeranylated cellular protein required for hepatitis C virus RNA replication. *Mol. Cell.* **2005**, 18, 425-34.

77. Wotske, M.; Wu, Y.; Wolters, D. A. Liquid chromatographic analysis and mass spectrometric identification of farnesylated peptides. *Anal. Chem.* **2012**, 84, 6848-55

78. Bakkestuen, A. K.; Gundersen, L. L.; Petersen, D.; Utenova, B. T.; Vik, A. S. Synthesis and antimycobacterial activity of agelasine E and analogs. *Org. Biomol. Chem.* **2005**, 3, 1025-33.

79. Griffiths-Jones, C.M.; Knight, D.W. Carbenium ion trapping using sulfonamides: an acid-catalysed synthesis of pyrrolidines by intramolecular hydroamination. *Tetrahedron*. **2010**, 66, 4150-4166.

80. Brown, M. J.; Milano, P. D.; Lever, D. C.; Epstein, W. W.; Poulter, C. D. Prenylated Proteins - a Convenient Synthesis of Farnesyl Cysteinyl Thio*ethers. J. Am. Chem. Soc.* **1991**, *113*, 3176-3177.

81. Lichtor, P. A.; Miller, S. J. Combinatorial evolution of site- and enantioselective catalysts for polyene epoxidation*. Nature Chem*. **2012**, 4, 990-995.

82. Chowdhury, S. M.; Munske, G. R.; Ronald, R. C.; Bruce, J. E. Evaluation of low energy CID and ECD fragmentation behavior of mono-oxidized thio-ether bonds in peptides. *J. Am. Soc. Mass Spectrom.* **2007**, 18, 493-501.

83. Chowdhury, S. M.; Munske, G. R.; Bruce, J. E. Collisionally activated dissociation and electron capture dissociation of several mass spectrometry-identifiable chemical cross-linkers*. Anal. Chem.* **2006**, 78, 8183-93.

84. Thissen, J. A.; Gross, J. M.; Subramanian, K.; Meyer, T.; Casey, P. J. Prenylation-dependent association of Ki-Ras with microtubules. Evidence for a role in subcellular trafficking. *J. Biol. Chem.* **1997**, 272, 30362-70.

85. Kassai, H.; Satomi, Y.; Fukada, Y.; Takao, T. Top-down analysis of protein isoprenylation by electrospray ionization hybrid quadrupole time-of-flight tandem mass spectrometry; the mouse Tgamma protein. *Rapid Comm. Mass Spectrom.* **2005**, 19, 269-74.

86. Kim, C., T.G. Traylor, and C.L. Perrin, *MCPBA epoxidation of alkenes: Reinvestigation of correlation between rate and ionization potential. J. Am. Chem. Soc.* **1998**, *120*, 9513-9516.

87. Roskoski, R., Jr. Protein prenylation: a pivotal posttranslational process. *Biochem Biophys. Res. Comm.* **2003**, 303, 1-7.

88. Zhang, F.L.; Casey, P.J. Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem*. **1996**, 65, 241-69.

89. Seidler, J. De novo sequencing of peptides by MS/MS. *Proteomics*. **2010**, 10, 634-49.

90. Bhawal, R.P.; Sadananda, S. C.; Bugarin, A.; Laposa, B.; Chowdhury, S.M. Mass spectrometry cleavable strategy for identification and differentiation of prenylated peptides. *Anal. Chem.* **2015**, 87, 2178-86.

91. Zhang, J.; Sprung, R.; Pei, J.; Tan, X.; Kim, S.; Zhu, H.; Liu, C. F.; Grishin, N. V.; Zhao, Y. Lysine acetylation is a highly abundant and evolutionarily conserved modification in Escherichia coli. *Mol. Cell Proteomics*. **2009**, 8, 215-25.

92. Zhan, X.; Desiderio, D.M. Nitroproteins from a human pituitary adenoma tissue discovered with a nitrotyrosine affinity column and tandem mass spectrometry. *Anal. Biochem.* **2006**, 354, 279-89.

93. Vocadlo, D.J.;Hang, H. C.; Kim, E. J.; Hanover, J. A.; Bertozzi, C. R. A chemical approach for identifying O-GlcNAc-modified proteins in cells. *Proc. Natl. Acad. Sci. U S A*. **2003**, 100, 9116-21.

94. Sprung, R.; Nandi, A.; Chen, Y.; Kim, S. C.; Barma, D.; Falck, J. R.; Zhao, Y. Tagging-via-substrate strategy for probing O-GlcNAc modified proteins. *J. Proteome Res.* **2005**, 4, 950-7.

95. Rikova, K., et al., Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell*, **2007**, 131, 1190-203.

96. Zhang, Q.; Qian, W. J.; Knyushko, T.V.; Clauss, T. R.; Purvine, S. O.; Moore, R. J.; Sacksteder, C. A.; Chin, M. H.; Smith, D. J.; Bigelow, D. J.; Smith, R. D. A method for selective enrichment and analysis of nitrotyrosine-containing peptides in complex proteome samples. *J. Proteome Res.* **2007**, 6, 2257-68.

97. Roth, A.F.;Wan, J.; Bailey, A. O.; Sun, B.; Kuchar, J. A.; Green, W. N.; Phinney, B. S.; Yates, J. R.; Davis, N. G. Global analysis of protein palmitoylation in yeast. *Cell*. **2006**, 125, 1003-13.

98. Kostiuk, M.A., et al., Identification of palmitoylated mitochondrial proteins using a bio-orthogonal azido-palmitate analogue. *FASEB J*, **2008**, 22, 721-32.

99. Larsen, M.R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell Proteomics*. **2005**, 4, 873-86.

100. Matsuoka, S., et al., ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*. **2007**, 316, 1160-6.

101. Lee, J.; Xu, Y.; Chen, Y.; Sprung, R.; Kim, S. C.; Xie, S.; Zhao, Y. Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Mol. Cell Proteomics*. **2007**, 6, 669-76.

102. Kokubu, M.; Ishihama, Y.; Sato, T.; Nagasu, T.; Oda, Y. Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis. *Anal. Chem.* **2005**, 77, 5144-54.

103. Villen, J.; Gygi, S. P.The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat. Protoc.* **2008**, 3, 1630-8.

104. Ong, S. E.; Mittler, G.; Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods*. **2004**, 1, 119-26.

105. Pandey, A.; Podtelejnikov, A. V.; Blagoev, B.; Bustelo, X. R.; Mann, M.; Lodish, H. F. Analysis of receptor signaling pathways by mass spectrometry: identification of vav-2 as a substrate of the epidermal and platelet-derived growth factor receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 179-84.

Biographical Information

Ruchika Bhawal earned her B.S in Chemistry from University of Calcutta, India. Then she continued her studies and earned her M.S in Analytical Chemistry from M.S. University of Baroda, India. She managed to secure first position in her master degree. Then she moved to Arlington, Texas with her husband and a small little one year old daughter. She joined the University of Texas at Arlington in spring 2011 and started her research with Dr. Purnendu Dasgupta. She worked on cavity enhanced absorption spectrometry for about two years. Then she joined Dr. Saiful Chowdhury group in spring 2013 and started her journey towards PhD in Analytical Chemistry. Her PhD works focused on the novel methods for identification of posttranslational lipid modified proteins by tandem mass spectrometry. She was awarded Dissertation fellowship for summer 2015 from UTA and got an award for Graduate oral presentation from American Chemical Society Meeting in Miniature Conference. She received travel stipends from American Chemical Society for Mass spectrometry for the conferences in 2014 and 2015.