ABSTRACT

Title of Dissertation: INTEGRATION OF ^{18}O LABELING AND SOLUTION ISOELECTRIC FOCUSING IN A SHOTGUN ANALYSIS OF MITOCHONDRIAL PROTEINS

Jinshan Wang, Doctor of Philosophy, 2007

Dissertation Directed by: Professor Catherine Fenselau

Department of Chemistry and Biochemistry

The coupling of efficient separations and mass spectrometry instrumentation is highly desirable to provide global proteomic analysis. When quantitative comparisons are part of the strategy, separation and analytical methods should be selected, which optimize the isotope labeling procedure. Enzyme-catalyzed ^{18}O labeling is considered to be the labeling method most compatible with analysis of proteins from tissue and other limited samples. The introduction of label at the peptide stage mandates that protein manipulation be minimized in favor of peptide fractionation post-labeling. In the present study, forward and reverse ^{18}O labeling are integrated with solution isoelectric focusing and capillary LC-tandem mass spectrometry to study changes in

mitochondrial proteins associated with drug resistance in human cancer cells.

A total of 637 peptides corresponding to 278 proteins were identified in this analysis. Of these, twelve proteins have been demonstrated from the forward and reverse labeling experiments to have abundances altered by greater than a factor of two between the drug susceptible MCF-7 cell line and the MCF-7 cell line selected for resistance to mitoxantrone. Galectin-3 binding protein precursor was detected in the resistant cell line, but was not detected in the drug susceptible line. Such proteins are challenging to ^{18}O and other isotope strategies and a solution is offered, based on reverse labeling. These twelve proteins play a role in several pathways including apoptosis, oxidative phosphorylation, fatty acid metabolism and amino acid metabolism. For some of these proteins, their possible functions in drug resistance have been proposed.

INTEGRATION OF ^{18}O LABELING AND SOLUTION ISOELECTRIC FOCUSING IN A SHOTGUN ANALYSIS OF MITOCHONDRIAL PROTEINS

by

Jinshan Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Dr. Catherine Fenselau, Chair
Dr. Neil Blough
Dr. Jonathan D. Dinman
Dr. David Fushman
Dr. Douglas Julin

Dedication

This dissertation is dedicated to my parents

Wenxiu Zhang and Mingsheng Wang.

Acknowledgements

I would first like to thank my advisor Dr. Catherine Fenselau for introducing me to the field of proteomics and for her encouragement, ideas and advice throughout the years. I have benefited greatly from her never-ending support and developed into a better and more independent scientist.

I also want to thank the many former and current members of the Fenselau group for their warm help both in my research and life. Special thanks to Dr. Yanming An, Dr. Faith Hays, Dr. Rachael Strong, Ms. Natasha Smith for being such good friends and those happy hours we spent.

I am thankful to Dr. Peter Gutierrez for his assistance in operation of the LCQ mass spectrometry instrument and database searching, Dr. Nathan Edwards for his kind help in bioinformatics.

I would like to thank my parents and my sister for all their love, understanding and support throughout my life. They have encouraged me in every way imaginable, and I love them dearly.

Finally, I would like to thank my husband, Wei and my daughter, Jialei for their love and support to make this possible.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACN: acetonitrile

CHAPS: 3-[(3-cholamidopropyl) dimethyl-ammonio]-1-propanesulfonate

DTT: dithiothreitol

EDTA: ethylenediaminetetraacetic acid

PBS: phosphate buffered saline

EGTA: ethylene glycol-bis(2-aminoethyl ether)-N, N, N’, N’-tetraacteic acid

HEPES: N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid

IAA: iodoacetamide

IPG: immobilized pH gradient

MCE: multicompartment eletrolyzer

MEM: Eagle’s minimal media

PBS: phosphate buffered saline

SCX: strong cation exchange

Solution IEF: solution isoelectric focusing

TFA: trifluoroacetic acid

Tris: tris[hydroxymethyl]aminomethane

Chapter 1: Introduction

Part A

Drug resistance in cancer chemotherapy

Cancer is the second leading cause of death among all ages in the United States. Scientists are exploiting all ways to prevent, detect, diagnose and treat the disease. Surgery and radiotherapy cures about 40% of all cancer patients, leaving the remaining to rely on systemic chemotherapeutic treatment (1). Chemotherapy, which was introduced into the clinic more than fifty years ago, is an approach to cancer treatment that uses drugs to stop the growth of cancer cells either by killing the cells or by stopping the cells from dividing (2).

In reality, the effectiveness of chemotherapy has suffered from a range of puzzling factors including systemic toxicity and drug resistance. The latter problem has been the least understood and presents major obstacles to the successful treatment of tumors. Drug resistance can be either intrinsic or acquired. Intrinsic resistance results in failure to respond to the first chemotherapy which is given to patients. Acquired resistance occurs in successive treatments following initial response (2-5). When tumor cells develop resistance, they become resistant to not only the drug with which the patient has been treated but also to a broad spectrum of structurally and functionally unrelated drugs, in a phenomenon referred to as multidrug resistance (MDR). Clearly, if drug resistance could be overcome, the impact on

survival would be highly significant.

A variety of mechanisms at molecular and cellular levels for MDR have been intensively studied since the seventies (6-10). They include reduced drug accumulations by enhancing efflux of anti-cancer drug from the cell or preventing drug influx, increasing activity or expression of detoxifying systems, disruptions in the apoptotic pathway, alterations in the drug targets, up-regulation of DNA repair systems.

The major mechanism of multidrug resistance is the alteration in drug efflux due to overexpression of ATP-binding cassette (ABC) transporter proteins such as P-glycoprotein (P-gp) and multidrug resistance protein (MRP) (4, 8, 11). P-gp acts as a plasma membrane drug efflux pump that actively exports drugs from cancer cells, thereby allowing the cell to accumulate fewer drugs and thus survive higher doses. The normal function of MRP has been identified as a carrier of negatively charged natural-product drugs. This has led to the development of P-gp or other ABC transporter inhibitors in clinical treatment to restore, enhance or prolong drug sensitivity (12). Another important ATP binding cassette family member is breast cancer resistance protein (BCRP), which has been discovered in the breast cancer cell lines resistant to mitoxantrone, daunorubicin and doxorubicin, etc (13).

Another mechanism of MDR which has been well studied is the alteration in the amount, structure or activity of molecular targets (2, 6). A typical example for such target-related MDR is resistance to topoisomerase II (Topo II)

inhibitors and poisons. DNA topoisomerases are essential enzymes which are involved in the processes of replication and transcription by binding to DNA, forming a transient DNA break followed by DNA strand passage and then releasing the DNA. A class of anticancer drugs (e.g., mitoxantrone, doxorubicin and others) targets these topoisomerases to kill tumor cells. These drugs prevent the releasing process by freezing and accumulating the cleavable DNA-enzyme complexes, leading to cell death. Clinical studies have shown an association between low cellular levels of Topo II and poor treatment outcomes.

Numerous studies to date strongly suggest that most forms of chemotherapy kill cancer cells by inducing them to undergo apoptosis (14, 45). Thus an important mechanism which can prevent cancer cells from dying, despite adequate drug-induced damage at the molecular target, is the processes that block apoptosis (15-17). For example, functional loss of p53, which is a critical initiator of the apoptosis pathway, correlates with multidrug resistance in many tumor types (6). Also, mutations or altered expression of Bcl-2 related proteins are associated with drug resistance in human cancers (18-20).

To date the mechanisms of drug resistance are poorly understood. They are likely to be multifactorial and complex in most cancer patients (7-8). Many proteins are involved and work together to confer and promote drug resistance. Understanding the causes could improve the efficiency of existing therapies

and potentially reveal new treatment strategies. Even though previous studies have suggested that many proteins such as P-gp, MRP and Bcl-2, are differentially expressed in drug-resistant tumor cells, global protein pattern changes in these tumor cells have not yet been determined. With recent developments in mass spectrometric and separation technologies, along with genomic and protein bioinformatics, the complex status of protein expression, structure and function, defined as proteomics, can be analyzed. The proteomics approach has provided a powerful tool to study drug resistance.

Mitochondria and cancer chemotherapy

Owing to the complexity of human cells, in order to assemble a list of the total complement of proteins in the cells, proteomic studies of subcellular compartments and organelles has become a major focus (21-22). Thus smaller and more manageable subsets of proteins are involved. In this study, a proteomic analysis of mitochondrial fraction has been used to study alterations in the protein abundance profile between drug susceptible and drug resistant human MCF-7 breast cancer cells.

The mitochondrial genome encodes only 13 polypeptides. Therefore the vast majority of mitochondrial proteins are encoded by the nuclear genome, synthesized in the cytosol, and then imported into mitochondria by a specific transport system (23-24). The number of distinct proteins in human mitochondria is estimated to be approximately 2000 (24-25).

Mitochondria are attractive targets for subcellular proteomics because

they play vital roles in several cellular functions, including energy production by oxidative phosphorylation, fatty acid metabolism, citric acid cycle and programmed cell death (apoptosis) (26-28).

In the last decade, there has been a surge of interest in investigating the mitochondrial role in apoptosis. Apoptosis is defined as genetically programmed autonomous cell death. It occurs in healthy cells at varying rates (46). Changes in the genetics of apoptotic regulation may result in an increase in cell numbers, which begins the process of tumorigenesis (47). There are two major apoptosis pathways, intrinsic and extrinsic, in mammalian cells. Both of them are mediated through the mitochondria, as seen in figure 1 (29). The extrinsic pathway is triggered by members of the death receptor superfamily, such as CD95. Binding of CD95 ligand to CD95 results in receptor clustering and formation of a death inducing signaling complex (DISC). This complex recruits multiple procaspase-8 molecules via the adaptor molecule Fas-associated death domain protein (FADD), which will then rapidly activate the initiator caspase 8. The intrinsic pathway occurs when various stress-induced apoptotic stimuli, including DNA damage that results from chemotherapy, trigger the release of cytochrome c from mitochondria. The exit of cytochrome c is regulated partially by Bcl2 family members, with anti-apoptotic (e.g. Bcl2) and pro-apoptotic (e.g. Bax, Bak) members inhibiting or promoting the release, respectively. Cytochrome c then interacts with apoptotic protease-activating factor 1 (Apaf1) and caspase-9 to form the

“apoptosome complex”, resulting in activation of the initiator caspase-9. The extrinsic and intrinsic pathyways converge at the level of caspase cascade (e.g. caspase-3, 6, 7) activation, which is induced by caspase-8 or 9. Those caspase cascades are responsible for the cleavage of important cellular substrates leading to cell death (30-33).

Because the importance of mitochondrial dysfunction during apoptosis induced by anticancer drugs has been illustrated (41-42) and preventing apoptotic death process contributes to drug resistance (15-17), mitochondrial targeting strategies have become a unique potential for the design of anticancer drugs (17, 34). For example, the antitumoral drug Lonidamine, which is used in combination with standard chemotherapy, induces apoptosis via a direct effect on mitochondria by disrupting the mitochondrial transmembrane potential, thus helping release of cytochrome c (35-36). All of the evidence support our rationale that information about mitochondrial proteins, identified or quantitated by proteomics, could provide insight towards a better understanding of drug resistance.

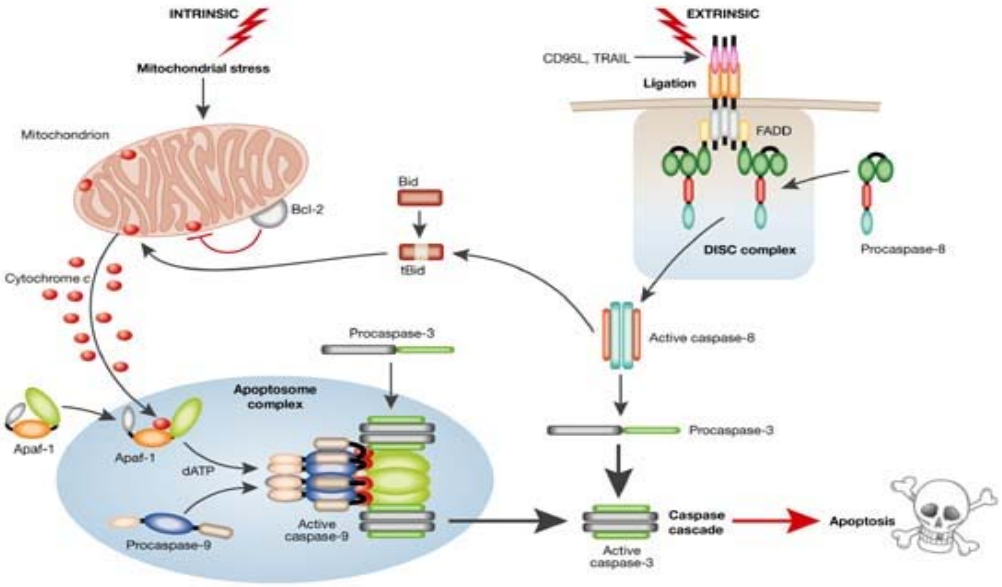


Figure 1. Intrinsic and extrinsic pathways of apoptosis (33)

There are intrinsic and extrinsic apoptosis pathways in mammalian cells. Both of them are mediated through the mitochondria. Triggering of cell surface death receptors including CD95 results in rapid activation of the initiator caspase 8 after its recruitment to a trimerized receptor-ligand complex (DISC) through the adaptor molecule Fas-associated death domain protein (FADD). In the intrinsic pathway, stress-induced apoptosis results in perturbation of mitochondria and the ensuing release of proteins, such as cytochrome *c.* The release of cytochrome *c* is regulated in part by Bcl2 family members. Once released, cytochrome *c* binds to apoptotic protease-activating factor 1 (Apaf1), which results in formation of the Apaf1–caspase 9 apoptosome complex and activation of the initiator caspase 9. The activated initiator caspases 8 and 9 then activate the effector caspases 3, 6 and 7, which are responsible for the cleavage of important cellular substrates resulting in the apoptotic phenotype.

MCF-7 cells as a model system

In the present study, a comparative proteomic study has been carried out between a drug-susceptible MCF-7 human breast cancer cell line (parental cell line) and the MCF-7 cell line that has been selected for resistance to mitoxantrone.

The MCF-7 cell line was derived in 1970 from the pleural effusion of a patient with metastatic breast cancer (37). The cell line is one of the most widely used cell lines in laboratories as an investigative tool because of multiple characteristics, including its stability (38).

The drug resistant cell line used in this study is the mitoxantrone resistant MCF-7 cell line (MCF-7/MX). Mitoxantrone is commonly used in treatment of leukemia, lymphomas and breast cancer (40). The structure of this drug is shown in figure 2, which is a synthetic anthraquinone. The exact mechanism of its action is unknown but includes intercalation with DNA to cause inter- or intra-strand DNA cross-linking. It also inhibits DNA topoisomerase II. All these prevent DNA synthesis and repair, thus leading to cell death (43-44). Furthermore, since mitoxantrone contains a quinine functional group in its structure it has been considered to undergo activation by metabolic reduction (48). Quinones can undergo reduction to form semiquinone radicals, which can produce various reactive oxygen species (ROS), such as hydrogen peroxide, in the presence of molecular oxygen. This may ultimately lead to oxidative stress and cell death. Researchers have found that there is low but

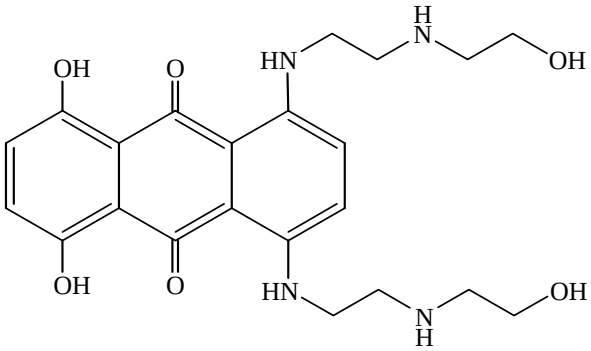


Figure 2. Mitoxantrone structure

some production of ROS by mitoxantrone (49). MCF-7/MX cells, which were provided by Dr. Ken Cowan at National Institutes of Health, were isolated by serial passage of the parental MCF-7 cells in stepwise increasing concentrations of the mitoxantrone (39). MCF-7/ MX cell line is approximately 4000-fold more resistant to mitoxantrone, and also 10-fold cross-resistant to doxorubicin and etoposide. The most well defined mechanism for resistance has been associated with enhanced drug efflux due to overexpression of a characterized ATP-binding cassette transporter protein, BCRP (50).

It was known that mitoxantrone damages the mitochondrial membrane and mitochondrial energy metabolism (51-52). In addition, mitoxantrone can cause mutation and loss of mitochondrial tDNA (mtDNA) (53). Collectively, mitochondria may play a role in mitoxantrone function, as well as the process of mitoxantrone-resistance process.

Part B

Proteome and Proteomics

The human genome has been completely sequenced, and 20,000~25,000 protein-encoding genes have been identified (54). However, researchers are realizing that much of the complexity of the human organism and biological function must rely on the proteins these genes encode—the “proteome” (56-58). The proteome, which was first coined in 1995, was defined as the time- and cell-specific protein complement of the genome (55). The expressed products of a single gene in reality represent a protein population, which may

be regulated at the level of translation, or more often, the accumulation, degradation and posttranslational modification during biological processes. Protein expression is also dynamic in response to external and internal stimuli such as pathological, pharmacological and aging conditions (59). With accumulating evidence that mRNA levels frequently do not reflect the protein levels (57-58), proteomics has become a systematic and indispensable discipline for separating and “visualizing” the protein components within a cell. These include the study of proteins, protein levels, protein-protein interactions, protein modifications of an organism. One of the most interests in proteomic application is disease investigation including development of novel biomarkers for diagnosis, identification of new therapeutics targets (60-62).

The number of different protein molecules in a mammalian organism ranges from 1 million to 20 million (63). Moreover, the dynamic range of protein abundance in biological samples can be as high as 10^6 (64). These complexities require the development of rapid and easily utilized analytical techniques, among which, mass spectrometry has been the most powerful tool in proteomics.

Principles of mass spectrometry in proteomics

Mass spectrometry measures the masses of individual molecules and atoms. The first essential step in mass spectrometry analysis is to convert the analyte molecules into gas-phase ionic species in the ionization source. The

excess energy transferred to the molecule during its ionization process will lead to fragmentation of that molecule. Next, a mass analyzer separates these molecular ions or their charged fragments according to their m/z (mass over charge) ratio. Finally, the number of ions at each m/z value are detected by a suitable detector and displayed in the form of a mass spectrum.

Since the 1990s, mass spectrometry has undergone tremendous technological improvements, which made it applicable to proteins, peptides and other biological related molecules. Among them, electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI) are now the most common ionization sources for biomolecular mass spectrometry (65). They are also referred to as “soft” ionization because they can generate ions from nonvolatile macromolecules such as proteins and peptides without significant fragmentation and even maintain noncovalent interactions under some specific conditions.

Electrospray ionization (ESI) mass spectrometry was developed by Fenn and co-workers in the 1980s (66). In ESI, the sample solution containing a protic primary solvent is sprayed from the tip of a metal nozzle under a strong electric field. The solution is then dispersed into a find spray of charged droplets. Either heat or dry gas is applied to the droplets at atmospheric pressure, causing the solvent to evaporate from each droplet. The charge density on its surface increases as the size of the charged droplet decreases, leading to multiple charged analyte ions. Thus very large molecules (up to

70,000 Da) could be detected by mass analyzers with a relatively small mass range, because mass spectrometry measures the mass-to-charge ratio (m/z). Another advantage of ESI, which made it immediately popular, is that it could be interfaced with liquid-based (e.g. chromatographic and electrophoretic) separation tools. A variation of ESI was developed by Wilm and Mann (67), called nanospray (nanoESI). In this method, the spray needle has been made very small (flow rates are 10~100 nl/min) and is positioned close to the entrance to mass analyzer. The simple adjustment, however, increases sensitivity and efficiency dramatically.

MALDI was first demonstrated for proteins in 1988 by Karas, Hillenkamp (68) and Tanaka (69). In MALDI analysis, the analyte is first co-crystallized with a large molar excess of a matrix compound, which is usually a UV-absorbing organic acid. Irridiation of this analyte-matrix mixture by a laser results in the vaporization of the matrix, which carries the analyte into gas phase. During the desorption process, the analyte molecules are protonated and form singly charged ions predominately. The efficient and direct energy transfer during MALDI ionization provides high sensitivity and a high mass range (up to 300,000 Da). MALDI-MS is normally used to analyze relatively simple mixtures.

As for mass analyzer, the three most commonly used analyzers are the ion trap, time-of-flight (TOF) and quadrupole. In the ion trap technique, the analyte ions are trapped in a region consisting of two end-cap electrodes and a

ring electrode under a variable radio-frequency voltage. Ions are then ejected separately from the ion trap to the detector when the radio frequency is scanned to resonantly excite the ions. Because the ions are all captured for a certain time interval before ejecting to detector, ion traps are very sensitive.

The linear time-of –flight analyzer is the simplest analyzer. When a group of ions are given the same amount of energy through an accelerating poteintial and accelerated to a detector, the lighter and heavier ions reach the detector at different times. The different traveling times are dependent on the mass-to-charge ratios of different ions. TOF has a significant mass range which can be higher than 300,000 m/z. The time-of-flight analyzer is often coupled with a MALDI ionization source as shown in figure 3.

The quadrupole mass analyzer selects ions by varying electric fields between four rods to permit a stable trajectory only for ions of a particular m/z value (70).

The separated ions eventually impinge on a detector. The mass spectrum is a graph of signal intensity against the m/z value. It contains a series of peaks, each corresponding to a particular ion (e.g. peptide or protein), and peak heights that show the relative abundances of the ions. The ions in these peaks can be further fragmented to give a second mass spectrum, referred to as a tandem mass spectrum (MS/MS spectrum). One of the primary fragmentation processes is known as collision-induced dissociation (CID). CID is achieved by generating the ion of interest and selecting it with an analyzer. The ion then

collides with inert gas molecule, leading to the fragments, which are then analyzed as regular ions. In a multiple quadrupole mass analyzer, the subsequent MS/MS analysis is performed by consecutive analyzer. In an ion trap analyzer, however, MS/MS is performed with the same analyzer. Tandem mass spectrometry allows the analysis of a heterogeneous mixture of peptides, because each peptide can be selected for individual fragmentation. The fragment ions, referred to as product ions, can be separated into two classes (Figure 4). One class retains the charge on the N-terminal where cleavage occurs at one of three different positions. There are a-, b- and c- ions. The other class retains the charge on the C-terminal in ions formed by cleavage at three different positions, named as x-, y- and z- ions. Among these, b- and y-ions are the predominant ions in MS/MS spectra because the amide bond is most easily broken. The mass difference between two conjunct b- or y-series ions reflects the mass of an amino acid.

The main instrument used for this thesis is a quadrupole time-of-flight mass spectrometer, which is shown in figure 5. The ESI quadrupole TOF combines the stability of a quadrupole analyzer with the high sensitivity and accuracy of a time-of-flight mass analyzer.

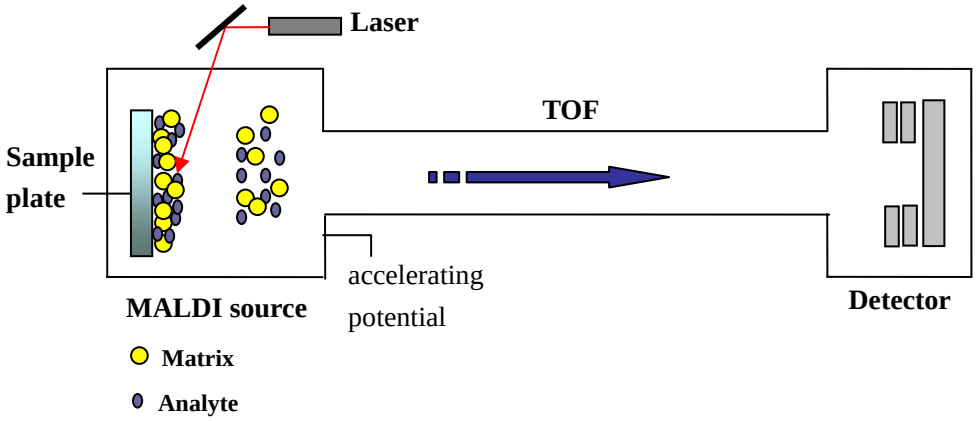


Figure 3. A MALDI time-of-flight mass spectrometer

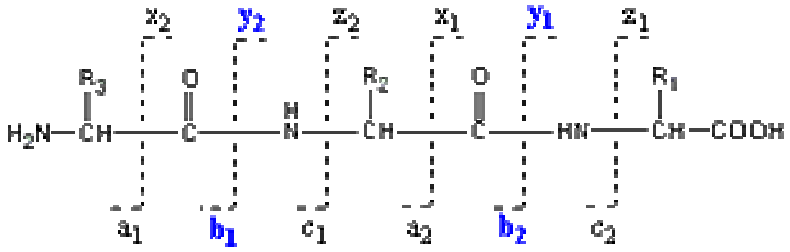


Figure 4. Nomenclature for fragmentation of peptides

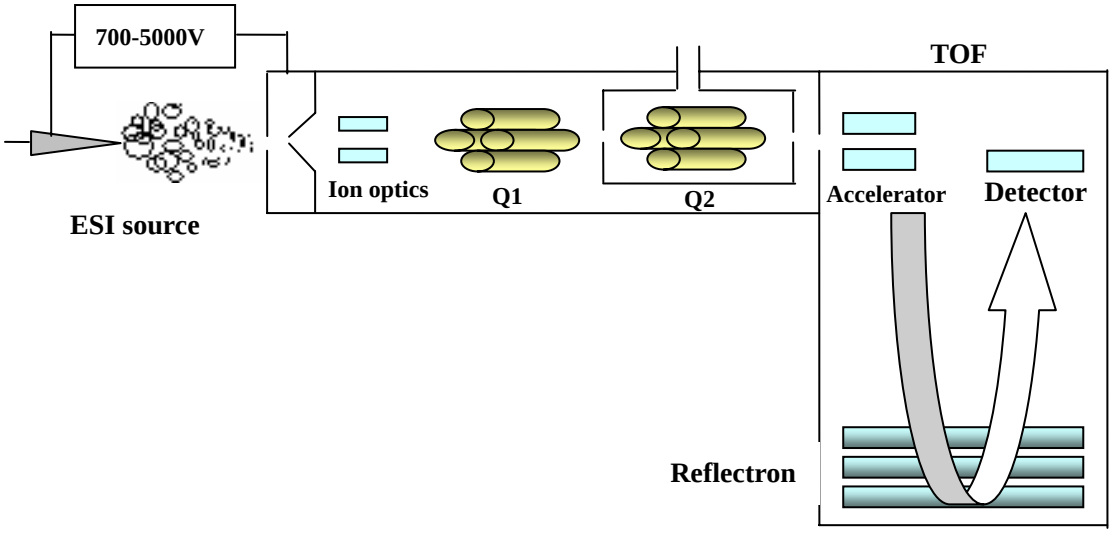


Figure 5. An electrospray ionization quadropole time-of-flight mass spectrometer

The quadrupole TOF instrument combines quadrupole instrument with a TOF section for measuring the Mass of ions. Ion optics focuses the ions, Q1 is the first quadrupole for ion selection and Q2 is the second quadruple for collision induced dissociation of selected ions.

Protein identification and bioinformatics

Protein identification in proteomics refers to distinguishing different proteins by their different amino acid sequences. Conventional sequencing strategies include Edman degradation and *de novo* sequencing (71). In those methods, partial sequences are normally obtained either by automated, stepwise chemical degradation of proteins/peptides or fragmenting peptides by MS. These partial sequences were frequently used for the generation of probes for the isolation of the gene coding for the protein from a gene library (65). With the rapid improvement of high mass accuracy of MS and completion of sequencing of the human and many other genomes, high-throughput protein identification can be achieved by correlating mass spectrometric data with sequences in the databases with the aid of novel search algorithms.

Two mass spectrometric methods are now widely used in protein identification: peptide mass mapping (peptide mass fingerprinting) and peptide tandem spectrometry. In peptide mass mapping, the protein of interest is either enzymatically or chemically cleaved and the resulting peptide masses are measured as accurately as possible in a mass spectrometer. The obtained "peptide mass fingerprint" of the protein is compared to calculated peptide masses obtained by theoretical cleavage of protein sequences stored in databases. In the comparison, one of those available computer algorithms, such as probability based matching (MASCOT) can be applied to assign a score to each match that ranks the quality of the matches (72-73). The

top-scoring candidate can be considered as the identified protein. It has been shown that only a small number of accurately measured peptide masses are required for unambiguous protein identification (74). This approach, however, requires a relatively purified target protein. Thus prior protein fractionation is often necessary.

Protein identification using tandem mass spectrometry is more specific than that achieved by peptide mass mapping, because the pattern of fragment ions in the MS/MS spectrum provides fairly unique sequence information about individual peptides from the peptide mixture. Those observed fragment ions of each peptide are matched against predicted fragment ions of all peptides derived from the proteins in the database using the computer algorithm. The best peptide and thus protein candidate can be determined. This approach can typically only require a single to at most several peptide MS/MS spectra to identify a protein from a complex mixture in a constrained database (75). In addition, it can be easily automated and can also be adapted to find peptides carrying specified post-translational modification by instructing the program to expect modifications at specific residues (76).

Part C

Separation of protein and peptide mixtures

There are two important analytical problems existing in proteomics analysis: dynamic range of protein abundance and diversity of protein expression, such as multiple protein isoforms (59). Although mass

spectrometry including MS and tandem MS offers a powerful separation tool because individual peptide can be selected, isolated and sequenced among the co-detected peptides, to overcome these problems, one of two approaches is usually taken prior to mass spectrometry: 1) Proteins are prefractionated and the intact proteins are introduced into the mass spectrometer for protein identification and analysis (this may be characterized as “top-down” proteomics (77-79)); 2) A complex protein mixture is either separated first and then digested into peptides or first digested and then resolved (both may be characterized as “bottom-up” proteomics and the latter one is often referred to as “shotgun” proteomics (80)). In both approaches, routine and reproducible separation or prefractionation techniques need to be developed. Distinct characteristics of proteins such as molecular weight, shape, solubility and hydrophobicity can guide the design of separation techniques. Figure 6 shows the general flow chart of proteome analysis including the commonly employed separation tools.

1. Separation of a specific protein or a group of proteins

Affinity liquid chromatography or capillary electrophoresis (CE) is effective method to specifically isolate and enrich specific target protein(s) (81). The principle of affinity columns is based on the ability of biologically active substances (affinant) to bind specifically and reversibly with complimentary substances (protein of interest). Those affinants attached to the surface of the columns may include metals, lectins, antibodies, etc. When protein mixture is

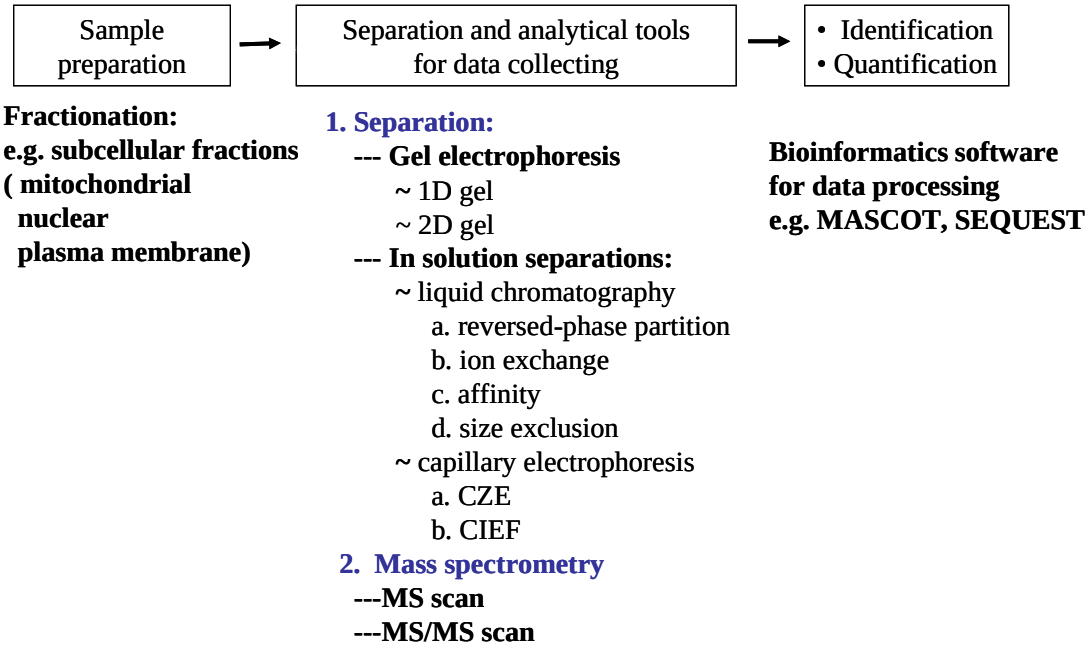


Figure 6. Schematic of proteomic analysis

subjected to the column, the proteins of interest will be captured by the affinant and all the other molecules will pass through the column. The captured protein can later be eluted by changing the properties of the buffer, pH or temperature, etc. Many affinity columns are designed to isolate post-translationally modified proteins/peptides, which are normally present at very low abundance in complex mixture (82). For example, immobilized metal affinity columns (IMAC) loaded with Fe (III) or Ga (III) ions has been successfully used for the isolation of phosphorylated proteins/peptides (83-84). Immunoaffinity chromatography is another powerful technique for isolation of phosphorylated molecules (85). Glycoproteins, which also play an essential role in the protein functions and cellular processes, can be separated by lectin affinity or antibody chromatography (86).

2. Multidimensional separations of proteins/peptides

Resolving a complex mixture from a cell, tissue or organism can not be achieved by a single chromatographic or electrophoretic method. It is clear that multidimensional separation tools should be combined to fractionate proteins for mass spectrometric analysis.

A. 2-D gel electrophoresis

Two-dimensional gel electrophoresis (2-D gel), a technique 30 years old, is still the most used multidimensional separation technique for cell and tissue proteins. Proteins are separated in the first dimension by isoelectric focusing on an immobilized pH gradient strip and in the second dimension by their

molecular weights using sodium dodecylsulphate-polyacrylamide gels (SDS-PAGE). Inorganic or organic dyes such as silver stain and Commassie blue may be applied to visualize the proteins in 2D gel after separation. Individual protein spots on the gel can be excised and subjected to proteolytic digestion (usually trypsin). After that, the resulting peptides can be removed from the gel matrix and analyzed with subsequent peptide mass mapping or tandem MS, leading to protein identification and other characterization. Thousands of proteins including distinct post-translationally modified protein isoforms can be separated, visualized and quantitated in a single 2D gel run (58, 87). However, it is widely recognized that 2D gel method also suffers from several technical limitations. Specific classes of proteins, including proteins with extremes in pI, molecular weight, hydrophobicity, and low abundant proteins have been known to be excluded in 2D gel. In addition, 2D based methods are relatively tedious and time consuming.

B. Multidimensional separations in-solution

Alternative two-dimensional separation systems in solution (“gel-free” approaches) have emerged to interface protein and peptides separations directly to mass spectrometers. One of the most promising approaches is termed the “shotgun” technique. The components of a proteomic sample are digested with a suitable protease and the resulting peptides are resolved by multidimensional separation techniques, which employ a combination of two or more different separation steps before introduction into a mass spectrometer.

One advantage is that peptides are more soluble and easier to separate than the parent proteins. The disadvantage is that the increase in the number of peptides makes the sample mixture more complicated than the original protein pool, which must be overcome by subsequent high resolution and comprehensive analytical techniques. Currently in-solution separation strategies are based on two techniques: liquid chromatography (HPLC) and gel-free electrophoresis.

B-1: HPLC separations

High performance liquid chromatography is highly compatible with mass spectrometry and there is a broad selection of stationary and mobile phases, which makes LC a versatile and fundamental tool in proteomics. There are four main types of LC: reversed-phase, ion-exchange, affinity and size-exclusion chromatography.

Currently, most LC separations in proteomics are achieved by reversed-phase partitioning HPLC (commonly termed reversed-phase HPLC or RPLC). Proteins/peptides stick to reverse phase HPLC columns in high aqueous (usually water) mobile phase and are eluted from the columns with continuously increased organic component such as acetonitrile, methanol in the mobile phase. Analytes are separated based on their hydrophobic characteristics. The least polar molecules are flushed out last. The specificities of HPLC columns are affected by the stationary phases, which are generally made up of hydrophobic alkyl chains that interact with the analytes. There are

three common chain lengths, C4, C8 and C18. The larger protein molecule will have more hydrophobic regions to interact with the column and thus a shorter chain length is more appropriate. So C8 and C18 are used for peptides while proteins are usually separated by a C4 column.

To improve resolution, reversed phase HPLC is often coupled with another LC column(s) to give multidimensional separation. In these methods, ion-exchange chromatography is used most often as the first chromatographic dimension.

Ion-exchange chromatography relies on charge-charge interactions between the sample and the functional charged sites immobilized on the resin. Ion-exchange chromatography can be classified into cation exchange chromatography, in which positive-charged ions bind to a negative-charged resin such as $–SO_3^-H^+$ and anion exchange chromatography, where the binding ions are negative and the immobilized functional group such as $–N(CH_3)_3^+$ OH is positive. The analytes are bound to the column with the starting buffer (e.g. NaCl) of low ionic strength, and then the bound molecules are eluted off using a gradient of a second buffer which steadily increases the ionic strength. Proteins/peptides with lower charge strength elute earlier.

Each fraction from strong Cation-exchange chromatography is often directly coupled with reversed-phase LC-MS (on-line SCX-RPLC-MS). Yates and his co-workers (80, 88-89) have developed the multidimensional protein identification technology (MudPIT). In this approach, digested peptides from a

protein mixture are separated using a “mixed-bed” microcapillary column packed with moieties of strong cation-exchange and reversed-phase stationary phase, which are interfaced back to back (Figure 7). Peptide fractions were eluted from the SCX beads onto the reversed-phase material using a salt gradient. Between salt steps, an organic solvent gradient was used to elute peptides from the RP column into the electrospray MS. The MudPIT has emerged as an effective and robust tool.

Affinity chromatography, as mentioned above, may also be combined with RPLC to study a particular set of proteins, such as low abundant post-translationally modified proteins/peptides, or to remove high-abundance components that mask other proteins (86, 90-91). For example, Apffel et al. interfaced concanavalin A affinity chromatography to reversed-phase RPLC in order to separate tryptic digests of serum glycoproteins prior to mass spectrometric detection (90).

Size-exclusion chromatography (SEC) separates proteins and other biological macromolecules on the basis of their molecular size. The column is packed with small porous polymer beads designed to have pores of different sizes. As the analytes travel down the column some particles enter the pores. The larger the particles, the fewer pores can they enter. Thus they will have less overall volume to pass through, resulting in faster elution time. A few reports have used a SEC-RPLC-MS technique for proteome research (92-94).

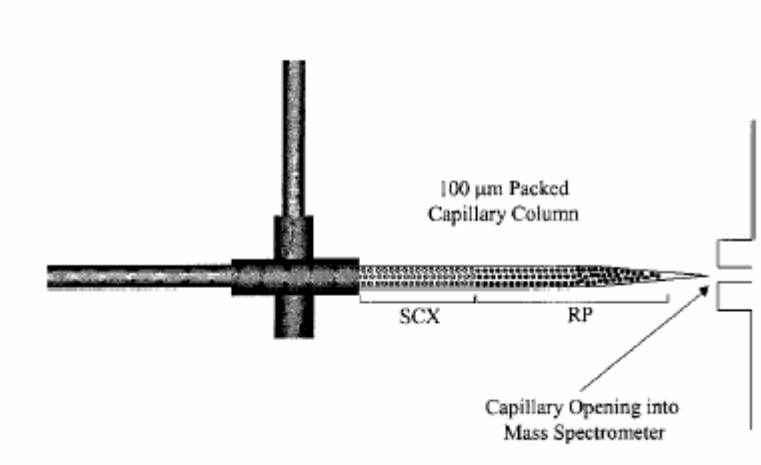


Figure 7. Multidimensional protein identification technology (MudPIT) electrospray interface including a biphasic microcapillary column packed with strong cation-exchange and reversed-phase packing material (80).

Gao et al. use size exclusion to separate whole cell lysates, and following proteolysis of these fractions, each was separated by RPLC, which was coupled on-line via electrospray to an ion trap mass spectrometer (92). The smaller number of applications of SEC is due to its low resolution and limited loading capacity (82).

B-2: Electrophoretic separations

Electrophoresis separates proteins in an electric field based on the differences in their charge, or mass-to-charge ratio. There are gel-based (mentioned above in the 2D gel separation) and liquid-based systems that use this approach. Among the liquid-bases systems, capillary electrophoresis and solution isoelectric focusing are the most two promising methods in MS-based proteomics.

i) Capillary electrophoresis (CE)

CE separation depends on the different rates at which analyte ions migrate through narrow-bore capillaries under an electric field. The mobility of each ion is determined by its charge and size. It is well known that CE provides high speed and high resolution separation (95-97) of proteome-wide proteins and their peptides. There are two frequently applied modes in CE: capillary zone electrophoresis (CZE) and capillary isoelectric focusing (CIEF).

Capillary zone electrophoresis is the simplest and most widely-used mode in CE. In CZE the cationic and anionic analytes are attracted towards different directions. However, they are all moving to the cathode because the

electroosmotic flow is usually significantly higher than the analyte velocity due to its electrophoretic mobility. Cations elute first because the direction of their migration is the same as the direction of EOF. Neutral molecules elute next without being resolved as they only move with EOF, while anionic analytes elute last. Among same-charge-sign ions, those with the highest charge/mass ratio will migrate first.

CZE-MS has shown as a high efficient method in proteomics (98-100). However, the requirement for a background electrolyte in CZE results in a lower dynamic range of measurements than LC-MS (101). In multidimensional separations, CZE is often used as a final dimension before MS due to its fast separation and limited loading capacity (95).

Capillary isoelectric focusing (CIEF) separates amphiprotic molecules based on differences between their isoelectric points (pI). CIEF is performed by filling the capillary with the sample and a mixture of carrier ampholytes with a certain pH range. An electric field is applied across the capillary with a basic solution at the cathode and an acidic solution at the anode. The molecules will then migrate to the point where their net charges are zero (pH=pI). As a result, analytes are separated into narrow zones of different pH through the capillary.

CIEF-MS has been successfully applied to a number of complex proteomic systems (102-105). Larger loading capacity than CZE can be achieved, leading to good sensitivity and sample is concentrated in the process by a factor of two or three orders (102). But it was found that the

carrier ampolytes might interfere with MS interfacing (95). In multidimensional separations, CIEF is suitable to be the first dimension since it is an equilibrium technique.

ii) Solution isoelectric focusing

Methods of solution isoelectric focusing are electrokinetic strategies, which are performed in free solution relying on isoelectric focusing steps similar to CIEF. The applied devices include multichamber apparati, such as the multicompartment electrolyzer with immobilized membranes (MCEs) and rotationally stabilized focusing apparatus (Rotofor) (106).

The Rotofor system has a long history in solution isoelectric focusing methodologies. It was developed by Bier (107). The Rotofor cell consists of a cylindrical focusing chamber that holds a plastic core dividing the chamber into 20 compartments separated by polyester screens. The screens are resistant to fluid convection but they do not prevent the flow of current or the migration of proteins. Proteins and carrier ampholytes are loaded into the assembled chambers and then migrate in response to an electrical field to the compartments that are at pH values nearest to their pIs. The whole chamber rotates in order to avoid overheating and gravity effects during the separation process. Each fraction with narrow pI range can be subjected to further separation. Researchers have reported combining Rotofor separation with HPLC to analyze protein mixture (108-110).

Righetti described a multicompartment eletrolyzer (MCE) where each

compartment is separated by a polyacrylamide gel membrane with a specific pH (111-112) maintained by immobilized ampholytes. Proteins/peptides in the solution are fractionated under an electric field by migrating into different compartments according to their pIs. That is, only peptides with pIs between the limited pH values of the boundary membranes remain in each chamber. Some advantages of such a separation will be: 1) the isoelectric membrane can be specially designed to select the pH range and degree of fractionation desired (113-115) 2) the chance of precipitation is small in solution 3) sample loading is quite flexible and even a peptide mixture with a large dynamic range is expected to be well separated and concentrated. The limitation is that automation is hard when MCE is coupled with other separation methods on line. This dissertation outlines a shotgun method with the combination of MCE in the first dimension and RPLC-MS as the second dimension for analysis of mitochondrial peptides.

Part D

Quantitation

Although identification of proteins is often a necessary first step in proteomics approaches to elucidate protein function, measuring protein abundance levels is crucial to obtain a complete picture of many biological processes, and is referred to as quantitative differential proteomics. Cells are dynamic. The abundance of proteins within cells is not only regulated by transcription but also by translation and post-translation at events. Since

studies showed that there is no good correlation between mRNA quantities and protein quantities (116-118), globally profiling protein abundances becomes an essential part of proteomics. Previously, quantitative studies were limited to smaller numbers of proteins and were mainly dependent on the use of antibodies, which may be considered as “semiquantitative” because binding affinities of antibodies vary and antibodies against all proteins are not available (119). With well-established mass spectrometry and rapid methodological developments, large-scale quantitative proteomics strategies are able to produce numerous analyses that would have taken decades to measure with classical methods.

1. Relative quantitation

It is often not necessary to measure the absolute amount of protein present but rather to compare the relative abundances of proteins among related states. This is referred to as comparative proteomics (120-121). An important goal in this field is to provide a snapshot of protein expression within a cell in response to biological perturbations such as disease state, drug treatment, and aging (121) and thus understand the dynamics of living organisms.

A. Gel-based methods

For a number of years, comparative proteomic studies have been implemented by contrasting the position and intensities of protein spots among parallel gels derived from different samples. Computer software (e.g.

Compugen Z3) can make digitized images of the stained gel arrays and evaluate protein abundance changes by differential densitometry (122-123). Another approach involves labeling of two samples with two different fluorescent dyes (e.g. Cy-3 and Cy-5) and differentially imaging the fluorescent emission with optical filters (124-125). Although scanning of spots in 2-D gels is a straightforward process, which avoids possible downstream interference and is able to detect intact proteins and may distinguish protein isoforms (126). There still are some inherent problems in the method. For example, more than one protein might be present in a single spot. The cost of the dyes and the equipment used for visualization restrict their use.

B. Non-gel-based methods

In the last five years, there are widespread efforts to develop mass-spectrometry based approaches for comparative proteomics. Among these, chemical, metabolic or enzymatic stable isotope labeling of proteins/peptides in concert with isotope ratio measurement by MS has been the most rapidly advancing. A common theme in these methods is that chemically identical but mass-differentiated stable isotope tags are introduced into the proteins/peptides in two or more sample mixtures. The mixtures are combined either before or after proteolysis. Mass spectrometric analysis is then performed and differences in abundances of proteins are derived by comparing the ion intensities of isotopically labeled peptide pairs. There are two main classes in these MS based approaches: *in vivo* labeling and *in vitro*

labeling. A third MS-based approach intends to provide absolute and relative quantitation without the use of isotope-labels.

B-1: in vivo labeling

Incorporation of stable isotopes (e.g. ^{2}H, ^{13}C, ^{15}N) into newly synthesized proteins using isotope-substituted culture media is a well-established technique. Two groups of cells are grown in two separate culture media identical in all respects except that one of the media contains the “heavy” stable isotopes (e.g. ^{2}H, ^{13}C, ^{15}N). After multiplying the cell population, eventually, proteins synthesized in each cell group will be incorporated with “heavy” or “light” stable isotopes. Metabolic labeling has a higher fidelity than labeling *in vitro* (incorporating nearly 100% efficiency) and eliminates sample-to-sample variability derived from subsequent biochemical experiments (119). However, this method only works in cell culture systems that allow the incorporation of isotope-substituted media, and may not be suitable for animal tissue or body fluids. Furthermore, in the case of ^{15}N labeling, both the backbone and side-chain nitrogens are labeled, which makes the mass difference between labeled and non-labeled peptides unpredictable without a previous knowledge of the peptide sequence (127). Several laboratories introduced a modified version of the same overall approach, termed SILAC by Mann (128), for stable isotope labeling by essential amino acids in cell culture, which makes the mass shift between tryptic peptide pairs predictable. For example, Fenselau et. al have reported

that growing the parental MCF-7 cell line in ^{13}C6-arginine and ^{13}C6- lysine medium resulted in C-terminal labeling of all tryptic peptides and a 6 Dalton difference compared to their ^{12}C-counterparts (129).

B-2: in vitro labeling

There are far more versions of *in vitro* labeling than *in vivo* labeling. The former is mostly achieved by chemical derivatization of primary amine or C-terminal carboxyl groups in proteolytic peptides. Here I focus on those strategies being used most extensively.

i) Acylation

One of the initial labeling approaches involved the global coding of tryptic peptides through alkylation of their N-terminal amino group and ε-amino group on lysine. For example, the technique, also known as global internal standard technology (GIST), used $^{2}H_3$- and $^{1}H_3$- forms of N-acetoxysuccinimide ester for differentially labeling (130). The MALDI-MS detection sensitivity with C-terminal lysine-containing peptides, however, is much more limited due to the neutralization of the positive charge by acylation (120, 131). Furthermore, because peptides coded with $^{1}H_3$-and $^{2}H_3$-acetate differ by only 3 Dalton there is small overlap between the M+3 isotope peak of the $^{1}H_3$-acetate labeled and the monoisotope peak of the $^{2}H_3$-actate labeled peptide. The overlap needs to be considered to enable accurate isotope ratio measurements (120).

ii) Isotopically coded affinity tags (ICAT)

This approach involves labeling cysteine-containing peptides from different samples with “light” versus “heavy” forms of a reactive chemical, which differ by 8 Dalton (132). The ICAT reagent consists of a biotin group followed by a linker and terminated with a cysteine-reactive group. The only difference between the light and heavy tags is the presence of eight hydrogen or deuterium atoms in the linker region. The samples are combined and enzymatically digested, and the labeled peptides (cysteine-containing peptides) are selectively enriched via bio-avidin affinity chromatography. The specificity and advantage of this technique is that it largely reduces the complexity of the peptide mixture. Some successful applications in the study of low-abundant proteins and membrane bound proteins have been reported (133-135).

The major limitation of ICAT is that it is only useful to examine concentration or structural changes in cysteine-containing peptides, and usually protein identifications must be based on a single peptide because cysteine is generally present in only 10-20% of the peptides derived from a proteome (120, 127). In addition, the peptide isoforms from labeled and unlabeled ICAT regents are partially resolved during HPLC separation. A second generation of ICAT regents was designed to overcome these problems (136-137).

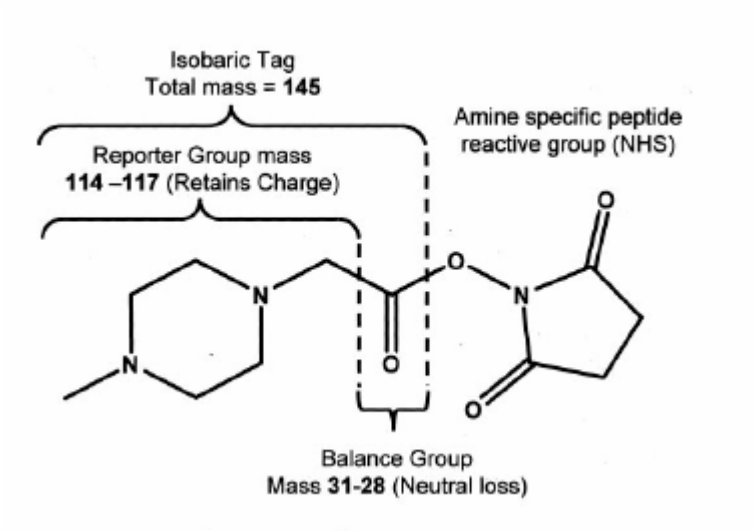
iii) Multiplexed, isobaric stable isotope tags (iTRAQ)

The reagents (total of four) were designed to measure up to four samples simultaneously. They consist of a charged reporter group that is unique to

each of the four reagents (mass range 114 -117 Da), a neutral balance group (mass range 31-28 Da) to maintain an overall mass of 145 for each reagent and the same peptide reactive group (Figure 8A) (151). The reactive group forms an amide linkage to any amino-termini and the ε-amino group of lysine side chains. Figure 8B depicts the general labeling procedure. Each individual protein sample is digested and labeled with one of the multiplex set separately. Then the four peptide mixture are combined and analyzed by LC-MS/MS. Four identical peptides from the four different samples will give a single, unresolved precursor ion in MS because the overall mass of each labeling reagent is identical. But when they are subjected to fragmentation (e.g.CID), the neutral balance group is lost from each peptide, leaving four different charged reporter groups (m/z range from 114-117 Da) and all other sequence-informative fragment ions (e.g. y-, b- ions) which remain as additive isobaric signals. The relative concentration of the peptides is thus derived from the relative intensities of their corresponding reporter ions. In contrast to ICAT and similar mass-different labeling methods, quantitaion of iTRAQ is performed at the MS/MS stage rather than in MS.

There are many advantages of the iTRAQ approach: Up to four samples can be compared at the same time; isobaric peptides were generated in MS without increasing the complexity of mixture; sensitivity is highly increased in MS/MS spectra since all the peptide backbone fragments' ions are isobaric and additive, leading to better identification. One drawback of this approach is

A:



B:

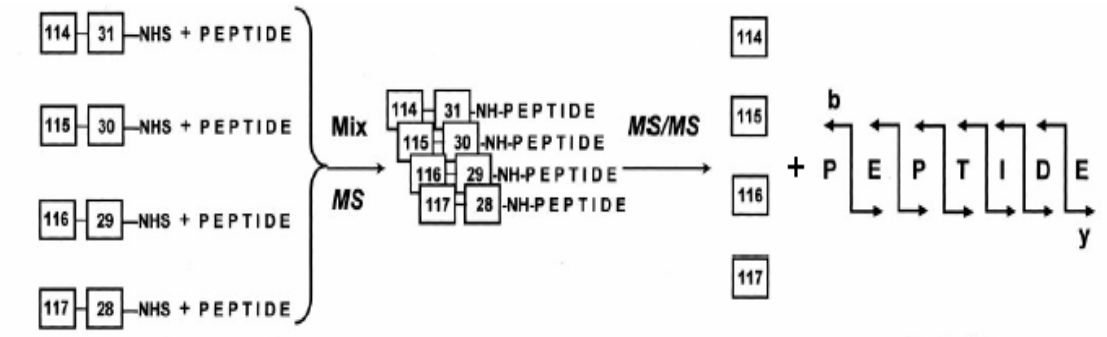


Figure 8. Diagrams of iTRAQ reagents and workflow (151). A, Diagram showing the components of the multiplexed isobaric tagging chemistry; B, Illustration of the isotopic tagging used to arrive at four isobaric combinations with four different reporter group masses

that MS/MS spectra must be obtained, which requires more analysis time than performing result-dependent analysis only on differentially expressed peptide pairs in MS (e.g. ICAT) (151) and good tandem mass spectra can not be obtained for all the peptides.

iv) ^{18}O labeling

Another simple and robust global labeling strategy is to incorporate two atoms of ^{18}O from $H_2{}^{18}O$ into new carboxy-termini of peptides during protein proteolysis, resulting in a 4 Da mass increase (138-139). All peptides will be labeled except the peptide originating from the C-terminus of the protein. A variety of serine proteases including trypsin, endoprotease Glu-C, endoprotease Lys-C are able to catalyze the exchange of oxygen atoms (140). These enzymes form a covalent tetrahedral intermediate with the peptide bond of the protein, and the first ^{18}O atom is incorporated through the hydrolysis of the peptide bond. After proteolysis is complete, these enzymes continue to form reversible covalent intermediates with the carbonyl group of the nascent peptide product and then a second ^{18}O atom is introduced by pseudohydrolysis (140-141). The repeating binding/hydrolysis cycles result in complete equilibration of both oxygens in the C-terminus of the peptides with oxygens from solvent water. This mechanism can be seen in figure 9. This method has been shown to have at least femtomolar sensitivity (127, 131). Chemical back-exchange is reported to happen only under extreme acidic and basic conditions (142). The catalytic enzyme must be removed before the

sample is exposed to $H_2{}^{16}O$. This is easily done if immobilized trypsin or other protease is used.

^{18}O labeling was originally applied to aid peptide *de novo* sequencing via mass spectrometry (143-144). Fenselau and co-workers developed and applied it to quantitative proteomic studies (141), which has also been used by other groups (145-149). An improved variation conducts the proteolysis step in ^{16}O water and a postproteolysis labeling incubation in ^{18}O water (150). Thus the proteolytic peptides can be labeled with a limited amount of ^{18}O water and both digestion and labeling conditions can be optimized separately.

^{18}O labeling has a number of advantages over other metabolic and chemical labeling methods: it is a natural consequence of protein proteolysis involving only one step with none of the exotic reagents or side reactions inherent to chemical labeling (145). The concentration of ^{18}O water, approaching 55 M, drives the reaction towards completion even for low-level proteins; every proteolytic peptide (except those containing the parental C-termini) is labeled, which renders global information for protein identification and quantitation. As the other labeling strategies, ^{18}O labeling also has its downsides including that the ^{18}O exchange will be inhibited by high concentrations of urea; the rate of labeling differs with peptide sequence (130).

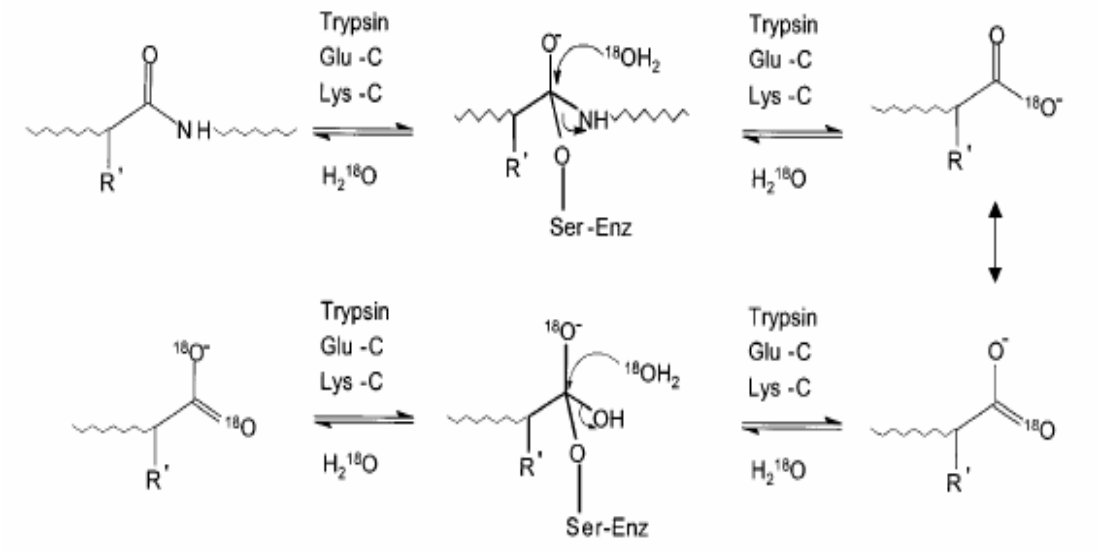


Figure 9. Mechanism of incorporation of ^{18}O into proteolytic fragments (26)

B-3: Label-free approach

Even though stable isotope coding continues to be the most powerful approach in comparative proteomics, there are disadvantages including the cost of isotopic labeling reagents, time-consuming labeling and the requirement for pairwise comparisons among samples, which complicates multiple and retrospective comparisons(152). Recently, label-free protein quantitation techniques have been proposed. Some studies have shown that peak intensities of peptide ions in mass spectra correlate well with protein abundances in complex samples (152-155). For example, Bondarenko *et. al.* described an approach of adding peak areas of identified peptides from one protein to define the total reconstructed peak area (153). The total reconstructed peak is further normalized to the peak area of an internal standard protein digest present in the mixture at a constant level. They tested this method using human plasma and demonstrated linear responses of peptide ion peak areas between 10 and 1000 fmol of internal standard with a relative standard deviation <11%. In another approach, Liu *et. al.* found that spectral counting, which compares the number of MS/MS spectra assigned to each protein accurately reflects relative abundance with a liner correlation over a 2 order of magnitude linear dynamic range (155). Later, Old *et. al.* combined these two methods and showed that protein ratios determined by spectral counting agreed well with those determined from peak area intensity measurements, and both agreed with independent measurements based on gel staining intensities (152). But it seems that for real biological samples, such

as serum only the abundant proteins can be analyzed using these methods. Enrichment of the low abundant proteins is probably required prior to analysis.

2. Absolute quantitation

An important application of absolute quantification is as an alternative to immunoassays. This is much more difficult than relative quantitation. It is normally achieved by spiking known amounts of a synthetic, isotopically coded proteolytic peptide as an internal standard (157-161). Thus prior identification of the studied peptide/protein and estimation of its absolute amount are often needed. Especially, when analyzing a complex mixture, a large investigation is required for choosing more than one standard to “mimic” the proteins to be analyzed (162).

Part E

Hypothesis and objectives

The development of drug resistance presents a major problem in the chemotherapy treatment of cancer patients. Understanding the mechanisms of drug resistance is a key step to improve clinical treatment. It is now believed that any mutation that disables apoptosis can produce drug resistance. Mitochondria play a central role in the apoptosis process. Many proteins associated with mitochondria are involved in promoting or evading apoptosis. We hypothesize that an investigation of the mitochondrial proteome of drug susceptible and drug resistant breast cancer cells should obtain insight into the involvement of the mitochondria in the mechanism of acquired drug resistance.

The present research aims to study changes in mitochondrial proteins associated with drug resistance in human cancer cells. Our specific objectives include:

1) Develop a reproducible method to extract soluble mitochondrial proteins form MCF-7 cancer cell mitochondria.
2) Identify mitochondrial proteins using shotgun analysis consisting of a two-dimensional peptide separation based on solution isoeletric focusing and reversed-phase HPLC-MS. This methodology can be compared with a gel-based strategy previously executed in our laboratory.
3) Integrate forward and reverse ^{18}O labeling with the two-dimensional separation method to identify proteins with altered abundances between a drug-susceptible and a mitoxantrane-resistant MCF-7 cancer cell line.
4) Consider mechanisms of drug resistance based on the functions of those altered proteins.

Chapter 2: Experimental

Materials:

The drug susceptible and mitoxantrone resistant MCF-7 cell lines were provided by Dr. Ken Cowan (Eppley Institute, University of Nebraska Medical Center, Omaha, NE). Cell culture flasks were obtained from Corning (Corning, NY). MEM was from ATCC (Manassas, VA). Distilled water came from Milli-Q water (Billerica, MA). Fetal bovine serum was obtained from Atlanta Biologicals (Lawrenceville, GA). Ultracentrifuge tubes were from Beckman Coulter (Fullerton, CA). Percoll and IPG buffer (pH 3-10) were purchased from Amersham Biosciences (GE Healthcare, Piscataway, New Jersey). Formic acid (88%), o-phosphoric acid (85% HPLC) and ammonium bicarbonate were obtained from Fisher Scientific (Hampton, NH). Micro Bio-Spin 6 chromatography columns and protein assay dye reagent concentrate were obtained from Bio-Rad (Hercules, CA). PepClean^{TM} C-18 spin columns and bovine serum albumin (2mg/ml) were purchased from Pierce (Rockford, IL). Modified porcine trypsin (sequence grade) was purchased from Promega (Madison, WI). Ultrafiltration membrane (500 Da molecular weight cut off) came from Millipore (Billerica, MA). Multi-compartment eletrolyzer (MCE) kit was from Proteome Systems (Woburn, MA). Isotopically enriched $H_2{}^{18}O$ (>95% ^{18}O) was from Isotech, Inc. (Miamisburg, OH). Poroszyme bulk immobilized trypsin was purchased from Applied Biosystems (Foster City, CA).

Centrifugal devices (0.45 μ m membrane) were obtained from Pall Corporation (East Hills, NY). Sucrose (>99%), penicillin streptomycin solution, 25% trypsin-EDTA solution, 0.4% trypan blue solution, HEPES (minimum 99.5%), D-manitol (A.C.S. reagent), sodium hydroxide (A.C.S.), Lysine-L, Argine-L (98% TLC), TFA (99%), ACN (HPLC), α -cyano-4-hydroxycinnamic acid, Trizma base (minimum 99.9%), PBS (pH 7.4), urea, thiourea, DTT (99%), CHAPS, EGTA (minimum 97%), chicken egg white were obtained from Sigma Aldrich (St. Louis, MO).

Equipment

The tissue grinder (7ml) was obtained from Kimble Kontes, Fisher Scientific (Hampton, NH). The SpeedVac was from Thermo-Savant (Holbrook, NY). The Optima LE-80K preparative ultracentrifuge and DU 530 UV-Vis spectrometer were from Beckman Coulter (Brea, CA). The sonicator cell disruptor (Model W 185 F) was from Heat Systems-Ultrasonics Inc. (Plainview, NY). The separation device and electrophoresis tank for solution Isoelectric focusing were from Amika Corp. (Columbia, MD). Power supply (power PAC 3000) for IEF was from Bio-Rad (Hercules, CA). The LCQ-TOF mass spectrometer was from ThermoFinnigan (San Jose, CA) LCQ Deca XP ion-trap mass spectrometer equipped with a nano-LC electrospray ionization source. The ESI-TOF mass spectrometer was from Applied Biosystems (Foster City, CA) Qstar Pulsar i with a nanospray ion source from Protana

(Odense, Denmark), and the AXIMA-CFR MALDI-TOF mass spectrometer was from Kratos, Shimadzu Corporation (Chestnut Ridge, NY). The HPLC with UV detector was from Shimadzu SCL-10 AVP instrument (Columbia, CA) with a Phenomenex 250mm × 4.6mm, 5μm, C_{18} column (Torrance, CA). Nano-LC-MS/MS was performed on ESI-TOF connected on-line with the nano HPLC system, which was consist of Ultimate HPLC, famous autosampler and Switchos II all from LC Packing/Dionex (Sunnyvale, CA). The HPLC columns used in nano-LC were also from LC packing. The SilicaTip needle was from New Objectives (Woburn, WA). BioAnalyst software was from Applied Biosystems. MASCOT in-house search engine was purchased from Matrix Science (London, UK).

Methods

Cell culture and harvest

MCF-7 cell lines were cultured in house. The cells were grown on 150 cm^2 flasks in MEM solution containing 10% fetal bovine serum and 1% penicillin streptomycin solution at 37°C and 5% CO_2. The cells were harvested at 95% confluence (20 flasks per harvest). After removal of the growth medium, cells from each flask were washed twice with 25 ml PBS. Then 3ml trypsin-EDTA solution was added and incubated at 37℃ to dislodge cells from the flask. After 3 minutes 13 ml MEM was added to stop tryptic activity. The resulting mixture was transferred to pre-weighed centrifugation tubes and centrifuged at 500g for 5 minutes. Cell pellet was obtained and then washed twice with PBS.

Isolation of mitochondria

All work in this part was done on ice unless specified otherwise. The method of isolating crude mitochondria in this study is modified from the Mitochondria Isolation Kit (Sigma) as follows. The cell pellet was suspended with 10 volumes of extraction buffer containing 10mM HEPES, pH 7.5, 70mM sucrose, 200mM mannitol and 1mM EGTA. The suspension was homogenized using a 7 ml Tenbroeck tissue grinder. This cell lysate was stained with 0.4% trypan blue solution to monitor the degree of lysis. Upon 75% lysis the homogenate was then centrifuged at 600g for 5 minutes. The supernatant liquid was transferred to another centrifuge tube. The pellet was re-suspended in 5 volume of extraction buffer and homogenization was repeated. Another supernatant was obtained after spinning the homogenate at 600g for 5 minutes. The two supernatants were combined and centrifuged at 10000g for 20 minutes to pellet the crude mitochondria fraction.

The crude mitochondrial pellet was re-suspended in extraction buffer (1mg/ml). One milliliter aliquot of the suspension was loaded in centrifuge tube containing 20ml of 30% percoll in extraction buffer (163). The mixture was spun down at 95000g for 30 minutes in a preparative ultra-centrifuge. Mitochondria were collected from the lower fraction and washed twice by being diluted into a ten-fold volume of extraction buffer and centrifuged at 10000g for 20 minutes. The final resulting pellet was the purified mitochondrial fraction.

Extraction of mitochondrial proteins

All work in this part was also done on ice unless specified otherwise. The purified mitochondria pellet was re-suspended in 10 mM HEPES buffer (pH 7.4) in a ratio of 5 ml solution every 1 gram of mitochondria pellet (163). The suspension was vortexed vigorously for 15 seconds every 5 minutes, for a total 25 minutes. An equal volume of 1.4 M sucrose was then added to the sample with further incubation for 20 minutes. The resulting mixture was then subjected to treatment for 15 seconds with Sonicator cell disruptor. After one minute rest, another 15-second sonication was repeated. The sonicated suspension was centrifuged at 15000g for 15 minutes. The obtained pellet containing unbroken mitochondria and mitoplasts was discarded. The supernatant was diluted with an equal quantity of 100 mM Tris buffer (pH 7.5), then centrifuged at 10500g for 50 minutes. The resulting pellet contained mitochondrial insoluble protein and the supernatant constituted the soluble fraction. Both fractions were stored at -80℃.

Protein assay

The protein assay for the mitochondrial soluble fraction was based on the Bradford method (164). A series of bovine serum albumin dilutions: 0, 1.5, 2.5, 5, 7.5, 10 μg/ml was used to generate the standard curve. The unknown sample (from mitochondrial soluble fraction) was diluted properly to obtain a concentration which is within the range of standard curve. Two sample

dilutions were assayed in triplicate for accuracy and reproducibility. Bio-Rad protein assay dye reagent was added to each dilution solution at a volume ratio of 1:4 with incubation for at least 5 minutes and no more than 1 hour (165). Absorbance was read at 592nm on a DU UV/Vis spectrometer. The protein concentration was calculated based on absorbance and dilution factor.

Digestion of Soluble Protein

The soluble protein fraction was denatured with 6M urea, reduced with two 45 minute incubations with 10mM DTT and alkylated by reaction with 100mM IAA in the dark for 1 hour. The resulting sample was then desalted using Bio-Spin 6 size exclusion columns. Each column was centrifuged at 1000g for 2 minutes to remove the packing buffer. A solution of protein sample (no more than 100 μl) was loaded directly into the center of the column. The column was then centrifuged at 1000g for 4 minutes. The eluate contained the desalted protein in 10mM Tris buffer (pH 7.4). A protein assay was performed again to measure the recovery of sample. Protein was then digested by modified porcine trypsin (1μg trypsin per 50μg protein) at 37°C for 16 hours.

Proteolytic ^{18}O Labeling of mitochondrial soluble proteins

Immobilized trypsin was washed three times with a five-fold volume of distilled water and centrifuged at 1000g for 1 minute each time. Supernatants were discarded. To each digested peptide pool a volume of the clean immobilized trypsin was added equal to 20% of the sample volume. The

mixture was then dried completely by SpeedVac. The pool was redissolved in 20 % acetonitrile and 80% $H_2{}^{18}O$. The unlabeled peptide counterpart was redissolved in 20% acetonitrile and 80% $H_2{}^{16}O$. The solution was rotated at 37°C for approximately 5 hours on a bench-top rotator. The resulting peptide pool was then filtered with a centrifugal filtering device (0.45 μ m membrane) to remove immobilized trypsin. The labeled and unlabeled peptide filtrates were combined with a 1:1 ratio. The mixture was then dried by SpeedVac and stored at -80℃ to await subsequent separation.

Proteolytic ^{18}O Labeling of lysozyme from chicken egg white

About 200 μg lysozyme from chicken egg white was dissolved and denatured in 200 μl 50 mM Tris (pH 8) and 4 M urea. Then the sample was reduced, alkylated and digested as previously mentioned. The digested lysozyme solution was divided equally into two pools. Twenty μl immobilized trypsin was added to each pool, which were then dried completely in a Speed Vac. One pool was redissolved in 80% $H_2{}^{18}O$ and 20 % acetonitrile. The other pool was redissolved in 80% $H_2{}^{16}O$ and 20% acetonitrile. Two pools were incubated at 37°C for five hours and then centrifuged at 1,000 g for 2 minutes. Two supernatants were mixed (1:1) and analyzed by MALDI-TOF mass spectrometry.

Forward and Reversed ^{18}O labeling

In the forward ^{18}O labeling experiment, peptides from drug susceptible cell line were labeled with ^{18}O water. In contrast, peptides from drug resistant cell line were labeled with ^{18}O water in the reversed ^{18}O labeling experiment.

Solution Isoelectric Focusing (sIEF) Separation

Six Teflon dialysis chambers including four separation chambers and two terminal electrode chambers (500μl volume each) were connected in tandem. Five 12mm diameter immobiline gel membranes having different pH values 3, 5, 6.5, 8 and 11 from an MCE kit were assembled between these chambers (Figure 12). Two ultrafiltration membranes (500Da cut-off) were used for ending the two terminal electrode chambers. O-rings were used between the chambers to assist in sealing the chamber compartments. A sample of the dried peptide mixture (0.5~1mg) was redissolved in the Chamber Buffer from MCE kit and loaded into the middle two chambers. The two terminal chambers were filled with Electrode Buffer from MCE kit while the other chambers were all filled with Chamber Buffer. The assembled chambers were put into the elctrophoresis tank, which was divided into two parts: cathode and anode compartments. Seven mM phosphoric acid was used to fill the anode compartment while 20mM lysine and 20mM arginine solution was used for the cathode compartment. A Bio-Rad power supply was used to drive isoelectric focusing. The running program was: 100V for 10 min, 200V for 20 min, 500 V for 40 min and 1000 V for about 100 min, till the current decreased to around

0.3~0.6 mA. After focusing, fractionated samples from different chambers were collected separately, the surface of the gel membranes and the inside wall of the separation chambers were rinsed with 250μl Chamber Buffer and these rinses were combined with the appropriate sample. Each solution was dried by SpeedVac and stored at -80℃ to await chromatography.

Reversed-Phase HPLC (RPHPLC) analysis with UV Detector

RP-HPLC was performed with a 250 mm×4.6 mm C_{18} Column assembled with a Shimadzu SCL-10 AVP instrument. The injection volume was 0.5 ml and the flow rate was 1ml/min. Buffer A was 0.1% TFA in water and buffer B was 0.1% TFA in acetonitrile. Each tryptic peptide fraction from the solution IEF was loaded into and eluted from the column with an 80 min gradient as follows: 10% B for 5 min, 10%-60% B for 55 min, 60%-90% B for 10 min, 90% B for 5 min and 10% B for 5 min. The HPLC profile was detected by UV absorption at 214 nm.

Peptide desalting

Each dried peptide sample from solution IEF fractionation was purified and concentrated using PepClean C-18 spin columns according to the manufacturer's instructions. Each column was activated with 50% ACN and then equilibrated with 0.5% TFA in 5% ACN solution. Each peptide sample was redissolved in 10% ACN with 1% TFA and loaded onto the column. The column was washed with 0.5% TFA in 5% ACN four times to remove high

levels of contaminants (i.e., urea, thiourea). The peptide mixture was finally eluted with 70% ACN. The eluate was dried by SpeedVac and stored at -80℃ before further analysis.

MALDI-TOF Mass Spectrometry

One μl sample solution was spotted on the MALDI plate and allowed to dry. One μl of 25 mM α- cyano-4-hydroxycinnamic acid dissolved in 70% acetonitrile in 0.1% TFA was spotted on top of the dried sample. The sample was then analyzed on the Kratos AXIMA-CFR MALDI-TOF mass spectrometer in reflectron mode.

NanoLC-QqTOF mass spectrometry and protein identification

The dried peptides recovered from solution IEF were redissolved in 0.1% formic acid and separated through LC packing HPLC system. Aliquot (2.2 μl) of the peptide mixture was first desalted through a C18 300μm ID × 5 mm precolumn for 10 minutes and then introduced to a C18 PepMap 75 μm ID × 15 cm separation column. Peptide elution was accomplished with 3%-60% B for 60 min, 60%-97% B for 10 min, 97% B for 10 min. The effluent at a flow rate of 200μl/min was introduced into the ESI-TOF mass spectrometer via an uncoated 10μm ID SilicaTip needle hold at 2400V. The mass spectrometer automatically scanned the m/z range 300-1500 Da at 1second/scan. Data-dependent control was used to automatically perform MS/MS on the most three intensive doubly or triply charged peptides detected in each MS

scan. Previously selected precursor ions were excluded for 30 seconds. BioAnalyst software submitted the resulting MS/MS spectra to the SwissProt human database through the MASCOT in-house search engine (Matrix Science, London, UK) to identify peptides and thus their original proteins. Criteria used for the searches allow for carbamidomethylation of cysteine and oxidation of methionine as variable modifications, one missed cleavage of trypsin, and fragment mass tolerance at ±0.8Da. Peptides matched with more than 95% confidence are considered to be identified. A scheme for peptide and thus protein identification using tandem mass spectra is shown in figure 10.

μLC-ion trap mass spectrometry and protein identification

The desalted peptides were redissolved in 0.1% formic acid. A 10 μl sample solution was then injected to the Surveyor HPLC system. The reversed-phase separation column was a 75μm ID×10cm fused capillary packed with 300Å BioBasic C18 particles, which was connected to a nanoelectrospray tip. The peptides were eluted with the same gradient used in PR-HPLC-UV described above, at a flow rate ~275 nl/min. The spray voltage was 2.0 kV. Acquisition of each full mass spectrum was followed by acquisition of MS/MS spectra of the three most intense peaks. Peptide and protein identifications were made the same way as mentioned above.

Quantitation of the relative abundance of proteins

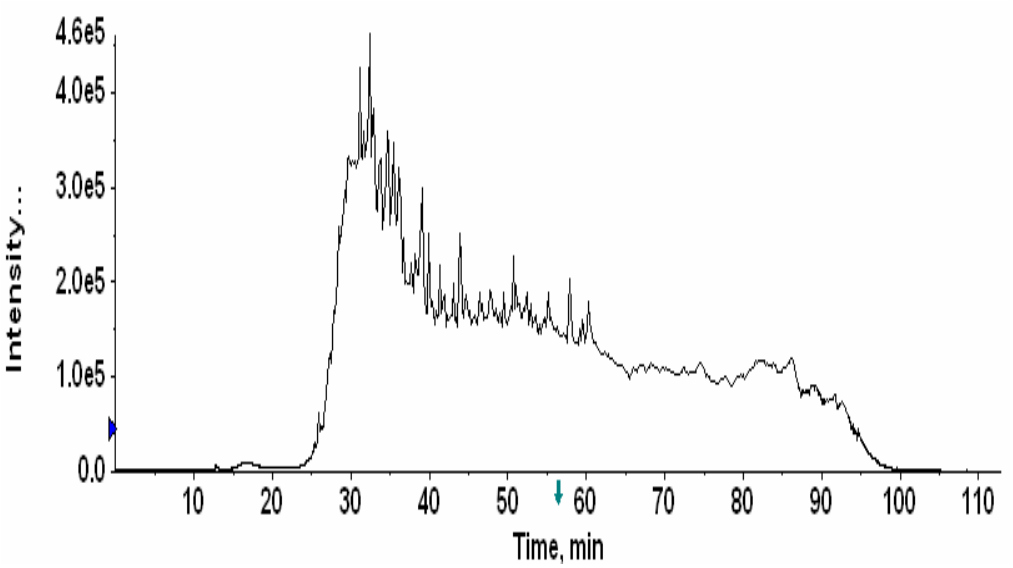
The abundance ratios of proteins in drug susceptible and the mitoxantrone resistant MCF-7 cells was calculated based on peak area ratios in mass spectra of labeled and unlabeled digested peptide pairs in the mixed sample. First, an extracted ion chromatogram (XIC) was extracted by summing ions in the m/z range of interest using the Extract Ion function on the Analyst QS software. XIC gave the pair sets of isotopic clusters with ions 4 Da apart (2 Da for doubly charged peptides or 1.3 Da for triply charged peptides). Second, the observed peak areas for the monoisotope ions for the peptides without ^{18}O labeling (I_0), the area of peaks with 2 Da higher (I_2), and the areas of peaks with 4 Da higher (I_4) were recorded from the spectrum. Third, the sequence of this identified peptide (from earlier identification) was used to calculate the theoretical isotope envelope distribution using MS-Isotope program (166). M_0 represents the theoretical peak area for the monoisotope peak for the unlabeled peptide, M_2 was the peak area for the peaks with masses 2 Da higher and M_4 was the area of the peak with 4 Da higher. Figure 11 shows an extracted ion chromatogram of peptide MEEFKDQLPADECNK (+carbamidomethylation, m/z 618.6) from the mixture of ^{18}O labeled and ^{16}O labeled peptides, the partial mass spectrum of the pair of doubly charged peptides with I_0, I_2, I_4 assignments, and the theoretical isotope clusters of the unlabeled peptide (166). The natural isotope distribution pattern for the ^{18}O labeled peptide was assumed to be the same. Finally, those experimental data

and theoretical parameters were inserted into the equation (141) below to calculate $^{18}O/^{16}O$ ratio:

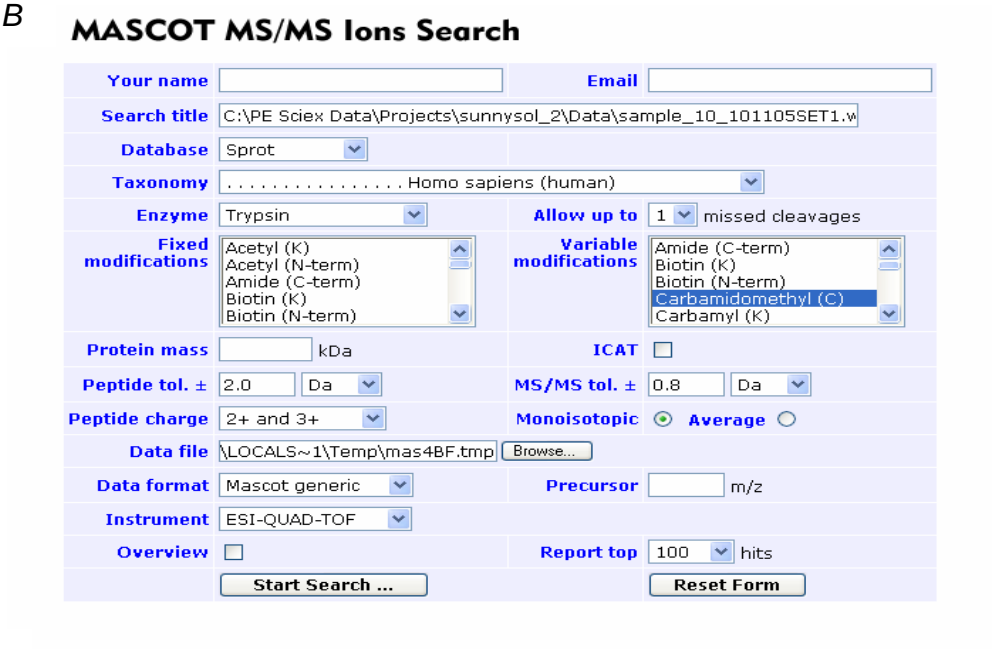
$$ratio = \frac{I_4 - \frac{M_4}{M_0} I_0 - \frac{M_2}{M_0}\left(I_2 - \frac{M_2}{M_0} I_0\right) + \left(I_2 - \frac{M_2}{M_0} I_0\right)}{I_0} \qquad \text{(Equation 1)}$$

Incomplete ^{18}O labeling (single incorporation of ^{18}O at the C-terminal of a peptide) was considered in this equation in order to obtain a more accurate ratio. Ratios from several peptide pairs from the same protein were averaged, as well as the ratios for the same protein from three cell harvests to provide standard deviations.

A


4.6e5
4.0e5
3.0e5
2.0e5
1.0e5
0.0
Intensity...
10
20
30
40
50
60
70
80
90
100
110
Time, min

B



C

Figure 10. Protein identification using tandem MS. *A*: Total ion chromatogram representing summed MS and MS/MS spectra. *B*: Criteria were chosen for the search. *C*: Proteins with more than 95% confidence are considered to be identified.

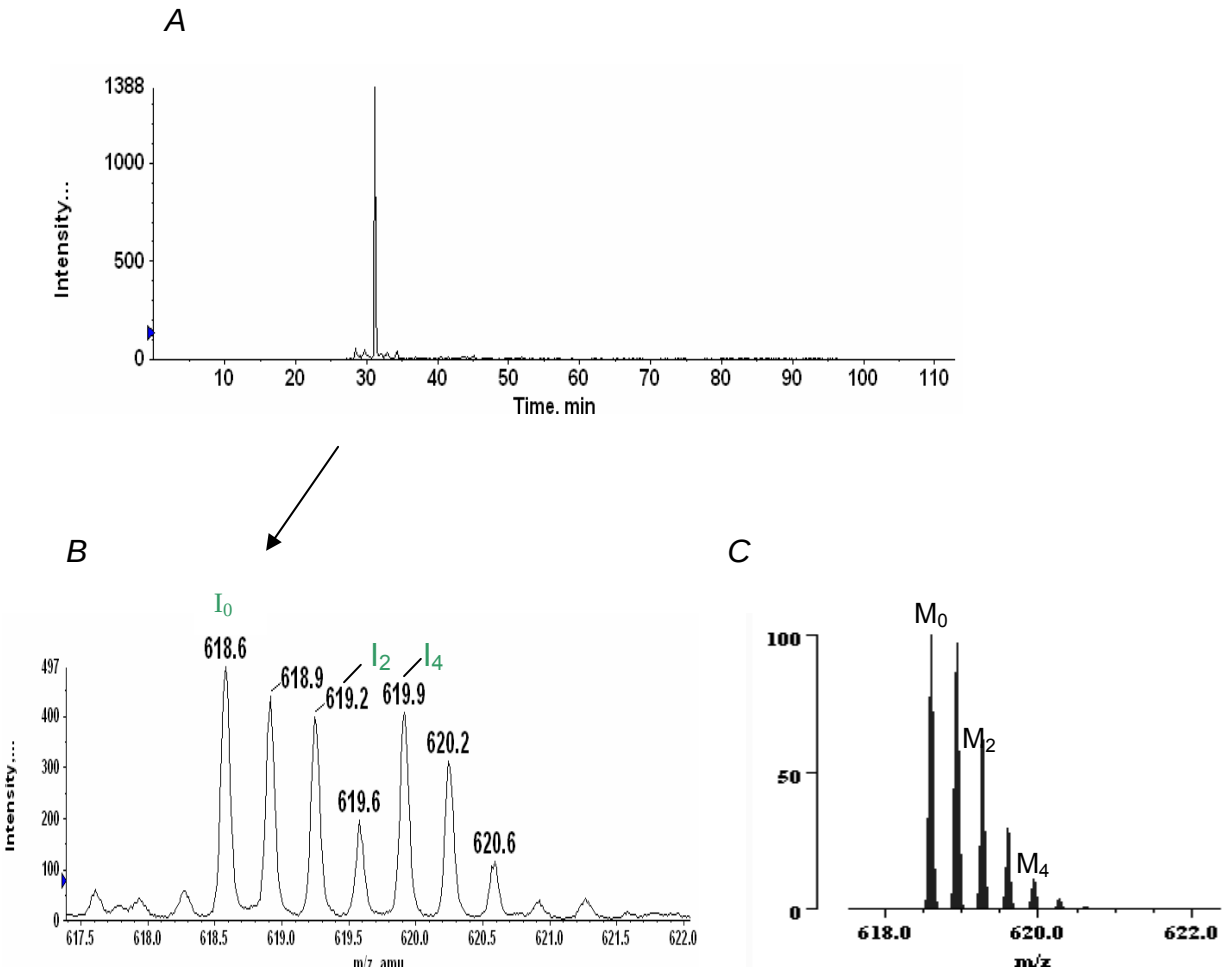


Figure 11. Relative quantitative analysis of digests of mitochondrial proteins prepared in $H_2{}^{16}O$ and $H_2{}^{18}O$ and combined. A: An extracted chromatogram reconstructed by summing ions of m/z 618.6 (MEEFKDQLPADECNK with carbamidomethylation, +57 Da) assigned to the triply charged peptide. B: Partial mass spectrum of the peak eluting from 31.09 to 31.38 min with I_0, I_2, I_4 assigned. C: Theoretical isotope envelope (166) of unlabeled peptide m/z 618.6.

Chapter 3: Results and Discussion

Mitochondrial isolation and protein extraction

The mitochondrial fraction was isolated from MCF-7 human cancer cells through a reproducible optimized method combining homogenization, differential centrifugation and percoll gradient purification. Typically, 3~4 g cell pellet was obtained from twenty 150 cm^2 flasks, from which ~2 g crude mitochondrial pellet could be isolated. Following purification using percoll gradient (163), the purified pellet usually had about one-fifth of the weight of the original organelle. A balance must be found between sample purity and recovery.

The mitochondrion is a membrane-bound organelle. Swelling of mitochondria with hypotonic medium (10mM Hepes, pH=7.3), followed by shrinking of mitochondria with isotonic medium (1.4M sucrose), resulted in rupture of the outer and inner membranes respectively and thus released proteins mostly into the soluble fraction.

Peptide fractionation with solution isoelectric focusing as the first dimensional separation

The proteomic analysis of complex mitochondrial proteins requires the separation of the original protein mixture or their digested peptides. The former separation was done with a 2-D gel strategy by Dr. Strong, another member in our research group (174). As an alternative approach, the latter one, also

called shotgun proteomics, is applied in this project. The most common strategy of the shotgun approach uses strong cation exchange (SCX) chromatography in the first dimension and reversed-phase HPLC as the second dimensional separation. But redundancy and relatively small range of pI value in peptide identification between fractions were also reported (167). Another valuable tool is the separation of proteins/peptides according to their pI. In this research, we applied solution isoelectric focusing (sIEF) using a small volume device (Amika) as the first dimensional separation to fractionate peptide mixture.

An experimental setup for sIEF is reported in figure 12, which consists of a series of Teflon chambers separated by thin porous acrylamide gel membranes containing immobilines at specific pH values. In this study, a setup made of six chambers was divided by five membranes with pH values of 3.0, 5.0, 6.5, 8 and 11. The assembled chambers were placed into an electrophoresis tank where the anode chamber was immersed into an anode buffer and cathode chamber was inside the cathode solution. It can also help heat dissipation during the focusing. The peptide sample was dissolved in buffer containing high concentrations of urea and thiourea before loading into the chambers, which reduced chances of peptide precipitation during focusing run. Immobilized pH gradient buffer (IPG buffer, 0.5% V/V) was also added to increase the conductivity of solution, especially for the low concentration sample.

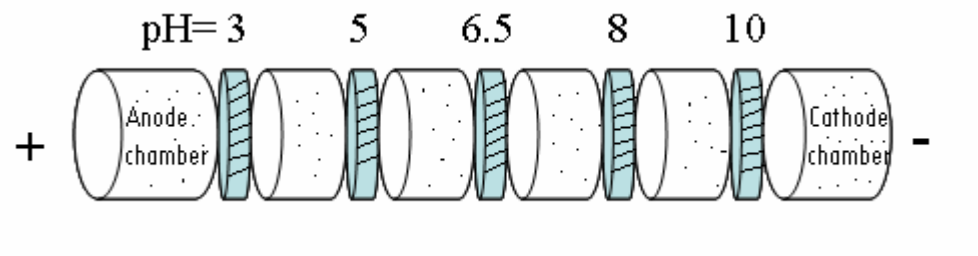


Figure 12. Schematic illustration of the solution isoelectric focusing device (106)

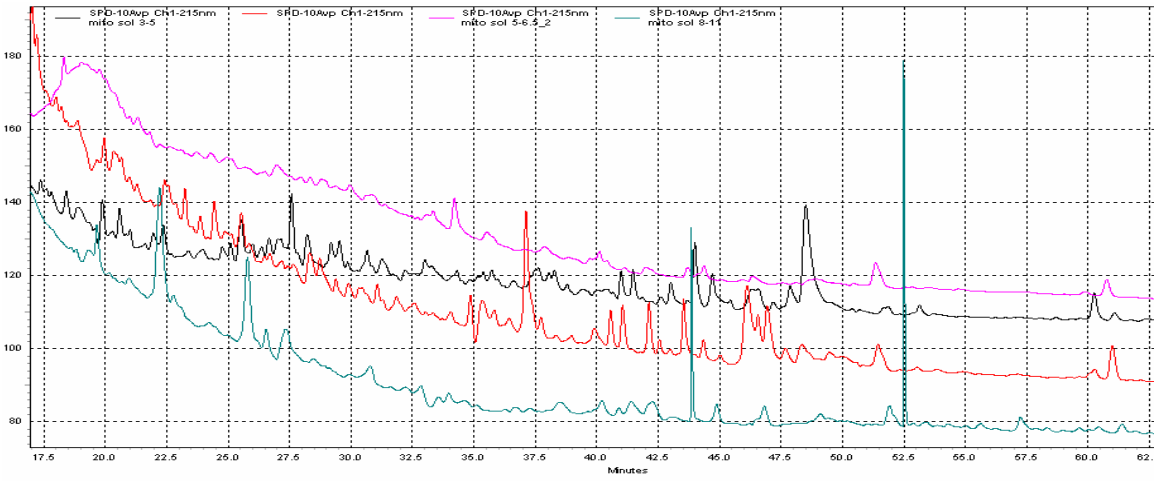
As studied previously in our lab (115), a faster running program (about 150 min) was set up for focusing peptide mixtures rather than proteins, which may also minimize peptide diffusion. Under an electric field, peptides were moved and finally trapped in chambers delimited by membranes which encompass their pIs. Eventually the electric current decreased to near 0. The resulting pI ranges of the six chamber fractions after focusing were: pH = below 3, 3-5, 5-6.5, 6.5-8, 8-11 and above 11 respectively.

Since the peptide mixture in each chamber should have a different composition after focusing, reversed phase liquid chromatography with UV detector was used to evaluate this separation. Figure 13A shows distinct HPLC profiles for different fractions, indicating the separation is effective. The reproducibility of this method was examined by performing the separation three times for the samples from three cell harvests (Figure 13B). Those profiles are similar and most of the peaks appeared in all of the three runs.

Peptide separation with reversed-phase LC/MS/MS as the second dimensional separation

Each chamber fraction was then separated with HPLC coupled on-line for MS and tandem MS analysis. Figure 14 shows the “base peak” chromatogram of mitochondrial peptides from each fraction of sIEF. The three most intensive peptides from the full MS scan were selected automatically for tandem MS analysis. Unlike a UV detector, the mass spectrometer is not able to monitor LC effluent continuously although the MS scan speed may be as short as 1

A:



B:

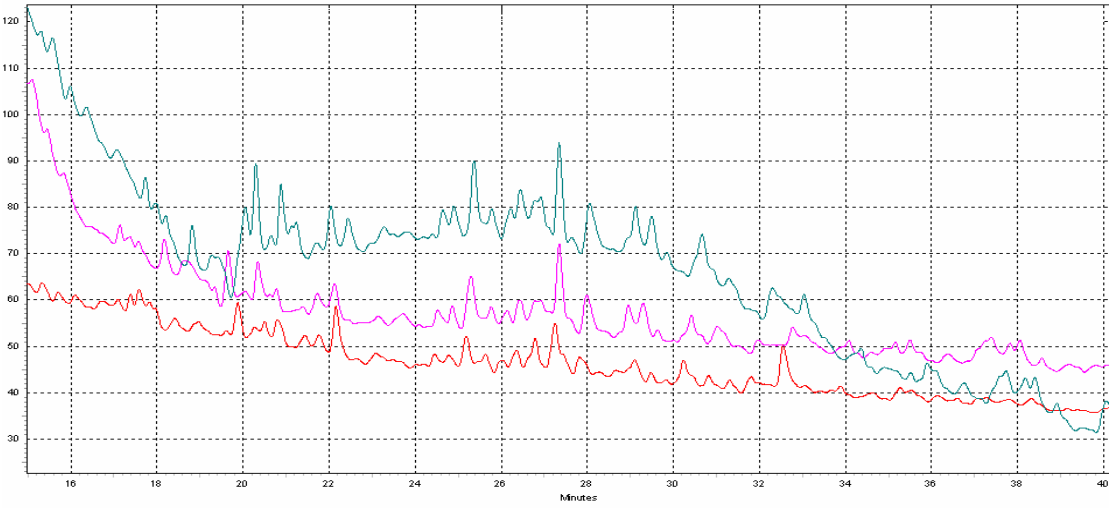
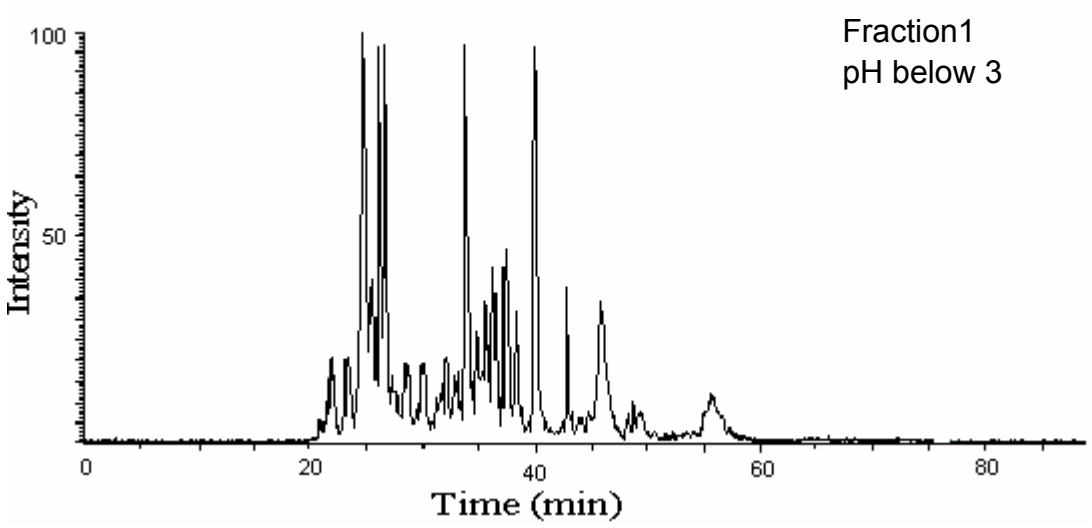
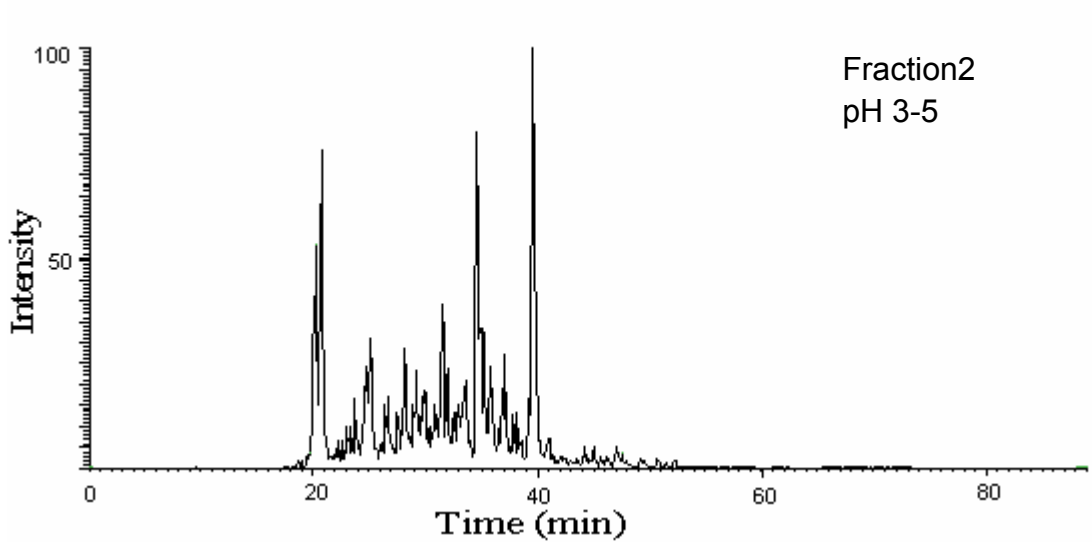
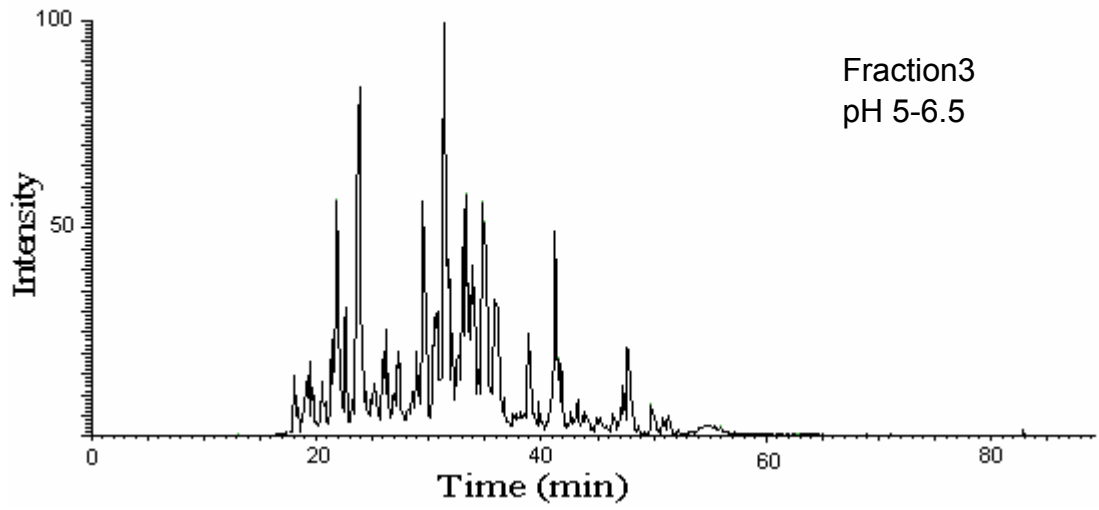


Figure 13. UV (214nm) HPLC chromatograms of liquid fractions from solution IEF separations. A. UV (214nm) HPLC chromatogram of four pI-focused liquid fractions (pH 3-5, black trace; 5-6.5, red; 6.5-8, pink; 8-11, green); B. UV (214nm) HPLC chromatogram of a single fraction from three solution IEF separations (only fraction pH 3-5 is shown here).


Fraction1
pH below 3
Intensity
100
50
0
20
40
60
80
Time (min)


Fraction2
pH 3-5
Intensity
100
50
0
20
40
60
80
Time (min)


Fraction3
pH 5-6.5
Intensity
100
50
0
20
40
60
80
Time (min)

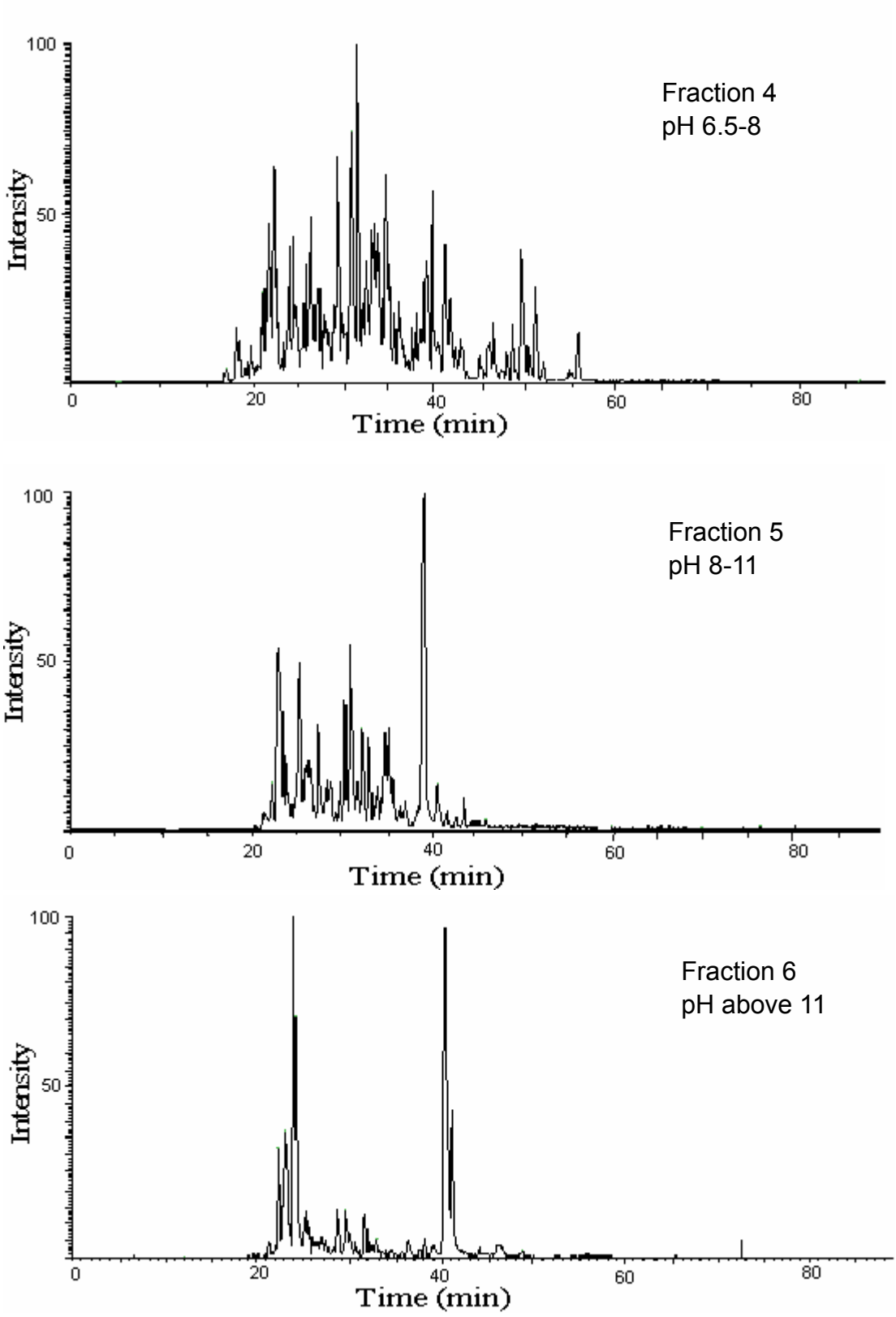


Figure 14. Chromatograms of mitochondrial peptides from sIEF fractions recorded on an LC-ion trap MS.

second/scan. Thus each scan time may only catch portions of eluent at a certain time. More important, MS/MS scans are obtained using computer-controlled data acquisition and largely rely on ion abundance levels. The selection of ions for MS/MS is dependent on the width of the chromatographic peaks or the concentration of peptides. Thus there is bias toward highly abundant peptides. Repeating the analysis several times may analyze new ions, which were lost in previous trials and improve the dynamic range of analysis (89, 155). In this work, each sample was analyzed three times by LC-MS/MS.

Peptide identification using tandem mass spectrometry

The MS/MS spectra were submitted to the SwissProt human database through a MASCOT in-house search engine with the criteria detailed in the Experimental chapter. MASCOT uses probability-based scoring algorithm (72), which calculates the probability that the observation of a match between the experimental data and the sequence database entry is by chance. Thus lower probability is reported as a better match. A widely used threshold for peptide identifications is that the probability of a random match is less than 0.05. Figure 15 shows examples of MS and MS/MS spectra of a doubly charged peptide from chamber 2. The peptide was identified as ILGADTSVDLEETGR from its tandem MS spectrum through MASCOT.

Values of pI for all identified peptides were determined by pI tool in Swiss-Prot database (www.expasy.org). A summary of all identified peptides

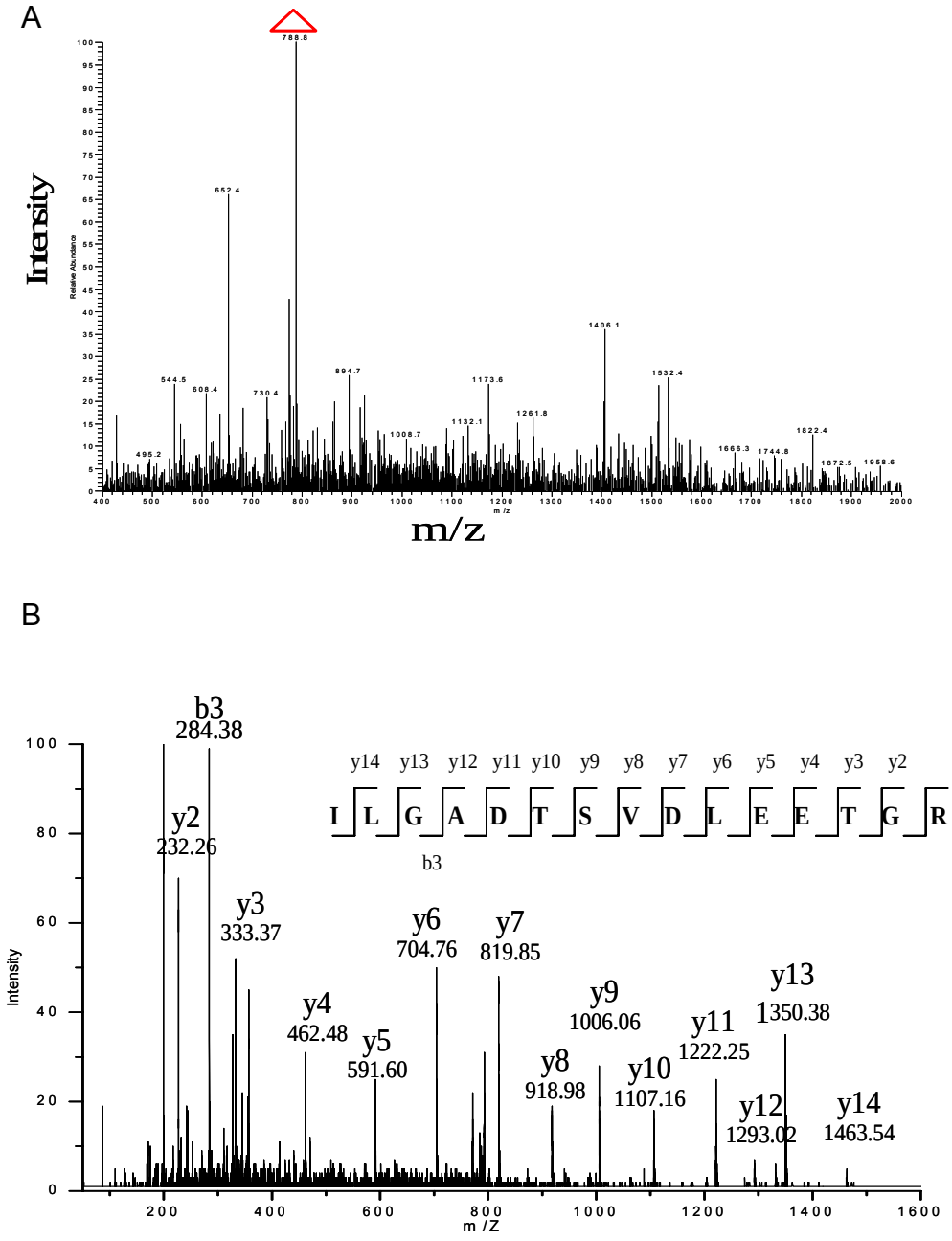


Figure 15. The MS and MS/MS scans of the ion with m/z 788.8. (A) The MS scan of ions eluted at 24.56 min in fraction 2. (B) MS/MS scan from the third most abundant ion with m/z 788.8. Peptide ILGADTSVDLEETGR from ATP synthase alpha chain was identified by Mascot using this spectrum.

with their calculated pIs and masses is shown in table 1.

The correlation between peptide pI and chamber pH defined by boundary immobiline membranes is presented in figure 16. Each bar shows the pI range of identified peptides in each fraction. Fifty percent of the peptides are in the box. The line extensions run about between 5% and 95%. The outliers represent the peptides having pI values beyond the 5% and 95%. It can be seen that the pI ranges from fractions 2 and 5 are located within the pH ranges. Peptides in fraction 1 are more basic, while pI values of peptides from chambers 3, 4 and 6 are lower than the defined pH ranges. Several reasons may contribute to the results. The two terminal fractions come from the cathode and anode chambers, which are connected to the amphoteric cathode and anode buffers. This might interfere with the focusing of peptides there. The intrinsic properties of peptides may play a role in the peptide separation and distribution. For example, the actual peptide pIs may be shifted from the calculated ones due to their interaction or the property of buffer solution.

The whole pI range is from ~3-10. There are few tryptic digest peptide having pI smaller than 3 or bigger than 11, because trypsin cuts proteins at C-termini of arginine and lysine, which have pKa of 9.74 and 10.76 respectively. This peptide range is similar to the reported result (80), which used strong cation exchange chromatography coupled with LC-MS to study phosphorylase B. Peptides eluted off SCX by increasing pI. Their pI range can be seen in figure 17. In addition, it is interesting to observe that there are fewer

Peptide sequence	pI	MW
VPLPSLSPTMQAGTIAR	9.72	1739.1
LLGGVTIAQGGVLPNIQAVLLPK	9.72	2271.8
DVPIGAIICITVGKPEDIEAFK	4.32	2328.8
DFLAGGVAAAISK	5.84	1219.4
TITLEVEPSDTIENVK	4	1788.0
ESTLHLVLR	6.85	1067.3
IQTQPGYANTLR	8.75	1361.5
NAGVEGSLIVEK	4.53	1215.4
VGGTSDVEVNEK	4.14	1233.3
TLNDELEIIEGMK	4	1504.7
CIPALDSLTPANEDQK	4.03	1714.9
IQEIIEQLDVTTSEYEKEK	4.14	2295.5
IMQSSSEVGYDAMAGDFVNMVEK	3.92	2508.8
LVQDVANNTNEEAGDGTTTATVLAR	3.92	2560.7
TVIIEQSWGSPK	5.66	1344.5
ISSIQSIVPALEIANAHR	6.75	1919.2
KPLVIIAEDVDGEALSTLVLNR	4.32	2365.8
DPGMGAMGGMGGGMGGGMF	3.8	1675.0
VGEVIVTK	5.97	844.0
LSDGVAVLK	5.84	901.1
GVMLAVDAVIAELK	4.37	1428.8
AAVEEGIVLGGGCALLR	4.53	1627.9
ALMLQGVDLLADAVAVTMGPK	4.21	2113.6
QSKPVTTPEEIAQVATISANGDK	4.68	2384.6
TALLDAAGVASLLTTAEVVVTEIPK	4.14	2482.9
GVMLAVDAVIAELKK	6.07	1556.9
VGLQVVAVK	8.72	912.1
GYISPYFINTSK	8.5	1389.6
KISSIQSIVPALEIANAHR	8.75	2047.4
TALLDAAGVASLLTTAEVVVTEIPKEEK	4.25	2869.3
DDAMLLK	4.21	805.0
GIIDPTK	5.84	742.9
EIGNIISDAMK	4.37	1190.4
IQEIIEQLDVTTSEYEK	3.91	2038.2
TLNDELEIIEGMKFDR	4.18	1923.2
ISSIQSIVPALEIANAHR	6.75	1919.2
IQEAGTEVVK	4.53	1073.2
TIIPLISQCTPK	7.89	1313.6
VDFPQDQLTALTGR	4.21	1560.7
MISDAIPELK	4.37	1116.3
GCDVVVIPAGVPR	5.83	1281.5

Peptide sequence	pI	MW
GCDVVVIPAGVPRKPGMTR	9.5	1952.4
LTLYDIAHTPGVAADLSHIETK	5.21	2365.7
AVLGASGGIGQPLSLLLK	8.8	1694.1
GYLGPEQLPDCLKGCDVVVIPAGVPR	4.56	2696.2
ANTFVAELK	6.05	992.1
FDGILGMAYPR	5.84	1239.5
DPDAQPGGELMLGGTDSK	3.84	1787.9
VVDALGNAIDGK	4.21	1171.3
TGTAEMSSILEER	4.25	1423.6
ILGADTSVDLEETGR	3.92	1575.7
TGAIVDVPVGEELLGR	4.14	1624.9
NVQAEEMVEFSSGLK	4.25	1667.9
YEELQSLAGK	4.53	1137.3
AEAESMYQIK	4.5	1169.3
SLDMDSIIAEVK	4.03	1320.5
ASLEAAIADAEQR	4.14	1344.4
LEGLTDEINFLR	4.14	1419.6
LSELEAALQR	4.53	1129.3
LVSESSDVLPK	4.37	1173.3
LESGMQNMSIHTK	6.75	1475.7
TEMENEFVLIK	4.25	1352.6
QLYEEEIR	4.25	1079.2
LEAELGNMQGLVEDFK	4	1793.0
YSTDVSVDEVK	4.03	1241.3
IIAEGANGPTTPEADK	4.14	1583.7
GFIGPGIDVPAPDMSTGER	4.03	1916.1
IIAEGANGPTTPEADKIFLER	4.41	2242.5
HGGTIPIVPTAEFQDR	5.32	1737.9
AKPYEGSILEADCDILIPAASEK	4.18	2433.8
NIMVIPDLYLNAGGVTVSYFEWLK	4.37	2743.2
NILGGTVFREPIICK	8.22	1660.0
DQTDDQVTIDSALATQK	3.77	1848.9
VCVETVESGAMTK	4.53	1353.6
TIEAEAAHGTVTR	5.37	1355.5
LNEHFLNTTDFLDTIK	4.54	1921.1
VAKPVVEMDGDEMTR	4.32	1676.9
YFDLGLPNR	5.84	1094.2
NILGGTVFR	9.75	976.1
DIFQEIFDK	4.03	1154.3
LEVAPISDIIAIK	4.37	1381.7
GTSFDAAATSGGSASSEK	4.37	1630.6
ASSTSPVEISEWLDQK	4.14	1776.9

Peptide sequence	pI	MW
AAVDAGFVPNDMQVGQTGK	4.21	1905.1
VFSVRGTSFDAAATSGGSASSEK	6.04	2219.4
VVPEMTEILK	4.53	1158.4
GLLPEELTPLILATQK	4.53	1736.1
LLYDLADQLHAAVGASR	5.21	1813.0
TIVAINKDPEAPIFQVADYGIVADLFK	4.23	2948.4
IVAPELYIAVGISGAIQHLAGMK	6.75	2351.8
LGGEVSCLVAGTK	5.99	1233.5
GLVYETSVLDPDEGIR	3.92	1762.9
EGSGIGAIDSNLDWSHNFTNMLGYTDHQFTELTR	4.42	3828.1
TVVGQITVDMMYGGMR	5.5	1758.1
HLPNDPMFK	6.74	1098.3
DILADLIPK	4.21	997.2
YWELIYEDSMDLIAK	3.92	1889.2
ALLTAAARLLGTK	11	1298.6
IDYGEYMDK	4.03	1133.2
SGDSEVYQLGDVSQK	4.03	1611.7
NPVTIFSLATNEMWR	6	1779.0
RQATTIIADNIIFLSDQTK	5.96	2148.4
IDVSIEAASGGK	4.37	1146.3
NLGLEELGIELDPR	4	1567.8
SEEQLKEEGIEYK	4.32	1581.7
NETLGGTCLNVGCIPSK	5.99	1706.0
IPNIYAIGDVVAGPMLAHK	6.74	1979.4
VWDDGIIDPADTR	3.77	1472.6
VLDSITTEILK	5.97	1213.5
QGTIFLAGPPLVK	8.75	1340.6
FEEEGNPYYSSAR	4.25	1548.6
KLEAAEDIAYQLSR	4.68	1606.8
ILFRPVASQLPR	12	1396.7
FDAGELITQR	4.37	1149.3
IFTSIGEDYDER	3.92	1444.5
LDVTIEPSEEPLFPADELYGIVGANLK	3.77	2930.3
IEYDTFGELK	4.14	1214.3
AAAEVNQDYGLDPK	4.03	1490.6
AIEMLGGELGSK	4.53	1204.4
IYELAAGGTAVGTGLNTR	6	1764.0
THTQDAVPLTLGQEFSGYVQQVK	5.29	2546.8
ETAIELGYLTAEQFDEWVKPK	4.25	2467.8
DDIENMVK	4.03	963.1
TTPSVVAFTADGER	4.37	1450.6
EQQIVIQSSGGLSK	6.1	1473.7

Peptide sequence	pI	MW
VINEPTAAALAYGLDK	4.37	1645.9
MEEFKDQLPADECNK	4.18	1797.0
VEAVNMAEGIIHDTETK	4.4	1857.1
STNGDTFLGGEDFDQALLR	3.84	2056.2
EQQIVIQSSGGLSKDDIENMVK	4.32	2418.7
CELSSVQTDINLPYLTMDSSGPK	4.03	2498.8
QAVTNPNNTFYATK	8.59	1568.7
ASNGDAWVEAHGK	5.32	1341.4
SDIGEVILVGGMTR	4.37	1446.7
NAVITVPAYFNDSQR	5.84	1694.9
SQVFSTAADGQTQVEIK	4.37	1809.0
ERVEAVNMAEGIIHDTETK	4.57	2142.4
EMAGDNKLLGQFTLIGIPPAPR	6.17	2338.8
MKETAENYLGHTAK	6.51	1333.4
LYSPSQIGAFVLMK	8.59	1553.9
DAGQISGLNVLR	5.84	1242.4
RDYASEAIK	6.07	1052.2
AANDAGYFNDEMAPIEVK	3.92	1955.1
TNVNGGAIALGHPLGGSGSR	9.44	1835.0
QTMQVDEHARPQTTLEQLQK	5.45	2381.7
DGTVTAGNASGVADGAGAVIIASEDAVK	3.84	2516.7
VSPETVDSVIMGNVLQSSSDAIYLAR	4.03	2752.1
INFDDNAEFR	4.03	1240.3
FFVADTANEALEAAK	4.14	1596.8
DIFAMDDKSENEPIENEAAK	3.95	2266.4
SCNGPVLVGSPQGGVDIEEVAASNPELIFK	4	3027.4
LVAGEMGQNEPDQGGQR	4.14	1785.9
LGDPAEYAHLVQAIIENPFLNGEVIR	4.4	2879.3
VDVAVNCAGIAVASK	5.8	1416.7
LVGQGASAVLLDLPNSGGEAQAK	4.37	2195.5
DVQTALALAK	5.84	1029.2
IQEENVIPR	4.53	1097.2
GDVENIEVVQK	4.14	1229.4
VIEEQLEPAVEK	4.09	1383.6
MEEANIQPNR	4.53	1201.3
VEDALNLK	4.37	901.0
IPENIYR	6	904.0
DLPVTEAVFSALVTGHAR	5.32	1883.1
GDEELDSLIK	3.92	1118.2
TVQSLEIDLDSMR	4.03	1506.7
GGMGSGGLATGIAGGLAGMGGIQNEK	6	2261.6
LLEDGEDFNLGDALDSSNSMQTIQK	3.66	2740.9

Peptide sequence	pI	MW
LEAEIATYR	4.53	1065.2
TITLEVEPSDTIENVK	4	1788.0
LQLETEIEALKEELLFMK	4.33	2177.6
DWSHYFK	6.74	982.1
VIHDNFGIVEGLMTTVHAITATQK	5.99	2596.0
IISNASCTTNCLAPLAK	8.06	1720.0
VPTANVSVVDLTCR	5.8	1473.7
GGMGSGGLATGIAGGLAGMGGIQNEK	6	2261.6
TVQGPPTSDDIFER	4.03	1561.7
VLPMNTGVEAGETACK	4.53	1619.9
AFYNNVLGEYEEYITK	4.25	1953.1
ALQDPNVAAFMVEPIQGEAGVVVPDPGYLMGVR	3.92	3441.0
VGDAIPAVEVFEGEPGNK	4	1828.0
VGDAIPAVEVFEGEPGNKVNLAELFK	4.25	2743.1
GVVDSEDIPLNLSR	4.03	1513.7
APALAAVPGGKPILCPR	9.51	1631.0
AFLDALQNQAEASSK	4.37	1592.7
EGIVTATEQEVK	4.25	1303.4
VTIAQGGVLPNIQAVLLPK	8.72	1931.4
VETTEDLVAK	4.14	1104.2
LSVISVEDPPQR	4.37	1339.5
GIHVEVPPAEAER	4.75	1403.6
IEVIKPGDLGVDLTSK	4.56	1684.0
KIEVIKPGDLGVDLTSK	6.12	1812.1
EIDGGLETLR	4.14	1102.2
LPAVVTADLR	5.84	1054.3
IDTIEIITDR	4.03	1188.3
TLADAEGDVFR	4.03	1193.3
VNAGDQPGADLGPLITPQAK	4.21	1962.2
EEDATLSSPAVVMPTMGR	4.14	1891.1
GLQVVEHACSVTSLMMGETMPSITK	5.4	2650.1
SSDPDYLAAVDK	3.93	1280.4
TTLPQDCSNPAPLSSPLNGVHDR	5.18	2419.7
VVDFIDEGVNIGLEVK	3.92	1746.0
GWTGQESLSDSDPEMWELLQR	3.83	2464.7
IMGLDLPDGGHLTHGYMSDVK	5.13	2256.6
AALEALGSCLNNK	6.04	1303.5
AHLLADMAHISGLVAAK	6.96	1718.1
GYSLVSGGTDNHLVLVDLRPK	6.75	2240.5
EVCDEVKAHLLADMAHISGLVAAK	5.32	2520.9
VLELVSITANK	5.97	1186.4
LMCPQEIVDYIADK	4.03	1637.9

Peptide sequence	pI	MW
LMCPQEIVDYIADKK	4.56	1766.1
GELFWDDGESLEVLER	3.77	1894.0
APSPLYSVEFSEEPFGVIVHR	4.75	2360.7
TVGIDDLTGEPLIQR	4.03	1626.8
YEEIDNAPEER	3.91	1364.4
KYEEIDNAPEER	4.25	1492.6
LLDAVDTYIPVPAR	4.21	1542.8
GEETPVIVGSALCALEGR	4.25	1801.0
TIGTGLVTNTLAMTEEEK	4.25	1908.2
NMITGTAPLDGCILVVAANDGPMPQTR	4.21	2756.2
QIGVEHVVVYVNK	6.75	1483.7
GITINAAHVEYSTAAR	6.75	1673.9
DLEKPFLLPVEAVYSVPGR	4.68	2129.5
ALEAANGELEVK	4.25	1243.4
FVSSSSSGGYGGGYGGVLTASDGLLAGNEK	4.37	2795.0
SLLEGQEDHYNNLSASK	4.65	1905.0
ILGATIENSR	6	1073.2
AFVDFLSDEIK	4.03	1283.4
VEEQEPELTSTPNFVVEVIK	3.98	2287.6
MSGGWELELNGTEAK	4.25	1621.8
ILTMDGLIEDIK	4.03	1360.6
NVEAMNFADIER	4.14	1408.6
TPAFAESVTEGDVR	4.14	1478.6
GLVVPVIR	9.75	852.1
AKPAEAPAAAAPK	8.64	1192.4
TGAAPAKAKPAEAPAAAAPK	9.7	1789.1
SYELPDGQVITIGNER	4.14	1791.0
SENGLEFTSSGSANTETTK	4.25	1960.0
TDEFQLHTNVNDGTEFGGSIYQK	4.31	2600.7
VNNSSLIGLGYTQTLKPGIK	9.7	2103.5
LNNDDNVDGLLVQLPLPEHIDER	3.96	2628.9
EAAGEGPALYEDPPDQK	3.83	1786.9
PEFLEDPSVLTK	4.14	1374.6
ALNALCDGLIDELNQALK	4.03	1914.2
NNTVGLIQLNRPK	11	1466.7
SLAMEMVLTGDR	4.37	1322.6
ICPVETLVEEAIQCAEK	9.6	1283.4
VLVEPDAGAGVAVMK	4.37	1455.7
NPPVNSLSLEFLTELVISLEK	4.25	2342.7
DADVQNFVSFISK	4.21	1469.6
TYEQVLENLESK	4.25	1452.6
FLGTEPEPDAVGLDSGHIR	4.31	2010.3

Peptide sequence	pI	MW
AIQGGLEWLK	6.05	1114.3
QVVEEPSPQLPADK	4.14	1536.7
VECVGDDIAWMR	4.03	1393.6
VLSGDLGQLPTGIR	5.81	1425.7
LGTPVLQALGDGDFVK	4.21	1629.9
SYLTEQVNQDLPK	4.37	1534.7
QCPIMDPAWEAPEGVPIDAIIFGGR	3.92	2683.1
DEGWLAEHMLILGITSPAGK	4.65	2138.5
KDPEPEDEVPDVK	4.02	1496.6
DPEPEDEVPDVK	3.71	1368.4
GDVVNQDDLYQALASGK	3.93	1792.9
QLEVEPEEPEAENK	3.9	1640.7
VVDNPIYLSDMGAALTGAESHELQDVLEETNIPKR	4.19	3826.3
VPTAAGAWLLR	9.72	1154.4
TENPLILIDEVDK	3.92	1498.7
AKLSSDVLTLLIK	8.64	1400.7
TPLAVELEVLDGHDPDPGR	4.1	2030.2
QAEVANQETKEDLPAENGETK	4.14	2301.4
TPDGTENGDFLALDLGGTNFR	3.84	2210.3
AQAELVGTADEATR	4.14	1431.5
SYELPDGQVITIGNER	4.14	1791.0
YPIEHGIITNWDDMEK	4.31	1961.2
VAPEEHPVLLTEAPLNPK	4.75	1954.3
QEYDESGPSIVHR	4.65	1516.6
SYELPDGQVITIGNER	4.14	1791.0
LCYVALDFEQEMATAASSSSLEK	4	2493.8
VTINTAIGHINR	9.73	1308.5
AFGGQSLK	8.8	806.9
CCCVADR	5.82	768.9
ISKLYGDLK	8.5	1036.2
HVNGQDQIVPGLYACGEAACASVHGANR	5.99	2838.1
NTVVATGGYGR	8.75	1094.2
GEGGILINSQGER	4.53	1329.4
VPPIKPNAGEESVMNLDK	4.68	1938.2
LRLEVNLQAMK	8.75	1314.6
GALAK	8.75	458.6
LEATQLEGVAR	4.53	1186.3
DMPAAGSLGSSSRNR	9.6	1505.6
YSTDVSVDEVK	4.03	1241.3
DDGSWEVIEGYR	3.92	1425.5
SHSLKDLAMVE	5.48	1229.4
GDGSCDVR	4.21	807.8

Peptide sequence	pI	MW
DAGEGGLSLAVEGPSK	4.14	1486.6
CTYRPAMEGPHTVHVAFAGAPITR	8.24	2583.0
QVAEVNLWGTVRMTK	8.75	1732.0
LRTVQLNVCSSEEVEK	4.79	1834.1
LFLTMEINPK	6	1205.5
TSAPITCELLNK	5.66	1289.5
ELLGAAGHR	6.85	923.0
GHALRDTETTLR	6.75	1369.5
VLINLYCCAAEDAR	4.37	1553.8
AAAAK	8.8	430.5
NETEIELGSLLR	4.25	1373.5
QHMAQMEEMKTR	6.76	1519.8
QEATESLKCQEELR	4.49	1663.8
SPGSGCILAHCMGLGK	7.82	1530.8
GMYQPVAGGMQPPPLQR	8.75	1827.2
QKTELMK	8.59	877.1
ADFEEQLWKK	4.68	1293.4
VLDSGAPIKIPVGPETLGR	6.04	1919.3
TIAMDGTEGLVR	4.37	1262.4
IMNVIGEPIDER	4.14	1385.6
IPSAVGYQPTLATDMGTMQER	4.37	2266.6
TVLIMELINNVAK	5.66	1457.8
LVLEVAQHLGESTVR	5.4	1650.9
IGLFGGAGVGK	8.75	975.2
SLQDIIAILGMDELSEEDKLTVSR	4.02	2676.0
IPVGPETLGR	6	1038.2
VALVYGQMNEPPGAR	5.97	1601.8
LLKMAVGMR	11	1018.3
SPLAQRPLR	12	1037.2
ISSPCLKADSGACGPDSCPYCAR	5.93	2301.6
GKGGEIQPVSVK	8.59	1198.4
VLQATVVAVGSGSK	8.72	1315.5
VLLPEYGGTK	5.97	1076.3
GGEIQPVSVK	6	1013.2
FLPLFDR	5.84	907.1
GAPTTSLISVAVTK	8.75	1344.6
FQNEEEVFAWNNEVK	4.09	1883.0
GEVITTYCPANNEPIAR	4.53	1848.1
VNLLSFTGSTQVGK	8.72	1450.7
LQQAPNQPK	8.75	1023.2
FVTVQTISGTGALR	9.75	1449.7
ISVAGVTSSNVGYLAHAIHQVTK	8.61	2352.7

Peptide sequence	pI	MW
IAAAILNTPDLR	5.84	1267.5
DDNGKPYVLPSVR	5.96	1459.6
VNNSSLIGVGYTQTLRPGVK	9.99	2103.4
WNTDNTLGTEIAIEDQICQGLK	3.92	2462.7
QEYDESGPSIVHR	4.65	1516.6
SYELPDGQVITIGNER	4.14	1791.0
VAPEEHPVLLTEAPLNPK	4.75	1954.3
AQQATPGGAAPTIFSR	9.79	1572.7
ISQAEEEDQQLLGHLLLVAK	4.4	2234.5
APALGGSFAGLEPMGLLWALEPEKPLVR	4.79	2920.5
TMETLHLEYFEEAMNYLLSHPEVK	4.62	2925.3
KHPDASVNLSEFSK	6.75	1558.7
LGEMWNNTAADDKQPYEK	4.32	2110.3
APEQEQAAPGPAAGGEAPK	4.25	1775.9
AEPPKAPEQEQAAPGPAAGGEAPK	4.49	2298.5
AQGPAASAEEPKPVEAPAANSDQTVTVK	4.41	2764.0
AAEAAAAPAESAAPAAGEEPSKEEGEPK	4.12	2636.8
ITDLANLSAANHDAAIFPGGFGAAK	5.21	2441.7
GVEVTVGHEQEEGGKWPYAGTAEAIK	4.64	1554.6
GGAEVQIFAPDVPQMHVIDHTK	5.21	2389.7
GFLNSSELSGLPAGPDR	4.37	1716.9
LGQHVVGMAPLSVGSLDDEPGGEAETK	4.17	2694.0
LIAINANDPEASK	4.37	1355.5
ILGILALIDEGETDWK	3.92	1786.1
NVTGHYISPFHDIPLK	6.92	1838.1
NDEYENLFNMIVEIPR	4	1996.2
IEQLSPFPFDLLLK	4.37	1660.0
YAELLVSQGVVNQPEYEEEISK	3.98	2524.8
VIPEDGPAAQNPENVK	4.14	1677.8
GSGDPSSSSSSGNPLVYLDVDANGKPLGR	4.43	2834.0
LPAARACSK	9.51	916.1
VGNGFEEGTTQGPLINEK	4.25	1890.0
ISFTGSTTTGK	8.75	1099.2
FSHEEIAMATVTALR	5.4	1675.9
YTPSGQAGAAASESLFVSNHAY	5.24	2228.4
ADDGRPFPQVIK	6	1342.5
STELLIR	5.72	831.0
TVTAMDVVYALK	5.5	1310.6
DNIQGITKPAIR	8.75	1325.5
VVQVSAGDSHTAALTDDGR	4.41	1899.0
ALGSVGPVDLLVNNAAVALLQPFLEVTK	4.37	2849.4
VNAVNPTVVMTSMGQATWSDPHK	6.71	2470.8

Peptide sequence	pI	MW
NEQDAYAINSYTR	4.37	1544.6
QAVLGAGLPISTPCTTINK	8.22	1884.2
ENGTVTAANASTLNDGAAALVLMTADAAK	4.03	2762.0
FGNEVIPVTVTVK	6	1402.7
AFQYVETHGEVCPANWTPDSPTIKPSPAASK	5.45	3329.7
HLSVNDLPVGR	6.74	1206.4
NGGLGHMNIALLSDLTK	6.74	1754.0
DYGVLLEGSGLALR	4.37	1462.7
VLVHPPQDGEDEPTLVQK	4.31	2001.2
EPVEAAPAAEPVPAST	3.67	1535.7
GQLTTDQVFPYPSVLNEEQTQFLK	4.14	2783.1
GFGGITHGPPEK	6.75	1196.3
LVEIVGMHDLGVGITLGAHQSIGFK	5.99	2592.1
TPVTDPATGAVK	5.5	1156.3
ELGAFGLQVPSELGGVGLCNTQYAR	4.53	2579.9
IGIIDGEYVVNPTR	4.37	1545.8
TLNDRSSIVMGEPISQSSSNSQ	4.37	2337.5
SETAPAAPAAAPPAEK	4.53	1478.6
LLQYSDALEHLLTTGQGVVLER	4.65	2455.8
VEGSFPVTMLPGDGVGPELMHAVK	4.65	2467.9
IESEGLLSLTTQLVK	4.53	1630.9
SETAPAETATPAPVEK	4.25	1598.7
ATGPPVSELITK	6.05	1212.4
KATGPPVSELITK	8.59	1340.6
ATGPPVSELITKAVAASK	8.64	1740.0
TVAGIIVEPIQSEGGDNHASDDFFR	4.1	2674.9
REDLLNNAAHAGK	6.75	1408.5
MLDLYSQISSVPIGYSHPALLK	6.49	2432.9
VDVEFDYDGPLMK	3.84	1527.7
GTFCSFDTPDDSIR	3.92	1560.7
IDIPSFDWPIAPFPR	4.21	1771.1
NLLLAEVINIIK	6	1352.7
HELQANCYEEVKDR	4.83	1733.9
HGYPLIIYDVFPDACK	5.21	1851.2
TPVGFIGLGNMGNPMAK	8.41	1704.0
MGAVFMDAPVSGGVGAAR	5.59	1693.0
KPAEQAEETAPIEATATKEEEGSS	4.12	2503.6
VNDNKTAAEEALR	4.68	1430.5
TIAQGNLSNTDVQAAK	5.5	1630.8
VTSEELHYFVQNHFTSAR	6	2165.4
AVAFQNPQTHVIENLHAAAYR	6.96	2350.6
QKAAAAFAQLQGAMEMLGISESEQR	4.79	2666.0

Peptide sequence	pI	MW
NMDAHKVMLDLLQIPYDK	5.3	2144.5
CVVEPAAGDLDNPPK	4.03	1524.7
LYQGHLQEESGPPPESMPK	4.75	2124.4
YLESEEYQER	4.09	1345.4
NGPLEVAGAAVSAGHGLPAK	6.75	1816.1
SIAFPSIGSGR	9.47	1091.2
ASGPPVSELITK	6.05	1198.4
VAGHPNIVINNAAGNFISPTER	6.72	2291.6
DPDMVQNTVSELIK	4.03	1588.8
ILMAAPGMAIPPFIMNTLEK	6	2158.7
GMTTLLSSLGAQCVIASRK	9.51	1936.3
VAFTGSTEVGHLIQK	6.72	1586.8
LAPALATGNTVVMK	8.75	1385.7
VAEAHENIIHGSGATGK	6	1690.8
TQHHVEALVEHQNGK	6.19	1726.9
GPGGSSLLIEALSNSSHK	6.75	1753.9
ALEMAIEAGAEDVKETEDEEER	3.84	2464.6
AADLQLEMTQKPHK	6.8	1609.9
KPGPGEPLVFGK	8.59	1225.5
LGGNYGPTVLVQQEALKR	8.59	1943.2
LFDFQGLQHQVAHVATQLEAAR	5.99	2479.8
VGSFCLSEAGAGSDSFALK	4.37	1846.0
IGTIYEGASNIQLNTIAK	6	1906.2
LDSPAGTALSPSGHTK	6.74	1538.7
IQEPNTFPAILR	6	1398.6
QIGLLLK	8.75	784.0
MSLWGLVSK	8.5	1020.3
LGGDLGTYVINK	5.83	1249.4
FAAEHTIFASNTSSLQITSIANATTR	6.75	2753.0
TPMTSQKTFESLVDFSK	5.73	1946.2
TFESLVDFSK	4.37	1172.3
TLSTIATSTDAASVVHSTDLVVEAIVENLK	4.31	3085.5
HVTVIGGGLMGAGIAQVAAATGHTVVLVDQTEDILAK	5.21	3614.2
DTPGFIVNR	5.84	1018.1
GNIETNHNLPPSHK	6.92	1557.7
IGGIGTVPVGR	9.75	1025.2
VMPNAIVQSVGVSSGK	8.72	1572.8
TLQYLSQGNVVFK	8.26	1496.7
TGQAPGYSYTAANK	8.17	1428.5
ADLIAYLK	5.88	906.1
ASGPPVSELITK	6.05	1198.4
SETAPAAPAAPAPAEK	4.53	1478.6

Peptide sequence	pI	MW
RGDIIGVQGNPGK	8.75	1310.5
QTLSTPGTIILGTIPVPK	8.75	1836.2
ISALQSAGVVVSMSPAQLGTTIYK	8.59	2421.8
VSGVMDNNILVLVPDPHAK	5.21	2018.4
VPFCLQSCVKPLK	8.9	1461.8
GSTHPQPGVSPPAAPAAPGPK	8.76	1921.1
VRLAWAALAR	12	1126.4
IAEEFEVELER	3.98	1663.5
AEAEAQAEELSFPR	4.09	1547.6
HGEEVTPEDVLSAAMYPDVFAHFK	4.49	2690.0
VAVEEVDEEGK	3.91	1203.3
MQQQLDEYQELLDIK	3.92	1894.1
TLEGELHDLR	4.65	1182.3
QELIECVANSDEQLGEMFLEEK	3.77	2554.8
YLEATGQLPVK	6	1218.4
TGIEQGSDAGYLCESQK	4.14	1785.9
TGIEQGSDAGYLCESQKFGELVMTK	4.41	2692.0
EDLPAENGETKTEESPASDEAGEK	3.89	2533.6
IQDVGLVPMGGVMSGAVPAAAAQEAVEEDIPIAK	3.83	3334.9
AALEAVGGTVVLE	4.24	1228.4
GVVDSEDLPLNISR	4.03	1513.7
EGLELPEDEEEK	3.77	1416.5
EDQTEYLEER	3.91	1311.3
GDVGMAGVAIDTVEDTK	3.84	1677.8
VADPWGGSYMMECLTNDVYDAALK	3.84	2650.0
SVTEQGAELSNEER	4.09	1548.6
DSTLIMQLLR	5.84	1189.4
EAAGEGPALYEDPPDQK	3.83	1786.9
SVTNEDVTQEELGGAK	4	1676.8
EVDVGLAADVGTLER	3.92	1543.7
YQETFNVIER	4.53	1298.4
LIAEGPGETVLVAEEEAAR	3.98	1954.2
GEPAAAAAPEAGASPVEK	4.25	1622.8
ALYETELADAR	4.14	1251.4
QASIQHIQNAIDTEK	5.32	1695.9
YGLIPEEFFQFLYPK	4.53	1891.2
EISEVFPDQFIHLGGDEVEFK	4.08	2435.7
VAPTVETSDPYADDPVR	3.84	1832.0
DAINQGMDEELERDEK	3.95	1892.0
DIIFAIK	5.84	819.0
TNHLVTVEGGWPQFGVGAEICAR	5.34	2441.8
DVQGTDASLDEELDR	3.66	1662.7

Peptide sequence	pI	MW
TEGDEEAEEEQEENLEASGDYK	3.55	2501.4
HYLPLSSILDTLDVMAYNK	5.21	2193.5
GSYNPVTHIYTAQDVK	6.74	1793.0
TPYTDVNIVTIR	5.5	1391.6
GEFDPGQDTYQHPPK	4.54	1715.8
IVYGHLDDPASQEIER	4.31	1842.0
LEAPDADELPK	3.92	1197.3
EDIANLADEFK	3.92	1264.4
DLEDLQILIK	4.03	1199.4
LPHLPGLEDLGIQATPLELK	4.65	2154.5
STELLIR	5.72	831.0
KEEILLIK	6.14	985.2
HLNFLTSEQALADFAELIK	4.65	2160.5
GLDTVVALLADVVLQPR	4.21	1779.1
GKPAVAALGDLTDLPTYEHIQTALSSK	5.38	2797.2
CPGESSHICDFIR	6	1463.7
TPLFDQIIDMLR	4.21	1461.7
PGLVDSNPAPPESQEK	4.14	1664.8
APSVPAAEPEYPK	4.53	1355.5
HEWVTTENGIGTVGISNFAQEALGDVVYCSLPEVGTK	4.25	3922.3
NMITGTSQADCAVLIVAAGVGEFEAGISK	4.14	2853.3
IITLEEGDIILTGTPK	4.14	1713.0
IPGIYVLSLEIGK	6	1401.7
CPGESSHICDFIR	5.32	1463.5
VVNSETPVVVDFHAQWCGPCK	5.32	2315.7
DEDQLEAFLK	3.92	1207.3
SPALLLSQLLPYMENR	5.72	1845.2
EPGTVALVSK	6.1	1000.2
DLLEVADVLEK	3.92	1243.4
FDPYEHEALFHTPVEGK	4.8	2016.2
VVPLVQMGETDANVAK	4.37	1670.9
FWITNGPDADVLIVYAK	4.21	1922.2
QLGNLGVLGITAPVQYGGSGLGYLEHVLVMEEISR	4.75	3671.2
IPAANILGHENK	6.75	1276.5
DITDHMDR	4.41	1002.1
LEQNGSPLGRGR	9.6	1283.4
LGCQEEGAAEVK	4.25	1233.4
EVEELILTESK	4.09	1289.5
AVEILADIIQNSTLGEAEIER	6.05	992.1
SPAGLQVLNDYLADK	4.21	1603.8
ADEGISFR	4.37	894.0
LIFKPDLTLEEVQAENPK	4.41	2084.4

Peptide sequence	pI	MW
IEPLSPELVAAASAVADSLPFDK	3.92	2340.7
VPINDVLAEDK	4.03	1212.4
EGICGSCAMNINGGNTLACTRR	5.21	2389.7
GDVTAQIALQPALK	5.84	1424.7
SEHPGLSIGDTAK	5.3	1311.4
GTPEQPQCGFSNAVVQILR	5.99	2044.3
VNILTRLAAELNK	8.72	1454.7
GIHSAIDASQTPDVVFASILAAFSK	5.21	2545.9
ESVNYLVSQQNMLLIPTSFSPLK	6.1	2609.0
SLLCSLICYR	7.79	1170.5
ALQDRLVATNLK	8.79	1341.6
ALWQAAEVERDR	4.68	1443.6
DGNGTIYPMAKDCMGGIR	5.95	1899.2
KTASPEDSDMPDHDLEPPR	4.23	2137.3
DGMTPKGPNHPVQVMPK	8.6	1833.2
EDPNLVPSISNK	4.37	1312.4
MKPLWLVYNNK	9.7	1405.7
DCEGSALLK	4.37	935.1
ALDAMLDLLK	4.21	1102.4
ATTDLGRSLGPVELLLR	6.12	1811.1
DLKLDNLLLDTEGYVK	4.23	1849.1
TPVTQVNEVTGTLR	5.66	1514.7
DVEDFLSPLLGK	4.03	1332.5
IEVEKPFAIAK	6.14	1244.5
APVTCTPGQPGQQRVLHLELK	8.27	2272.7
PGIVELPTLEELKVDEVK	4.25	2008.3
VGAGAPVYMAAVLEYLTAEILELAGNAAR	4.25	2934.4
DGADIHSDLFISIAQALLGGTAR	4.41	2341.6
SIVEEIEDLVAR	4	1372.5
AMGIMNSFVNDIFER	4.37	1744.0
THTDTESEASILGDSGEYK	4.17	2040.1
LGAGYPMGPFELLDYVGLDTTK	4.03	2357.7
EEEAIQLDGLNASQIR	4	1785.9
QAITQVVVSR	9.75	1100.3
VVQVVKPHTPLIR	11	1485.8
KDINNIVK	8.59	943.1
AAELIANSLATAGDGLIELR	4.14	1998.3
LAAAFAVSR	9.75	905.1
CRSGQCVLASR	9.02	1179.4
LSNTQGVVSAFSTMMSVHR	9.76	2052.4
AVTAIMTR	9.79	862.1
ELISNASDALDK	4.03	1275.4

Peptide sequence	pI	MW
LTLSALIDGK	5.84	1030.2
VLQHYQESDKGEELGPGNVQK	4.83	2355.6
TPALVNAAVTYSKPR	9.99	1587.8
FSLFAGGMLR	9.75	1098.3
VNLLQIVR	9.72	954.2
GINTLVTYDMVPEPK	4.37	1677.0
WVTYFNKPDIDAWELR	4.56	2053.3
NFSEVFQK	6	998.1
VPDFSEYR	4.37	1012.1
AGLVDDFEK	4.03	993.1
FLQDTIEEMALK	4.14	1437.7
GTGGVDTAATGGVFDISNLDR	3.93	2023.1
SEVELVQLVIDGVNYLIDCER	3.83	2406.7
MPKFSMPSLK	10	1165.5
VGQGWSYSAVQDIPARR	8.72	1890.1
HRVSLFGTDAPAVVNCLHILAR	8.27	2389.8
KPECFGPALRGEGGSGLLAAIEEAIR	4.95	2642.0
EPGLFDVVIINDSLDQAYAELK	3.77	2449.7
DIEEIIDELK	3.83	1216.4
DDTIYEDEDVK	3.66	1341.4
ELSEALGQIFDSQR	4.14	1592.7
LDAITDEENDMLDLAYGLTDR	3.57	2383.6
NLLLSGAQLEASR	6	1371.6
GPTGVIATTMTDSFLTGQMLLQDLK	4.21	2639.1
LPNQTHPDVPVGDESQAR	4.54	1960.1
LEEGPPVTTVLTR	4.53	1411.6
SEMEMAHLYSLCDAAHAQTEVAKK	5.35	2664.0
ILLDQVEEAVADFDECIR	3.71	2078.3
MSVLSLDLTAIK	5.59	1290.6
VPEFDGKMSVLSLDLTAIK	4.56	2063.4
KLLTLDMYNAVMLGWAR	8.59	1995.4
VMVDANEVPIQKMFEK	4.68	1878.2
LPVPRVSATIQR	12	1336.6
EILLSADHIIIATGGRPR	6.85	1932.3
GAAAGQRDYDLLVVGGGSGGLACAK	5.95	2306.6
APPWVPAMGFTLAPSLGCFVGSR	8.29	2361.8
SPATVEAQPLPAS	4	1267.4
GETEELQANACTNPAVHEK	4.48	2041.2
ATSLGRPEEEEDELAHR	4.41	1939.0
NASLISALSTGR	9.75	1189.3
HLFCLYVVSK	8.21	1208.5
TRLQLQDAGPAR	9.26	1325.5

Peptide sequence	pI	MW
LPPLPLTLALGAFLNHR	9.76	1843.3
LTAASVGVQGSGWGWLGFNK	8.75	2035.3
VPADLGAEAGLQQLLGALR	4.37	1892.2
FVTHVSDWGALATISTLEAVR	5.32	2273.6
IYAGQMAVLGR	8.75	1178.4
AEGSDVANAVLDGADCIMLSGETAK	3.77	2437.7
IEGTPLETIQK	4.53	1228.4
LWGLTEMFPER	4.53	1378.6

Table 1. List of peptides identified from mitochondrial fraction by the two-dimensional separation strategy. Isoelectric point (pI) and molecular weight (MW) of each peptide are provided. Threshold for peptide identifications through MASCOT is that the probability of a random match is less than 0.05.

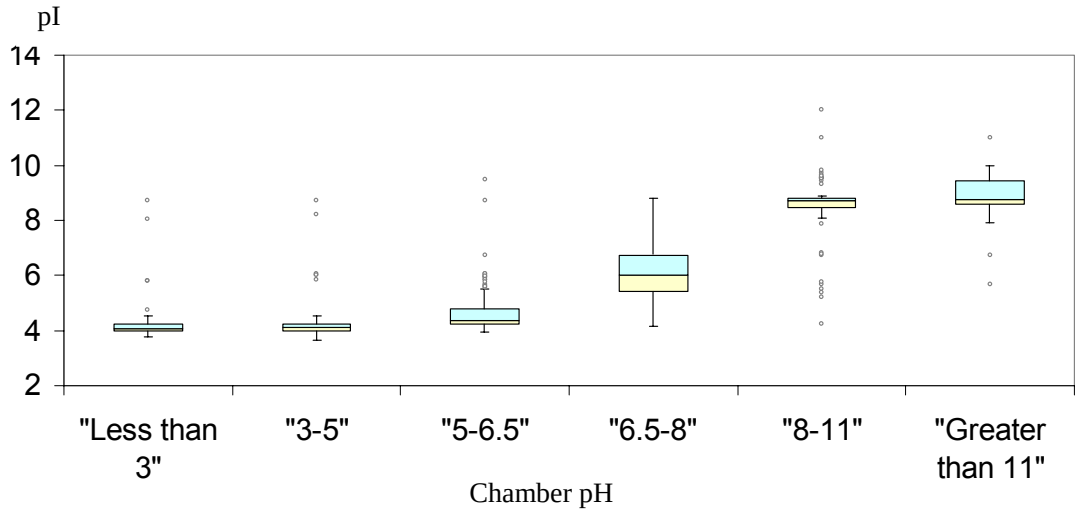


Figure 16. pI Range of peptides from different chamber fractions

The correlation between peptide pI and chamber pH defined by boundary immobiline membranes is presented. Each bar shows the pI range of identified peptides in each fraction. Fifty percent of the peptides are in the box. The top of the box cuts off lowest 25% of the data and the bottom of the box cuts off highest 25% of the data. The middle line inside the box shows the median value of the data. A horizontal line is connected a small horizontal mark (whisker) with the box. The lower whisker represents 1.5×range of the box lower than the 25% of the data and the higher whisker represents 1.5×range of the box higher than the 75% of the data. The closed dots indicate outliers which are lower or higher than the whiskers.

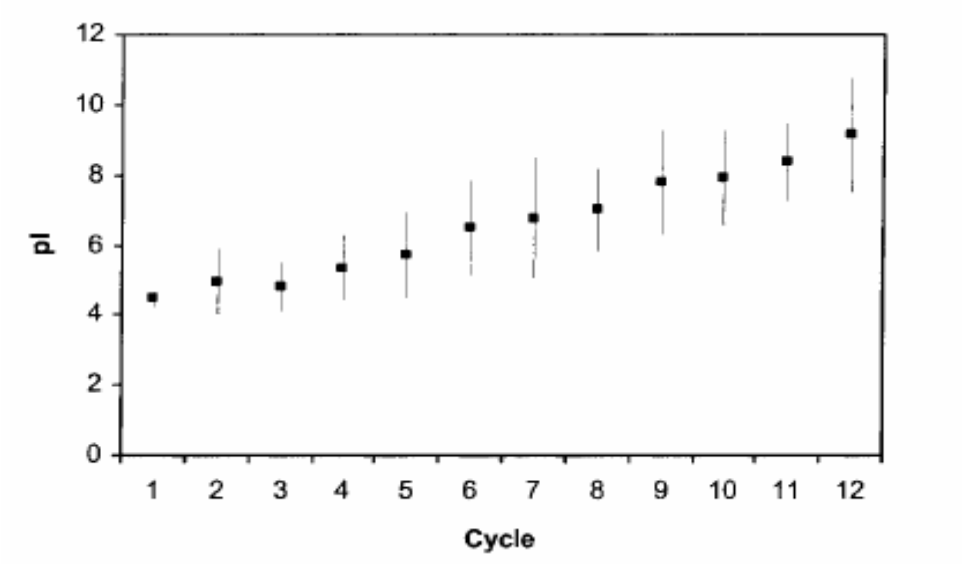


Figure 17. Average pI values of peptides separated from the digest of phosphorylase B using cation exchange chromatography (80). A total of 12-step gradients with increasing salt concentrations was performed on cation exchange columns.

peptides in the range of pI 7 to 8 (Figure 16), which is in agreement with the previous research in our lab (115) and the results from Bundy and coworker, who integrated gel based isoelectric focusing with LC-MS/MS and found the similar pI distribution for their yeast and *E. coli* digest peptides (167-168).

Figure 18 shows the total percentage of peptides identified in each fraction. The middle two fractions have the largest number of peptides. Their pI range is about from 4 to 8. This is not surprising since only those larger peptides have a higher chance of having negatively or positively charged amino acids (glutamic acid, aspartic acid or histidine), which make the peptide more acidic or basic. These peptide identifications are not necessarily unique since there is a small portion of peptides identified in multiple chambers. As shown in figure 19, about 80% of the peptides appear in one fraction, and only 5% peptides present in more than two fractions. This also demonstrates the good resolution of sIEF in our research. In addition, it seems that this overlap is much smaller than when SCX separation is used as the first dimensional separation (80).One change which can be made to further reduce the overlap is to wash the HPLC and desalting columns extensively to remove possible high abundant peptides leaving in the column, such as going through the same gradient elution twice with deionized water before the next fraction is injected.

As previously mentioned, the analysis of the sample can be repeated several times on LC-MS/MS to catch new ions each time, obtain more peptide identifications and thus increase the overall proteome coverage. The

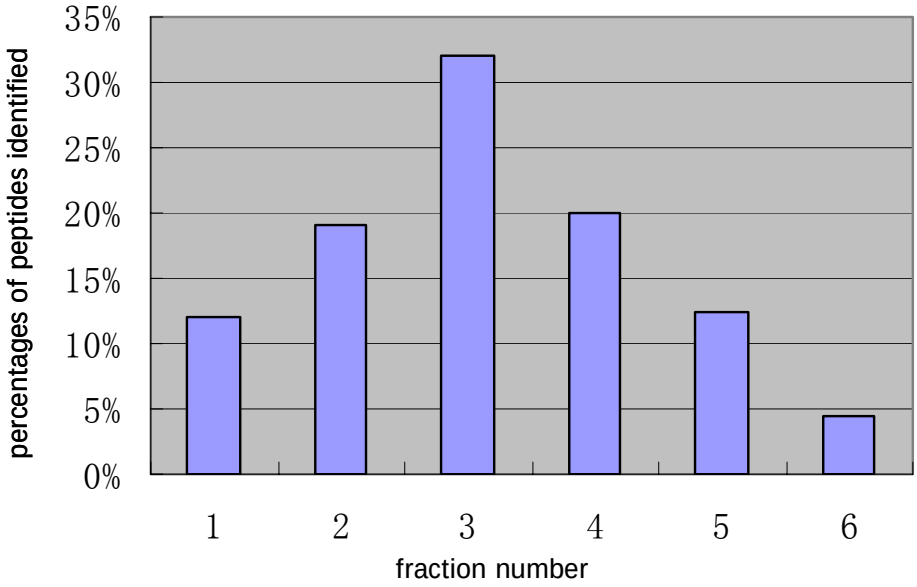


Figure 18. The percentage of identified peptides in each chamber fraction from solution IEF separation

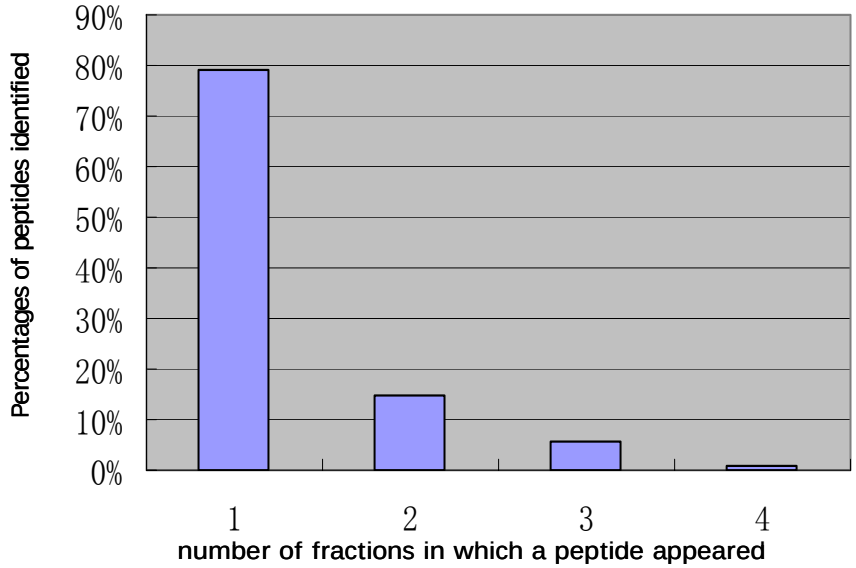


Figure 19. The percentage of identified peptides presenting in a unique fraction or in multiple fractions from solution IEF separation

cumulative number of identified peptides increased with increasing runs (Figure 20). A total of 637 peptides was identified from the two-dimensional separation of mitochondrial soluble fraction (Table 1).

Protein identification and classification

Each protein was identified based on sequences from identified peptides using MASCOT software. The assembling of the unique peptides results in a total of 278 distinct proteins identified from 637 peptides in the mitochondrial soluble fraction (Table 2). Forty-five percent of the proteins were identified based on at least two peptides and for each protein identified with only one peptide, then MS/MS spectra was manually checked.

The proteins are assigned to sub-cellular groups annotated by SwissProt database or using Mitoprot (score between 0.8 and 1.0) (169) and PSORT II Prediction (170) softwares. Figure 21 shows a graphical representation of the subcellular locations of those identified proteins, along with the number of 637 sequenced peptides that originated from those proteins (Figure 22). Of the 278 proteins identified, 166, or 60%, have been categorized as mitochondrial proteins. There are also many mitochondrial-associated proteins identified which are not annotated as mitochondrial proteins by Swissprot or other prediction software. For example, galectin-3 binding protein, which is annotated by SwissProt as a secreted protein, was reported to be enriched in mitochondria and prevent mitochondrial damage and cytochrome c release (171-172). Another protein, cathepsin D, which was known to be required for

Bax insertion into the mitochondrial outer membrane during the apoptosis process (173), is classified as lysosomal protein, however. The list of identified proteins in this study may offer evidence for other researchers when they study compartmental locations of proteins or specific protein functions.

The distribution of pI and molecular weight of identified mitochondrial proteins is demonstrated in figure 23. More than half proteins have pI greater than 7, which is in agreement with the observation from other researches (177), and may indicate that cationic properties are needed for import into mitochondrial locations (e.g. inner membrane or matrix). It was also found that proteins with relatively low molecular weights (<50 kDa) dominate in our database, which probably presents another characteristic of the mitochondrial proteome. But it should be noted that the values of pI and molecular weight reported here are obtained from the protein precursors. The mature forms of proteins may undergo posttranslational processing during translocation into the mitochondria.

The protein list was also compared with the work of Dr. Strong, a former graduate student of the Fenselau's group. She studied mitochondrial proteins from MCF-7 cell lines using the two-dimensional gel electrophoresis technique (174). Figure 24 compares the mitochondrial proteins identified in her study and this research. About half of the proteins detected here were analyzed using both methods while the other half was not detected with 2D-gel method, Many proteins with high pI values, such as cytochrome c (pI: 9.59) and some

mitochondrial ribosomal proteins, were not detected using 2D-gels, but were found in this study. The pI range of 2D-gel is normally from 3 to 10 and very basic proteins are difficult to focus near or beyond the edge of the range. In addition, since Dr. Strong used rehydration buffer containing a high concentration of urea and detergent to solubilize the mitochondrial pellet, many mitochondrial membrane proteins may be extracted and be shown on the 2D map even though membrane proteins can not be focused well. It can be seen that the shotgun strategy and the traditional 2D gel can be complementary in analyzing complex protein mixtures.

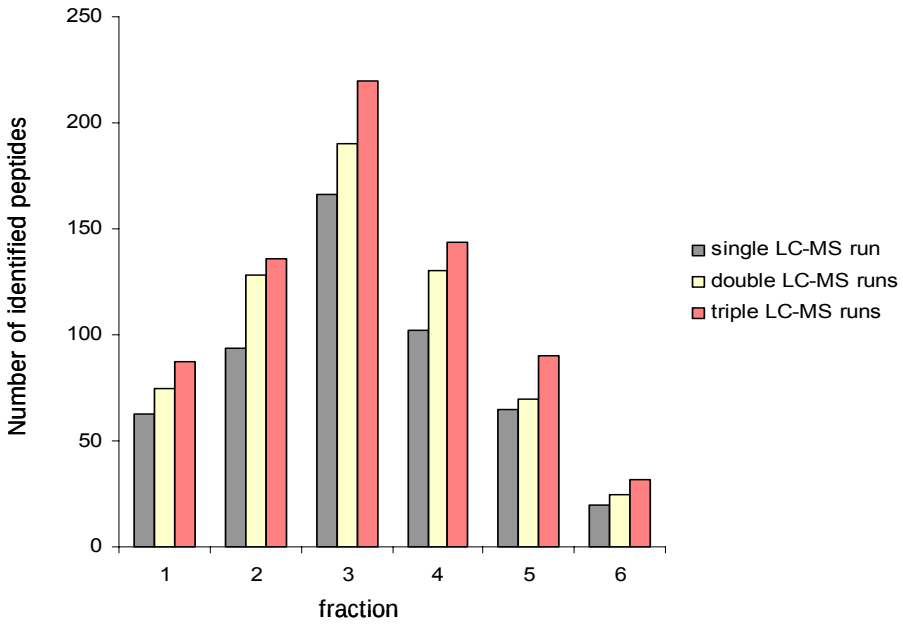


Figure 20. The cumulative number of peptides identified in each fraction as a function of injection times on LC-MS/MS. Gray bars: numbers of identified peptides in different fractions for the first LC-MS/MS analysis; Yellow bars: numbers of identified peptides in different fractions for the first two analyses; Red bars: numbers of identified peptides in different fractions for the total three analyses.

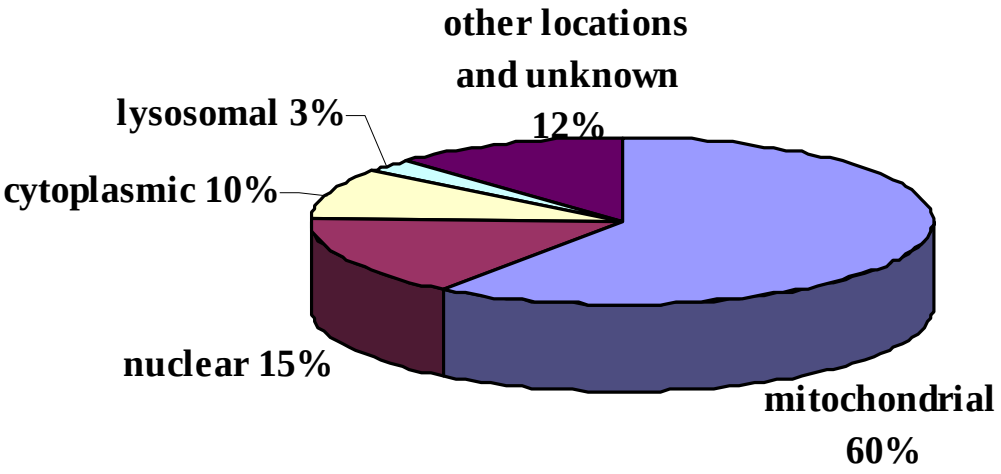


Figure 21. Subcellular distribution of the identified proteins from the drug susceptible MCF-7 cancer cells.

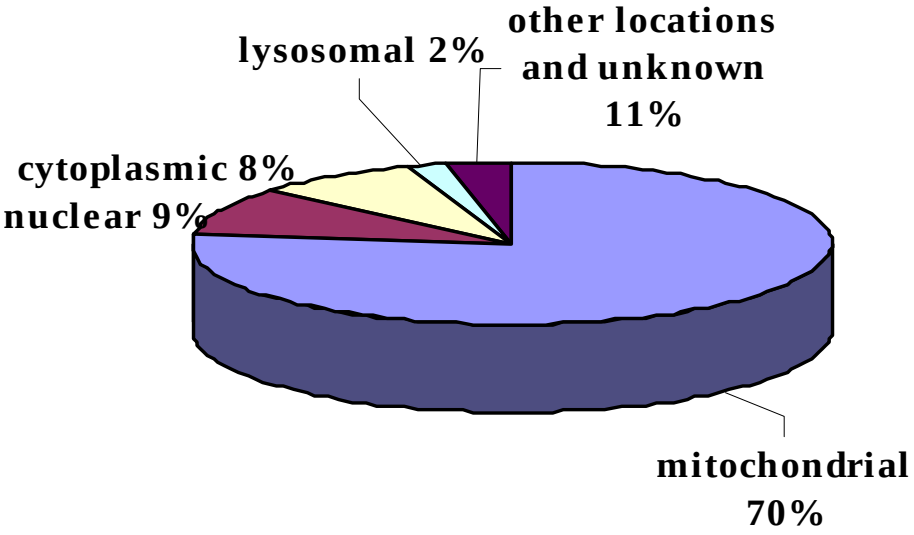


Figure 22. Subcellular distribution of the identified peptides using for those protein identifications in the MCF-7 cells.

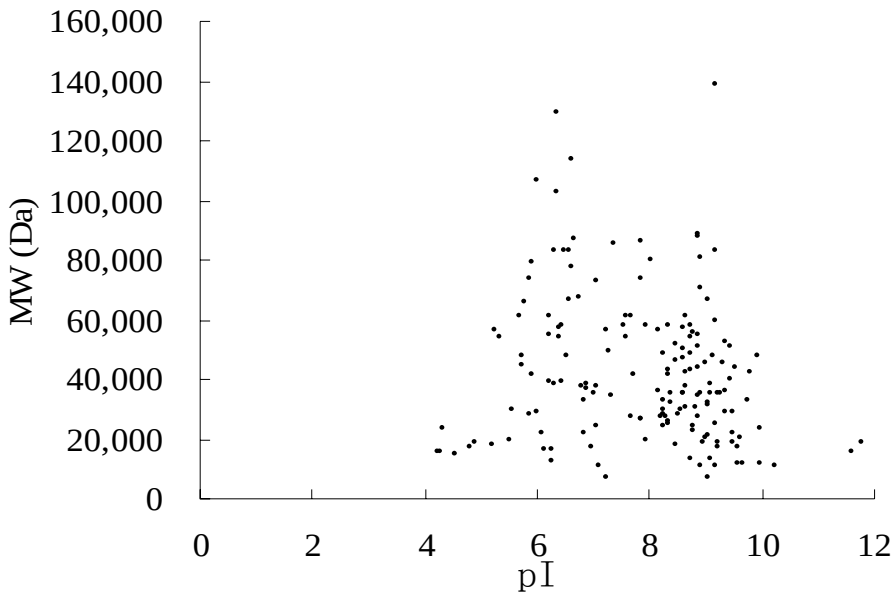


Figure 23. Identified mitochondrial proteins plotted according to pI and molecular weight.

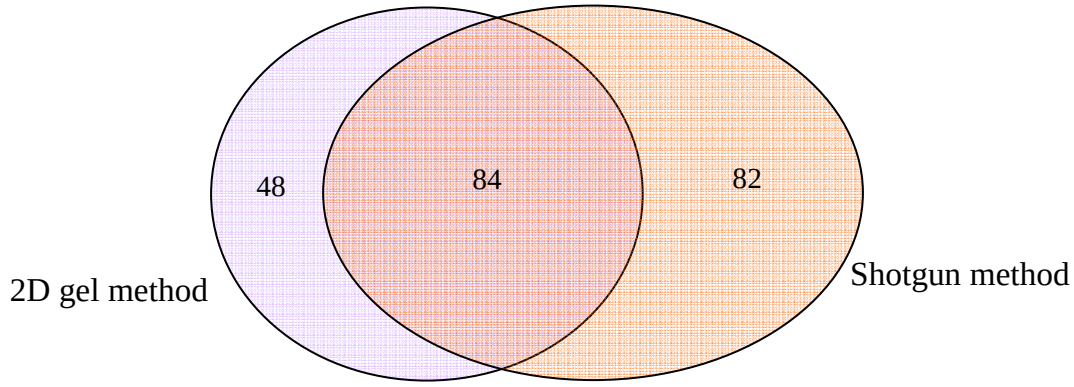


Figure 24 Comparison of the mitochondrial proteins identified in this study with proteins identified using 2D gel electrophoresis and mass spectrometry (174)

Accesion Number	Protein name	pI	MW	Subcellular Location	# of peptides
O00411	DNA-directed RNA polymerase	9.19	138620	M	1
O14561	Acyl carrier protein	4.82	17417	M	2
O15382	Branched-chain-amino-acid aminotransferase	8.88	44288	M	3
O43497	Voltage-dependent T-type calcium channel alpha-1G subunit	6.14	262472	unknown	3
O43676	NADH-ubiquinone oxidoreductase B12 subunit	9.19	11271	M	1
O43837	Isocitrate dehydrogenase [NAD] subunit beta	8.64	42212	M	1
O60513	Beta-1,4-galactosyltransferase 4	9.18	40041	Golgi	1
O75311	Glycine receptor alpha-3 chain precursor	8.6	53800	unknown	1
O75367	Core histone macro-H2A.1	9.8	39486	N	2
O75390	Citrate synthase	8.45	51712	M	7
O75438	NADH-ubiquinone oxidoreductase MNLL subunit	9.03	6961	M	1
O75439	Mitochondrial processing peptidase beta subunit	6.38	54366	M	1
O75629	CREG1 protein precursor	7.06	24075	M	1
O75691	Down-regulated in metastasis protein	7.07	318426	N	1
O75947	ATP synthase D chain	5.22	18360	M	1
O75964	ATP synthase g chain	9.65	11428	M	1
O94826	Mitochondrial precursor proteins import receptor	6.75	67455	M	2
O94925	Glutaminase, kidney isoform, mitochondrial precursor	7.85	73461	M	2
O95071	Ubiquitin--protein ligase EDD	5.59	309352	unknown	2
O95202	Leucine zipper-EF-hand containing transmembrane protein 1	6.3	83354	M	1
O95299	NADH-ubiquinone oxidoreductase 42 kDa subunit	6.87	37147	M	1
O95613	Pericentrin 2	5.39	378081	unknown	2
O95831	Programmed cell death protein 8	9.04	66901	M	1
O96008	Probable mitochondrial import receptor subunit TOM40 homolog	6.79	37893	M	1
P00367	Glutamate dehydrogenase 1	7.66	61398	M	7
P00505	Aspartate aminotransferase	9.14	47476	M	4
P02545	Lamin A/C	6.57	74139	N	3
P04075	Fructose-bisphosphate aldolase A	8.39	39289	unknown	3
P04179	Superoxide dismutase [Mn]	8.35	24722	M	1
P04181	Ornithine aminotransferase	5.72	44808	M	4
P04406	Glyceraldehyde-3-phosphate dehydrogenase	8.58	35922	C	4
P04844	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase	5.44	69284	ER	1
P05114	Nonhistone chromosomal protein HMG-14	9.61	10528	N	1

P05114	Nonhistone chromosomal protein HMG-14	9.61	10528	N	1
P05141	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2)	9.76	32764	M	1
P05165	Propionyl-CoA carboxylase alpha chain	6.63	77354	M	1
P05166	Propionyl-CoA carboxylase beta chain	7.56	58206	M	1
P05783	Keratin, type I cytoskeletal 18	5.34	47897	keratin	7
P05787	Keratin, type II cytoskeletal 8	5.52	53573	keratin	11
P06576	ATP synthase beta chain	5.26	56560	M	10
P06865	Beta-hexosaminidase alpha chain precursor	5.04	60689	L	2
P07339	Cathepsin D precursor	6.1	44552	L	2
P07686	Beta-hexosaminidase beta chain precursor	6.29	63111	L	1
P07900	Heat shock protein HSP 90-alpha	4.94	84529	C	3
P07954	Fumarate hydratase	8.85	54637	M	6
P08238	Heat shock protein HSP 90-beta	4.97	83133	C	1
P08559	Pyruvate dehydrogenase E1 component alpha subunit	8.35	43296	M	1
P08727	Keratin, type I cytoskeletal 19	5.04	44079	keratin	4
P09622	Dihydrolipoyl dehydrogenase	7.59	54150	M	5
P10109	Adrenodoxin	5.51	19393	M	1
P10253	Lysosomal alpha-glucosidase precursor	5.62	105338	L	2
P10412	Histone H1.4	11.03	21734	N	2
P10515	Dihydrolipoyllysine-residue acetyltransferase component	5.79	65781	M	3
P10606	Cytochrome c oxidase polypeptide Vb	9.07	13696	M	1
P10809	60 kDa heat shock protein	5.7	61055	M	29
P10911	Proto-oncogene DBL	5.72	107673	C	2
P11047	Laminin gamma-1 chain precursor	5.01	177607	extracellular	1
P11177	Pyruvate dehydrogenase E1 component beta subunit	6.2	39219	M	3
P11498	Pyruvate carboxylase	6.37	129634	M	3
P12532	Creatine kinase	8.6	47037	M	2
P12956	ATP-dependent DNA helicase II	6.23	69712	N	1
P13804	Electron transfer flavoprotein alpha-subunit	8.62	35080	M	11
P13995	Bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase	7.34	34137	M	1
P14618	Pyruvate kinase, isozymes M1/M2	7.95	57806	M	1
P14625	Endoplasmin precursor	4.76	92469	ER	1
P14927	Ubiquinol-cytochrome c reductase complex 14 kDa protein	8.75	13399	M	1
P15586	N-acetylglucosamine-6-sulfatase precursor	8.6	62082	L	1
P16278	Beta-galactosidase precursor	6.1	76091	L	2
P16401	Histone H1.5	10.91	22449	N	4
P16402	Histone H1.3	11.02	22219	N	1

P16403	Histone H1.2	10.94	21234	N	1
P16471	Prolactin receptor precursor	5.23	69506	unknown	1
P16836	Short chain 3-hydroxyacyl-CoA dehydrogenase	11.61	15909	M	1
P17317	Histone H2A.z	10.58	13422	N	1
P18754	Regulator of chromosome condensation	7.18	44969	N	1
P19367	Hexokinase, type I	6.36	102486	M	2
P19404	NADH-ubiquinone oxidoreductase 24 kDa subunit	8.22	27392	M	1
P20674	Cytochrome c oxidase polypeptide Va	5.89	79468	M	2
P20700	Lamin B1	5.11	66277	N	1
P21796	Voltage-dependent anion-selective channel protein 1	8.63	30623	M	3
P21817	Ryanodine receptor 1	5.18	565176	Membrane	3
P21912	Succinate dehydrogenase [ubiquinone] iron-sulfur protein	9.03	31628	M	1
P22033	Methylmalonyl-CoA mutase	6.48	83120	M	2
P22570	NADPH:adrenodoxin oxidoreductase	8.72	53837	M	1
P22626	Heterogeneous nuclear ribonucleoproteins A2/B1	8.97	37430	N	1
P22695	Ubiquinol-cytochrome-c reductase complex core protein 2	8.74	48443	M	3
P23284	Peptidyl-prolyl cis-trans isomerase B precursor	9.33	22742	ER	1
P23434	Glycine cleavage system H protein	4.91	18911	M	1
P23528	Cofilin-1	8.26	18371	N/C	1
P24534	Elongation factor 1-beta	4.5	24633	unknown	1
P24539	ATP synthase B chain	9.37	28909	M	2
P24752	Acetyl-CoA acetyltransferase	8.98	45200	M	4
P25705	ATP synthase alpha chain	9.16	59751	M	5
P26440	Isovaleryl-CoA dehydrogenase	8.45	46319	M	3
P26583	High mobility group protein 2	7.77	23903	N	1
P27695	DNA-(apurinic or apyrimidinic site) lyase	8.42	35423	N	1
P27695	DNA-(apurinic or apyrimidinic site) lyase	8.42	35423	N	1
P27708	CAD protein	6.02	242984	C	1
P28001	Histone H2A.a	11.05	14004	N	1
P29375	Retinoblastoma-binding protein 2	6.42	195816	N	1
P29966	Myristoylated alanine-rich C-kinase substrate	4.47	31423	unknown	1
P30042	ES1 protein homolog	8.5	28170	M	3
P30044	Peroxiredoxin 5	6.96	16899	M	2
P30048	Thioredoxin-dependent peroxide reductase	7.68	27693	M	4
P30049	ATP synthase delta chain	4.53	15019	M	1
P30084	Enoyl-CoA hydratase	5.88	28354	M	4
P30405	Peptidyl-prolyl cis-trans isomerase	9.49	22040	M	2
P30536	Peripheral-type benzodiazepine receptor	9.23	18779	M	1

P30837	Aldehyde dehydrogenase X	6.41	57217	M	2
P31040	Succinate dehydrogenase [ubiquinone] flavoprotein subunit	7.06	72692	M	7
P31937	3-hydroxyisobutyrate dehydrogenase	8.38	35329	M	3
P32322	Pyrroline-5-carboxylate reductase 1	7.18	33361	C	1
P34897	Serine hydroxymethyltransferase	8.76	55993	M	8
P34947	G protein-coupled receptor kinase 5	8.39	67787	unknown	1
P35232	Prohibitin	5.57	29804	M	6
P35270	Sepiapterin reductase	8.25	28048	M	1
P35579	Myosin heavy chain, nonmuscle type A	5.5	226401	N	1
P36551	Coproporphyrinogen III oxidase	8.59	50152	M	1
P36776	Lon protease homolog	6.01	106422	M	5
P36957	Dihydrolipoyllysine-residue succinyltransferase component	5.89	41349	M	5
P38117	Electron transfer flavoprotein beta-subunit	8.29	27712	M	4
P38646	Stress-70 protein	5.87	73681	M	20
P40926	Malate dehydrogenase	8.92	35531	M	10
P40939	Trifunctional enzyme alpha subunit	9.16	83000	M	2
P42126	3,2-trans-enoyl-CoA isomerase	6	28736	M	3
P42167	Lamina-associated polypeptide 2, isoforms beta/gamma	9.39	50539	N	1
P42226	Signal transducer and activator of transcription 6	5.84	94135	C	1
P42338	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit, beta isoform	6.69	122762	unknown	1
P42704	130 kDa leucine-rich protein	5.49	145201	C	7
P42765	3-ketoacyl-CoA thiolase	8.32	41924	M	5
P42785	Lysosomal Pro-X carboxypeptidase precursor	6.76	55800	L	1
P43897	Elongation factor Ts	8.62	35391	M	2
P45880	Voltage-dependent anion-selective channel protein 2	6.32	38069	M	2
P45954	Acyl-CoA dehydrogenase, short/branched chain specific	6.53	47485	M	3
P46100	Transcriptional regulator ATRX	6.23	282566	C	2
P46531	Neurogenic locus notch homolog protein 1 precursor	4.99	272500	unknown	1
P47985	Ubiquinol-cytochrome c reductase iron-sulfur subunit	8.55	29652	M	1
P48047	ATP synthase oligomycin sensitivity conferral protein	9.97	23277	M	1
P48735	Isocitrate dehydrogenase [NADP]	8.88	50909	M	9
P49411	Elongation factor Tu	7.26	49542	M	9
P49419	Aldehyde dehydrogenase family 7 member A1	6.24	55200	M	4
P49448	Glutamate dehydrogenase 2	8.63	61434	M	2

P49748	Acyl-CoA dehydrogenase, very-long-chain specific	8.92	70390	M	5
P49753	Peroxisomal acyl-coenzyme A thioester hydrolase 2a	8.93	53257	P	2
P50213	Isocitrate dehydrogenase [NAD] subunit alpha	6.46	39592	M	1
P50416	Carnitine O-palmitoyltransferase I	8.85	88368	M	1
P50440	Glycine amidinotransferase	8.26	48455	M	1
P50897	Palmitoyl-protein thioesterase 1 precursor	6.07	34194	L	2
P51553	Isocitrate dehydrogenase [NAD] subunit gamma	8.75	42794	M	1
P51649	Succinate semialdehyde dehydrogenase	8.62	57215	M	2
P51687	Sulfite oxidase	5.35	53885	M	1
P51800	Chloride channel protein ClC-Ka	7.62	75285	unknown	1
P51970	NADH-ubiquinone oxidoreductase 19 kDa subunit	7.93	19974	M	1
P52564	Dual specificity mitogen-activated protein kinase kinase 6	7.01	37492	unknown	1
P52815	39S ribosomal protein L12	9.05	21348	M	2
P53597	Succinyl-CoA ligase [GDP-forming] alpha-chain	9.11	35047	M	1
P54819	Adenylate kinase isoenzyme 2	7.85	26347	M	1
P54819	Adenylate kinase isoenzyme 2	7.85	26347	M	1
P54886	Delta 1-pyrroline-5-carboxylate synthetase	6.66	87302	M	2
P55084	Trifunctional enzyme beta subunit	9.45	51294	M	1
P60709	Actin, cytoplasmic 1	5.29	41737	C	4
P61604	10 kDa heat shock protein	8.91	10801	M	5
P62736	Actin, aortic smooth muscle	5.24	42009	C	2
P62805	Histone H4	11.36	11236	N	2
P62807	Histone H2B.a/g/h/k/l	10.32	13688	N	1
P62988	Ubiquitin	6.56	8565	N/C	2
P63104	14-3-3 protein zeta/delta	4.73	27745	C	2
P63261	Actin, cytoplasmic 2	5.31	41793	C	2
P68104	Elongation factor 1-alpha 1	9.1	50141	C	1
P68133	Actin, alpha skeletal muscle	5.23	42051	C	1
P80404	4-aminobutyrate aminotransferase	8.17	56439	M	7
P80723	Brain acid soluble protein 1	4.64	22562	N	4
P82650	Mitochondrial 28S ribosomal protein S22	7.7	41280	M	1
P82663	Mitochondrial 28S ribosomal protein S25	8.99	20116	M	1
P82675	Mitochondrial 28S ribosomal protein S5	9.93	48006	M	2
P82909	Mitochondrial 28S ribosomal protein S36	9.99	11466	M	1
P82921	Mitochondrial 28S ribosomal protein S21	10.23	10742	M	1
P84243	Histone H3.3	11.27	15197	N	1
P99999	Cytochrome c	9.59	11618	M	2
Q00325	Phosphate carrier protein	9.45	40095	M	1
Q01826	DNA-binding protein SATB1	6.1	85957	N	1

Q02218	2-oxoglutarate dehydrogenase E1 component	6.62	113476	M	3
Q02252	Methylmalonate-semialdehyde dehydrogenase	8.72	57840	M	4
Q02338	D-beta-hydroxybutyrate dehydrogenase	9.1	38157	M	2
Q04837	Single-stranded DNA-binding protein	9.59	17260	M	4
Q05639	Elongation factor 1-alpha 2	9.11	50470	N	1
Q06830	Peroxiredoxin 1	8.27	22110	C	1
Q07021	Complement component 1, Q subcomponent binding protein	4.32	23783	M	3
Q08257	Quinone oxidoreductase	8.56	35207	C	1
Q08380	Galectin-3 binding protein precursor	5.13	65331	secreted	3
Q09666	Neuroblast differentiation associated protein AHNAK	6.29	312493	N	1
Q10713	Mitochondrial processing peptidase alpha subunit	6.45	58253	M	2
Q12931	Heat shock protein 75 kDa	8.05	79961	M	4
Q13011	Delta3,5-delta2,4-dienoyl-CoA isomerase	8.16	35816	M	2
Q13268	Dehydrogenase/reductase SDR family member 2	8.9	27307	N	1
Q13405	Mitochondrial 39S ribosomal protein L49	9.47	19198	M	2
Q13535	Serine-protein kinase ATR	7.17	301367	N	1
Q14061	Cytochrome c oxidase copper chaperone	7.22	6784	M	1
Q14204	Dynein heavy chain	6.01	532408	C	2
Q14315	Filamin C	5.68	290959	C	3
Q14517	Cadherin-related tumor suppressor homolog precursor	4.84	506278	membrane	1
Q14573	Inositol 1,4,5-trisphosphate receptor type 3	6.04	304038	ER	2
Q15046	Lysyl-tRNA synthetase	5.94	68048	C	1
Q16512	Protein kinase N1	5.93	103990	C	2
Q16595	Frataxin	8.8	23135	M	1
Q16698	2,4-dienoyl-CoA reductase	9.35	36068	M	3
Q16740	Putative ATP-dependent Clp protease proteolytic subunit	8.26	30180	M	2
Q16762	Thiosulfate sulfurtransferase	6.83	33277	M	2
Q16774	Guanylate kinase	6.11	21594	M	1
Q16795	NADH-ubiquinone oxidoreductase 39 kDa subunit	9.81	42510	M	1
Q16822	Phosphoenolpyruvate carboxykinase	6.58	67004	M	6
Q16836	Short chain 3-hydroxyacyl-CoA dehydrogenase	8.88	34278	M	6
Q6FI13	Histone H2A.o	10.9	13964	N	1
Q6P587	Fumarylacetoacetate hydrolase domain containing protein 1	6.96	24843	unknown	1
Q7Z4W1	L-xylulose reductase	8.33	25913	M	2
Q7Z589	EMSY protein	9.37	141468	N	1

Q86SX6	Glutaredoxin-related protein C14orf87	6.28	16628	M	1
Q86TX2	Cytosolic acyl coenzyme A thioester hydrolase, inducible	6.9	46277	C	1
Q8HXP0	Superoxide dismutase [Mn]	6.86	22248	M	1
Q8IUG5	Myosin XVIIIB	6.49	285185	C	1
Q8IYE1	Coiled-coil domain containing protein 13	8.93	80854	M	2
Q8N4Q1	Coiled-coil-helix-coiled-coil-helix domain containing protein 4	4.23	15996	M	1
Q8N8R3	Mitchondrial carnitine/acylcarnitine carrier protein CACL	9.03	32062	M	1
Q8TCS8	Polyribonucleotide nucleotidyltransferase 1	7.87	85951	M	2
Q92523	Carnitine O-palmitoyltransferase I	8.86	87801	M	1
Q92552	Mitochondrial 28S ribosomal protein S27	5.72	47669	M	1
Q92665	28S ribosomal protein S31	9.32	45318	M	1
Q93081	Histone H3/b	11.13	15273	N	1
Q96EY1	DnaJ homolog subfamily A member 3	9.36	52538	M	1
Q96EY8	Cob(I)yrinic acid a,c-diamide adenosyltransferase	8.86	27388	M	1
Q96I99	Succinyl-CoA ligase [GDP-forming] beta-chain	6.15	16511	M	4
Q96MM6	Heat shock 70 kDa protein 12B	8.81	75688	unknown	1
Q96RP9	Elongation factor G 1	6.58	83471	M	2
Q96T58	Msx2-interacting protein	7.35	402248	N	1
Q99714	3-hydroxyacyl-CoA dehydrogenase type II	7.87	26792	ER	5
Q99757	Thioredoxin	8.46	18383	M	2
Q99798	Aconitate hydratase	7.36	85425	M	5
Q99807	Ubiquinone biosynthesis protein COQ7 homolog	8.77	24307	M	1
Q99996	A-kinase anchor protein 9	4.95	453667	C	2
Q9BSD7	Probable UPF0334 kinase-like protein	9.61	20713	unknown	1
Q9BSH4	UPF0082 protein PRO0477	8.37	32477	M	2
Q9BSY4	Coiled-coil-helix-coiled-coil-helix domain containing protein 5	6.28	12395	M	1
Q9BVK6	Transmembrane emp24 domain containing protein 9 precursor	6.67	25105	ER	1
Q9BWM7	Sideroflexin 3	9.26	35503	M	1
Q9BX68	Histidine triad nucleotide-binding protein 2	9.2	17162	M	3
Q9BXB5	Oxysterol binding protein-related protein 10	8.56	83970	N	1
Q9BXU1	Serine/threonine-protein kinase 31	5.06	115730	unknown	1
Q9BZZ5	Apoptosis inhibitor 5	5.84	57561	N	1
Q9H0U6	39S ribosomal protein L18	9.63	20577	M	1
Q9H2U2	Inorganic pyrophosphatase 2	7.07	37920	M	4
Q9H4K7	Putative GTP-binding protein 5	9.52	43955	M	1
Q9H9B4	Sideroflexin 1	9.22	35488	M	2
Q9H9J2	39S ribosomal protein L44	8.65	37535	M	1

Q9H9Q2	COP9 signalosome complex subunit 7b	5.83	29622	N/C	1
Q9HAV7	GrpE protein homolog 1	8.24	24279	M	3
Q9HCC0	Methylcrotonoyl-CoA carboxylase beta chain	7.58	61333	M	4
Q9NNW7	Thioredoxin reductase 2	7.23	56460	M	2
Q9NP81	Seryl-tRNA synthetase	8.35	58283	M	1
Q9NRC6	Spectrin beta chain, brain 4	6.23	416835	C	2
Q9NS69	Mitochondrial import receptor subunit TOM22 homolog	4.27	15390	M	1
Q9NZE8	39S ribosomal protein L35	11.8	19212	M	1
Q9NZJ6	Hexaprenyldihydroxybenzoate methyltransferase	7.12	10998	M	1
Q9P0J1	[Pyruvate dehydrogenase [Lipoamide]]-phosphatase 1	6.2	61054	M	2
Q9UBQ7	Glyoxylate reductase/hydroxypyruvate reductase	7.01	35668	M	1
Q9UFN0	NipSnap3A protein	9.21	28467	C	1
Q9UGV6	High mobility group protein 1-like 10	6.99	24218	N	2
Q9UIJ7	GTP:AMP phosphotransferase mitochondrial	9.16	25434	M	1
Q9UJ96	Potassium voltage-gated channel subfamily G member 2	8.33	51240	unknown	1
Q9UJZ1	Stomatin-like protein 2	6.87	38534	M	1
Q9UN73	Protocadherin alpha 6 precursor	4.93	102716	secreted	1
Q9UPY3	Endoribonuclease Dicer	5.45	217628	N	1
Q9UQE7	Structural maintenance of chromosome 3	6.77	141542	N	1
Q9Y277	Voltage-dependent anion-selective channel protein 3	8.84	30659	M	1
Q9Y3E5	Peptidyl-tRNA hydrolase 2	8.95	19194	M	1
Q9Y4A5	Transformation/transcription domain-associated protein	8.49	437600	N	1
Q9Y5Q5	Atrial natriuteric peptide-converting enzyme	4.87	116565	unknown	1
Q9Y676	28S ribosomal protein S18b	9.47	29396	M	2
Q9Y6C9	Mitochondrial carrier homolog 2	8.25	33331	M	1

Table 2. Proteins identified in the mitochondrial fraction. SwissProt accession numbers are provided. Subcellular location abbreviation: M-mitochondira, C-cytosol, N-nucleus, L-lysosome, ER-endoplasmic reticulum. The isoelectric point (pI), molecular weight (MW) and evaluation of the reliabilities of identification of each protein are provided.

Integration of ^{18}O labeling with solution isoelectric focusing (sIEF)

Since two ^{18}O atoms can be incorporated into a peptide C-terminus resulting in a mass difference of 4 Da, quantitiation of mitochondrial proteins can be achieved by enzyme-catalyzed ^{18}O labeling of mitochondrial peptides, coupled with efficient separation strategy and a mass spectrometer with adequate resolution.

Quantitative proteomics was evaluated using a model protein, lysozyme from chicken egg white. Lysozyme (200 μg) was digested and labeled with $H_2{}^{16}O$ or $H_2{}^{18}O$ both in the presence of immobilized trypsin. The use of immobilized trypsin allowed for a highly efficient catalysis because the enzyme concentration can be very high and trypsin autolysis will be minimized (175). The extent of labeling at room temperature and 37℃ was examined. Using MALDI-TOF/MS, figure 25 presents the spectra of singly charged lysozyme peptide WWCNDGR with a molecular weight of 933.3 Da. Panel A shows the isotopic distribution of this unlabeled peptide. There is a 4 Da shift between panel A and panel B, which means the peptide was labeled with two atoms of ^{18}O. The sample in panel C was labeled under the same conditions as panel B, except that the labeling process took place at 37℃ for 5 hours instead of room temperature for 10 hours. It can be seen that there is incomplete labeling in panel B even after 10-hour labeling while there are no detectable ^{16}O peaks in panel C. This indicates that the choice of 37℃ was more efficient for ^{18}O labeling due to the optimal enzyme activity at this temperature. Thus we

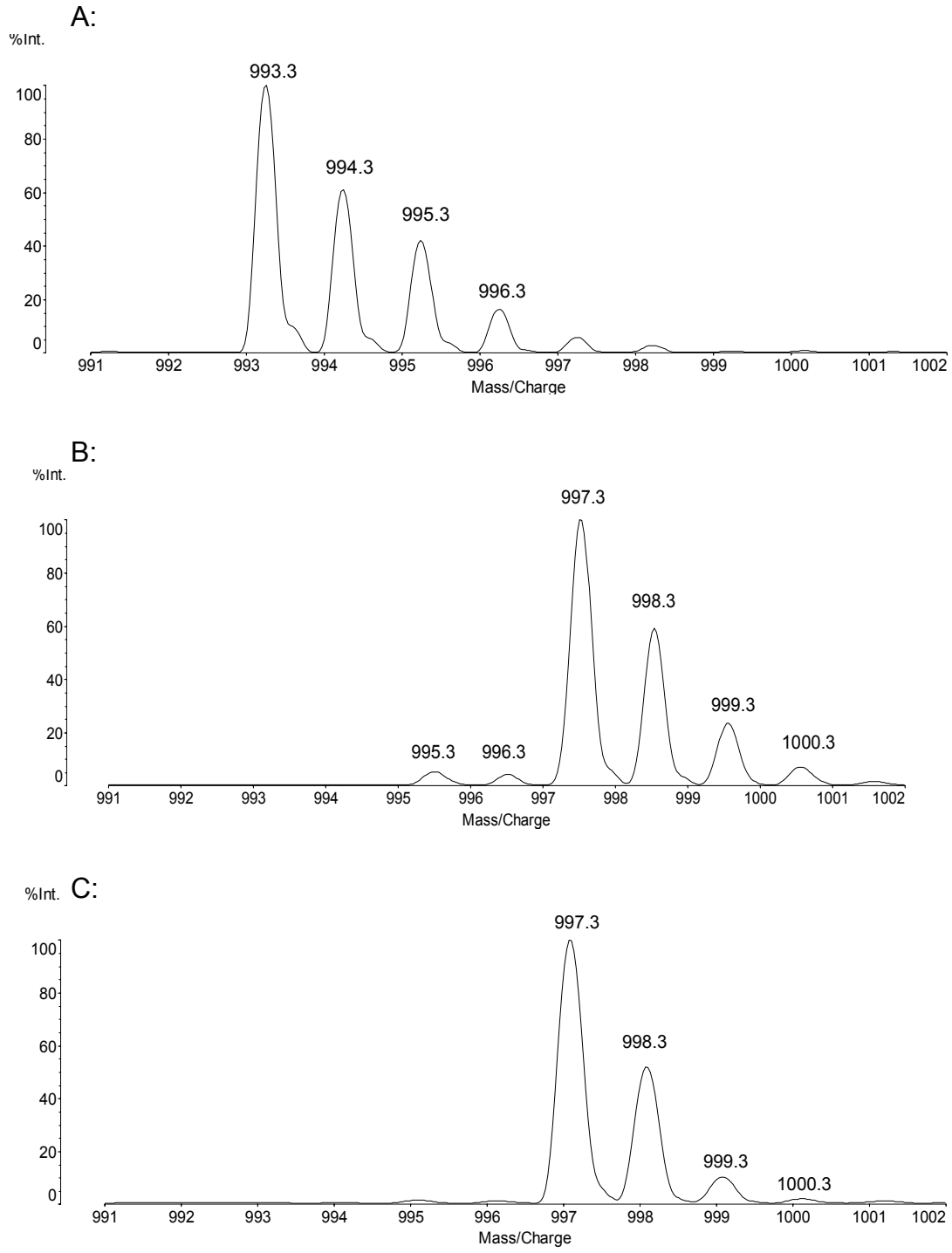


Figure 25. MALDI-TOF mass spectra of the singly-charged peptide WWCNDGR (carbamidomethylated). A: the isotopic distribution of the unlabeled peptide with monoisotopic peak at m/z 993.3. B: Peptide labeled by ^{18}O at room temperature for 10 h. C: Peptide labeled by ^{18}O at 37℃ for 5 h.

selected 37℃ and 5 h as the labeling condition for subsequent studies.

Molar equivalents of labeled and unlabeled lysozyme peptides were then combined and analyzed using MALDI-TOF mass spectrometry. Figure 26 presents the peptide mass map of the mixture of labeled and unlabeled lysozyme peptides. Enlarged partial spectra of singly-charged peptides WWCNDGR and FESNFNTQATNR are shown in figure 27. The isotopic ratio of $^{18}O/^{16}O$ was calculated from each monoisotopic labeled and unlabeled peak area, which should be very closed to the results calculated from equation 1 (see Experimental part). Table 3 lists the theoretically and experimentally digested peptides. The average ratio ($^{18}O/^{16}O$) from all detected peptide pairs is 0.97, in agreement with the mixing ratio.

Next, ^{18}O labeling was integrated with solution isoelectric focusing and capillary LC-tandem mass spectrometry to study changes in mitochondrial proteins associated with drug resistance in MCF-7 human cancer cells. The overall scheme of sIEF and LC-MS/MS analysis of $^{16}O/^{18}O$ labeled mitochondrial peptides from MCF-7 cells is shown in figure 28. Briefly, the proteins were extracted from mitochondrial pellets of MCF-7 drug susceptible and drug resistant cells and digested into peptides separately. The resultant peptides were $^{16}O/^{18}O$ labeled (drug resistant sample labeled by ^{18}O and drug susceptible sample labeled by ^{16}O). The pooled sample (1:1) was fractionated by solution IEF and then each fraction was submitted for nanoLC-ESI-TOF analysis for protein quantitative study.

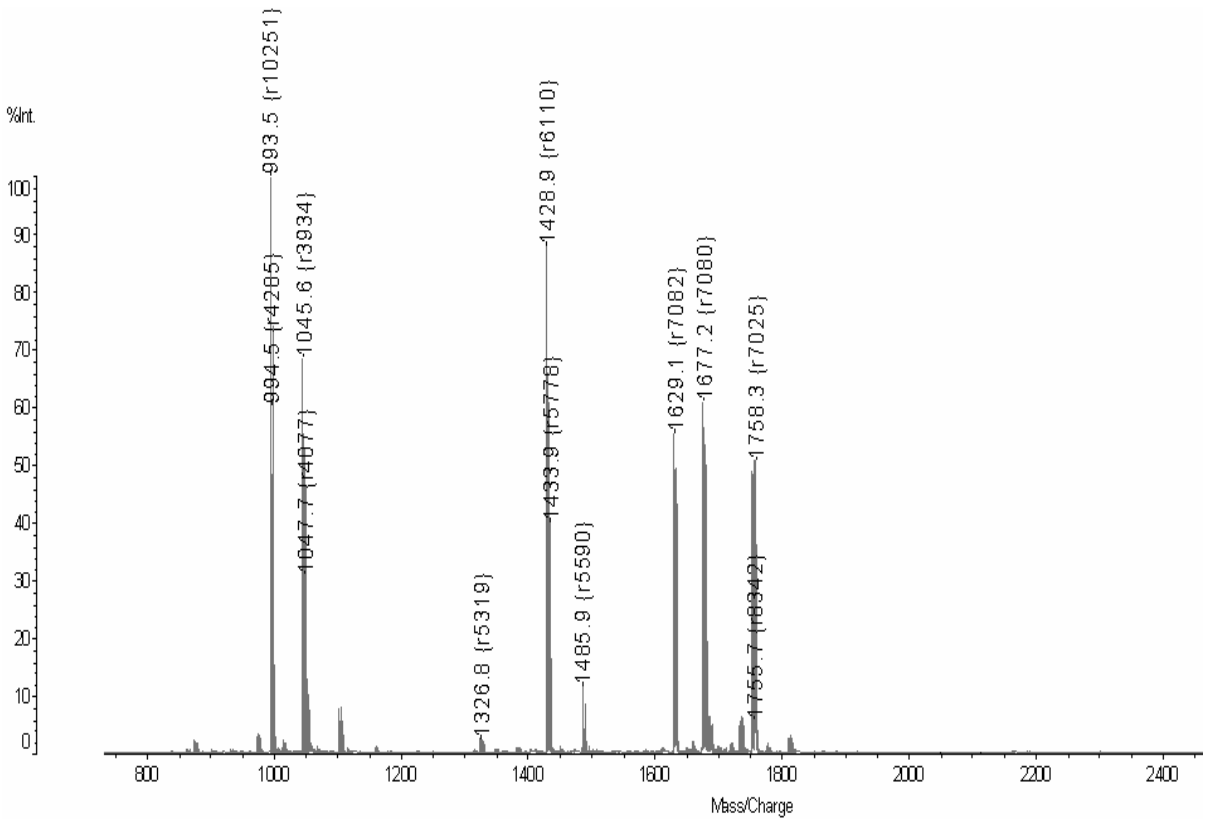


Figure 26. MALDI-TOF mass spectrum of MS scan of digested labeled and unlabeled lysozyme peptides (1:1 mixing). R represents resolution.

MH^+	Sequence	Observed peptides	$^{18}O/^{16}O$ observed ratio
993.4	WWCNDGR	yes	0.94
1045.5	GTDVQAWIR	yes	0.96
1325.6	GYSLGNWVCAAK	yes	0.95
1428.6	FESNFNTQATNR	yes	1.02
1675.8	IVSDGNGMNAWVAWR	yes	1.00
1754.8	NTDGSTDYGILQINSR	yes	0.95
2508.2	NLCNIPCSALLSSDITASVNCAK	no	N/A

Table 3. The theoretical and observed tryptic-digested lysozyme peptides and quantification results from MALDI-TOF MS.

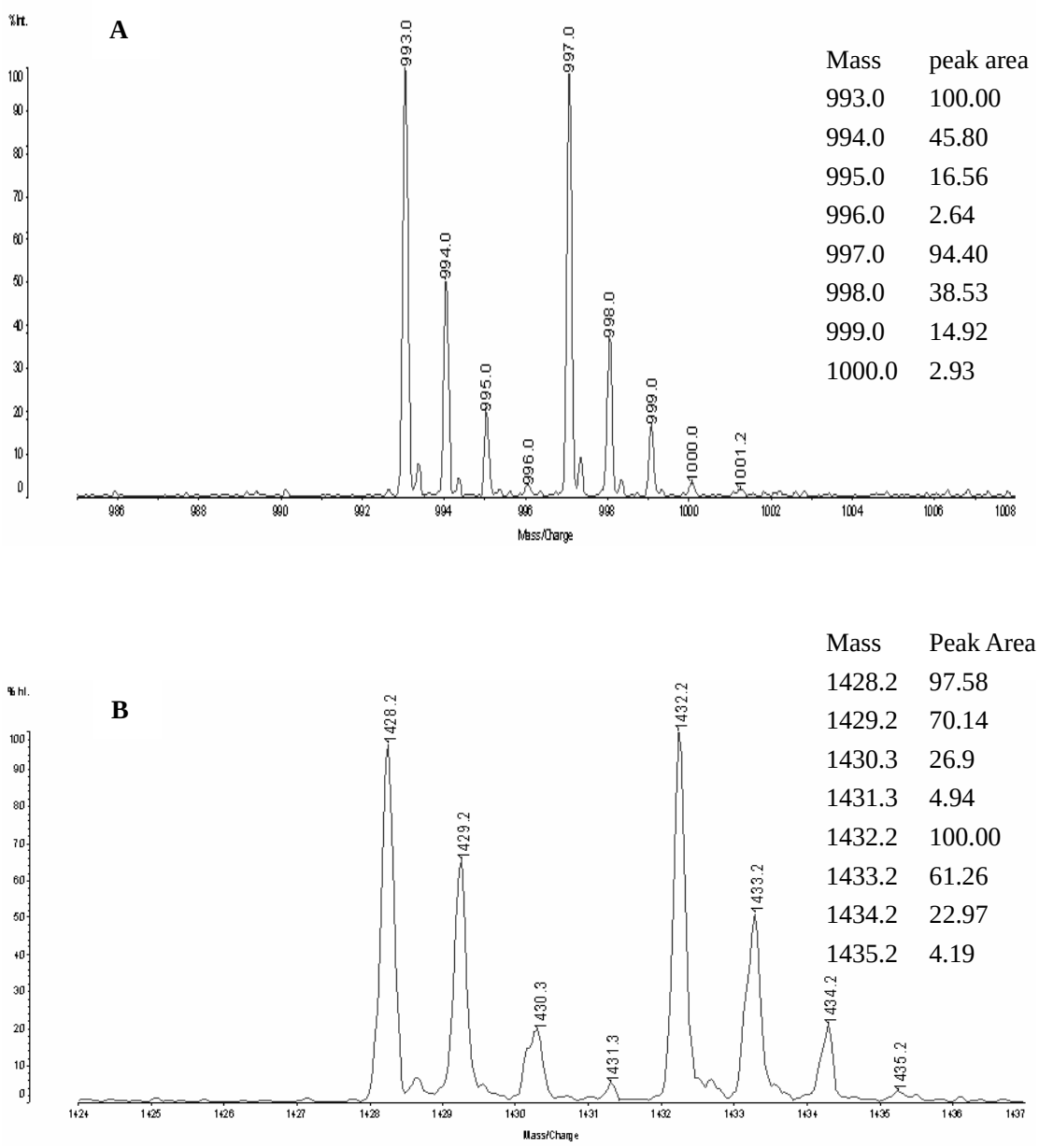


Figure 27. Mass spectra of isotope pairs of two digested lysozyme peptides recorded on a MALDI-TOF mass spectrometer. A: isotope pair of peptide WWCNDGR with monoisotopic unlabeled peak at 993.0. B: isotope pair of peptide FESNFNTQATNR with monoisotopic unlabeled peak at 1428.2. The insets list the relative peak area of each isotopic peak.

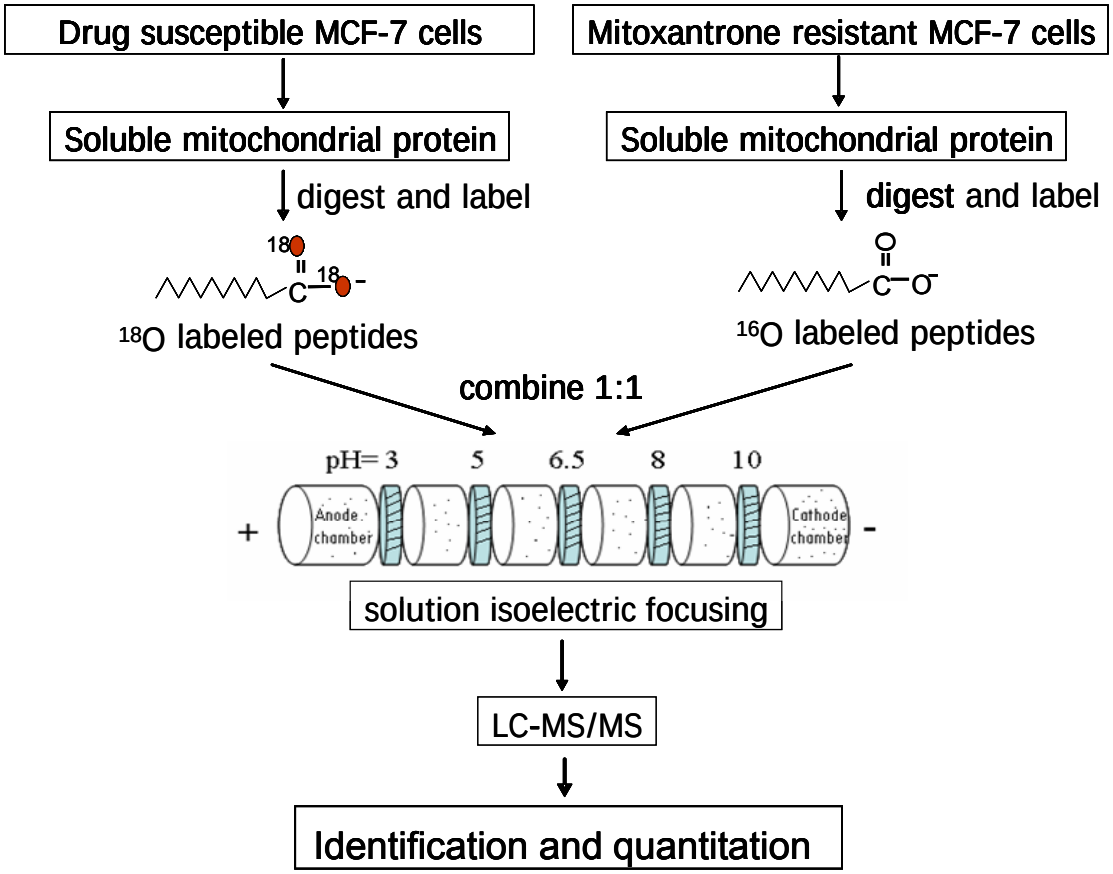


Figure 28. Overall scheme of integration of ^{18}O labeling strategy with sIEF for comparative proteomics.

Extensive peptide MS/MS fragmentations were obtained for the most three abundant MS peaks from each MS scan during an LC run. Peptides, and thus proteins, were first identified from all those MS/MS scans. The experimental peak areas for the isotopic peaks of peptides were derived from the reconstructed mass spectra of the MS scans across the chromatographic peaks. In order to account for single incorporation of an ^{18}O atom at the C-terminus of a peptide and to obtain more accurate ratio measurements, the theoretical isotopic distribution for each peptide can be obtained based on its sequence using the MS-isotope program (166). Those measured and theoretical peak areas were then used to determine the relative ratio of ^{16}O and ^{18}O quantities for each peptide isotope pair using equation 1 (shown in Experimental part). Finally, the overall ratio of a protein was calculated from an average of all identified peptide ratios. We usually consider the ratio of $^{16}O/^{18}O$ lower than 0.5 or higher than 2.0 to be a significant change biologically. The majority of the identified proteins did not show altered abundances in mitoxantrone-resistant MCF-3 cells. Figure 29 shows partial mass spectra of peptides from the stress-70 protein as an example of those proteins whose abundances do not change. The partial MS spectra of the isotopic pairs of doubly-charged labeled and unlabeled peptides TTPSVVAFTADGER and DAGQISGLNVLR are presented in panel A and B respectively. Each labeled and unlabeled peptide pair co-eluted from the LC and appeared as isotopic doublets with 2 Da apart in the mass spectrum. These two peptides are both

identified as belonging to mitochondrial protein stress-70 protein. Figures 30 and 31 illustrate partial MS spectra of peptides from two proteins that have altered abundances. They show the increase in the abundance of elongation factor Tu by 2.3 fold and the decrease in the abundance of dihydrolipoyl dehydrogenase by 2.2 fold, respectively.

It was noted that there is some suppression of peaks immediately to the right of intense peaks in the spectrum of isotopic envelope of peptide. For example, in figure 32, the isotope distribution of +2 charge state of unlabeled peptide SDLAVPSELALLK is distorted by a high ion count. The first and second ^{13}C isotopic peak (A) was lower than expected by comparison with the theoretical isotope envelope (B). This is a detector saturation effect, which happened when the detector is saturated by high ion counts in a certain situation. One of the results is that when two ions reach the detector in sequence within a short interval, some suppression of the second ion count happens because of the detector saturation by the first intense ion (176). Most mass spectrometer instruments have a correction program to minimize this saturation effect. But we still encountered some of these problems in our spectra. But fortunately, this detector saturation effect will not affect our results significantly, because only the monoisotopic peak and the peak with masses 2 Da higher are required for relative quantitation calculation for each labeled and unlabeled peptide pair.

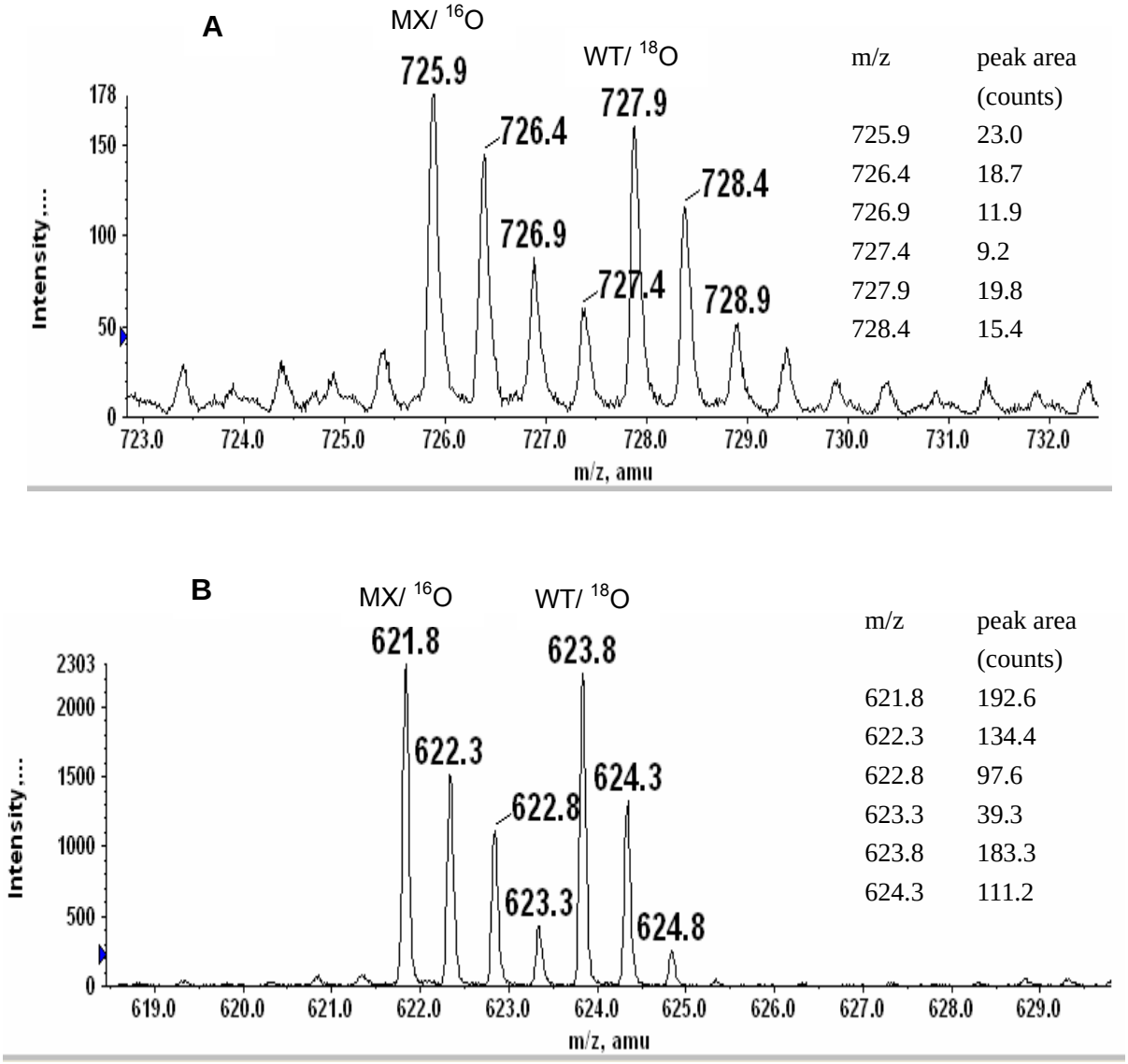


Figure 29. Partial ESI-TOF mass spectra of two peptides from stress-70 protein. Isotopic pairs of the unlabeled peptide from the mitoxantrone resistant cell line (MX) and the ^{18}O labeled peptide from the drug susceptible cell line (WT) are shown. A: the isotopic pair of the doubly-charged peptide TTPSVVAFTADGER. B: the isotopic pair of the doubly-charged peptide DAGQISGLNVLR. The insets list the relative peak area of each isotopic peak.

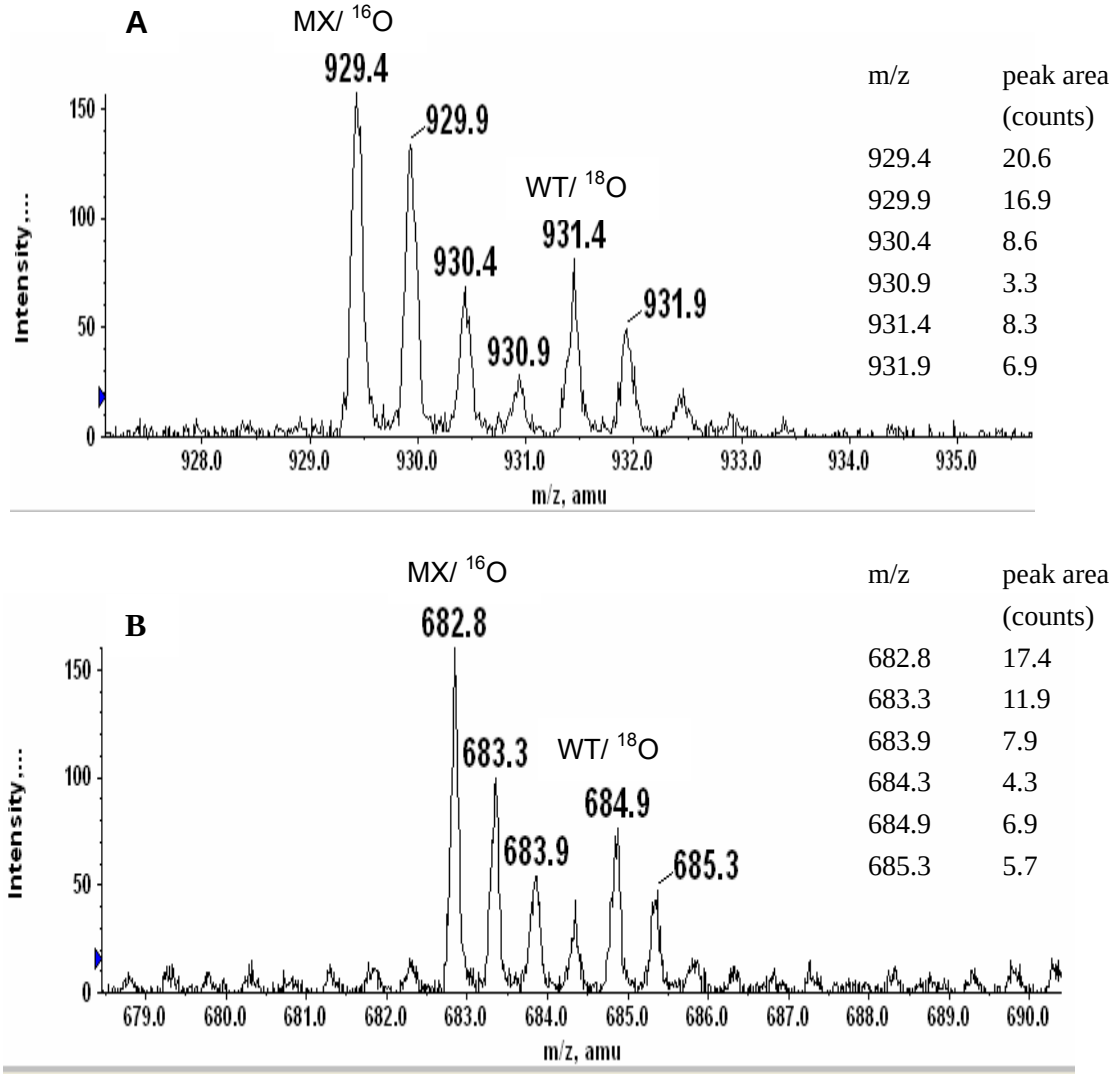


Figure 30. Partial ESI-TOF mass spectra of two peptides from mitochondrial protein elongation factor Tu. Isotopic pairs of the unlabeled peptide from the mitoxantrone resistant cell line (MX) and the ^{18}O labeled peptide from the drug susceptible cell line (WT) are shown. A: the isotopic pair of the doubly-charged peptide GEETPVIVGSALCALEGR. B: the isotopic pair of the doubly-charged peptide YEEIDNAPEER. The insets list the relative peak area of each isotopic peak.

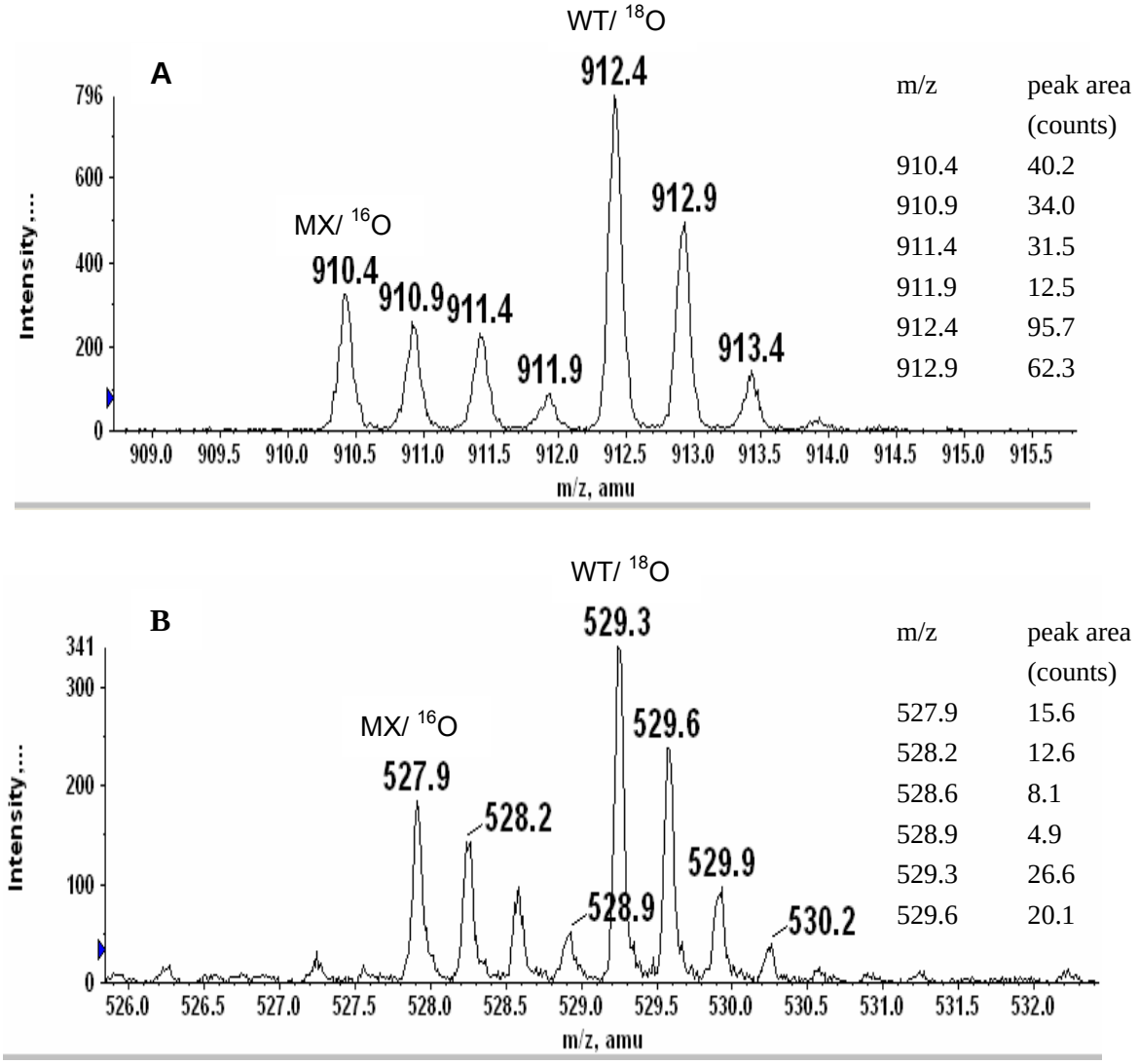
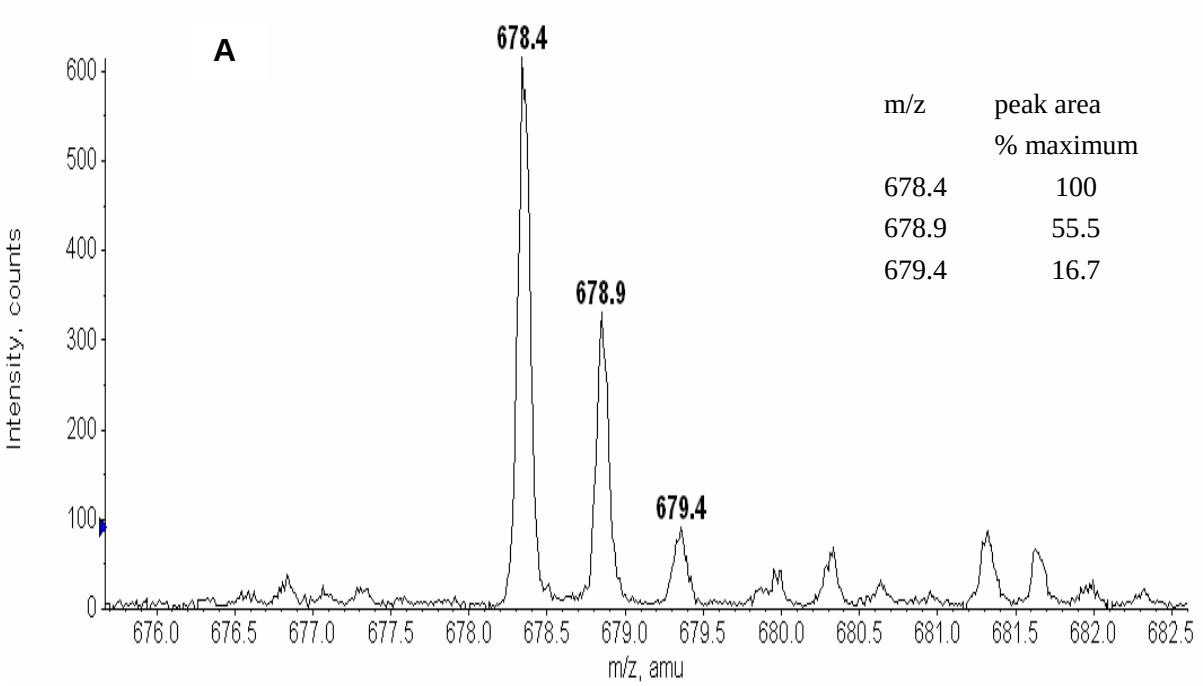
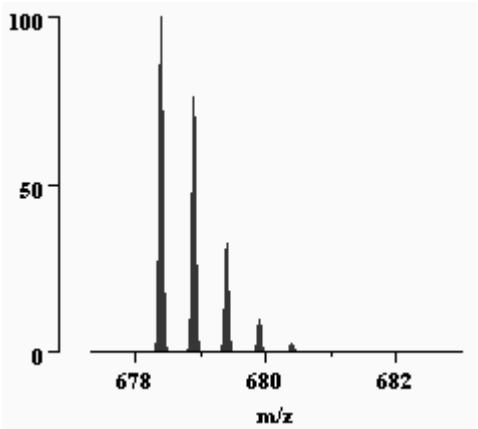


Figure 31. Partial ESI-TOF mass spectra of two peptides from dihydrolipoyl dehydrogenase protein. Isotopic pairs of the unlabeled peptide from the mitoxantrone resistant cell line (MX) and the ^{18}O labeled peptide from the drug susceptible cell line (WT) are shown. A: The isotopic pair of the doubly-charged peptide SEEQLKEEGIEYK. B: the isotopic pair of the triply-charged peptide NETLGGTCLNVGCIPSK (carbamidomethylated). The insets list the relative peak area of each isotopic peak.



m/z	peak area % maximum
678.4	100
678.9	55.5
679.4	16.7

B



m/z	peak area % maximum
678.4	100
678.9	75.8
679.4	32.4

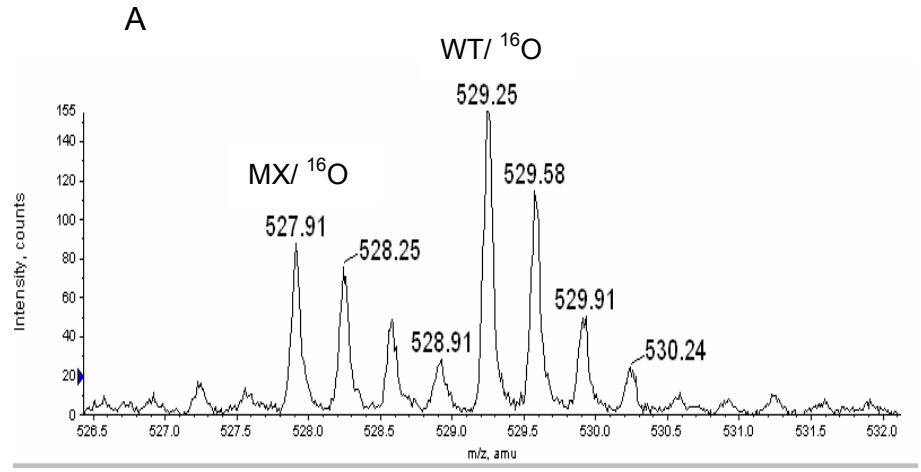
Figure 32. Experimental and theoretical isotope pattern of peptide SDLAVPSELALLK. A. Partial spectrum recorded by ESI-TOF. B. Theoretical isotope distribution of this peptide calculated and visualized by MS-Isotope program (166). The tables on the right side show the relative peak area for each isotopic peak.

Evalution of forward and reverse ^{18}O labeling

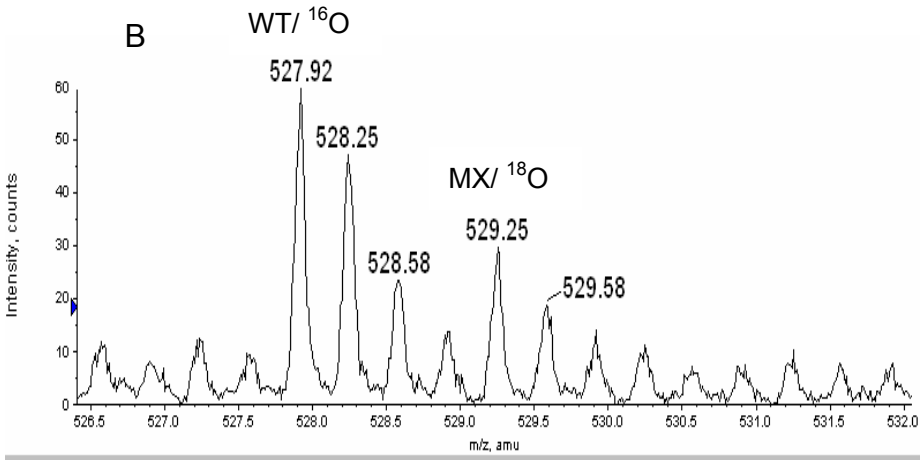
To achieve accuracy and reproducibility in protein quantitation by mass spectrometry, a reverse labeling method was proposed by Wang et al (178-179). This strategy has also been evaluated in the present project. In the labeling process shown in figure 28, referred to as forward labeling, peptides stemming from drug susceptible MCF-7 cells were labeled in $H_2{}^{18}O$. In the reverse labeling method, conversely, the peptides stemming from drug resistant cells were labeled with ^{18}O. The two inverse labeling experiments were performed in parallel. Each peptide mixture was fractionated and analyzed by the same strategy as previously described. As a result, an inverted labeling pattern will be produced between the two parallel experiments. For example, figure 33 shows MS scans from the inverse labeling experiments. The characteristic pattern of a 2 Da mass shift for doubly-charged ions can be seen from both spectra. The signal intensity ratios ($^{16}O/^{18}O$) are approximately reciprocal between two labeling strategies. It is not possible to get 100% complete labeling for every peptide in the mixture, for example, different peptide might have different ^{18}O exchange rate. Thus reverse labeling strategy is expected to reduce such ambiguity in data interpretation, especially for some short peptides, whose mass spectra do not typically provide unambiguous completion of labeling (150). In addition, reverse labeling experiment is especially useful for those situations where a protein is only detected in one sample while not in the other. Figure 34 and 35

are examples that demonstrate this benefit. Two peptides ELSEALGQIFDSQR and SDLAVPSELALLK both originate from galetin-3 binding protein precursor. In each forward labeling experiment, ^{18}O labeled peptides from drug susceptible cells were not detected in MS scans. This phenomenon was confirmed by the reverse labeling process, where those undetected peptides (labeled with ^{16}O this time) were not found either. It should be noted that the dynamic range of measuring $^{16}O/^{18}O$ in our mass spectrometer is about 10:1. A low abundant protein with the dynamic range beyond 10 will behave as an undetected protein in our study.

In addition, the double experiments can provide identification of additional peptides and thus proteins. An example is shown in figure 36 with two tandem mass spectra of a peptide from two inverse labeling experiments. This peptide was not detected in drug susceptible cells, thus there were only MS/MS spectra for peptides from drug susceptible cells (^{16}O coded in forward labeling and ^{18}O coded in reverse labeling) for both experiments. Tandem mass spectrum from forward labeling was submited to database search to identify the peptide sequence (ELSEALGQIFDSQR) and its original protein (galetin-3 binding protein). The peptide from reverse labeling was not identified due to the y-ion mass shift by 4 Da. By comparing the two tandem mass spectra, it can be confirmed that they characterized the same peptide.

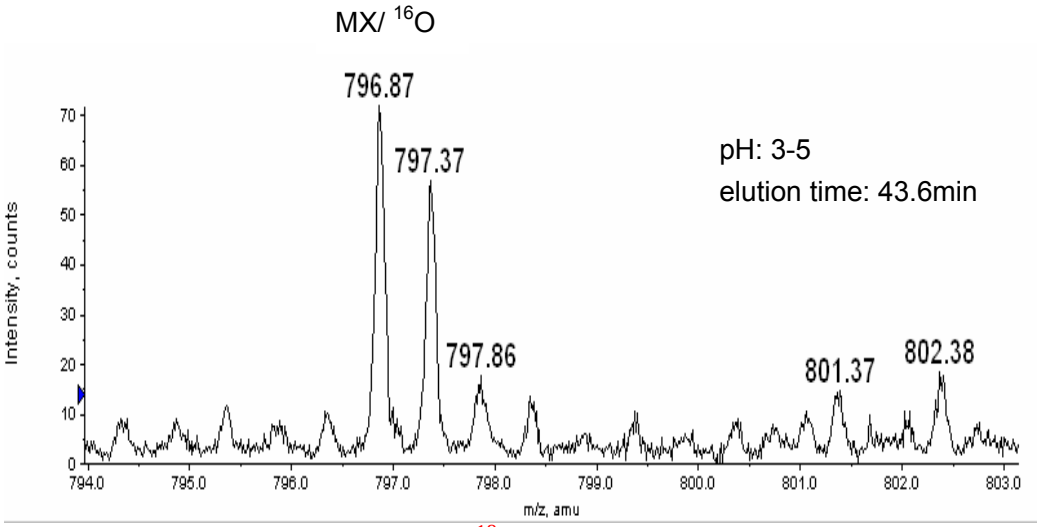


Forward ^{18}O labeled

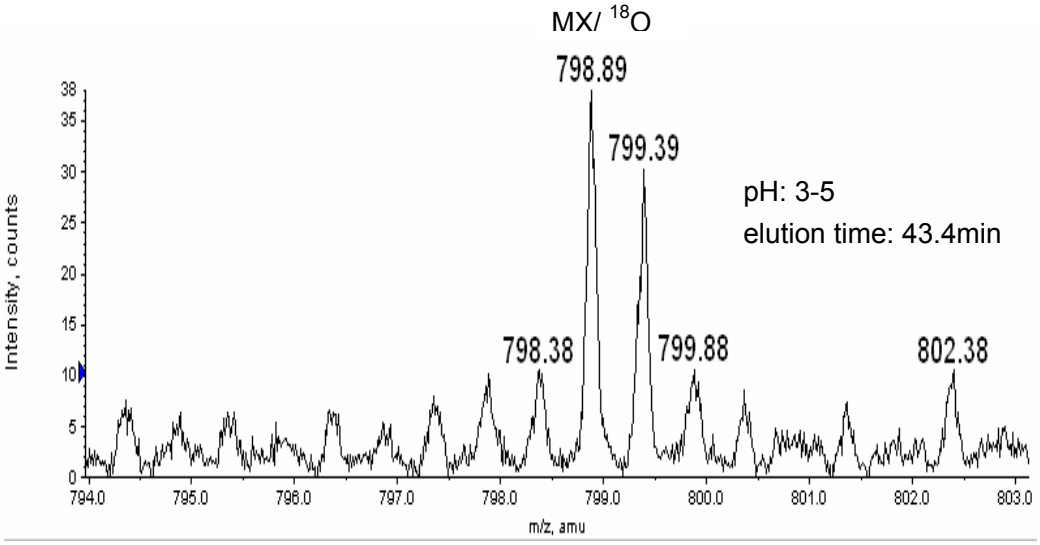


Reverse ^{18}O labeled

Figure 33. Partial MS spectra of the same peptide pair from two inverse labeling experiments. This doubly-charged peptide was identified as SEEQLKEEGIEYK, originating from mitochondrial protein: dihydrolipoyl dehydrogenase. A: In the forward labeling procedure, peptides stemming from drug susceptible cells (WT) were labeled with ^{18}O. B: In the reverse labeling method, peptides from drug resistant cells (MX) were labeled with ^{18}O.

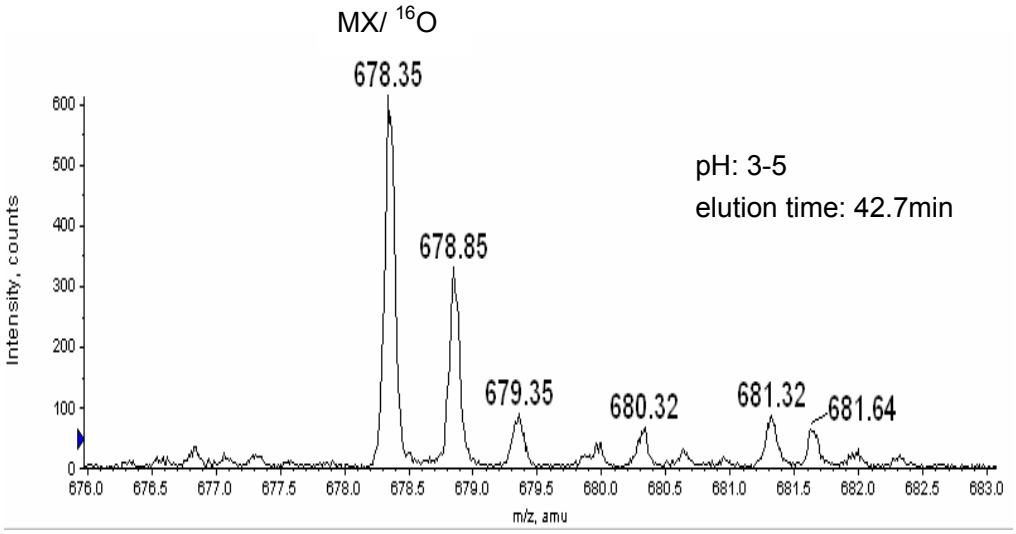


Forward ^{18}O labeled

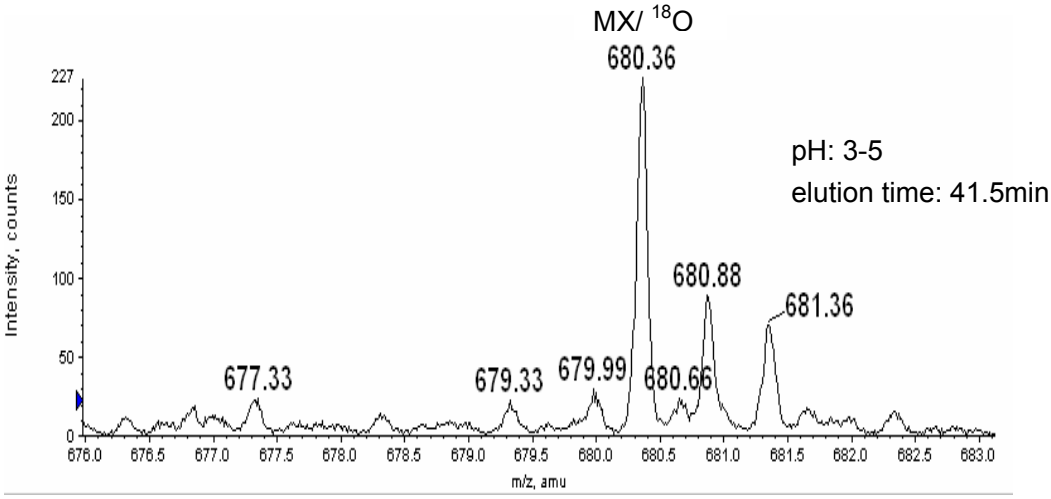


Reverse ^{18}O labeled

Figure 34. Partial MS spectra of the same peptide pair from two inverse labeling experiments. The peptide was identified as ELSEALGQIFDSQR, originating from mitochondrial protein: galectin-3 binding protein. This peptide is not detected in WT cells in either labeling experiments. The insets show the fraction chamber where the peptide presented in the sIEF separation and its elution time from LC-MS.

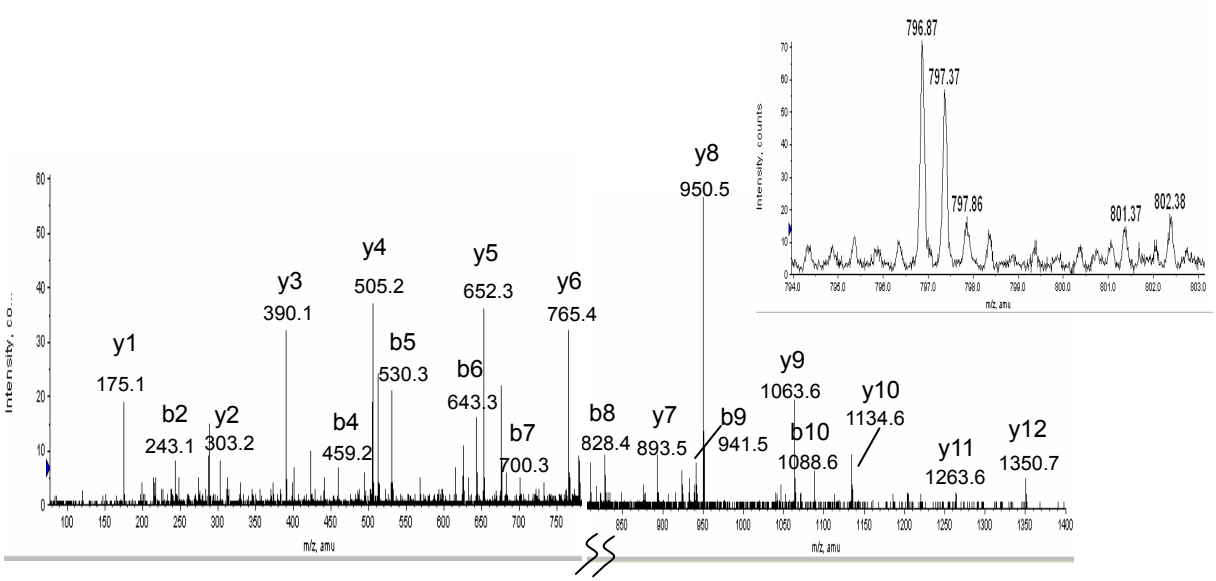


Forward ^{18}O labeled

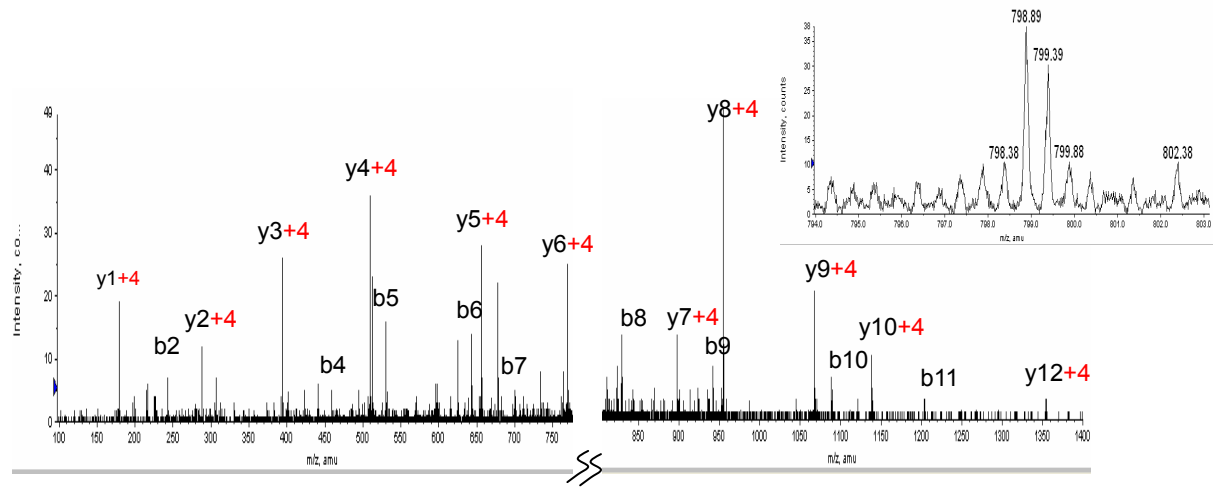


Reverse ^{18}O labeled

Figure 35. Partial MS spectra of the same peptide pair from two inverse labeling experiments. The peptide was identified as SDLAVPSELALLK, originating from mitochondrial protein: galectin-3 binding protein. This peptide is not detected in WT cells in either labeling experiments. The insets show the fraction chamber where the peptide presented in the sIEF separation and its elution time from LC-MS.



Forward ^{18}O labeled



Reverse ^{18}O labeled

Figure 36. Tandem mass spectra of precursor peptide ELSEALGQIFDSQR from two inverse labeling experiments. This peptide was not detected in drug susceptible cells. Thus MS/MS spectra are obtained only for peptides from drug susceptible cells (^{16}O coded in forward labeling and ^{18}O coded in reverse labeling). The two insets show the MS scans of the precursor peptide.

Mitochondrial protein abundance profile in the MCF-7 cell line resistant to mitoxantrone

Forward and reverse ^{18}O labeling experiments integrated with solution isoelectric focusing and LC-MS were used to compare the protein abundance profiles between the drug-susceptible MCF-7 cell line and the cell line selected for resistance to mitoxantrone. For each harvest (including three-replicate LC-MS runs after solution IEF), ratios from several peptides and different replicates from the same proteins of the origin were averaged. Three ratios for the same protein from three harvests were averaged again to obtain the final ratio and standard deviation for each protein. Table 4 lists the proteins that show alterations in their abundance profiles between these two cell lines. It can be seen that for some proteins, there are differences in ratios between forward labeling and reverse labeling experiments. This observation may be partially due to the incomplete labeling for some specific peptides which may need extended time for ^{18}O labeling. That is also an important benefit from performing the reverse labeling experiments, which can reduce such ambiguity from incomplete labeling in data interpretation.

It is obvious that efficient software should be developed for the task of data analysis and interpretation in both the identification and quantitation processes. For example, the same peptide from different experiments could be matched according to its physical properties (e.g. elution time, pI) or its MS and MS/MS spectra. Those peptides can then be further quantitated based on

their isotope patterns (in stable isotope method) or absolute intensities (in absolute quantitation).

Accession number	Protein name	Abundance ratio (MX/WT) in forward labeling	Abundance ratio (MX/WT) in reverse labeling
Q9HCCO	Methylcrotonoyl-CoA carboxylase beta chain	2.0±0.6	1.9±0.1
P49411	Elongation factor Tu	2.4±0.4	2.4±0.5
P54886	Delta 1-pyrroline-5-carboxylate synthetase	2.5±0.6	2.6±0.5
P35232	Prohibitin	2.9±0.3	2.4±0.1
P82650	Mitochondrial 28S ribosomal protein S22	2.9±0.2	3.2±1.2
Q92665	Mitochondrial 28S ribosomal protein S31	3.3±0.5	2.5±0.5
Q9BZZ5	Apoptosis inhibitor 5	3.3±0.6	2.4±0.2
P20674	Cytochrome C oxidase polypeptide Va	3.5±0.6	2.0±0.4
Q16698	2, 4-dienoyl-CoA reductase	3.6±0.3	2.1±0.1
Q08380	Galectin-3 binding protein precursor	Only present in MX	Only present in MX
P09622	Dihydrolipoyl dehydrogenase	0.5±0.1	0.4±0.1
Q9HAV7	GrpE protein homolog 1	0.5±0.1	0.4±0.1

Table 4. Proteins with altered abundances between the mitoxantrone resitant (MX) and drug susceptible (WT) MCF-7 cell lines analyzed using both forward and reversed labeling experiments. Only changes >2 are considered significant. The average and standard deviation of each ratio are calculated from three harvests of each cell line.

Biological implications of abundance changes

The proteins summarized in table 4 can be classified into seven sets according to their known or predicted functions based on published literature, as well as from the mitochondrial protein resource: the MitoProteome database (180):

1. Cell death, defense, rescue and aging (including apoptosis)

- Apoptosis inhibitor 5
- Prohibitin
- Galectin-3 binding protein precursor

2. Fatty acid metabolism

- 2, 4-dienoyl-CoA reductase

3. Oxidative phosphorylation (OXPHOS)

- Cytochrome C oxidase polypeptide Va

4. Tricarboxylic acid cycle (TCA) cycle

- Dihydrolipoyl dehydrogenase

5. Protein synthesis

- Mitochondrial 28S ribosomal protein S22
- Mitochondrial 28S ribosomal protein S31
- Elongation faction Tu

6. Amino acid metabolism

- Delta 1-pyrroline-5-carboxylate synthetase
- Methylcrotonoyl-CoA carboxylase beta chain

7. Transport:

- GrpE protein homolog 1

The goal of this study was to identify the proteins with altered abundances between the two cell lines. Here we consider the biological implications of these abundance changes, which might elucidate possible mechanisms for resistance to mitoxantrone in MCF-7 cells (MX MCF-7).

Apoptosis

As previously described, most chemotherapeutic agents kill cancer cells by inducing them to undergo apoptosis (14, 45). Mitoxantrone is also known to be a potent inducer of apoptosis in MTLn3 breast cancer cells (181). Blocking apoptosis contributes an important mechanism of drug resistance (15-17). In our studies, three proteins which can be involved in apoptosis were observed to have higher abundances in the cells resistant to mitoxantrone: prohibitin, galectin-3 binding protein and apoptosis inhibitor 5.

Galectin-3 binding protein, also known as Mac-2 binding protein, was originally described as a tumor-secreted antigen in human breast cancer cells (182) and then reported to participate in the immune defense against cancer and other pathogens (183). It binds to galectins (e.g. galectin 1 and 3), collagens and fibronectin and may relate to cell-cell and cell-extracellular adhesion (184). Galectins and their binding ligands have been implicated to play a role in cancer metastasis and their abundance may correlate with survival of cancer patients (184-186). Experimental data showed that patients

with lymphoma who have high galecin-3 binding protein levels displayed significantly lower responses to chemotherapy than did patients with low levels of this protein (187). There are few reports elucidating the mechanism and subcellular location of galectin-3 binding protein. But extensive studies have been carried out for its most important receptor, galectin-3, which is a member of the β-galactoside-binding lectin family (188). Galectin-3 was reported to regulate tumor proliferation, angiogenesis and apoptosis and has shown antiapoptotic function in certain cell types (189-190). There is a report demonstrating that galectin-3 is translocated to the perinuclear mitochondrial membranes from the cytoplasm and inhibits cytochrome c release following a variety of apoptotic stimuli (191). It has already been suggested that galectin-3 may confer chemotherapy and apoptosis resistance (192-194). In our research galectin-3 binding protein was found in the MX MCF-7 cell line but was not detected in the drug susceptible cell line. This may suggest that galectin-3 binding protein is upregulated in MX cells, contributes to apoptosis evasion and allows the cells survive. Unfortunately, galectins, including galectin-3 were not detected in our studies. Further work could target these specific proteins.

We also observed a significant increase in the abundance of apoptosis inhibitor 5 (API-5) in MX cells. API-5 was originally recognized as a putative nuclear inhibitor of apoptosis and has been little-studied so far. It was shown that both the levels of API-5 mRNA expression and its protein were higher in some cancer cells and the effect of API-5 on cultured cervical cancer cells was

associated with antiapoptotic process (195-197). More recently, API-5 was revealed to be a potent suppressor of E2F-dependent apoptosis (198). E2-promoter binding factor (E2F) family members are not only involved in cell proliferation, but also sensitize cells to apoptosis. The latter effect is associated with E2F1 protein (199). It was also found that E2F1 overexpression is related to a mitochondrial pathway to induce cell death (200). E2F-induced apoptosis causes the release of mitochondrial apoptosis-inducing factor (AIF), which can induce chromatin condensation and DNA fragmentation. As a suppressor of E2F, API-5 may also be present in mitochondria in cancer cells. Based on these conclusions and our data, in which we see an increased abundance of API-5 in MX cells, we suggest that DNA damage, caused by the inhibition of topoisomerase II by mitoxantrone drug initiates E2F-dependent apoptosis in MCF-7 cells. In the drug resistant cells higher abundance of apoptosis inhibitor 5 deregulates E2F -1 activity thus blocking E2F-dependent apoptosis pathway.

Prohibitin is the third protein relating to apoptosis whose abundance was found here to be increased in MX cells. The best-described function of prohibitin is as a chaperone protein in the mitochondria to stabilize newly synthesized subunits of mitochondrial respiration enzymes (201). Prohibitin also plays a role as a tumor suppressor and a regulator protein in apoptosis (202). It has been shown that over-expressed prohibitin in human B cells prevented apoptosis induced by the topoisomerase I inhibitor, camptothecin

(203). The same author also reported that prohibitin played a protective role in breast cancer cells treated with chemotherapy drugs and suggested that it could be used to determine the chemosensitivity of cancer cells. They later suggested that prohibitin regulated E2F function as a suppressor in the apoptotic pathway, and that it may also interact with p53, which is an important apoptotic inducer in the nucleus (204). A significant portion of prohibitin, which was co-localized with both E2F1 and p53, was found to be localized in the nucleus of MCF-7 cells and it remained in cytoplasm after apoptosis induced by camptothecin. An earlier study in our lab also showed that prohibitin was present in the nuclear fraction of MCF-7 cells (205) and was detected with an increased abundance in MX MCF-7 cells. The literature, together with our observations, indicates that prohibitin is up-regulated and acts as an anti-apoptotic factor in drug resistance.

We should also notice that since chemoresistance may be multi-factorial (7-8), the destiny of the cancer cell after chemotherapeutic treatment is a result of the overall apoptotic capacity of that cell. Several factors are expected to be involved in the pro-apoptotic or anti-apoptotic pathways in drug resistance mechanism.

Fatty acid oxidation

Fatty acid oxidation (β -oxidation) occurs in mitochondria, where fatty acid molecules are degraded through the sequential removal of two carbon units by oxidation at the β -carbon position of the fatty acyl-CoA molecules. Each

round of β-oxidation produces NADH, $FADH_2$ and acetyl-Co A molecules. Acetyl-Co A will then enter TCA cycle, where it is further oxidized to CO_2 with generation of NADH, $FADH_2$ and ATP molecules. Those NADH and $FADH_2$ molecules from both fatty acid oxidation and the TCA cycle may enter the respiratory pathway for the production of ATP. In our work with the MX MCF-7 cell line we detect an increased abundance of a key enzyme involved in fatty acid oxidation, 2, 4-dienoyl –CoA reductase.

The degradation of unsaturated fatty acids, which generally contain *cis* double bonds, requires auxiliary enzymes in addition to the enzymes necessary for oxidation of saturated fatty acids. 2, 4-Dienoyl –CoA reductase is a key mitochondrial enzyme in the metabolism of unsaturated fatty acids. It catalyzes 2,4-dienoyl-CoA to yield trans-3-enoyl CoA, which is then converted into trans-2-enoyl CoA by 3,2-trans-enoyl CoA isomerase, a common intermediate in the β-oxidation of saturated fatty acids. An increase in 2, 4-dienoyl –CoA reductase has not been previously associated with drug resistance. It should be noted that some other enzymes involved in fatty acid oxidation were identified in our studies, but were not detected to be changed. Warburg proposed in 1930 that cancer cells rely on the glycolytic pathway to convert glucose to ATP (206). It was first demonstrated by proteomics that elevated glycolysis occurred in renal cancer tissue (207). Furthermore, fatty acid oxidation has been shown to be reduced in cancer cells (208). The molecular mechanism of this less efficient metabolic phenotype in cancer cells

is still under debate. Some authors support the view that the phenotype shift to glycolysis results from a damaged mitochondrial function which includes energy transduction (209-211). This damage may be due to frequent mitochondrial DNA mutation in cancer cells. In addition, since the mitochondrial respiration chain is the major endogenous source of reactive oxygen species (212), which may damage DNA and promote apoptosis, alteration to high-rate but inefficient glycolysis to produce ATP has been proposed to protect newly synthesized and exposed DNA in cancer cells.On the other hand, a report has demonstrated that there is higher use of fatty acids for fuel in mitochondria, accompanied by a higher rate of cytosolic glycolysis in various drug-resistant cancer cells than in drug-susceptible cancer cells (213). These authors suggested that drug-resistant cells may be better at repairing damage induced by ROS. Taking these observations and our data into consideration, we suggest that the increase in 2, 4-dienoyl –CoA reductase supports the higher use of fatty acid by drug resistant MCF-7 cells.

Oxidative phosphorylation

As a cell's energy factory, mitochondria play a crucial role in the production of high-energy phosphate ATP. During oxidative phosphorylation, electrons are transferred from NADH (or $FADH_2$) to molecular oxygen, eventually forming H_2O. A series of protein complexes located in the inner mitochondrial membranes are involved in this process. Meanwhile protons are pumped from the mitochondrial matrix into intermembrane space as a result of

this flow of electrons. The final phase of oxidative phosphorylation is carried out by ATP synthase (complex V), synthesizing ATP from ADP, which is driven by the flow of protons back into the mitochondrial matrix.

Cytochrome C oxidase (COX) is the terminal complex of the electron-transport chain and transfers electrons from reduced cytochrome C to molecular O_2, forming water. This is the rate determining step of the electron transport chain (214). The COX complex is composed of 13 subunits, three of them (I, II, III) comprise the catalytic core of the enzyme and are all coded by mitochondrial DNA. The function of the remaining ten subunits (including Va and Vb), which are encoded by nuclear DNA, has not been well studied. More recently, it was suggested that COX might act as a molecular switch that induces apoptosis under energy stress conditions (215-216). They found that decreased expression of COX subunit I was significantly related to apoptosis resistance. Other researchers found that the concentration ratio of nuclear encoded COX subunits to mitochondrial encoded subunits was increased in prostate cancer compared to normal cells (217-218). The authors implied that this increase may lead to altered metabolism in cancer cells. But how this ratio alteration might interfere with chemotherapy has not been studied yet. Cytochrome C oxidase Va subunit was detected to increase in abundance in MXR cells in our study. This alteration of the COX subunit ratio may also indicate alterations in metabolism pathways in drug resistant cells. It should be noted that some other nuclear-encoded subunits, such as COX Vb, were

identified in our study but were not found to be changed.

Tricarboxylic acid cycle (TCA)

The central function of TCA is the oxidation of acetyl-CoA to CO_2 and H_2O. Acetyl-CoA is derived from the metabolism of fuel molecules such as fatty acid and carbohydrates. This oxidation accounts for about two thirds of the total oxygen consumption and ATP production in humans. The citric acid cycle also participates in some important synthetic reactions, such as amino acid synthesis.

As the starting point of TCA cycle, the formation of acetyl-CoA from glycolysis is less direct than from fatty acid oxidation. The pyruvate from glycolysis is transported into the mitochondria matrix, where pyruvate is oxidatively decarboxylated by the pyruvate dehydrogenase complex (PDH). This irreversible reaction is the key link between glycolysis and the citric acid cycle. PDH is a large, highly integrated complex of three kinds of enzymes: pyruvate dehydrogenase component (E1), dihydrolipoyl transacetylase (E2) and dihydrolipoyl dehydrogenase (E3). It has been demonstrated that the flavin of the dihydrolipoyl dehydrogenase (Dld) component, which is abundant in mitochondria and has a sufficient redox potential to allow for superoxide production, can generate superoxide (219-220). More recently, there is a report detecting Dld as a source of reactive oxygen species when there is limited glucose (221). They suggested Dld instead of the electron transport chain as the main ROS source when glucose is limited. Taking these

observation and previous information together with our data that E3 has a decreased abundance in MX cells, we consider that MX cells may have an impaired TCA cylcle as well as a lower rate of respiration. Together, these reduce ROS and prevent mitoxantrone-induced cytotoxicity.

Protein synthesis

Mitochondrial DNA (mtDNA) encodes 13 polypeptides, each of which is a subunit of one of four respiratory enzyme complexes localized to the inner mitochondrial membrane. They include seven subunits of respiratory enzyme complex I, one subunit of complex III, three subunits of complex IV, and two subunits of complex V. Since these proteins that are essential to oxidative phosphorylation have high hydrophobicity, nuclear synthesis and cytoplasmic transport are precluded.

Mitochondrial ribosome is the factory for synthesis of mitochondrial-encoded proteins. It is a ribonucleoprotein particle made of a small (28S) and a large subunit (39S). It was discovered that the key sites in the ribosome are composed almost entirely of RNA and contributions from the ribosomal proteins are minor. Many of the proteins having elongated structure extending into the rRNA core to stabilize its structure, and they may play a role in streamlining the process of protein synthesis (222). The amino acid sequence is known for 30 proteins in the small subunit and 48 proteins in the large subunits. We detected 13 (16%) out of those 78 proteins. Two of them have increased abundance in MX cells, mitochondrial 28S ribosomal protein

S22 and mitochondrial 28S ribosomal protein S31. It has been reported that two proteins, death-associated protein 3 and PDCD 9, known to be involved in promoting apoptosis in mammalian cells have been identified as mitochondrial ribosomal proteins (223). But their precise function is not known. The authors suggested that components of the mitochondrial proteins' biosynthetic system might play a pivotal role in apoptosis. How the ribosomal protein might function in cancer drug resistance needs further investigation.

Elongation factors are a set of proteins that facilitate the translational elongation, including the formation of all peptide bonds. Elongation factor Tu (EF-Tu) is responsible for the selection and binding of the aminoacyl-tRNA and GTP to the active site of the ribosome. After binding, GTP is hydrolyzed to GDP and released from the ribosome with EF-Tu. Elongation factor Ts (EF-Ts), a second elongation factor, joins the EF-Tu complex and induces the dissociation of GDP. Once EF-Tu binds to another GTP, EF-Ts is concomitantly released. There are few studies associating EF-Tu with cancer. One report stated that a mitochondrial elongation factor-like protein is over-expressed in tumors (224). That protein has 50-70% homology to eukaryotic-EF-Tu. Our observation of a higher abundance of EF-Tu in MX cells might relate to the increase of the protein translation rate, which might be associated with drug resistance in an indirect way.

Amino acid metabolism

The primary uses of amino acids are as building blocks for protein and

peptide synthesis and as a source of nitrogen for the synthesis of other amino acids. Amino acids in excess of those needed for biosynthesis are used as metabolic fuel through catabolism. There are two proteins, which are detected to have increased abundance in MX cells here that are enzymes in proline biosynthesis and leucine catabolism, respectively.

Delta1-pyrroline-5-carboxylate synthetase is located in the mitochondrial inner membrane, where it catalyzes the ATP and NAD(P)H–dependent conversion of L-glutamate to pyrroline-5-carboxylate. The latter is converted into L-proline by pyrroline-5-carboxylase reductase in the cytosol. Proline can then enter mitochondria where it is oxidized back to pyrroline-5-carboxylate catalyzed by proline oxidase (POX). Reports have shown that catalytic cycling between pyrroline-5-carboxylate and proline can mediate redox transfers between mitochondria and cytosol and can regulate apoptosis in cells (225-226). The expression of POX has been found to be up-regulated in a human colon cancer cell line using adriamycin to initiate p53-dependent apoptosis. Increased reactive species were found by addition of proline in cells expressing POX, which induced apoptosis through the mitochondrial pathway. More recently, they found evidence suggesting that POX may also induce extrinsic apoptotic pathways (226). We saw an increase of delta 1-pyrroline-5-carboxylate synthetase in MX cells, which may indicate that production of pyrroline-5-carboxylate is increased. As a hypothesis, the equilibrium reaction from proline to pyrroline-5-carboxylate may be reduced,

which will also repress POX expression as a feedback and thus reduce ROS-induced apoptosis in drug resistant cancer cells. But meanwhile, that means the increased pyrroline-5-carboxylate has to go through some other pathway to be used other than producing “toxic” proline. One possibility is that it is converted into L-ornithine and then arginine.

Methylcrotonoyl-CoA carboxylase beta chain is an enzyme involved in leucine catabolism. Unfortunately, no literature could be found on the association between leucine catabolism and cancer or drug resistance.

Transport

The majority of mitochondrial proteins are encoded by nuclear genes and then targeted to the mitochondria by specific transport systems (227). GrpE protein homolog 1 cooperates with mitochondrial heat shock protein 70 (Hsp70) in the import of proteins from the cytoplasm. Hsp70 has been identified as a potent anti-apoptotic factor (228-230), which was suggested to prevent recruitment of procaspase-9 to the apoptosome complex or to inhibit caspase-3 activation, therefore hindering initiation of caspase cascade. A decrease abundance of GrpE protein homolog 1 was detected in MX cells from this research. How GrpE protein homolog 1 interacts with the cell death pathway to regulate response to chemotherapy needs further investigation.

Of the twelve proteins exhibiting significant abundance changes, most of them are increased in the mitoxantrone-resistant MCF-7 cells compared to mitoxantrone-susceptible MCF-7 cells. As many of these proteins may block

one of the pathways leading to apoptosis induced by drugs as mentioned above, it is suggested here that anti-apoptotic factors play more important roles in drug resistance in MCF-7 cells rather than pro-apoptotic factors do.

Chapter 4: Conclusions

Proteomics techniques enable comprehensive studies of proteins from complex biological systems. There has been an increased interest in mitochondrial proteomics over the last decade. The objectives of the present study were to evaluate a novel integrated proteomic strategy applicable to mitochondrial proteomics, as to identify mitochondrial proteins whose abundances are altered between drug-susceptible MCF-7 breast cancer cells and MCF-7 cells selected for resistance to mitoxantrone, in order to consider how these proteins contribute to drug resistance.

Efficient separation techniques are required in mass spectrometry to provide global proteomic analysis. In this study, solution isoelectric focusing separation was applied to fractionate soluble mitochondrial proteins and thus reduce sample complexity. Each fraction was then further fractionated with reversed-phase capillary HPLC. This orthogonal two-dimensional separation technique is demonstrated here to have high reproducibility, resolution and flexible sample capacity. A total of 637 peptides corresponding to 278 proteins were identified. About 100 more proteins were detected by this method compared to the result with traditional 2D-gel performed also on the soluble mitochondrial protein sample in our lab (174). Among those, are many proteins with high pI values or hydrophobic character, which are usually excluded in 2D gel methods. In addition, it is estimated that half of the time was required for our shotgun method compared to the 2D-gel method, when we used the same

amount of sample.

This shotgun method can be used in combination with stable isotope labeling techniques to determine changes in protein abundances between clinical or other limited samples. In the present study, we integrated enzyme-catalyzed ^{18}O labeling with the shotgun method to study changes in mitochondrial proteins associated with drug resistance in human cancer cells. This combination appeared well-suited for sensitive quantitative comparison. Furthermore, reversed ^{18}O labeling experiments were carried out to provide complementary pairs of isotopically labeled peptides. These were used to estimate the precision and dynamic range of the method, and to confirm the behavior of those proteins which were only detected in one cell line. The overall strategy identified twelve mitochondrial proteins with abundances altered significantly in drug resistant cells. These proteins are actively involved in apoptosis, oxidative phosphorylation, fatty acid metabolism and amino acid metabolism. None of them has been previously confirmed to contribute to drug resistance. Because we know that drug resistance in cancer is a multi-factorial process, we suggest that these proteins might play different but related roles. This study also provides a list of target proteins for further investigation, including molecular biology to confirm the proteins' functions in drug resistance.

References

1. Liem, A. A.; Chamberlain, M. P.; Wolf, C. R.; Thompson, A. M. *Eur. J. Surg. Oncol.* **2002**, *28*, 679-684.
2. Simon, S. M.; Schindler, M. *Proc. Natl. Acad. Sci. U.S.A*. **1994**, *91*, 3497-3504.
3. Longley, D.; Johnston, P. *J. Pathol.* **2005**, *205*, 275-292.
4. Gutierrez, P. L.; Desai, T. T. *BioMedicina* **1999**, *2*, 235-239.
5. Bates, S. E. *Clin. Cancer Res*.**1999**, *5*, 3346-48.
6. Lehnert, M. *Eur. J. Cancer* **1996**, *32A*, 912-20
7. Gottesman, M. M. Annu. *Rev. Med.* **2002**, *53*, 615-27
8. Gottesman, M. M.; Fojo, T.; Bates, S. E. *Nat. Rev. Cancer* **2002**, *2*, 48-58.
9. Fojo, T. *J. Natl. cancer Inst.* **2001**, *93*, 1434-1436.
10. Litman, T.; Druley, T. E.; Stein, W. D.; Bates, S. E. *Cell Mol. Life Sci.* **2001**, *58*, 931-59.
11. Endicott, J. A.; Ling, V. *Annu. Rev. Biochem.* **1989**, *58*, 137-171
12. Thomas, H.; Coley, H. M. *Cancer Control* **2003**, *10*, 159-165.
13. Doyle, L. A.; Yang, W.; Abruzzo, L. V.; Krogmann, T.; Gao, Y.; Rishi, A. K.; Ross, D. D. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 15665-70.
14. Lowe, S. W.; Lin, A. W. *Carcinogenesis* **2000**, *21*, 485-495.
15. Lasorella, A.; Lavarone, A.; Israel, M. A. *Cancer Res.* **1995**, 55, 4711-16.
16. Lowe, S. W.; Bodis, S.; McClatchey, A.; Remington, L; Reley, H. E.; Fisher, D. E.; Housman, D. E.; Jacks, T. *Science* **1994**, *266*, 807-810
17. Johnstone, R. W.; Ruefli, A. A.; Lowe, S. W. *Cell* **2002**, *108*, 153-164.
18. Wallace-Brodeur, R. R.; Lowe, S. W. *Cell. Mol. Life Sci.* **1999**, *55*, 64-75.

19. Soengas, M. S.; Lowe, S. W.; *Oncogene* **2003**, *19*, 3138-51.

20. Cheng, Q.; Lee, H. H.; Li, Y.; Parks, T. P.; Cheng, G. *Oncogene* **2000**, *5*, 4936-40.

21. Taylor, S. W.; Fahy, E.; Ghosh, S. S. *Trends in Biotech.* **2003**, *21*, 82-88.

22. Jung, E.; Heller, M.; Sanchez, J.; Hochstrasser, D. F. *Electrophoresis* **2000**, *21*, 3369-77.

23. Schatz, G. *J. Biol. Chem.* **1996**, *271*, 31763-66.

24. Westermann, B.; Neupert, W. *Nat. Biotech.* **2003**, *21*, 239-40.

25. Warnock, D. E.; Fahy, E.; Taylor, S. W. *Mass Spectrom. Rev.* **2004**, *23*, 259-280.

26. Yaffe, M. P. *Science* **1999**, *283*, 1493-96.

27. Mootha, V. K.; Bundenborg, J.; Lisen, J. V.; Hjerrild, M.; Wisniewski, J. R.; Stahl, E.; Bolouri, M. S.; Ray, H. N.; Sihag, S.; Kamal, M.; Patterson, N.; Lander, E. S.; Mann, M. *Cell* **2003**, *115*, 625-40.

28. Scheffler, I. E. *Mitochondrion 1* **2001**, 3-31.

29. Mak et al. Arthritis Res. **2002**, *4* (Suppl 3), S243

30. Hengartner, M. O. *Nature* **2000**, *407*, 770-76.

31. Bernardi, P.; Petronilli, V.; DiLisa, F.; Forte, M. *Trends in Biochem. Sciences* **2001**, *25*, 112-17.

32. Danial, N. N.; Korsmeyer, S. J. *Cell* **2004**, *116*, 205-219.

33. MacFarlane, M.; Williams, A. C. *EMBO reports* **2004**, *5*, 674-78.

34. Bouchier-Hayes, L.; Lartigue, C.; Newmeyer, D. *J. Chin. Rev.* **2005**, *115*, 2640-47.

35. Ravagnan, L.; Marzo, I.; Costantini, P.; Susin, S.A.; Zamzami, N.; Petit, P. X.; Hirsch, F.; Goulbern, M.; Poupon, M. F.; Miccoli, L.; Xie, Z.; Reed, J. C.; Kroemer, G.

Oncogene **1999**, *18*, 2537-46.

36. Armstrong, J. S. *British J. Pharm.* **2006**, *147*, 239-48.
37. Soule, H. D.; Vazquez, J.; Long, A.; Albert, S.; Brennan, M. *J. Natl. Cancer Inst.* **1973**, *51*, 1409-16.
38. Hathout, Y.; Gehrmann, M. L.; Chertov, A.; Fenselau, C. *Cancer Letters* **2004**, *210*, 245-53.
39. Nakagawa, E.; Schneider, E.; Dixon, K. H.; Horton, K.; Kelley, C.; Morrow, C.; Cowan, K. H. *Cancer Res.* **1992**, *52*, 6175-81.
40. Parker, B. S.; Cullinane, C.; Phillips, D. R. *Nucleic. Acids Res.* **1999**, *27*, 2918-23.
41. Solary, E.; Bettaieb, A.; Dubrez-Daloz, L.; Corcos, L. *Leuk. Lymphoma* **2003**, *44*, 563-74.
42. Szewloczyk, A.; Wojtczak, L. *Phamacol. Rev.* **2002**, *54*, 101-27.
43. Lenk, H.; Muller, U.; Tanneberger, S. *Anticancer Res.* **1987**, *7*, 1257-64.
44. Kluza, J.; Marchetti, P.; Gallego, M. A.; Lancel, S.; Fournier, C.; Loyens, A.; Beuvillain, J. C.; Baily, C. *Oncogen* **2004**, *23*, 7018-30.
45. Ferrer, A.; Marce, S.; Bellosillo, B.; Villamor, N.; Bosch, F.; Lopez-Guillermo, A.; Espinet, B.; Sole, F.; Montserrat, E.; Campo, E.; Colomer, D. *Oncogene* **2004**, *23*, 8941-9.
46. Wu, J. *Anticancer Res* **1996**, *16*, 2233-40.
47. Simstein, R.; Burow, M.; Parker, A.; Weldon, C.; Beckman, B. *Exp. Bio. Med.* **2003**, *228*, 995-1003.
48. Li, S. J.; Rodgers, E. H.; Grant, M. H. *Chemico-Biological Interactions* **1995**, *97*,

101-18.

49. Koutinos, G.; Stathopoulos, G.; Dontas, I.; Perrea-Kotsarelis, D.; Couris, E.; Karayannacos, P.; Deliconstantinous, G. *Anticancer Res.* **2002**, *22*, 815-20.
50. Diah. S. K.; Smitherman, P. K.; Aldridge, J.; Volk, E. L.; Schneider, E.; Townsend, A. J.; Morrow, C. S. *Cancer Res.* **2001**, *61*, 5461-67.
51. Bachmann, E.; Weber, E.; Zbinden, G. *Cancer Treat. Rep.* **1987**, *71*, 361-66.
52. Kule, C.; Ondrejickova, O.; Verner, K. *Mol. Pharmacol.* **1994**, *46*, 1234-40.
53. Kim, G.; Sikder, H.; Singh, K. K. *Mutagenesis* **2002**, *17*, 375-81.
54. International Human Genome Sequencing Consortium *Nature* **2004**, *431*, 931-45.
55. Rappsilber, J.; Mann, M. *Trends in Biochem. Sci.* **2002**, *27*, 74-78.
56. Mann, Matthias *Nature Biotechnology* **1999**, *17*, 954-955.
57. Anderson, L.; Seilhamer, J. *Electrophoresis* **1997**, *18*, 533-537.
58. Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. *Mol. Cell Biol.* **1999**, *19*, 1720-1730.
59. Peng, Junmin; Gygi, S. P. *J. Mass Spectom.* **2001**, *36*, 1083-1091.
60. Hanash, S. *Nature* **2003**, *422*, 226-32.
61. Figeys, D. *Anal. Chem.* **2002**, 413A.
62. Petricoin, E.; Wulfkuhle, J.; Espina, V.; Liotta, L. A.; *J. Proteome Res.* **2004**, *3*, 209-17.
63. Stults, J. T.; Arnott, D. *Methods in Enzymology* **2005**, *402*, 253.
64. Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837-46.
65. Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-95.
66. Fenn, B.; Mann, M.; Meng, C. K.; Wong, S. F. Whitehouse, C. M. *Science* **1989**, *246*,

64-71.

67. Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.

68. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

69. Takana, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-53.

70. Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198-207.

71. Standing, K. G. *Curr. Opion Struc. Biol.* **2003**, *13*, 595-601.

72. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrrell, J. S. *Electrophoresis* **1999**, *20*, 3551-67.

73. Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom* **1994**, *5*, 976-89.

74. Gevaet, K.; Vandekerckhove, J. *Electrophoresis* **2000**, *21*, 1145-54.

75. Ferguson, P. L.; Smith, R. D. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 399-424.

76. Yates, J. R. III; Eng, J. K.; McCormack, A. L.; Shcieltz, D. *Anal. Chem.* **1995**, *67*, 1426-36.

77. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaerud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 806-12.

78. Zhai, H.; Han, X.; Breuker, K.; McLafferty, F. W. *Anal. Chem.* **2005**, *77*, 5777-84.

79. Demirev, P. A.; Feldman, A. B.; Kowalski, P.; Lin, J. S. *Anal. Chem.* **2005**, *77*, 7455-61.

80. Wolters, D. A.; Washburn, M. P.; Yates, J. R. *Anal. Chem.* **2001**, *73*, 5683-90.

81. Issaq, H. J. *Electrophoresis* **2001**, *22*, 3629-38.

82. Shi, Y.; Xiang, R.; Horvath, C.; Wilkins, J. A. *J. Chromatogr A* **2004**, *1053*, 27-36.

83. Posewits, M. C.; Cooper, J. *Anal. Chem.* **1999**, *71*, 2883-92.

84. Chaga, G. S.; *J. Biochem. Biophys. Methods* **2001**, *49*, 313-34.

85. Burgess, R. R.; Thompson, N. E. *Curr. Opin. Biotechnol.* **2002**, *13*, 304-08.

86. Xiong, L.; Andrews, Regnier, F. E. *J. Proteome Res.* **2003**, *2*, 618-25.

87. Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Boucherie, H.; Mann, M. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 14440-45.

88. Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Barvik, B. M.; Yates, J. R. *Nature Biotech.* **1999**, *17*, 676-82.

89. Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242-47.

90. Apffel, A.; Chakel, J.; Hancock, W.; Souders, C.; Timkulu, T.; Pungor, E. *J. Chromatogr. A.* **1996**, *732*, 27-42.

91. Geng, M.; Ji, J.; Regnier, F.; *J. Chromatogr. A* **2000**, *870*, 295-313.

92. Gao, J.; Opiteck, G. J.; Friedrichs, M. S.; Dongre, A. R.; Hefta, S. R. *J proteome Res.* **2003**, *2*, 643-49.

93. Opiteck, G. J.; Ramirez, S.M.; Jorgenson, J. W.; Moseley, M. A. *Anal. Biochem.* **1998**, *258*, 349-61.

94. Blom, K. F.; Larsen, B. S.; McEwen, C. N. *J. Comb. Chem.* **1999**, *1*, 82-90.

95. Simpson, D. C.; Smith, R. D. *Eletrophoresis* **2005**, *26*, 1291-1305.

96. Servais, A.; Crommen, J.; Fillet, M. *Electrophoresis* **2006**, *27*, 2616-29.

97. Cooper, J. W,; Wang, Y.; Lee, C. S. *Electrophoresis* **2004**, *25*, 3913-26.

98. Kaiser, T.; Hermann, A.; Kielstein, J. T.; Wittke, S.; Bartel, S.; Krebs, R.; Hausadel, F.;

Hillmann, M.; Golovko, I.; Koester, P.; Haller, H.; Weissinger, E. M.; Fliser, D.; Mischak, H.; *J. Chromatogr. A* **2003**, *1013*, 173-81.

99. Von Neuhoff, N.; Kaiser, T.; Wittke, S.; Krebs, E.; Pitt, A.; Burchard, H. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 149-56.

100. Moni, M.; Demars, S. M.; Huang, H. *Anal. Chem.* **2002**, *74*, 3772-76.

101. Shen, Y.; Smith, R. D. *Electrophoresis* **2002**, *23*, 3106-24.

102. Chen, J.; Gao, J.; Lee, C. S. *J. Proteome Res.* **2003**, *2*, 249-54.

103. Jensen, P. K.; Pasa-Tolic, L.; Anderson, G. A.; Horner, J. A.; Lipton, M. S.; Bruce, J. E.; Smith, R. D. *Anal. Chem.* **1999**, *71*, 2076-84.

104. Storms, H. F.; van der Heijden, R.; Tjaden, U. R.; van der Greef, J. *Electrophoresis* **2004**, *25*, 3461-67.

105. Zhang, C. X.; Xiang, F.; Pasa-Tolic, L.; Anderson, G. A.; Veenstra, T. D.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 1462-68.

106. Righetti, P. G.; Castagna, A.; Herbert, B.; Reymond, F.; Rossier, J. S. *Proteomics* **2003**, *3*, 1397-1407.

107. Bier, M. *Electrophoresis* **1998**, *19*, 1057-63.

108. Wall, D. B.; Kachman, M. T.; Gong, S.; Hinderer, R. *Anal. Chem.* **2000**, *72*, 1099-1111.

109. Davidsson, P.; Paulson, L.; Hesse, C.; Blennow, W.; Nilsson, C. L. *Proteomics* **2001**, *1*, 444-52.

110. Wang, H.; Kachman, M. T. Schwartz, D. R.; Cho, K. R.; Lubman, D. M. *Electrophoresis* **2002**, *23*, 3168-81.

111. Righetti, P. G.; Wenisch, E.; Faupel, M. *J. Chromatogr.* **1989**, *475*, 293-309.

112. Herbert, B.; Righetti, P. G. *Electrophoresis* **2000**, *21*, 3639-48.

113. Zhu, Y.; Lubman, D. M. *Electrophoresis* **2004**, *25*, 949-58.

114. Zuo, X.; Speicher, D. W. *Anal. Biochem.* **2000**, *284*, 266-78.

115. An, Y.; Fu, Z.; Gutierrez, P. ; Fenselau, C. *J. Proteome Res.* **2005**, *4*, 216-32.

116. Anderson, L.; Seilhamer, J. *Electrophoresis* **1997**, *18*, 533-537.

117. Gygi, S. P.; Rochon, Y.; Franza, B. R.; Aebersold, R. *Mol. Cell Biol.* **1999**, *19*, 1720-1730.

118. Griffin, T. J.; Gygi, S. P.; Ideker, T.; Rist, B.; Eng, J.; Hood, L.; Aebersold, R. *Mol. Cell. Proteomics* **2002**, *1*, 323-33.

119. Ong, S.; Foster, L. J.; Mann, M. *Methods* **2003**, *29*, 124-30.

120. Regnier, F. E.; Riggs, L.; Zhang, R.; Xiong, L.; Liu, P.; Chakraborty, A.; Seeley, E.; Sioma, C.; Thompson, R. A. *J. Mass Spec.* **2002**, *37*, 133-45.

121. Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevingsky, J. R.; Resing, K. A.; Ahn, N. G. *Mol. Cell. Proteomics* **2005**, *4*, 1487-1502.

122. Herbert, B. R.; Harry, J. L.; Packer, N. H.; Gooley, A. A.; Pedersen, S. K.; William, K. L. *Trends Biotechnol.* **2001**, *19*, Suppl. 10, 3-9.

123. Hathout, Y.; Riordan, K.; Gehrmann, M.; Fenselau, C. *J. Proteome Res.* **2002**, *1*, 435-42.

124. Sechi, S.; Oda, Y. *Curr. Opion. Chem. Biol.* **2003**, *7*, 70-77.

125. Bondarenko, P. V.; Chelius, D.; Shaler, T. A. *Anal. Chem.* **2002**, *74*, 4741-49.

126. Strong, R.; Nakanishi, T.; Ross, D; Fenselau, C. *J. Proteome Res.* **2006**, *5*, 2389-95.

127. Bantscheff, M.; Dumpelfeld, B.; Kuster, B. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 869-76.

128. Ong, S.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376-86.

129. Gehrmann, M. L.; Hathout, Y.; Fenselau, C. *J. Proteome Res.* **2004**, *3*, 1063-68.

130. Ji, J.; Charkraborty, A.; Geng, M.; Zhang, X.; Amini, A.; Bina, M.; Regnier, F. E.; *J. Chromatogr. B. Biomed. Sci. Appl.* **2000**, *745*, 197-210.

131. Julka, S.; Regnier, F. *J. proteome Res.* **2004**, *3*, 350-63.

132. Gygi, S. P.; Rist, B.; Berber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999**, *17*, 994-99.

133. Gygi, S. P.; Rist, B.; Griffin, T. J. Eng, J.; Aebersold, R. *J. Proteome Res.* **2002**, *1*, 47-54.

134. Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotechnol.* **2001**, *19*, 946-51.

135. Gulna, T.; Purvine, S. O.; Yi, E. C.; Eng, J.; Goodlett, D. R.; Aebersold, R.; Miller, S. I. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2771-76.

136. Oda, Y.; Owa, T.; Sato, T.; Boucher, B.; Daniels, S.; Yamanaka, H.; Shinohara, Y.; Yokoi, A.; Kuromitsu, J.; Nagasu, T. *Anal. Chem.* **2003**, *75*, 2159-2165.

137. Hansen, K. C.; Schmitt-Ulms, G.; Chalkley, R.; Hirsch, J.; Baldwin, M. A.; Burlingame, A. L. *Mol. Cell Proteomics* **2003**, *2*, 299-314.

138. Bender, M.; Kemp, K. *J. Am. Chem. Soc.* **1957**, *79*, 111-16.

139. Sharon, N.; Grisaro, V.; Neumann, H. *Arch. Biochem. Biophys.* **1962**, *97*, 219-21.

140. Schnolzer, M.; Jedrzejewski, P.; Lehmann, W. D. *Electrophoresis* **1996**, *17*, 945-53.

141. Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 2836-42.

142. Murphy, R. C.; Clay, K. L. *Biomed. Mass Spectrom.* **1979**, *6*, 309-14.

143. Rose, K.; Simona, M. G.; Offord, R. E.; Pior, C. P.; Otto, B.; Thatcher, D. R. *Biochem. J.* **1983**, *215*, 273-77.

144. Shevchenko, A.; Chernushevich, I.; Ens, W.; Standing, K. G.; Thomson, B.; Wilm, M.; Mann, M. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1015-24.

145. Stewart, I. I.; Thomson, R.; Figey, D. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 2456-65.

146. Reynold, K. J.; Yao, X.; Fenselau, C. *J. Proteome Res.* **2002**, *1*, 27-33.

147. Zang, L.; Palmer, D.; Hancock, W. S.; Sgroi, D. C.; Karger, B. L. *J. Proteome Res.* **2004**, *3*, 604-12.

148. Mirgorodskaya, E.; Wanker, E.; Otto, A.; Lehrach, H.; Gobom, J. *J. proteome Res.* **2005**, *4*, 2109-16.

149. Wang, Y. K.; Ma, Z.; Quinn, D. F.; Fu, E. W. *Anal. Chem.* **2001**, *73*, 3742-50.

150. Yao, X.; Afonso, C.; Fenselau, C. *J. Proteome Res.* **2003**, *2*, 147-52.

151. Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Prukayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154-69.

152. Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resting, K. A.; Ahn, N. G. *Mol. Cell. Proteomics* **2005**, *4*, 1487-1502.

153. Bondarenko, P. V.; Chelius, D.; Shaler, T. A. *Anal. Chem.* **2002**, *74*, 4741-4749.

154. Chelius, D.; Bondarenko, P. V. *J. Proteome Res.* **2002**, *1*, 317-23.

155. Liu, H.; Sadygov, R. G.; Yates, J. R. *Ana. Chem.* **2004**, *76*, 4193-4201.

156. Wang, W.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderie, M.; Becker, C. H. *Anal. Chem.* **2003**, *75*, 4848-26.

157. Stemmann, O.; Zhou, H.; Gerber, S. A.; Gygi, S. P.; Kirschner, M. W. *Cell* **2001**, *107*, 715-26.

158. Barnidge, D. R.; Dratz, E. A.; Martin, T.; Bonilla, L. E.; Moran, L. B.; Lindall, A. *Anal. Chem*. **2003**, *75*, 445-451.

159. Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100,* 6940-6945.

160. Barnidge, D. R.; Hall, G. D.; Stocker, J. L.; Muddiman, D. C. *J. Proteome Res.* **2004**, *3*, 658-61.

161. Anderson, N. L.; Anderson, N. G.; Haines, L. R.; Hardie, D. B.; Olafson, R. W.; Pearson, T. W. *J. Proteome Res.* **2004**, *3*, 235-44.

162. Julka, S.; Regnier, F. *Briefings in functional genomics and proteomics* **2005**, *4*, 158-77.

163. Rigobello, M. P.; Donella-Deana, A.; Cesaro, L.; Bindoli, A. *Biochem. J.* **2001**, *356*, 567-70.

164. Bradford, M. *Anal. Biochem.* **1976**, *72*, 248-54.

165. Bio-Rad protein assay instruction.

166. http://prospector.ucsf.edu/prospector/4.0.7/html/msiso.htm

167. Cargile, B. J.; Sevinsky, J. R.; Essader, A. E.; Stephenson, J. L.; Bundy, J. L. *J. Biomol. Tech.* **2005**, *16*, 181-89.

168. Cargile, B. J.; Bundy, J. L.; Freeman, T. W.; Stephenson, J. L. *J. Proteome Res.* **2004**, *3*, 112-19.

169. http:// ihg.gsf.de/ihg/mitoprot.html.

170. http://psort.nibb.ac.jp/cgi-bin/runsport.pl.

171. Yu, F.; Finley, R. L.; Raz, A.; Kim, H. C. *J. Bio. Chem.* **2002**, *277*, 15819-27.

172. Yang, R.; Hsu, D. K.; Liu, F. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6737-42.

173. Jaattela, M.; Cande, C.; Kroemer, G.; *Cell Death Differ.* **2004**, *11,* 135-36.

174. Strong, R. F. Ph.D. dissertation, University of Maryland, **2005**.

175. Brown, K. J.; Fenselau, C. *J. Proteome Res.* **2004**, *3*, 455-462.

176. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. *J. Mass Spectrom.* **2001**, *36*, 849-65.

177. Taylor, S. W.; Fahy, E.; Zhang, B.; Glenn, G. M.; Warnock, D. E.; Wiley, S.; Murphy, A. N.; Gaucher, S. P.; Capaldi, R. A.; Gibson, B. W.; Ghosh, S. S. *Nature Biotech.* **2003**, *21*, 281-86.

178. Wang, Y. K.; Ma, Z.; Quinn, D. F.; Fu, E. W. *Anal. Chem.* **2001**, *73*, 3742-50.

179. Wang, Y. K.; Ma, Z.; Quinn, D. F.; Fu, E. W. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1389-97.

180. www.mitoproteome.org

181. Kluza, J.; Marchetti, P.; Gallego, M.; Lancel, S.; Fournier, C.; Loyens, A.; Beauvillain, J.; Bailly, C. *Oncogene* **2004**, *23*, 7018-30.

182. Iacobelli, S.; Arno, S.; D'Orazio, A.; Coletti, G. *Cancer Res.* **1986**, *46*, 3005-10.

183. Ullrich, A.; Sures, I.; D'Egidio, M.; Jallal, B.; Powell, T. J.; Herbst, R.; Dreps, A.; Azam, M.; Rubinstein, M.; Natoli, C.; Shawver, L. K.; Schlessinger, J.; Iacobelli, S. *J. Biol. Chem.* **1994**, *269*, 18401-7.

184. Marchetti, A.; Tinari, N.; Buttitta, F.; Chella, A.; Angeletti, C. A.; Sacco, R.; Mucilli, F.; Ullrich, A.; Iacobelli, S. *Cancer Res.* **2002**, *62*, 2535-39.

185. Grassadonia, A.; Tinari, N.; Iurisci, I.; Piccolo, E.; Cumashi, A.; Innominato, P.; D'Egidio, M.; Natoli, C.; Piantelli, M.; Iacobelli, S. *Glycoconj J.* **2004**, *19*, 551-56.

186. inari, N.; Kuwabara, I.; Huflejt, M.E.; Shen, P.; Iacobelli, S.; Liu, F. Int. J. *Cancer* **2001**, *91*, 167-72.

187. Fornarini, B.; D'Ambrosio, C.; Natoli, C.; Tinari, N.; Silingardi, V.; Iacobelli, S. *Blood* **2000**, *96*, 3282-85.

188. Akahani, S.; Nangia-Makker, P.; Inohara, H.; Kim, H. C.; Raz, A. *Cancer Res.* **1997**, *57*, 5272-76.

189. Yang, R.; Hsu, D. K.; Liu, F. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6737-42.

190. Akahani, S.; Nangia-Makker, O.; Inohara, H.; Kim, H.; Raz, A. *Cancer Res* **1997**, *57*, 5272-76.

191. Yu, F.; Finley, R. L.; Raz, A.; Kim, H. C. *J. Biol. Chem.* **2002**, *277*, 15819-27.

192. Takenaka, Y.; Fukumori, T.; Yoshii, T.; Oka, N.; Inohara, H.; Kim, H. C.; Bresalier, R. S.; Raz, A. *Mol. Cell. Biol.* **2004**, *24*, 4395-4406.

193. Fukumori, T.; Oka, N.; Takenaka, Y.; Nangia-Makker, P.; Elsamman, E.; Kasai, T.; Shono, M.; Kanayama, H.; Ellerhorst, J.; Lotan, R.; Raz, A. *Cancer Res.* **2006**, *66*,

3114-19.

194. Prieto, V. G.; Mourad-Zeidan, A. A.; Melnikova, V.; Johnson, M. M.; Lopez, A.; Diwan, H.; Lazar, A. J. F.; Shen, S. S.; Zhang, P. S.; Reed, J. A.; Gershenwald, J. E.; Raz, A.; Bar-Eli, M. *Clin. Cancer Res.* **2006**, *12*, 6709-15.

195. Liang, X.; Zhao, J.; Hajivandi, M.; Wu, R.; Tao, J.; Amshey, J. W.; Pope, M. R. J. *Proteome Res.* **2006**, *5*, 2632-41.

196. Kim. J.; Cho, H.; Kim, J.; Hur, S.; Kim, T.; Lee, J.; Kim, I.; Namkoong, S. *Lab Invest.* **2000**, *80*, 587-94.

197. Hguyen, H.; Sankaran, S.; Dandekar, S. *Virology* **2006**, *354*, 58-68.

198. Morris, E.; Michaud, W. A.; Ji, J.; Moon, N.; Rocco, J. W.; Dyson, N. J. *PloS Genetics* **2006**, *2*, 1836-48.

199. Bracken, A. P.; Ciro, M.; Cocito, A.; Helin, K. *Trends biochem. Sci.* **2004**, *29*, 409-17.

200. Vorburger, S. A.; Pataer, A.; Yoshida, K.; Liu, Y.; Lu, X.; Swisher, S. G.; Hunt, K. K. *Ann. Surg. Onc.* **2003**, *10*, 314-22.

201. Nijtmans, L. G. *EMBO J.* **2000**, *19*, 2444-51.

202. Mishra, S.; Murphy, L. C.; Nyomba, G. B.L.; Murphy, L. J. *Trends in molecular Med.* **2005**, *11*, 192-97.

203. Fusaro, G.; Wang, S.; Chellappan, S. *Oncogene* **2002**, *21*, 4539-48.

204. Fusaro, G.; Dasgupta, P.; Rastogi, S.; Joshi, B.; Chellappan, S. *J. Biol. Chem.* **2003**, *278*, 47853-61.

205. Fu, A.; Fenselau, C. *J. Proteome Res.* **2005**, *4*, 1583-91.

206. Warburg, O. *Science* **1956**, *123*, 309-14.

207. Unwin, R. D.; Craven, R. A.; Harnden, P.; Hanrahan, S.; Totty, N.; Knowles, M.; Eardley, I.; Selby, P. J.; Banks, R. E. *Proteomics* **2003**, *3*, 1620-32.

208. Ockner, R. K.; Kaikaus, R. M.; Bass, N. M. *Hepatology* **1993**, *18*, 669-76.

209. Parrella, P.; Xiao, Y.; Fliss, M. *Cancer Res.* **2001**, *61*, 7623-26.

210. Tan, D. J.; Bai, R. K.; Wong, L. J. *Cancer Res.* **2002**, *62*, 972-76.

211. Isidoro, A.; Casado, E.; Andres, R.; Acebol. P.; Espinosa, E.; Alonso, A.; Acebol, P.; Expinosa, E.; Aloso, A.; Cejas, P.; Hardisson, D.; Vara, J. A.; Belda-Iniesta, C.; Gonzalez-Baron, M.; Cuezval, J. M. *Carcinogenesis* **2005**, *26*, 2095-104.

212. Raha, S.; Robinson, B. H. *Trends Biochem. Sci.* **2000**, *25*, 502–508.

213. Harper, M.; Antoniou, A.; Villalobos-Menuey, E.; Russo, A.; Trauger, R.; Vendemelio, M.; George, A.; Bartholomew, R.; Carlo, D.; Shaikh, A.; Kupperman, J.; Newell, E.; Bespalov, E. A.; Wallace, S. S.; Liu, Y.; Rogers, J. R.; Gibbs, G. L.; Leahy, J. L.; Camley, R. E.; Melamede, R.; Newell, K. M. *FASEB* **2002**, *16*, 1550-57.

214. Villani, G.; Greco, M.; Papa, S. *J. Bio. Chem.* **1998**, *273*, 31829-36.

215. Payne, C. M.; Holubec, H.; Bernstein, C.; Bernstein, H.; Dvorak, K.; Green, S. B.; Wilson, M.; Dall'Agnol, M.; Dvorakova, B.; Warneke, J.; Garewal, H. *Cancer Epidemiol Biomarkers Prev.* **2005**, *14*, 2066-75.

216. Kadenbach, B.; Arnold, S.; Lee, I. *Biochem. Biophys. Acta* **2004**, *1655*, 400-08.

217. Herrmann, P. C.; Gillespie, J. W.; Charboneau, L.; Bichsel, V. E.; Paweletz, C. P.; Calvert, V. S.; Kohn, E. C.; Emmert-Buck, M. R.; Liotta, L. A.; Petricoin, E. F. *Proteomics* **2003**, *3*, 1801-10.

218. Krieg, R. C.; Knuechel, R.; Schiffmann, E.; Liotta, L. A.; Petricoin, E. F.; Herrmann, P.

C. *Proteomics* **2004**, *4*, 2789-95.

219. Starkov, A. A.; Fiskum, G.; Chinopoulos, C.; Lorenzo, B, J,; Browne, S. E.; Patel, M. S.; Beal, M. F. *J. Neuroscience* **2004**, *24*, 7779-88.

220. Bunik, V. I.; Sievers, C. *Eur. J. Biochem.* **2002**, *269*, 5004-15.

221. Tahara, E. B.; Barros, M. H.; Oliveira, G. A.; Netto, L. E. S.; Kowaltowski, A. J. *FASEB J.* **2007**, *21*, 274-83.

222. Maguire, B. A.; Zimmermann, R. A. *Cell* **2001**, *104*, 813-16.

223. Koc, E. C.; Ranasinghe, A.; Burkhart, W.; Blackburn, K.; Koc, H.; Moseley, A.; Spremulli, L. L. *FEBS letters* **2001**, *492*, 166-70.

224. Well, J.; Henkler, F.; Leversha, M. Koshy, R. *FEBS Letters* **1995**, *358*, 119-25.

225. Donald, S. P.; Sun, X.; Hu, C. A.; Yu, J.; Mei, J. M.; Valle, D.; Phang, J. M. *Cancer Res.* **2001**, 61, 1810-15.

226. Liu, Y.; Borchert, G. L.; Surazynski, A.; Hu, C. A.; Phang, J. M. *Oncogene* **2006**, *25*, 5640-47.

227. Zhong, L.; Peng, X.; Hialgo, G. E.; Doherty, D. E.; Stromberg, A. J.; Hirschowitz, E. A. *Proteomics* **2004**, *4*, 1216-25.

228. Fleming, M. D.; Campagna, D. R.; Haslett, J. N.; Trenor, C. C.; Andrews, N. C. *Genes & Development* **2001**, *15*, 652-7.

229. Navone, N. M.; Polo, C. F.; Frisardi, A. L.; Andrade, N. E.; Battle, A. M. *Int. J. Biochem.* **1990**, *22*, 1407-11.

230. King, F. W.; Wawrzynow, A.; Honfeld, J.; Zylicz, M. *EMBO J.* **2001**, *20*, 6297-305.