ABSTRACT

Title of Document: ALTERATIONS IN THE PRIMARY STRUCTURES OF RIBOSOMAL PROTEINS IN ACQUIRED DRUG RESISTANCE

Karen Lynn Lohnes, PhD, 2012

Directed By: Professor Catherine Fenselau, Department of Chemistry and Biochemistry

Acquired drug resistance is a multifactorial process that is one of the major causes for cancer treatment failure. The anticancer drug, mitoxantrone, was recently determined to inhibit ribosome biogenesis. Changes in ribosomal protein composition and efficiency with which the ribosomes incorporate ^{35}S-methionine has been noted in a mitoxantrone resistant MCF7 cell line when compared with a drug-susceptible parental cell line. This dissertation evaluated three proteomic workflows in order to successfully characterize the changes in the primary structures of cytoplasmic ribosomal proteins isolated from a mitoxantrone resistant breast cancer cell line that could serve some functional significance to the resistance when compared with a parental drug-susceptible cell line.

A combination of the data from the three workflows allowed for the identification of 76 of the 79 human ribosomal proteins with an average sequence coverage of 76%. The N-terminal ends of 52 of the ribosomal proteins were

identified using bottom-up and middle-down mass spectrometric approaches. An additional 7 N-terminal fragments were identified by top-down high resolution mass spectrometric analysis. Forty of the 52 N-terminal peptides were observed to have lost their N-terminal methionine and 19 were acetylated. Identification of the N-terminal peptides was most successful using the middle-down approach. Internal acetylations (on lysine) and phosphorylations were only noted with trypsin in-gel digestion and HPLC fraction analysis.

Gel arrays of the two ribosomal populations illustrated differences in the protein compositions. Comparative densitometry imaging software indicated the presence of two novel protein spots in the drug resistant cell line as well six additional spots with increased and decreased abundances. High coverage bottom-up mass spectrometric analysis allowed for these protein spots to be assigned as isoform pairs of RPS3, RPS10, RPL11 and RPL23A. Molecular masses and top-down analyses were used to define the alterations in the ribosomal proteins in conjunction with high coverage bottom up and middle-down analyses. The change in the primary structures of these four ribosomal proteins is believed to alter access to the mRNA tunnel in the ribosome. This suggests that these ribosomes may participate in differential selective translation to allow for the cell to produce the necessary proteins during cellular stress.

ALTERATIONS IN THE PRIMARY STRUCTURES OF RIBOSOMAL PROTEINS IN ACQUIRED DRUG RESISTANCE

By

Karen Lynn Lohnes

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
2012

Advisory Committee:
Professor Catherine Fenselau, Chair
Professor Douglas Julin
Professor Nicole LaRonde-LeBlanc
Professor Jonathan Dinman
Professor Nathan Edwards

Dedication

This thesis is dedicated to my husband and my parents who have believed in me even when I had lost all hope.

Acknowledgements

Above all, I would like to acknowledge my advisor Dr. Catherine Fenselau for providing me with the amazing opportunity to work in her laboratory. Her wealth of knowledge in proteomics and mass spectrometry and her generosity in sharing her time with all of the members of her laboratory including myself is something for which I am deeply indebted. Her guidance has been critical in my development as an effective scientist. I would like to acknowledge the support of the Fenselau lab members, both past and present who created a collaborative and friendly work environment which I always looked forward to being a part of, even on my worst days. I would also like to recognize the tireless support, patience and guidance from Dr. Nathan Edwards without whom the entire bioinformatics foundation of my thesis research would not have been possible. I have received invaluable research advice from Dr. Yan Wang for which I am extremely thankful and which I will carry with me for the rest of my career. I am very thankful for the thoughtful support and invaluable time of my other committee members, Dr. Douglas Julin, Dr. Jonathan Dinman and Dr. Nicole LaRonde-LeBlanc. I would also like to recognize the frequent assistance and guidance I have received from Dr. Peter L. Gutierrez regarding all aspects of my research. This study would not have been possible without the generous contributions of MCF7 breast cancer cells from Dr. Takeo Nakanishi and Dr. Peter L. Gutierrez from the University of Maryland Greenebaum Cancer Center. I would also like to acknowledge the NIH for their financial support of this research study. Finally I would like to thank my family and friends for their continued support and good sense of humor when I needed it most.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Akt……RAC-alpha serine/threonine-protein kinase (protein kinase B)

ASF……A-site finger

CK2……casein kinase 2

CLIPS……..chaperones linked to protein synthesis

c-myc/Myc……..…myc proto-oncogene protein

DTT…….dithiothreitol

ERK……..extracellular signal-regulated kinase

ESI-MS…….electrospray ionization mass spectrometry

FOX03/Fox03a…….Forkhead box protein 03

IAA……...iodoacetamide

IRES…….internal ribosome entry site

MALDI…….matrix-assisted laser desorption ionization

MDM2……E3 ubiquitin-protein ligase mdm2 (p-53 binding protein mdm2)

miRISC Ago2……..microRNA silencing complex argonaute 2

mLST8……target of rapamycin complex subunit LST8/G-protein beta subunit-like

mTOR…….serine/threonine protein kinase mTOR (FK506 binding protein)

mTORC2…….mammalian target of rapamycin complex 2

MXR…….mitoxantrone resistant

MXS…….mitoxantrone susceptible

NAC…….nascent polypeptide associated complex

NDRG1...……..protein NDRG1 (N-myc downstream –regulated gene 1 protein)

nLC…….nanoflow liquid chromatography

NOLC1/Nopp140……nucleolar and coiled body phosphoprotein 1

NSI…….nanospray ionization

p53……cellular tumor antigen p53

PI3K……..…phosphoinositide 3-kinase

PKCδ……..protein kinase C delta

PRMT5……..protein arginine methyltransferase 5

PTM…….post translational modification

RACK1……..receptor for activated C kinase/Guanine nucleotide-binding protein subunit beta-2-like 1

Rictor……..rapamycin insensitive companion of mTOR

ROS……reactive oxygen species

RP-HPLC……reverse-phase high performance liquid chromatography

RPL#……ribosomal protein of the large subunit (followed by numerical identifier)

RPS#.....ribosomal protein of the small subunit (followed by numerical identifier)

SGK…serine/threonine-protein kinase sgk1 (serum/glucocorticoid-regulated kinase 1)

SIN1…….stress-activated map kinase interacting protein 1

SRP…….signal recognition particle

TIC…….total ion chromatogram

vRNA……..vault ribonucleic acid

2-DGE……..…two-dimensional gel electrophoresis

Chapter 1: Background & Introduction

Structure and function of the ribosome

Understanding the ability of a single cell, whether prokaryotic or eukaryotic, to manufacture all the proteins necessary to sustain the life of that cell has been the focus of intense scrutiny in the life sciences for decades. Collaborative efforts from researchers around the world have determined that an organelle referred to as the ribosome is central to the execution of this task in the cell. In 2009, Venkatraman Ramakrishnan, Thomas A. Steitz and Ada E. Yonath were awarded the Nobel Prize in Chemistry for their contributions to the studies of the structure and function of the ribosome [3].

The ribosome is a large ribonucleoprotein complex that catalyzes the peptidyltransferase reaction in polypeptide synthesis. It is often referred to as a "molecular machine" which plays a fundamental role in the generation of the cellular proteome [4]. The human ribosome is composed of one molecule each of 79 different proteins and four different ribosomal RNAs (rRNA). Coordinated expression of the ribosomal protein genes is necessary to ensure equimolar accumulation of ribosomal proteins [5]. Deficiencies in any of the ribosomal protein mRNAs or ribosomal rRNAs have been associated with cellular abnormalities in human cell lines[6]. In fact, investigations using live cell imaging and quantitative mass spectrometry have shown that healthy eukaryotic cells import more ribosomal proteins into the nucleus than export ribosomal subunits, suggesting that an excess of ribosomal proteins are produced and the fraction remaining in the nucleus are degraded. In an experiment

with human cervical cancer cells where the transcription of rRNA was prevented, ribosomal protein synthesis did not immediately end however the ribosomal proteins were found to rapidly degrade with the average half-life being calculated between 30-90minutes[7; 8].

Although the active site of the ribosome is comprised of rRNA, classifying this organelle as a "ribozyme", ribosomal proteins play a crucial function in providing the structure for the rRNA [9]. In addition, multiple ribosomal proteins have been implicated in playing a regulatory role in cell differentiation and apoptosis. For example, ribosomal protein RPS5 has been shown to play a role in cell cycle arrest [10]. Certain ribosomal proteins, like *Escherichia coli* ribosomal protein RPL4, have also been shown to function as repressors of their own transcription [11].

Ribosomes are critical for the survival of the cell since they are responsible for the translation of transcripts encoded in the cellular genome. The ribosome is comprised of two subunits referred to as the large and the small subunit. However, the composition and size of these subunits differ between bacteria, animals, fungi and plants. Both of these subunits have rRNA and protein components. The intact and subunit components have been characterized by their sedimentation coefficients. In the case of the eukaryotic ribosome, the intact cytosolic ribosome sediments at 80S while its large and small subunit sediment at 60S and 40S respectively. The prokaryotic ribosome is referred to as the 70S ribosome and its respective large and small subunits the 50S and 30S subunits [4]. The larger sedimentation coefficient of the eukaryotic ribosome is due to its mass of ~4MDa compared with the 2.3MDa mass of the prokaryotic *Escherichia coli* ribosome[12; 13]. The additional mass of the

eukaryotic ribosome is the result of it containing a larger number of proteins (approximately 2 dozen more), an additional rRNA molecule, and longer rRNA chains known as expansion segments [14]. Cell growth is closely coupled to ribosome accessibility since ribosomes are responsible for the production of all the proteins in the cell, and thus ribosomes are very abundant in multiplying cells. In a typical human cell there are on average 4×10^6 ribosomes with approximately 5 to 10% of cellular protein and roughly 80% of all cellular RNA being ribosomal [15]. It is estimated that a growing HeLa cell synthesizes approximately 7500 ribosomal subunits per minute[16].

The structural differences between the ribosomes of different animal kingdoms relates back to differences in their protein and RNA molecular composition. On average, there are a total of 79 ribosomal proteins of the eukaryotic ribosome (79 in the human; not counting sex linked RPS4 individually) in comparison with the 57 proteins of the prokaryotic ribosome and 68 proteins of the archaeal ribosome. The large subunit of the eukaryotic human ribosome contains 47 proteins versus 40 for archaea and 34 for prokaryotes. The small subunit of human ribosomes possess 32 proteins compared to the 28 seen in archaea and 23 seen in prokaryotes [17]. Despite the discrepancy in the number of proteins each of these evolutionary domain possess, many of the ribosomal proteins are homologous between species. Orthologous counterparts exist between roughly 30% of prokaryotic ribosomal proteins and the proteins of eukaryotes and archaeal ribosomes. An additional 30% of the ribosomal proteins of archaea have shared counterparts in eukaryotes [18].

As previously mentioned, the rRNA components of the ribosomes of the three evolutionary domains also differ. The large subunit in prokaryotes is composed of two rRNA molecules, the 23S and the 5S, while the large subunit in eukaryotes consists of three rRNA molecules, 25S (yeast)/28S (humans), 5.8S and 5S[12; 19; 20]. The additional eukaryotic rRNA of the large subunit, 5.8S, is similar in sequence to the 5' end of the prokaryotic 23S rRNA and contains about 160 nucleotides [21]. The 28S rRNA is about 2300 nucleotides longer in humans than the 23S of prokaryotes (*E. coli*) [20]. The small subunit in all species contains only one rRNA molecule however in eukaryotes it sediments at 18S versus 16S in prokaryotes (~1550 nucleotides) due to an additional 300+ nucleotides for a total of ~1900 nucleotides [12; 20]. The eukaryotic ribosome with its additional proteins and rRNA molecules is considerably larger than its prokaryotic counterpart.

Ribosomal proteins were originally named (and numbered) by their respective position on a two dimensional polyacrylamide gel starting from the number one at the top of the gel and counting down to the bottom. As a result, a small basic protein will be assigned a large number while a large acidic protein is appointed a small number [18]. Given the differences among the ribosome of various species, the respective nomenclature assigned to a ribosomal protein from one organism does not always correspond with the protein of the same number in another organism. The average human ribosomal protein is relatively small containing 164 amino acid residues (range from 25 to 421) and having a molecular weight around 18.5 kDa (range from 3.5 kDa for RPL41 to 47.7 kDa for RPL4) [22].

Human ribosomal proteins (and ribosomal proteins in general) are very basic. Rat ribosomal proteins which are similar in sequence to humans were determined by Wool and colleagues to have an average isoelectric point (pI) of 11.05 with a pI range from 4.07 for RPLP1 to 13.46 for RPL41 (12.96 theoretical pI for human RPL41). The basic amino acids that contribute to this pI are often found in clusters of 3 or 4 consecutive residues. Additionally, a very small number of ribosomal proteins have clusters of acidic amino acid residues often at or near the carboxyl terminus. Repeats of 3 to 8 residue sequences is a common structural feature of ribosomal proteins although the reasons for these repeats are not well understood [22]. Zinc finger motifs coordinated with zinc have been found in numerous ribosomal proteins in many species including eukaryotes and archaea suggesting that they may contribute in the interaction between the rRNA and the proteins [22; 23; 24]. A leucine zipper-like motif has also been found in several ribosomal proteins, presumably to mediate nucleic acid binding[22; 25; 26]. These protein structural motifs most likely help ribosomal proteins to play their role in ribosome biogenesis (docking them in the correct position on the forming ribosome).

Human ribosomal proteins additionally contain nuclear localization signals to facilitate their entry into the nucleolus where they associate with pre-rRNA during ribosome biogenesis [22]. It is in the nucleolus where RNA polymerase I (Pol I) generates the primary precursors of 5.8S, 18S and 25S/28S rRNA. It is believed that these pre-rRNA precursors along with a large subset of the ribosomal proteins, non-ribosomal proteins and small nucleolar ribonucleoprotein particles (snoRNPs) amass in the nucleolus to fashion the 90S pre-ribosomal particles. The rRNA nucleotides

undergo extensive modifications carried out by the snoRNPs, the majority of which are 2' O methylations and pseudouridylation. These rRNA modifications cumulatively have been shown to optimize the rRNA structure to generate ribosomes that undergo translation with improved accuracy and efficiency[19; 27; 28]. The 90S particles undergo nucleolytic cleavages to produce the pre-40S and pre-60S particles which are processed separately prior to export from the nucleoplasm into the cytoplasm where the final assembly and maturation steps occur [19].

Translation

The ribosomal subunits are found independent of one another in the cell except during active translation of an mRNA transcript. The ribosome contains three specialized binding sites which are involved in the translation process; the A site which binds the aminoacyl-tRNAs, the P site that binds peptidyl-tRNA and the E site which binds the deacylated tRNA before it exits the ribosome. The process of translation is commonly described as occurring in three phases; initiation, elongation, and termination [13]. An increased level of complexity, with the requirement for additional protein cofactors, distinguishes eukaryotic translation from the simpler archaeal and prokaryotic processes [29].

Translation initiation in eukaryotes requires the involvement of more than 25 polypeptides. As part of the initiation process, a ternary complex must be formed consisting of a GTP-coupled eukaryotic initiation factor 2 (eIF2) and the methionine-loaded initiator tRNA which is responsible for the recognition of the AUG codon on the mRNA [30]. This ternary complex becomes a part of the 43S pre-initiation

assembly which also includes the 40S subunit and eIF3, eIF1, eIF1A and eIF5. Binding of this assembly to the mRNA requires the formation of a scaffold which first involves the ATP dependent unwinding of structures in the 5'-untranslated region (UTR) by the eIF4F complex which has assembled on the 5' 7-methylguanosine cap of the mRNA. A poly(A) binding protein bound to the 3'-poly(A) tail of the mRNA then interacts with components of the eIF4F complex in conjunction with eIF3 to load the mRNA onto the 43S complex. Bound to the 5' end of the mRNA, the 43S pre-initiation complex first scans the mRNA in a 5' → 3' direction until the AUG initiation codon is found. eIF1 is required for the selection of the correct initiation codon [29; 31].

When this AUG codon is encountered, embedded in the Kozak sequence or a variant thereof, codon-anticodon base pairing takes place between the initiator tRNA in the ternary complex and the initiation codon, forming the 48S complex. This event prompts the hydrolysis of GTP by eIF2 with the assistance of the GTPase-activating protein (GAP) eIF5, which allows for the deposit of the Met-tRNA into the P-site of the 40S subunit and dissociation of eIF2*GDP, eIF1, eIF1A, eIF3 and eIF5 from the complex. The last step in eukaryotic translation initiation requires the recruitment of the 60S subunit to the 40S pre-initiation complex. eIF5B*GTP, another GTPase in the initiation pathway, is required for coupling of the 60S subunit with the 40S*Met-tRNA*mRNA complex. Although GTP hydrolysis by eIF5B is not required for the subunit coupling, it is necessary for ribosomal peptide bond catalysis to occur. eIF5B*GDP dissociates prompting the elongation phase of translation [29; 30; 31].

Peptide chain elongation has been highly conserved across the three evolutionary kingdoms with each set of protein cofactors in eukaryotes having analogous sets in prokaryotes and archaea. Elongation begins with the delivery of the cognate aminoacyl tRNA to the vacant A site by the ternary complex eEF1A*GTP*aa-tRNA where conformational changes within the decoding center of the small ribosomal subunit and GTP hydrolysis help to ensure only the cognate aa-tRNA is selected [32]. The formation of the peptide bond between the incoming amino acid on the cognate aa-tRNA and the methionine on the peptidyl-tRNA is then catalyzed via a transfer reaction by the ribosomal peptidyl transferase center (PTC). This involves a nucleophilic attack of the carboxyl group of the methionine by the amino group of the amino acid bound to the aa-tRNA. The forming polypeptide chain is placed on the tRNA that was in the A site and as a result of this reaction this tRNA assumes a hybrid state where the acceptor end is now in the P site of the large subunit while the anticodon end remains in the A site of the small subunit. The deacylated tRNA which carried the methionine is also in a hybrid state with its acceptor end in the E site of the large subunit while the anticodon end remains in the P site of the small subunit. eEF2*GTP allows for the complete translocation of these tRNAs to open up the A site for another aa-tRNA and to move the mRNA three nucleotides to place the next codon into the A site. GTP is hydrolyzed by eEF2 during the translocation process. This elongation cycle repeats until a stop codon in the mRNA moves into the A site which initiates termination [31]. Recent studies with bacteria, yeast, and rabbit liver-derived ribosomes have illustrated that in addition to the control of elongation by the EF factors there is also an allosteric relationship

between the A site and E site where the affinity for the one site is decreased by the occupation of the other [33].

Termination is also catalyzed by the PTC but in this case since the stop codon (UAA, UAG, or UGA) does not have a cognate aa-tRNA, a release factor referred to as eRF1 is recruited to the ribosome. eRF1 promotes the nucleophilic attack by water of the carboxyl group on the polypeptide linked to the peptidyl-tRNA leading to the release of the polypeptide chain. eRF3 is another eukaryotic release factor believed to aide in the ejection of eRF1 from the ribosome after the hydrolysis reaction [31]. The recycling of the ribosomal subunits for another round of translation is promoted by a combination of eIFs 3, 1, 1A, and 3j [34].

Mitoxantrone

The chemotherapeutic agent referred to as mitoxantrone is an anthracenedione shown in Figure 1.1. Its intended mechanism of action is to intercalate DNA which has been found to lead to the inhibition of topoisomerase II and DNA strand breaks[35]. Additional effects of mitoxantrone have been determined. Mitoxantrone along with many other chemotherapeutic drugs has been discovered to inhibit ribosome biogenesis by interfering in the transcription of 47S rRNA at concentrations below clinical relevance[36]. Mitoxantrone and related quinones have also been discovered to undergo metabolism by cytochromes-P450 to form oxidized products[37]. The formation of these metabolites also produces reactive oxygen species (ROS) that contribute to the cytotoxicity of the drug. The oxidized metabolites of mitoxantrone bind covalently to thiol groups of peptides. The mitoxantrone metabolites undergo a

strong interaction with glutathione to form thioether conjugates[37]. Two other known interactions/actions of mitoxantrone have recently been reported. Mitoxantrone binds to nucleolar and coiled body phosphoprotein 1 (NOLC1) also known as Nopp140. Phosphorylated Nopp140 interacts with casein kinase 2 (CK2) to inhibit the catalytic activity of CK2. Mitoxantrone enhances the interaction between these two proteins. CK2 is known to play a role in the regulation of rDNA transcription and apoptosis and the inhibition of its activity is thought to suppress cancer cell survival. A final interaction that mitoxantrone is known to have is with two specific vault RNAs (vRNAs)[38]. A recent study of three human cell lines resistant to mitoxantrone has found that vRNAs were overexpressed in all three cell lines. In addition, the interaction between mitoxantrone and the vRNAs was found to assist in the detoxification of the cell as the vRNAs allow for the export of toxic compounds from the cell. When the vRNAs were suppressed with RNA interference, the resistance progressively decreased[39].

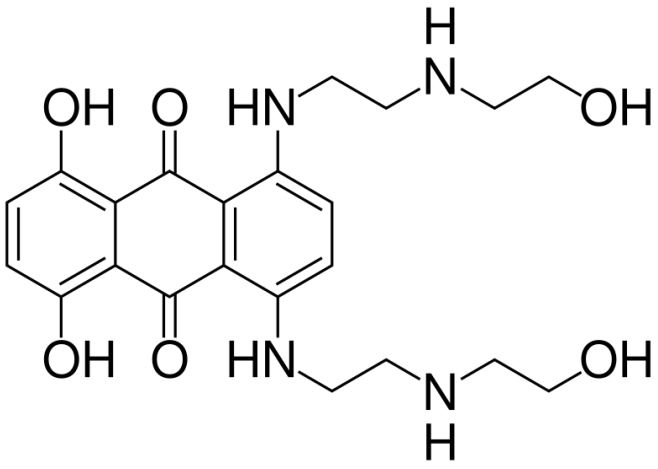

Figure 1.1 Structure of mitoxantrone

Drug Resistance and Ribosomal Proteins

Although the active role of translation is catalyzed by rRNA, ribosomal proteins have been shown to play an integral role in the transfer of information during the translation process. In particular, RPL2 (RPL8 in higher eukaryotes) and RPL3 have been demonstrated through mutant screenings in *Escherichia coli* to play critical roles in ribosomal function[40; 41]. Point mutations in either RPL2 or RPL3 in *Saccharomyces cerevisiae* have been shown to convey resistance to anisomycin, an A-site specific translational inhibitor[40; 41].

Changes in other ribosomal proteins have also been connected with drug resistance. Ribosomal protein RPL41 has been shown to convey resistance to cyclohexamide, which inhibits peptidyl elongation in various strains of yeast when Proline 56 is converted to Glutamine[42]. Point mutations discovered in RPS12, RPS5, and RPL22 have been shown to convey resistance in *E. coli* to streptomycin, spectinomycin and erythromycin respectively[43]. Changes of amino acids at position 37, 42, or 90 in ribosomal protein RPS12 in *Thermus thermophilus* led to various drug resistant or drug dependent phenotypes[44]. Carr *et al.* investigated further the link between streptomycin resistance and a β-methylthiolation modification found in this resistant bacterium at D88. Although they found the resistant phenotype in RPS12 could be retained in the absence of the β-methylthio-aspartic acid, it appeared to be under the condition that the modified residue at position 90 be a bulky residue[44; 45].

Multiple ribosomal proteins have been shown to have extra-ribosomal functions, interacting with other proteins either prior to the completion of ribosome

biogenesis or after being incorporated into the ribosomal machine. Some studies have even shown ribosomal proteins leaving the ribosome to perform these functions[25; 46]. Drug resistance, particularly multi-drug resistance (MDR), in humans has recently been shown to be associated with the interactions these ribosomal proteins have with non-ribosomal proteins/factors. A recent study of MDR in gastric cancer cells showed that over expression of RPS13 and RPL23 were associated with resistance to the anti-cancer drugs vincristine, adriamycin, 5-fluorouracil and in the case of RPL23 also resistance to cisplatin. This MDR was linked to suppression of drug-induced apoptosis and in the case of overexpression of RPL23, through the control of the glutathione-S-transferase-mediated drug-detoxifying system [47]. Drug resistance in humans is multi-factorial and involves many proteins with altered structures or abundances[47; 48].

Proteomics and the use of mass spectrometry

The use of proteomic strategies is the most practical approach to examine and compare a large set of proteins such as those from the ribosome complex. The term "proteome" was first used in print in 1995 by Wasinger *et al*. [49]. Proteome refers to the protein complement expressed by a given tissue type or genome under known environmental conditions or in a disease state [50]. The field of proteomics was defined by N. Leigh Anderson and Norman G. Anderson as "*the use of quantitative protein-level measurements of gene expression to characterize biological processes (e.g., disease processes and drug effects) and decipher the mechanisms of gene expression control*" [51]. According to Chhabil Dass, there are a number of subfields in proteome analysis including; (1) *functional*

proteomics, which investigates the identity of a group of proteins involved in a specific function, (2) *characterization proteomics*, which evaluates the proteins present in a tissue, cell or biofluid, and (3) *differential proteomics*, which distinguishes proteins that are differentially expressed in various physiological states, such as in drug resistant and drug susceptible cells [52].

The human genome is reported to consist of approximately 20,000 genes, however due to alternative splicing, one of the most conservative estimates of the number of human proteins is around 35,000 according to the UniProtKB/Swiss-Prot database[53] . The complexity of this set of proteins and the dynamic range of the concentrations in which these proteins exist poses major technical challenges in proteomics. This is further complicated by new research which suggests that the amino acid sequence of many human proteins do not match their underlying DNA sequence[54]. There are four basic stages incorporated into proteomic analysis to help overcome these seemingly insurmountable obstacles[52].

The first step for overcoming the large dynamic range of a sample, for example a human cell sample, is in the sample preparation. Subcellular fractionation as a part of sample preparation can drastically reduce sample complexity and allow for deeper proteome coverage. Sample preparation may also include depletion of highly abundant proteins to enrich low abundance proteins [52]. This strategy is often utilized in the analysis of blood serum by the depletion of albumin [55]. In the case of the ribosome, proteins are kept in approximately equimolar concentration, thus the dynamic range of the ribosome is believed to be only 1 [6].

The second step involves the separation and purification of proteins from the subcellular fraction(s) into individual proteins [52]. In the current study, the use of liquid chromatography and gel electrophoresis were essential for this purpose allowing for the enrichment of each protein in a respective fraction or gel spot.

The third stage of a proteomic workflow involves the end point analysis of the sample, which in most proteomic studies utilizes mass spectrometry, i.e. analysis of peptides [52]. Mass spectrometry is most often used because it provides sequence tags or peptide mass maps that form the basis for searches of the protein databases. It requires considerably less material and the sample need not be purified to homogeneity [56].

The fourth stage of a proteomic workflow involves the database search of the mass spectra. There are multiple publically available search engines that can be used to perform this task. The choice of an appropriate database depends on the nature of the sample analyzed (i.e. protein, DNA) and the type of mass spectrometric analysis used.

Mass spectrometric analysis can be approached using one of three strategies. The first approach is referred to as *bottom-up proteomics*. In this strategy, the protein(s) of interest is/are first digested using enzymatic or chemical cleavage. These peptides are then analyzed by one of several mass spectrometric platforms. One of the two most common workflows is *peptide mass fingerprinting* where the proteins are traditionally first separated by 2-D gel electrophoresis and then each individual gel spot is subjected to tryptic digestion and analyzed by mass spectrometry. In this case the mass measurements of the peptides provide a "mass

fingerprint" that can be used to identify the corresponding protein in the database. Another more comprehensive workflow requires digestion of the entire protein mixture, fractionation of the peptides via multi-step chromatography and analysis using *tandem mass spectrometry*. This step produces MS/MS spectra, which can be searched against a given database[52; 57]. Search algorithms have been developed to identify peptides on the basis of tandem mass spectra data in search engines such as MASCOT, SEQUEST, OMSSA, X!Tandem, and Myrimatch. The algorithms are used to match the arrangement of fragment ions detected in the spectrum with those calculated theoretically from the database entries. The most commonly used search engines are MASCOT and SEQUEST [52]

An alternative strategy being developed for proteomic analysis is the *top-down* approach. This approach uses mass spectrometry to weigh intact protein ions and multi-stage tandem mass spectrometry to produce sequence tags from large protein fragments. Post-translational modifications which play a crucial role in cell signaling can be more readily identified and quantified using top-down techniques due to the fact that intact proteins are less susceptible to the instrumental biases of small peptides (i.e. differences in ionization efficiency). Until recently however, this has been less widely used. This is in part due to the fact that larger sample quantities are required by current instrumentation. In addition, the analysis time has not, until recently been compatible with the chromatographic timescale[58; 59]. ProSight is currently the only commercially available algorithm for identifying protein forms from the tandem mass spectra of intact proteins[58; 60]. As the public database of protein modifications grows and the computational tools for assigning protein

identities and PTMs become more robust, the top-down proteomic approach is becoming the method of choice for investigating combinatorial PTMs [58; 61]. The most robust characterizations will include a combination of both top-down and bottom-up proteomics.

In lieu of top-down or in combination with it, another approach that has been widely favored in proteomics research in recent years is referred to as the *middle-down* approach. The middle-down approach typically takes advantage of enzymatic or chemical cleavage with selectivity for a single amino acid residue. The resulting proteolytic products produced from the protein(s) of interest are large, with polypeptides typically observed between 3kDa to 10kDa in size. Larger polypeptides have been observed to fractionate with improved resolution by HPLC. These peptides can be analyzed using a variety of mass spectrometric platforms, often through a combination of methods used in top-down and bottom-up proteomics based on the nature and complexity of the sample being investigated. A common approach is fractionation of the peptides via nanoLC interfaced with a high resolution instrument, such as a Thermo LTQ-Orbitrap, for high resolution (survey scans acquired at 30K resolution) *tandem mass spectrometry* of both precursor and product ions. Large polypeptides carry a higher number of charges when electrosprayed which enhances CID and ETD. The middle-down approach is especially favored in the investigation of proteins containing multiple PTMs or proteomes which contain multiple protein isoforms[2; 62; 63; 64; 65; 66]. The interpretation of the mass spectra is dependent on the overall size of the observed polypeptides but frequently utilizes

several search engines including ProSightPC, Mascot and SEQUEST to interrogate the data [2; 64; 66].

Application to the study of ribosomal proteins

Since the inception of proteomics, the characterization of the ribosomal proteome of both prokaryotic and eukaryotic organisms has been the focus of many research laboratories. The ribosome as a research focus posed an opportunity for scientists to develop experimental approaches for the study of non-covalent protein complexes. Interest in the ribosomal proteome was also facilitated by the fact that an equimolar contribution of each protein was expected in the ribosomal machine. As a result, concerns regarding dynamic range were thought to be negligible. Although researchers were aware that the ribosomal proteins could be modified, it was not until these studies were underway that researchers became aware of the extensive number and, on occasion, combination of PTMs found on the ribosomal proteins.

Research on the ribosome using proteomics tools and techniques has often used a combination of top-down and bottom-up mass spectrometry to identify ribosomal proteins and their associated PTMs with a more complete characterization than one technique alone can provide. The bottom-up strategy has frequently involved the use of two-dimensional gel electrophoresis for visualization of ribosomal protein spots or 1 or 2-D fractionation of the proteins with liquid chromatography (usually strong cation exchange followed by reverse phase) prior to proteolysis. Occasionally the whole ribosomal protein mixture has been digested without fractionation however it is worth noting that this can often preclude the ability of the

researcher to distinguish between protein isoforms. Proteolysis of the ribosomal proteome has been reported with trypsin, Lys-C, Glu-C or acid digestion among others. Proteolytic products are analyzed with peptide mass fingerprinting or tandem mass spectrometry (usually ESI/MS/MS) to identify protein components and associated PTMs on a particular peptide. The top-down strategy has either involved MS or MS/MS of intact ribosomal proteins. The use of more sensitive and/or high resolution mass spectrometers for these measurements such as an FT-ICR mass spectrometer, hybrid ion trap mass spectrometer (for example LTQ-Orbitrap) or hybrid Quadrupole mass spectrometer (Quadrupole-FT or the Q-TOF) has become standard in ribosomal proteome studies. Bioinformatic tools have also been advanced which was necessary for the identification of whole ribosomal proteins measured with these instruments (i.e. Thrash algorithm and ProSightPC) [64; 67; 68; 69; 70; 71; 72]. These practices enabled the discovery of many co- and post-translational modifications on both eukaryotic and prokaryotic ribosomal proteins and ribosome associated proteins such as RACK1. Several examples of proteomic research focused on the ribosome follow.

One of the earliest successful studies to characterize the ribosomal proteome using proteomic techniques was published by the Ahn lab in 1996[73]. This study used a combination of top-down and bottom-up mass spectrometry and focused on ribosomal proteins of the small subunit from Rat-1 fibroblasts. The ribosomal protein mixture was simultaneously fractionated for collection via RPLC monitored via UV detector and a portion of the HPLC effluent was directed to a triple quadrupole mass spectrometer equipped with a nebulization ESI source and high pressure collision cell

for intact protein mass measurements (determined from ion series across multiple scans). LC/MS/MS analysis of intact proteins was also achieved where further analysis was determined necessary for protein identification. Corresponding HPLC fractions were digested with Lys-C for protein identification using tandem mass spectrometry with a triple quadrupole instrument. Forty one proteins were observed in total, 36 of which corresponded with the 32 expected small ribosomal subunit proteins. Four ribosomal proteins were discovered to have two forms (isoforms); RPS3, RPS5, RPS7 and RPS24. Of the 32 ribosomal proteins observed, twelve had molecular masses identical to the predicted mass of the proteins; RPS4, RPS6, RPS7, RPS8, RPS13, RPS15a, RPS16, RPS17, RPS19, RPS27a, RPS29, and RPS30. An internal hydroxylation or methylation was proposed for mass changes observed in RPS23. A loss of the N-terminal methionine and/or acetylation was observed or proposed in thirteen of the remaining proteins; RPSa, RPS3a, RPS5, RPS11, RPS15, RPS18, RPS20, RPS21, RPS24, RPS26, RPS27, RPS28, and one RPS7 isoform. Additional modifications were found on two of these proteins, RPS5 and RPS27 which were proposed to be internally formylated or acetylated respectively. The remaining ribosomal proteins; RPS2 (+ 220Da), RPS3 (Isoform 1;-75Da; Isoform 2; -362Da), RPS9 (+86Da), RPS10 (+57Da), RPS12 (-100Da), RPS14 (-117Da), RPS25 (-103Da) showed changes in their molecular mass which were sometimes localized to portions of the sequence but could not be explained [73]. This study indicated that the majority of ribosomal proteins are co- or post-translationally modified, illustrating the value of using proteomic techniques to characterize ribosomal proteins.

In 2005, the Leary lab examined and compared the ribosomal proteome of the human 40S subunit with the proteome of a human 40S subunit complexed with Hepatitis C Virus (HCV) IRES (Internal Ribosome Entry Site) using a combination of top-down and bottom-up methods[74]. The isolated ribosomal protein mixture was fractionated via RP-HPLC into 120 fractions. HPLC effluent was lyophilized and resolubilized in 79:20:1 Acetonitrile:Water:Formic acid. Molecular mass measurements were obtained by infusion of the samples at 1μL/min into an FT-ICR mass spectrometer equipped with an ESI source. Bottom-up analysis was achieved by reducing, alkylating and digesting the entire ribosomal protein mixture with trypsin. The polypeptide mixture was then subjected to LC/MS/MS using a Q-TOF mass spectrometer. Of the expected thirty two proteins found in the small subunit, thirty one were observed between all observations. All of the proteins identified using top-down methods were found to contain PTMs which included; N-terminal Met loss, acetylation, methylation, dimethylation, and disulfide bond formation. Between the two proteomes, the native 40S proteome and the HCV IRES-complexed ribosomal proteome, two proteins were observed to contain differences in their PTMs, RPS25 and RPS29. In the case of RPS25, the native 40S protein was always observed to be dimethylated while the IRES-complexed RPS25 was observed to exist in both a mono- and dimethylated form. The native 40S RPS29 protein was observed to contain 2 disulfide bonds while the IRES-complexed version of the protein did not contain any disulfide bonds. Six proteins from the native 40S proteome, RPS11, RPS4, RPS6, RPS8, RPS26, and RPS3a, were not found using the top-down method but were found with the bottom-up approach. There is a possibility that these

proteins also differ between the two proteomes (since 100% sequence coverage was not achieved with polypeptide analysis). Proteins associated with the ribosomes were also observed. RACK1 was detected with both proteomes however nucleolin was only witnessed in association with the IRES-complexed proteome [74]. This study suggested that PTMs on the ribosomal proteins that are part of native 40S complex versus those involved in a HCV-IRES complex may be useful in distinguishing healthy from diseased state.

In 2007, James P. Reilly and colleagues used a top-down/bottom-up approach to study the ribosomal protein components in the *Caulobacter crescentes* bacterium[72]. Ribosomal proteins isolated from these organisms were fractionated using an intricate two-dimensional liquid chromatography (2D-LC) system that contained a strong cation-exchange column which fed into 20 reversed phase trap columns followed by 2 reversed phase analytical columns. Eluent from the second analytical column was split with a portion of the flow directed to an ESI source coupled with a Q-TOF mass spectrometer for intact protein measurements. Protein identification was confirmed by proteolysis of protein fractions collected from the 2D-LC. Three methods of proteolysis were used to increase sequence coverage. To determine the sequence at the C-terminal end of the protein, Carboxypeptidase Y and Carboxypeptidase P were used and proteolytic products analyzed using a Q-TOF mass spectrometer. Protein fractions were in certain cases selectively digested with Glu-C and analyzed with MALDI/MS peptide mass mapping. Additionally, trypsin digestion was used for peptide analysis on each protein fraction. Proteolytic products of the trypsin digests were analyzed using both MALDI/MS and capillary-LC-ESI-MS/MS with an ion-trap

mass spectrometer. Fifty three of the fifty four ribosomal proteins found in *C. cresecentes* were identified and their intact masses determined. The masses of approximately one half of the ribosomal proteins matched the theoretical mass with the loss of the N-terminal methionine, while one quarter matched the theoretical mass without Met loss. The remaining ribosomal proteins possessed various PTMs including combinations of Met loss, acetylation, methylation, and oxidation in the case of one protein. There were also discrepancies between the observed protein masses and the theoretical protein masses, supported at the peptide level, which could be explained by truncations of 13 and 11 residues at the N-termini of RPL3 and RPS21 respectively and termination of the sequence in RPL27 prior to the final C-terminal residue, Glu. Reilly and his colleagues argued that since random proteolysis of the proteins was not observed in any other cases, this was strong argument for an error in the interpretation of the genome sequencing data for those 3 proteins [72]. Not only did this study illustrate the great extent to which prokaryotic ribosomal proteins may be modified but it also suggested that for certain prokaryotic organisms in particular, the gene annotation may have been misinterpreted.

In 2008, Carroll *et al*. studied the ribosomal proteins of *Arabidopsis thaliana* using a multifaceted bottom-up approach[75]. Ribosomal proteins were isolated and 100µg of protein was run in parallel in three lanes of a large format gel. Protein bands were divided into 30 gel regions, individually excised and subjected to in-gel protease digestion with trypsin. Using duplicate gels, low molecular mass protein bands (≤ 20kDa) were also digested separately with chymotrypsin and pepsin. Additional gels were stained with a commercially available phosphoprotein selective

stain in order to detect proteins which were phosphorylated. Protein bands containing these phosphorylated proteins were excised and subjected to in-gel proteolysis followed by phosphopeptide enrichment with TiO_2 microcolumns. Extracted peptides were subjected to LC/MS/MS using a Q-TOF mass spectrometer fitted with an ESI source. More detailed analyses were conducted using a nanospray source (also with the Q-TOF instrument) to identify PTMs. *In-silico* approaches in addition to custom software for the filtering of peptide match information led to the identification of 87 individual proteins and 79 of the 80 predicted protein families of the *Arabidopsis* ribosome. Five of these ribosomal protein gene families had never previously been observed experimentally for *Arabidopsis*; RPS29, RPS30, RPL29, RPL36a and RPL39. Theoretical approaches suggested that 63 of the ribosomal protein gene families should have distinguishable peptide products. Of these, the authors found good specific gene matches in the following protein families; RPS3, RPS3a, RPS6, RPS24,RP L7, RPL13a,RP L35, RPL4, RPL5, RPL7a, RPL14, RPL17, RPL23a, RPL26,RP L28, RPL32 and RPLP3. Four non-ribosomal proteins were identified as being associated with the ribosomes; 2 guanine nucleotide-binding family proteins, a ferritin-like protein and a eukaryotic translation initiation factor. Strong MS/MS evidence was found for 30 unique covalently modified peptides with a total of 41 covalent modification events. These modifications included 15 instances of Met loss, 12 cases of N-terminal acetylation, 1 occurrence of N-terminal dimethylation, 9 instances of phosphorylation and 3 cases of N-methylation of Lys side chains. Conservation of modifications across eukaryotes is suggested based on the fact that, for example, N-methylation of Lys observed in this study has also been observed in

homologous proteins in yeast. Additionally patterns of Met removal and N-terminal acetylation and phosphorylation are also reported as being widely conserved across eukaryotic ribosomal proteins [75]. This study reinforced the fact that with the heterogeneity in ribosomal protein gene families and possible PTMs, considerable forethought must be put into the acquisition and filtering of MS/MS data in order to identify specific members of a ribosomal protein family (which by the authors' own account sometimes only vary by one residue). The suggestion that there may be conservation of modifications found on homologous ribosomal proteins requires further investigation.

The Ribosome is not Static

Ribosomal proteins and disease

Over the last few decades, researchers have uncovered a link between ribosomal protein expression levels and/or gene mutations of ribosomal proteins with human diseases. Certain inherited conditions are linked with ribosomal protein dysfunction and mutation including Diamond Black-fan Anemia syndrome (RPS19), Turner syndrome (RPS4X), Camurati-Engelmann disease (RPS18), Noonan syndrome (RPL6), and Bardet-Beidl syndrome (RPS30) to name a few[76]. In addition, numerous studies have cited an increase in ribosomal protein expression levels in association with cancer. This includes (but is not limited to); esophageal cancer (RPL15), gastrointestinal cancer (RPL13), cervical cancer (RPS12), prostate cancer (RPL37 and RPL7a), colorectal cancer (RPS3, RPS6, RPS8, RPS12, RPL5, RPL22, RPL35, RPL36) and hepatocellular cancer (RPL13, RPL36a, RPS8, RPL12, RPL23a, RPL27 and RPL30) [76]. A study on maternally inherited deafness associated with a

homoplasmic mutation in a mitochondrial rRNA gene confirmed a coordinated over-expression of most cytoplasmic ribosomal protein genes with microarray, flow-cytometry and quantitative RT-PRC, believed to be a compensatory mechanism for this mutation [77]. Comparative studies of different colorectal carcinomas suggested that the expression of certain ribosomal proteins could be correlated with the stage of tumor and the malignancy potential of the cancer [78].

There are generally two theories regarding the reason/role that ribosomal proteins are differentially expressed in disease states such as cancer. One; a disruption in the distribution of ribosomal proteins in the cell affects their function in the ribosome in protein biosynthesis which either precedes, follows, or is the cause of tumorigenesis or two; the extra-ribosomal functions of these ribosomal proteins directly leads to a signal pathway(s) causing tumor formation and/or growth [78; 79; 80; 81]. Both theories are supported in the literature, however at this time only select ribosomal proteins have known extra-ribosomal functions. Research in the laboratory of Paul Fox (among others) has shown that not all ribosomal proteins are essential for the ribosomal machine to perform its function in protein biosynthesis and in certain circumstances modification of a ribosomal protein may lead to its departure from the ribosomal complex to perform an extra-ribosomal function. For example, a study investigating the extra-ribosomal function of human RPL13a illustrated that phosphorylation of RPL13a allowed for its release from the ribosome and subsequent function as a transcript-selective, translational silencer of ceruloplasmin. This suggested that the ribosome may act as a "storage depot" for translational control proteins [25]. Although this explanation is a convenient way of combining 2 theories

behind the differential expression of ribosomal proteins, ribosomal proteins which are not easily able to "leave" the ribosomal complex to perform extra-ribosomal functions would be excluded.

Currently there is a hypothesis referred to as the "ribosome filter hypothesis" which provides an explanation for how ribosomal proteins embedded in the complex may perform a function in the translational control of other proteins[82]. The ribosome filter hypothesis proposes that mechanisms involving differential mRNA capture allow for the ribosomal subunits to affect the translation of particular mRNAs. This hypothesis was originally rooted in the observation that most mRNAs have regions of their sequence which are complementary to sequences in the 28S or 18S rRNAs suggesting a potential mechanism by which mRNA-rRNA pairing might occur [82; 83]. Support for this hypothesis has come not only from the fact that the literature documents differential expression of ribosomal proteins (in diseased states for example) but also from the discovery that PTMs on the ribosomal proteins (within the confines of the ribosome) differ during different stages of the cell cycle. For example, a study by Haselbacher *et al* in 1979 found that phosphorylation of RPS6 was influenced by insulin-growth factor in the transition from G_0 to G_1. A similar discovery was made by Spence *et al* in 2000 when they discovered that ribosome-associated RPL29 (RPL27a in humans) was multi-ubiquitinated in both yeast and human cells as a functional cell-cycle dependent modification. Porras-Yakushi *et al* illustrated in 2006 the role that methylation plays in the ribosomal proteins, particularly at different times in the complex[84; 85; 86]. Four interrelated views which accompany the ribosomal filter hypothesis are as follows; 1) Mechanisms within the

ribosome structure allow for the preferential translation of subsets of the mRNA population, 2) ribosomal interactions allow for regulatory effects on mRNA, 3) binding site competition in the ribosomal subunits may affect the rate of translation of different mRNAs and 4) ribosomal heterogeneity may allow for masking or altering of particular binding sites on the ribosome [82].

Testing by Mauro and Edelman of the hypothesis that complementarity between rRNA and mRNA leads to preferential translation examined a 9 nucleotide (9-nt) element in the mRNA sequence of the *Gtx* homeodomain protein. Functional and biochemical studies showed that translation was maximally enhanced with a specific 7-nt sequence in the 9-nt element. In experiments where there was poor complementarity between the *Gtx* element and rRNA (such as in *Saccharomyces cerevisiae*), translation of the protein was not enhanced however when the nucleotide sequence was altered to complement the rRNA, a dramatic increase in translation efficiency was observed [82].

Additional work by Mauro and Edelman which supports the ribosome filter hypothesis illustrated that ribosomes are heterogeneous as previous studies have supported. Ribosome protein composition/expression in disease states and during stages of the cell cycle has been observed with different ribosomal protein expression patterns. Duplicate genes for ribosomal proteins found in *S. cerevisiae* have been observed to serve functionally distinct roles [87]. In light of the fact that some studies have suggested that a select number of ribosomal proteins are unnecessary for the protein biosynthesis function of the ribosome, there exists a possibility that ribosomes originating from different cell type or tissues have different protein compositions[25; 82].

Modifications of rRNA have been observed to play a role in IRES-dependent translation control and deficiencies in these modifications are linked with human X-linked Dyskeratosis Congenital syndrome and pituitary tumor formation [88].

Perhaps one of the greatest strides taken in ribosomal research in recent years has been the discovery that the ribosome plays a regulatory rather than a constitutive role in the coordination of gene expression during embryonic development. This should not be surprising considering that all other molecular machinery such as chromatin associated histones and the spliceosome involved in gene expression confer specialized functions in gene regulation. Kondrashov *et al* have recently found that RPL38 plays a critical function in the axial skeletal patterning during embryonic development of mice. Mutations in RPL38 manifest themselves by skeletal patterning defects and shorter kinky tails. When transgenic mice were created which rescued the RPL38 mutation, normal phenotypes and typical RPL38 expression levels were observed. Axial skeletal morphology is regulated in mammals with HOX genes. Kondrashov *et al* examined the expression boundaries and transcript levels of the HOX genes in the mutant mice and found that they were unchanged in the mutants suggesting the mutant phenotypes were attributable to RPL38 and not due to transcript levels and/or expression boundaries of the HOX genes. Differences in the amount of general cap-dependent vs. IRES-dependent translational control were measured as well as the global protein synthesis assessed and both found to be unchanged from the WT embryo. Microscale polysome analysis revealed that the number of small to large subunits and the polysome distribution was unchanged between WT and mutant mice, supporting the determination that there was

no change in global protein synthesis. Quantitative PCR analysis of the HOX mRNAs found that there was no evidence of a perturbation in transcriptional control of the HOX genes. Despite this fact, the association of certain HOX mRNAs with both the light and heavy polysomes was decreased in the mutants. The corresponding proteins for the HOX mRNAs which were found to associate at a lower level with the polysomes were also found in lower abundance. This suggested that the control of the HOX mRNA by RPL38 was at a translational level. Protein expression levels of the translationally deregulated HOX mRNAs in mutants was restored to normal levels in transgenic mice and was also able to be controlled by an in-vitro translation system. Mouse mutants of 5 other ribosomal proteins ($RPS19^{DSK3/+}$, $RPS20^{DSK4/+}$, $RPL29^{+/-}$, $RPL29^{-/-}$, $RPL24^{BST/+}$) were examined and it was found that no changes in HOX mRNA expression/translation or axial skeletal morphology were observed even in cases when global protein synthesis was markedly affected. Several experiments were conducted to determine the role which RPL38 played in the translational control of certain HOX mRNAs and perceived that RPL38 is involved in the formation of the 80S ribosome on these mRNAs (perhaps a form of ribosomal "recruitment"). Ribosomal protein expression patterns in tissues throughout the mice revealed specific patterns which are believed to be relevant to the role these ribosomal proteins play in their translational control. The question remains to be answered whether "specialized ribosomes" such as these exist as a result of the role that specific ribosomal proteins play in transcript-specific translational-control or simply by ribosomal protein heterogeneity. There also exists the possibility that PTMs contribute to the properties of "specialized ribosomes" [89].

All of the studies discussed above illustrate that the ribosome is a complex, malleable organelle whose protein isoforms/modifications and sometimes composition vary with tissue type, stage in cell cycle, health of the organism and developmental state. Ribosomal research has shown that despite the plasticity of the ribosome, certain mutations can lead to disease and death. Answering the questions posed by the Kondrashov study regarding "specialized ribosomes" will help scientists to better understand the role which the ribosome and ribosomal proteins can play in translational regulation and disease. Modifications of ribosomal proteins such as truncation and PTMs were distinguishing features of certain ribosomal proteins in disease and studies of the cell cycle [25; 73; 74; 75; 79; 87; 90]. Given the role the ribosome plays in the regulation of the cell cycle and the link it has to some forms of drug resistance in prokaryotic organisms, it is experimentally relevant to investigate and compare the ribosomal proteome in chemotherapeutic resistant human cell lines with a drug susceptible cell line.

Objective & Specific Aims

Based on the aforementioned evidence illustrating the plasticity of the ribosome, it is our objective to develop an effective means by which modified ribosomal proteins may be isolated and identified. These methods will be applied to characterize modifications in the ribosome of a drug resistant cancer cell line compared to the drug susceptible precursor. With this goal, we have the following specific aims;

1 Develop methods that allow the rapid analysis of ribosomal proteins with high sequence coverage.

2 Characterize the primary structure of altered ribosomal proteins in MCF 7 human cancer cells selected for resistance to mitoxantrone.

3 Consider the possible impact on function of the ribosome played by these altered proteins based on the known structure of ribosome.

Chapter 2: Materials and Methods

Materials

Mitoxantrone resistant cells and drug susceptible cells were provided by Dr. Takeo Nakanishi and Professor Peter L. Gutierrez of the University of Maryland Greenebaum Cancer Center. Canted neck T-150cm^2 cell culture flasks were purchased from Corning, Inc. (Lowell, MA). Improved Minimal Essential Media (IMEM) supplemented with L-glutamine, fetal bovine serum and phosphate buffered saline (PBS) were purchased from Mediatech, Inc. (Manassas, Virginia). Acetic acid, ammonium bicarbonate, 3-[(3-cholamidopropyl) dimethylamonio] -1-propanesulfonate (CHAPS), glycerol, magnesium acetate, magnesium chloride, penicillin streptomycin antibiotic solution, potassium chloride, sodium dodecyl sulfate (SDS), sucrose, thiourea, trifluoroacetic acid (TFA), Trizma base, cell culture grade trypsin (.025%), and urea were obtained from Sigma Aldrich (St. Louis, Missouri). LCMS grade Acetonitrile (Optima), formic acid and HPLC-grade water was purchased from Fisher Scientific (Fairlawn, New Jersey). The Immobilized pH gradient (IPG) strips, pH 7-11 NL (non-linear gradient), and corresponding ampholytes (IPG buffer) were purchased from GE Healthcare (Piscataway, New Jersey). The ultra-clear ultracentrifuge tubes were obtained from Beckman-Coulter (Fullerton, California). Molecular weight cut-off filters (MWCO, 3 kDa) were purchased from Millipore (Microcon Ultracel YM-3, Billerica, MA). The RC/DC protein assay kit, bromophenol blue, cup-loading sample cups, Protean II precast gels (8-16%), electrode wicks, dithiothreitol (DTT), iodoacetamide (IAA) and Bio-Safe

Coomassie Blue were purchased from Bio-Rad (Hercules, California). Sequencing grade trypsin was purchased from Promega Corporation (Madison, Wisconsin). Additional large format gels (8-16% and 14%) were purchased from Jule, Inc. (Milford, Connecticut).

Equipment

The mechanical homogenizer was purchased from Kinematica (Littau, Lucerne; Switzerland). The DU-530 UV-Vis spectrophotometer and Optima LE-80K preparative ultracentrifuge are from Beckman Coulter (Fullerton, CA). The Orbital shaker was purchased from Fisher Scientific (Pittsburgh, PA). The lyophilizer was purchased from Labconco (Freezone 2.5 Liter Benchtop Freeze Dry system, Labconco Corp., Kansas City, MO). Isoelectric focusing device, second dimension gel apparatus, and GS-800 densitometer were purchased from Bio-Rad (Hercules, CA). An Accela HPLC, electrospray ionization source, nanospray ionization source and LTQ-Orbitrap XL were obtained from Thermo Fisher Scientific (San Jose, CA). An additional ionization source, the Advance CaptiveSpray Plug-and-Play source was acquired from Michrom Bioresources (Auburn, CA). Two HPLCs (microflow and nanoflow) and a MALDI-TOF instrument were all purchased from Shimadzu Scientific Instruments (Columbia MD). The Discoverer Benchmate microwave system with a fiber-optic temperature probe and a 45mL digestion vessel capable of holding ten 300µL glass vials was purchased from CEM Corporation (Mathews, NC).

Methods

Cell Culture

Mitoxantrone susceptible (henceforth referred to as ***MXS***) and mitoxantrone resistant breast cancer cells (henceforth referred to as ***MXR***) were grown until confluency in 150 cm^2 flasks (Corning, New York) in Improved Minimal Essential Media (IMEM) with L-glutamine supplemented with 1% penicillin-streptomycin antibiotic solution and 10% heat inactivated fetal bovine serum. To retain the mitoxantrone resistant phenotype, cells were grown with media that contains 250nM mitoxantrone as described by Nakagawa *et al*, 1992[35]. Cells were maintained at a temperature of 37°C in a water jacketed CO_2 incubator with 5% carbon dioxide. Ribosomes were harvested when a minimum of 20 flasks of a particular cell type had reached confluency.

Ribosome Isolation

To isolate the ribosomes, a modification of a method previously developed in the Fenselau lab was used [1]. MCF7 cells were released from the flask bed by washing with 15mL of cell culture grade 10mM phosphate buffered saline (PBS) followed by a 5 minute incubation with 3mL cell culture grade trypsin. Tryptic activity was stopped by the addition of 10 mL IMEM after the incubation. Cells were suspended in solution and transferred to a pre-weighed centrifuge tube for centrifugation at 500g for 5 minutes at 4°C in a benchtop centrifuge (Allegra 21R centrifuge, Beckman, Fullerton, CA). The cell pellet was resuspended in PBS and centrifuged two times before it was weighed.

Homogenization of the cell pellet on ice with a Kinematica mechanical homogenizer (Littau, Lucerne; Switzerland) in two volumes homogenization buffer (50mM Tris-HCL, pH 7.5; 5mM $MgCl_2$, 25mM KCl, 200mM Sucrose) was followed by centrifugation at 10,000g for 10 minutes at 4°C in the benchtop centrifuge. The supernatant was collected and the remaining pellet re-homogenized on ice and centrifuged. The supernatant was layered 1:1 over a sucrose cushion buffer (50mM Tris-HCl, pH 7.5; 5mM $MgCl_2$, 25mM KCl, 2M Sucrose) and the ribosomal pellet isolated by centrifugation (Optima LE-80K preparative ultracentrifuge, Beckman Coulter, Fullerton, CA) at 260,000g at 4°C for 2 hours in a swinging bucket rotor (SW60Ti). At this time point the samples were stored in 1mL homogenization buffer at -80°C until protein extraction.

Extraction/Isolation of Ribosomal Proteins

All procedures were carried out on ice unless otherwise indicated. Prior to protein extraction, the ribosomes were first precipitated by the addition of 0.7 volumes ice cold ethanol and this suspension centrifuged for ten minutes at 7000rpm. The pellet was then resuspended in 250µL homogenization buffer without sucrose. Ribosomal proteins were separated from the rRNA using a variation of the acid extraction procedure described by Hardy *et al*, 1969 [91]. In brief, one volume of the ribosomal suspension was mixed with 0.25 volumes of 1M $Mg(OAc)_2$ followed by the addition of 1 volume glacial acetic acid. Each solution was incubated for 1 hour. The precipitated rRNA was pelleted by centrifugation at 10,000rpm at 4°C for 10 min in a benchtop centrifuge. The supernatant containing the ribosomal proteins was then

collected for each ribosomal type. Solution digestion was carried out with both microwave accelerated acid cleavage and trypsin while the in-gel digestion utilized trypsin.

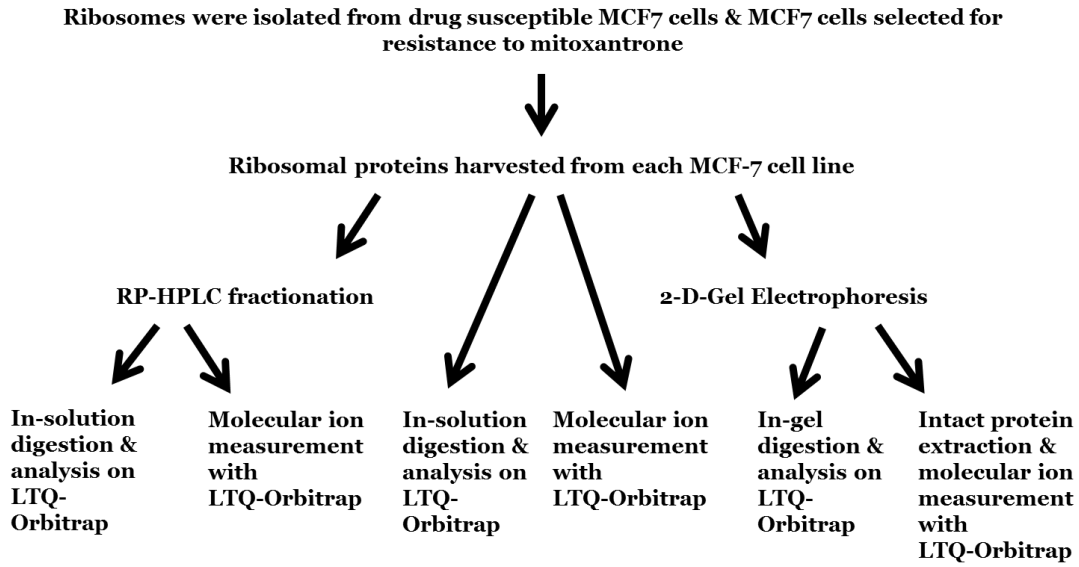

Figure 2.1 Experimental workflows used to evaluate the primary structure of the altered ribosomal proteins

Processing of the extracted ribosomal proteins depended on whether they were intended for use with an HPLC, gel electrophoresis or immediate cleavage of the whole ribosomal proteome with chemical digestion. In the case of both HPLC and cleavage of the whole ribosomal proteome, lyophilization and/or 3kDa MWCO filters were used to concentrate the samples and reduce the acid content. After desalting and concentrating the samples, the protein concentrations were determined using the Bio-Rad RC/DC protein assay (Bio-Rad, Hercules, CA). For samples intended for 2-D gel electrophoresis, acetone precipitation similar to that described by Barritault *et al*.,

1976 was used [92]. The proteins in the acetic acid extraction are mixed with 4 volumes ice cold acetone and incubated at -20°C for between 2 to 4 hours. The precipitated ribosomal proteins are then centrifuged in a benchtop centrifuge for 20 minutes at 10,500 g. The supernatant is quickly poured into another tube and the protein pellet washed with 1mL of acetone for two additional spins of 20 minutes at 10,500g. The supernatant from the original wash was also centrifuged to collect additional protein that did not pellet in the first run. This is washed with acetone as described for the main pellet. The acetone washed protein pellet was resuspended in a rehydration buffer, the volume and composition of which varied depending on the sample loading method for the first dimension. In the case of rehydration loading (both active and passive) the rehydration buffer consisted of 7M urea, 2M thiourea, 2% CHAPS, 0.5% IPG buffer, .05% bromophenol blue and 50mM DTT with a sample volume of approximately 240µL. In the case of cup-loading, the rehydration buffer consisted of; 7M urea, 2M thiourea, 15% isopropanol, 2.5% glycerol, 2% CHAPS, 0.5% IPG buffer, 0.05% bromophenol blue and 50mM DTT with a sample volume of approximately 100µL. Samples were incubated for a minimum of an hour in the rehydration buffer before an aliquot of 25uL was taken for a protein assay using the Bio-Rad RC/DC kit.

HPLC Fractionation of Ribosome Protein Mixture

(See Fig. 2.1)

Two methods were evaluated for pre-fractionation of the ribosomal proteome prior to bottom-up and molecular mass measurement. These were the use of 2-D gel electrophoresis or fractionation of the proteins using reverse phase-HPLC. These

methods fractionate based on different chemical characteristics of the proteins. In the case of the first stage of 2DGE, proteins are fractionated based on their isoelectric point while in the case of RP-HPLC, they are grouped based on their hydrophobic character.

Fractionation of the ribosomal proteome using HPLC first required the buffering of the protein mixture for reduction of the proteins. This was achieved by the addition of 2 volumes of 100mM ammonium bicarbonate. The sample was then reduced with 2mM DTT for 1 hour at 56° C. In some experiments, this was followed by alkylation with 4mM IAA in the dark at room temperature for 45 minutes. Since the injection loop on the HPLC used in these experiments holds 500µL, an appropriate volume of mobile phase A (10% acetonitrile/90% water/ 0.1% TFA), typically around 150µL, was added to each sample to aid with sample mixing upon injection.

The HPLC system used for these experiments was equipped with two Shimadzu LC-10 reciprocating pumps with a variable SPD-10 A_{VP} UV-VIS detector and a Rheodyne 500µL stainless steel sample loop and manual injection valve (Idex Health and Science, Rohnert Park, CA). A Phenomenex (Torrance, California) Jupiter 5µm particle size C-18 300Å 250 x 4.6mm i.d. column was used for these experiments. The column was run with a Phenomenex Security guard column equipped with a C-18 cartridge. The UV detector recorded chromatograms at 215nm. Solvents A and B were originally prepared as documented in earlier literature for the HPLC of ribosomal proteins with Solvent A consisting of 0.1% TFA in water and Solvent B, 0.1% TFA in acetonitrile [74]. When a steady baseline on the HPLC

instrument was not observed, different concentrations of acetonitrile (ACN) were evaluated for solvent A. Ten percent ACN containing 0.1%TFA was determined to generate the most reproducible results.

A set of four standard proteins chosen for their molecular weight and isoelectric points; hen egg white lysozyme (MW=14,307 Da, pI=11.35), horse heart cytochrome C (MW=12,384 Da, pI=10.5), bovine pancreas ribonuclease A (MW=13,690 Da, pI=9.6), and horse heart myoglobin (MW=17,641 Da, pI~7.2), was used to determine what HPLC conditions would generate reproducible separations. Stock solutions of each of these proteins were prepared as 0.5mg/mL in solvent A and kept frozen in between uses. Experiments determined that low nanomolar concentrations were the lower limit of protein concentration necessary for detection. A typical standard protein injection contained ~5 nanomoles of each protein. An average human ribosomal protein injection contained approximately 1.3mg/mL of protein. Taking into consideration that the average ribosomal protein weighs approximately 18.5kDa, there was approximately 70pmol/uL of each protein in these samples, meaning that for each 500µL injection in the sample loop there was ~10.5 nanomoles of each protein injected.

Several gradients were tested for the enriched ribosomal protein samples. Yeast ribosomal proteins were used initially as standards for testing HPLC conditions. The initial gradient tested was that reported by Yu *et al.* in 2005 for their work with the human 40S ribosomal subunit [74]. This gradient proceeded as follows; (1) 0 – 3 min 10% ACN, (2) 3 – 33 min 10 – 30% ACN, (3) 33 – 37 min 30 – 37% ACN, (4) 37 – 103 min 37 – 50% ACN, (5) 103 – 113 min 50 – 80% ACN, and (6)

113 – 123 min 80 – 95% ACN. The aforementioned gradient resulted in solvent mixing issues which led to poor resolution and peak capacity. Eleven additional gradients were tested before the following gradient was decided upon; (1) 0 – 83 min 15 – 95% ACN, (2) 83 – 93 min 95% ACN, (3) 93 – 94 min 95 – 15% ACN, (4) 94 – 113 min 15% ACN. Although different flow rates were attempted, 0.55mL/min produced the best results with the Phenomenex columns.

Fractions were manually collected every minute beginning at 5 minutes until 60 minutes. At 60 minutes, fractions were subsequently collected every 5 minutes until 80 minutes at which point one fraction was collected every ten minutes. Each fraction was split into two aliquots for bottom-up analysis and molecular mass analysis. Inspection of the whole proteins in each fraction using the MALDI-TOF revealed which fractions contained proteins. Molecular mass measurements were made with ESI-LTQ-Orbitrap on the whole protein fraction aliquots to determine the mass of individual proteins. Fractions containing proteins were digested in solution with trypsin to identify protein components.

Enzymatic digestion of these HPLC aliquots with trypsin was accomplished by first reducing the fractions with 2mM DTT for 1 hour at 56° C. Samples were then incubated in the dark with 4mM IAA at room temperature for 45 minutes. Each fraction was then incubated overnight with 0.6µg trypsin at 37° C. Peptides were dried down and redissolved in 60µL of 0.1% formic acid solution to stop the digestion and to prepare the samples for injection on the Thermo Accela HPLC coupled with the LTQ-Orbitrap for automated peptide analysis.

Whole Ribosome Protein Digestion Methods

Acid Digestion of Whole Ribosomal Protein Mixture

In cases where the ribosomal protein mixture was not fractionated prior to enzymatic or chemical cleavage, microwave accelerated acid digestion was evaluated. For the acid digestion 12.5% acetic acid and 5mM dithiothreitol were added to each sample. The Discover Benchmate microwave system was used. Microwave-accelerated acid cleavage on six 50µL aliquots of the ribosomal protein suspensions was carried out using methods previously described by this laboratory [64]. In brief, the digestion was carried out at a constant temperature of 140 ± 5°C with 300W for 20 minutes. Samples were allowed to cool before removal and combining them for mass spectrometric analysis.

Two-Dimensional Gel Electrophoresis

Isoelectric Focusing

Alkaline proteins are renowned for being difficult to focus in the first dimension of two-dimensional gel electrophoresis[93; 94; 95; 96]. This is due in part to the fact that the common reducing agent, DTT, is a weak acid with pka values of the thiol groups being 9.2 and 10.1 respectively. Dithiothreitol will thus ionize at basic pH and migrate towards the anode during IEF. This poses a problem for the focusing of the cysteine containing alkaline ribosomal proteins. In order to overcome this obstacle, researchers have taken many approaches including trying different sample loading methods, altering the composition of the rehydration buffer, and using a different reducing agent marketed by GE Healthcare as DeStreak™ [96; 97; 98].

Several methods of sample loading were tested for isoelectric focusing. Rehydration loading involves the loading of the sample onto the strip at the same time that the strip is rehydrated. Two versions of rehydration loading were tested. These are referred to as passive and active rehydration loading [99]. For each, 18cm IPG strips (7-11 NL, GE Healthcare) were used. In this procedure, 320µL rehydration solution containing 100µg sample was pipetted across a lane in the focusing tray. For passive rehydration, the IPG was laid face-down over this sample containing solution, covered with mineral oil to prevent the strip from drying out and incubated overnight before proceeding with the focusing. For the active rehydration tested in these experiments, the IPG strip was laid face down over the sample containing solution, the wetted strip covered with mineral oil and the focusing tray placed in the Bio-Rad Protean IEF cell at 50V overnight at 20° C to aide with sample entering the IPG strip. In both cases, rehydrated strips were then drained of mineral oil and placed in a clean focusing tray face-down. An electrode wick dipped in 15mM DTT is placed under the strip at the electrode on the cathode end while an electrode wick dipped in water is placed on the anode end. The strip is covered in mineral oil and the tray is placed in the apparatus for focusing. The best method for focusing with these forms of sample loading involved the following method; 500V for 2 hours with rapid ramping, 8000V for 30 min with linear ramping, and finally 8000V for 50kVhr with rapid ramping.

Neither form of rehydration loading resulted in reproducible and well resolved protein spots despite trying several focusing methods. Different rehydration buffer compositions from the literature were also tested which included the addition of

isopropanol and glycerol to suppress the reverse endoosmotic flow effect observed with each run [94; 100]. Though these changes to the buffer composition did improve the spot resolution on the gel, there were still problems with distinguishing individual spots.

As a result, a form of sample loading often recommended for basic proteins, referred to as cup-loading, was used [96]. Prior to cup-loading of the samples, the IPG strips needed to be rehydrated. This was done by incubating the IPG strips face down in 340µL rehydration buffer (7M urea, 2M thiourea, 2% CHAPS, 15% isopropanol, 2.5% glycerol, 0.5% IPG buffer, 0.05% bromophenol blue, 50mM DTT) overnight at 20° C covered with mineral oil. Rehydrated IPG strips were then picked-up with forceps to drain mineral oil, rinsed briefly (5 seconds) in Millipore water to remove any urea crystals, patted dry with filter paper (Bio-Rad) and placed face-up in the cup-loading isoelectric focusing tray. An electrode wick dampened with 15mM DTT was placed on the IPG strip at the cathode end while an electrode wick dampened with water was placed at the anode end. The platinum electrodes were placed over each end of the strip and 100µL sample cups placed firmly over the strip at the anode end. Rehydration buffer was used to test for a possible sample leakage from the cup. For each 18cm strip, 100µg of protein in approximately 50µL was added to each cup. Sample in each cup and strip were covered with mineral oil to prevent drying out and sample loss. The optimized method for focusing these proteins with cup-loading was found to be; 150V 1hr, 300V 2hr, 600V 1hr, 8000V 30min (linear ramping), 8000V 48kVhr. The total run averaged 60kVhr for 2 IPG strips. At least once during the run, the electrode wicks were replaced.

After optimization, the Destreak™ reagent (GE Healthcare, Piscataway, NJ) was tested as an alternative to DTT to simplify and possibly improve the method. The Destreak™ reagent is thought to be superior to DTT as instead of reducing the cysteine residues on the proteins, it oxidizes them resulting in "mixed disulfides" allowing the proteins to move into the basic region of the strip that posed a problem with DTT [96; 97]. For each 100µg sample, 1.2% Destreak™ reagent was added. The gels were run as above except that Destreak™ was used in place of DTT in the rehydration buffer, sample buffer and cathode wick solution. These gels were found to be of similar quality or poorer quality (more combined spots resulting in fewer spots in addition to sample loss) to the gels from the original method so it was determined that Destreak™ was not an improvement for this study.

Equilibration and Gel Electrophoresis

After isoelectric focusing it is necessary to equilibrate the IPG strips in equilibration buffer to allow the proteins to fully interact with SDS prior to electrophoresis. This is done in two steps in order to fully reduce and alkylate the proteins as well. However, since that is a slow process, the equilibration for each step must also be extended for at least 15min each [99]. Equilibration was performed using methods previously described by our lab with minor adjustments to account for the difference in sample loading [1]. In short, the IPG strip was removed from the cup-loading tray and drained of the mineral oil and placed in an equilibration tray (Bio Rad) containing Equilibration Buffer I (50Mm Tris-HCl (pH8.8), 6M Urea, 20% glycerol, 2% SDS and 2% DTT) for 30 minutes. The IPG strips were then removed

from this solution and incubated for 30 minutes in fresh equilibration buffer with 4% IAA replacing the DTT. The strip was removed, and quickly (5 seconds) rinsed in the electrophoresis running buffer (25mM Tris, 192mM glycine, 0.1% SDS) purchased from Bio Rad. It was then placed on top of an 8-16% Tris-HCl SDS-PAGE precast gel (Bio-Rad or Jule Inc.) and run using the same method previously used in our lab [1]. The strip was then covered with warm agarose solution purchased from Bio Rad and allowed to dry. Running buffer was added to the electrophoresis unit (Bio Rad) and the method used was; 16mA/gel for 30 minutes followed by 24mA/gel for 5 hours or until the Coomassie blue has reached the bottom of the resolving gel.

After the electrophoresis was completed the gel was removed and placed in a gel box (Corning) in fixing solution (45% Methanol/5%Acetic Acid/50% water) and left overnight on a shaker (Thermo Fisher). After fixing the gel was washed with water a minimum of 3 times for 15 minutes. The gel was then incubated on the shaker in the Bio Rad Bio-safe™ Coomassie blue staining solution for a period of at least 1 hour. Gels were then destained by soaking in water numerous times to remove background staining.

The gel image was obtained using a GS-800 Densitometer along with the associated Bio Rad software known as PDQuest™. PDQuest™ allows for comparison of individual gels as well as the composite of gels within that group (referred to as "automatching"). This is obtained by the software manually aligning gel spots (referred to as "landmark spots") from individual gels with one another. The researcher does have the ability to select additional landmark spots if the gels are

not aligned properly by the software. The user may also remove areas of background noise that are mistakenly identified as spots by the automatching process. A total of four gels for each harvest were combined for these analyses. Proteins from the two cell lines (MXR and MXS) were both harvested and separated by 2DGE on the same day in three of the four pairs. These gels were compared to one another as well as used in the composite gel for the MXR and MXS cell line comparison. Another software package called CompugenZ3™ (Compugen Limited, Tel Aviv, Israel) was used to analyze the gels in much the same way with a more automated system. The TIFF image of each gel obtained from the GS-800 densitometer was exported. The intensity of each spot in the gels was measured and background level values subtracted. Pairing of the gel spots between images was inspected manually using "zoomed" images. Spots whose abundances differed more than two fold from their matching spot in both software programs were selected for further investigation.

In-Gel Digestion

In order to identify the proteins of interest, a trypsin digestion of the gel spot was utilized. The protocol outlined by Shevchenko *et al*. in 2007 was used for this [101]. Gel spots of interest were excised from the gel under a laminar flow hood, cut into 1 x 1mm pieces and placed into a clean microcentrifuge tube. These gel spots were rinsed in 500µL milliQ water to remove any particulates. They were then spun down in a benchtop microcentrifuge for 30 seconds. Water was removed and gel pieces incubated with neat acetonitrile until they became white/opaque and stuck together.

Reduction and alkylation was accomplished by first removing the acetonitrile. Spots were then incubated at 56° with a fresh solution of 10mM dithiothreitol in 100mM ammonium bicarbonate (NH_4HCO_3) for 30 minutes. After the tubes had cooled down to room temperature, 100µL of neat acetonitrile was added to the solution and they were incubated for 10 minutes. All of the solution was removed and a fresh solution of 55mM iodoacetamide in 100mM NH_4HCO_3 was added and tubes incubated in the dark at room temperature for 20 minutes. As with the reduction step, 100µL of neat acetonitrile was added and tubes incubated for 10 minutes. After the gel pieces had shrunken, all of the liquid was removed and they were destained.

Destaining involved incubating the gel pieces with occasional vortexing in a 100 µL 1:1 (vol:vol) solution of 100mM NH_4HCO_3/neat acetonitrile for 30 minutes or longer. Five hundred microliters of neat acetonitrile was then added to this solution and the gels incubated with occasional vortexing until the gel pieces were almost entirely white/clear/opaque. This took between 10 minutes to 1 hour depending on the original intensity of the staining. The solution was removed and after assuring that the gel pieces were dry (from the acetonitrile), they were digested with trypsin.

Trypsin digestion was achieved by first adding sufficient trypsin buffer (13ng/µL trypsin in 10mM NH_4HCO_3 with 10% (vol/vol) neat acetonitrile) on the gel spot pieces to cover them (usually around 75µL buffer). They were incubated in this buffer in an ice bucket for 30 minutes. At this time, if there was area of the gel spot pieces that were no longer covered by trypsin buffer, additional buffer was added.

The gel spot pieces were incubated an additional 90 minutes on ice. At this time, 15µL of NH_4HCO_3 buffer was added and the gel spots placed in an incubator at 37° overnight. To extract the peptides, a 1:2 (vol/vol) of extraction buffer (usually 150µL extraction buffer, which consists of 1:2 (vol/vol) 5% formic acid/acetonitrile) was added to the digest. This mixture was then incubated at 37° on a shaker for 15 minutes. The microcentrifuge tubes containing the gel pieces were then spun briefly (30 seconds) in a benchtop microcentrifuge at 7000rpm. The extracted digest was removed with a fine gel-loader pipette tip and placed in a clean labeled tube and stored at -20°C until analysis on the LTQ-Orbitrap.

Extraction of Whole Proteins from Gels

Duplicate gels of each harvest and cell line were initially developed in order to have a gel available for protein extraction as well as one for trypsin digestion. Proteins were extracted following the protocol previously described by our laboratory[1] and originally developed by Mirza *et al* [1; 102]. This protocol involves first cutting the gel spots of interest from the gel with a clean razor blade and placing them in a labeled microcentrifuge tube. The spots are first washed for a minimum of one minute with 500µL HPLC grade water in order to remove any debris from the gel. Water was then removed and the excised gel pieces vortexed for 10 minutes in 10% acetic acid. The volume of this solution depended on the size of the gel spot however as a rule enough solution was used to cover the spot. After the removal of the acetic acid solution, the gel spots were washed 3 times with 500µL water for approximately 1 minute each. Following the water wash, the gel spots were washed with occasional

vortexing in neat acetonitrile for 20 minutes. The volume was again dependent on the size of the gel spot, but averaged 200µL per spot. The acetonitrile was removed after 20 minutes and the gel spots washed in 500µL water for approximately 1 minute each. The gel spots were then washed with HPLC grade methanol for 20 minutes with occasional vortexing. The solvent was again removed and the gel piece washed with HPLC grade water for 1 minute. The gel pieces were then each dipped into a solution of formic acid:water:isopropanol (FWI) (1:3:2, v/v/v) for between 30 minutes to 4 hours depending on the staining intensity of the spot. This final destaining solution was retained in clean labeled tubes for each spot to verify if there was protein loss in this step as this solution has also been shown to extract protein from gels. After the gel pieces were colorless, the gel spots were washed in water again and moved to clean labeled microcentrifuge tubes. They were then allowed to partially dry either on the bench top or under a laminar flow hood. The gel spots were crushed into small pieces and the proteins extracted by adding 30µL of extraction solution, 0.1% trifluoroacetic acid: acetonitrile (1:1, v/v) to each tube and vortexing them overnight (or a minimum of 5 hours). The following morning, the microcentrifuge tubes containing the extracted proteins were centrifuged in the benchtop microcentrifuge at 10,000rpm for 30 seconds to ensure the extraction solution could be collected at the bottom of the tube. A clean gel-loader pipette tip was used to collect the extracted protein solution from each spot and they were stored in a clean labeled Lo-bind™ Eppendorf tube at -20° until further analysis. The crushed gel spot material was also stored at -20° for each gel spot for further extraction if necessary.

When it became apparent that more material (protein) would be required for molecular mass determination with the ESI-Orbitrap, not only were additional 2-D gels prepared to collect replicate spots but also stored crushed gel spots were re-extracted with the extraction solution (no destaining or washing prior to additional extraction). Small amounts (2 to 3µL) of the material from each extraction were evaluated with MALDI to compare the new extraction to the older extractions. Extractions/re-extractions from a total of seven gels was combined for analysis on the ESI-Orbitrap.

Detection

MALDI (Intact Protein)

Intact proteins were evaluated with the Shimadzu-Axima CFR + MALDI-TOF instrument equipped with a nitrogen laser at a wavelength of 337nm to ensure that protein was in fact present. The methods described by this laboratory previously were used [1]. A MALDI matrix was prepared with 10mg/mL sinapinic acid in 50% acetonitrile/1% trifluoroacetic acid (matrix solution). Solutions were also prepared with 5% Triton-X 100. For each sample, 1µL protein extract was mixed with 1µL of either the standard matrix solution or the detergent-containing matrix solution in a clean, labeled tube. A method known as the sandwich method was used to spot these samples on a clean MALDI plate. This was achieved by first spotting the plate for each sample with 0.5µL of the MALDI matrix. After allowing this to dry, 1µL of the sample was spotted on top and allowed to dry and then followed by 0.5µL of the MALDI matrix. The same methods were used to calibrate the instrument prior to evaluating the samples with the SIGMA ProteoMass™ MALDI-MS protein

calibration kit which included protein standards which ranged in mass from approximately6kDa to 66kDa. The settings for the instrument were as follows; laser power: 100-115, over 200 profiles averaged; instrument method: linear mode, m/z range: 10,000 – 70,000.

Electrospray Ionization (ESI)

All ESI samples were transferred into HPLC sample vials and placed in the autosampler of either a Shimadzu Prominence HPLC or an Accela HPLC. Peptides from a preliminary investigation of the acid digested ribosomal proteins were analyzed by online ESI-MS/MS using a 1mm i.d. x 150mm column packed with reversed phase material (Biobasic-C-18, 300Å pore size, 5µ particle size) on the Accela HPLC running at a column flow rate of 40µL/min. The HPLC solvent gradient (solvent A 0.1% formic acid/2.5%ACN versus solvent B 0.1% formic acid/2.5% HPLC-grade water/97.5%ACN) was linear and began at 10% Solvent B and rose to 85% over 65 min. The HPLC solvent gradient was controlled by the XCalibur data system.

The LC-MSMS spectra were recorded on an LTQ-Orbitrap equipped with a Thermo electrospray ion source. The following ESI parameters were used: capillary temperature, 275°C; spray voltage, 4kV; capillary voltage, 21V; sheath gas flow, 35 arbitrary units; auxiliary gas flow, 8 arbitrary units. The Automated Gain Control (AGC) target and maximum injection time for precursor ions were set at $5x10^5$ and 250ms respectively for precursor scans, while for MS/MS they were set at $5x10^4$ and 100ms respectively. The three most abundant ions were selected for CID in the linear

ion trap and MS/MS analysis for every precursor scan. Precursor ions were scanned between *m/z* 350 and 2000. Precursor ions were isolated with a 3Da window and fragmented by low energy collisions with He gas for 30ms with normalized collision energy of 35 arbitrary units. Selected ions were excluded for the subsequent 10s for the *S. cerevisiae* sample and between 10-90s for the MCF7 samples (MXR and MXS). High resolution analysis (30,000 at m/z 400) of both precursor and product ions were determined using the Orbitrap.

Analysis of trypsin digests of the ribosomal proteome and the HPLC fractions were optimized under electrospray conditions of 50µL/min using the Shimadzu Prominence HPLC interfaced with the LTQ-Orbitrap via a Thermo ESI source. Samples were loaded onto a PepTrap 300Å C-18 pre-column at 5% solvent B for 10 minutes for desalting. Peptides were then eluted into an Agilent 5µ 300Å 15cm x 1mm ID C-18 column. ESI parameters used for the trypsin digested samples were: capillary temperature, 275°C; spray voltage, 4kV; capillary voltage, 35V; sheath gas flow, 35 arbitrary units; auxiliary gas flow, 10 arbitrary units. The AGC target and maximum injection time for precursor ions were set at $5x10^5$ and 500ms respectively. The five most abundant ions were selected for CID in the linear ion trap and MS/MS analysis for every precursor scan. Precursor ions were scanned between *m/z* 350 and 2000 for one full set of HPLC fractions and between *m/z* 400 and 2000 for the second full set. Precursor ions were isolated with a 2Da window and fragmented by low energy collision with He gas for 30ms with normalized collision energy of 35 arbitrary units. Selected ions were excluded for 45s. Precursor ions were measured

with high resolution analysis (60,000 at m/z 400) using the Orbitrap, while the linear ion trap was used for MS/MS analysis.

Intact protein analysis of the whole ribosomal proteome utilized the Accela HPLC system with a Waters X-Bridge C-18 (3.5µM particle size 300Å 4.6 x 250mm) column to separate the protein mixture over 130 minutes increasing acetonitrile from 15% to 80%, while a portion of the HPLC effluent was diverted to the LTQ-Orbitrap equipped with a Thermo ESI source. The remainder of the effluent was collected throughout the chromatographic separation for future analysis. When ion activation was used (for top-down analysis), the five most abundant ions in each precursor scan were automatically selected for CID in the linear ion trap and MS/MS analysis. Precursor ions were scanned between *m/z* 400 and 2000. Precursor ions were isolated with a 3Da window and fragmented by low energy collisions with He gas for 60ms with a normalized collision energy of 35 arbitrary units. Selected ions were excluded for 30s. High resolution analysis of both precursor (60,000 at m/z 400) and product (30,000 at m/z 400) ions were determined using the Orbitrap. ESI parameters were: capillary temperature, 275°C; spray voltage, 4kV; capillary voltage, 15V; sheath gas flow, 35 arbitrary units; auxiliary gas flow, 8 arbitrary units; sweep gas flow, 5 arbitrary units. The AGC target and maximum injection time for precursor ions were set at $5x10^5$ and 500ms respectively while for MS/MS they were set at $2x10^5$ and 100ms respectively.

Nanospray Ionization (NSI)

All NSI samples were transferred into HPLC sample vials and placed in the autosampler of the Shimadzu Prominence HPLC. Injections were made into the NanoLC interfaced with the LTQ Orbitrap mass spectrometer via either an Advance CaptiveSpray Plug-and-Play source or the ThermoFisher NSI Source.

Analyses of in-gel trypsin digests of gel spots were optimized for NSI conditions of 400nL/min. Samples were loaded onto a 0.3 x 5 mm^2 PepTrap 300Å C-18 precolumn in 5% solvent B (solvent A: 0.1% formic acid/2.5%ACN vs. solvent B: 0.1% formic acid/2.5% HPLC-grade water/97.5%ACN) for 10 minutes for desalting. Peptides were then eluted into an Agilent (5µ 100Å 15cm x 0.075mm ID) C-18 analytical column and separated with a linear gradient of solvent B (97.5% ACN/2.5%H_2O/0.1% formic acid) over 35 minutes. The HPLC gradient was controlled by the Thermo Fisher Scientific XCalibur® 2.0.7 data system. Precursor ions were scanned between *m/*z 400-2000. Survey scans were acquired in the Orbitrap with resolving power of 60,000 at m/z 400 and an AGC target level of $5x10^5$ and a maximum injection time into the Orbitrap of 500ms. The 5 most abundant ions were selected for CID in the ion trap. Precursor ions were isolated with a 2Da window and fragmented by low collision energy with He gas for 30ms with normalized collision energy of 35 (arbitrary units). NSI parameters used for the trypsin digested samples were: capillary temperature, 300°C; spray voltage, 1.8kV; capillary voltage, 35 arbitrary units. Selected ions repeated 4 times over 45s were excluded for the subsequent 90s.

Automated peptide analysis of acid digested ribosomal proteins was carried out by online NSI-MS/MS on the Shimadzu NanoLC interfaced with the LTQ-Orbitrap via the Advance CaptiveSpray Plug-and-Play source running at a column flow rate of 500nL/min. Sample injections were loaded onto a 0.3 x 5mm^2 Peptrap 300Å C-18 precolumn for 15min at 5% solvent B (0.1% formic acid/2.5% H_2O/97.5%ACN) for desalting. Peptides were eluted into a 150 x 0.1mm analytical column (Grace Vydac, Deerfield, IL) and separated with a linear gradient of solvent B; 5-15% in 5 minutes, 15-70%B in 115min. Survey scans were acquired in the Orbitrap with resolving power of 30,000 at *m*/*z* 400 and an AGC target level of $5x10^5$. The three most abundant ions were selected for fragmentation using CID in the linear ion trap. Precursor ions were isolated using a 3Da window and fragmented by low collision energy with He gas for 30ms with normalized collision energy of 35 (arbitrary units). The product ion AGC target level was set to $5x10^4$ and fragment ion scans were acquired in the Orbitrap with resolving power of 15,000 at *m*/*z* 400. NSI parameters used for the acid digested samples were: capillary temperature, 200°C; spray voltage, 1.6kV; sheath gas flow rate, 2 arbitrary units; capillary voltage, 32 arbitrary units. Dynamic exclusion parameters were set to exclude ions previously selected for fragmentation for 3 min.

Intact proteins extracted from gels spots were analyzed under NSI conditions of 800nL/min using the Shimadzu Prominence HPLC interfaced with the LTQ-Orbitrap via the Advance CaptiveSpray Plug-and-Play source. Samples were loaded onto a 0.3 x 2.5 mm^2 TARGA Piccolo 5µ C-18 precolumn (Higgins Analytical, Sunnyvale, CA) in 5% solvent B for 10 minutes for desalting. Proteins were then

eluted into a PLRP-S capillary column (150mm x 0.1mmID 5µ particle size, 1000Å pore size) and separated with a linear gradient of solvent B (solvent A: 0.1% formic acid/2.5%ACN vs. solvent B: 0.1% formic acid/2.5% HPLC-grade water/97.5%ACN) over 40 minutes. The HPLC gradient was controlled by the Thermo Fisher Scientific XCalibur® 2.0.7 data system. Ions were scanned between *m/*z 400-2000. Survey scans (4 microscans/survey scan) were acquired in the Orbitrap with resolving power of 60,000 at *m/z* 400. The 3 most abundant ions were selected for CID in the linear ion trap and fragment ion scans were acquired for each precursor ion in the Orbitrap with a resolving power of 30,000 at *m/z* 400. Precursor ions were isolated using a 5Da window and fragmented by low energy collision with He gas for 60ms with normalized collision energy of 35 (arbitrary units). NSI parameters used for the intact protein samples were: capillary temperature, 200°C; spray voltage, 1.6kV; sheath gas flow rate, 2 arbitrary units; capillary voltage, 32 arbitrary units. Reduced detection delay was used to improve protein detection. An in-source voltage of 6V was used to knock off salt adducts or impurities from the protein samples. Selected ions that were identified 3 times over 45s were excluded for the next 45s.

Bioinformatics

Analyses of .RAW data files using Mascot© (Matrix Science Ltd., London UK) searches required files to be converted first to .mgf files. This was accomplished in one of two ways; Thermo Proteome-Discoverer™ (ThermoFisher, San Jose CA) or via MM File Conversion Tools from MassMatrix.

Spectra obtained from the trypsin tandem mass spectral analysis of the ribosomal peptides were searched in collaboration with Professor Nathan Edwards at Georgetown University, on his network of computers using PepArML, a meta-search engine capable of machine learning which provides a single user interface to seven commonly used search engines. This system allows for confident peptide assignment as well as the detection of peptides that would otherwise be missed when using only one search engine, as it combines the results of searches across Mascot, Tandem, Tandem with K-score and S-Score scoring plugins, OMSSA, Myrimatch and InsPecT [103].

Mascot searches alone were also used for the bottom-up analyses of trypsin digests. Searches were carried out specifying trypsin. Up to 5 missed cleavages were allowed with precursor tolerance of 25 ppm and product ion tolerance of 0.6 Da. ESI-TRAP was selected for fragment specificity. Variable modifications were selected to include N-terminal acetylation, N,Q deamidation, M oxidation and S,T,Y phosphorylation.

Peptides generated from acid digestion were analyzed as previously described [2]. Spectral files were processed using ProSightPC 2.0 provided by Professor Neil Kelleher, University of Illinois, which is now commercially available from ThermoFisher. Each .RAW file was processed in High Throughput mode. Spectra were decharged with cRAWler using the THRASH algorithm. A FASTA format protein sequence database of 79 human ribosomal proteins was extracted from an in-house Ribosomal Protein Gene Database and the IPI human database and configured for acid-cleavage analysis with ProSightPC 2.0. Spectra were searched in Absolute

Mass mode using a 2.5 Da precursor window based on the peptide monoisotopic mass. An additional search, using a loose precursor window of 250 Da, was carried out to look for evidence of post-translationally modified peptide isoforms. ProSightPC's Sequence Gazer tool allowed for manual investigations of significant identifications, when sufficient *b* and/or *y* ions were matched despite discrepancies between the predicted precursor mass and the observed mass. The sequence positions of the matched *b* and/or *y* ions helped to localize the mass-shift from putative PTM's and single amino acid substitutions[104]. Mass tolerance for fragment ions was set at 15 ppm. False discovery rates (FDR) were calculated using a randomly shuffled version of the ribosomal protein sequence database previously described [64].

Mascot searches were also used to analyze the acid digest data. Searches were carried out specifying "no enzyme". Up to 9 missed cleavages were allowed for acid digests with precursor tolerance of 10 ppm and product ion tolerance of 0.05 Da. ESI-TRAP was selected for fragment specificity. Variable modifications were selected to include N-terminal acetylation, N,Q deamidation, M and H oxidation and S,T,Y phosphorylation.

Molecular ions and fragment data generated from top-down analyses of the ribosomal proteome were analyzed using similar methods as those used for acid polypeptides. ProSightPC 2.0 was used to process spectral files. Each .RAW file was processed in High Throughput mode. Spectra were decharged with cRAWler using the THRASH algorithm. A FASTA format protein sequence database of 79 human ribosomal proteins was extracted from an in-house Ribosomal Protein Gene Database and the IPI human database. Spectra were searched in Absolute Mass mode

using a 250 Da precursor window based on the protein monoisotopic mass to include the possibility of PTMs. ProSightPC's Sequence Gazer tool allowed for manual investigations of significant identifications, when sufficient *b* and/or *y* ions were matched despite discrepancies between the predicted precursor mass and the observed mass. The sequence positions of the matched *b* and/or *y* ions helped to localize the mass-shift from putative PTM's and single amino acid substitutions. Mass tolerance for fragment ions was set at 15 ppm.

This information was coordinated with molecular masses determined from whole protein analyses to identify and locate modifications. With analyses of the spectra produced from repeated injections, protein identification and high coverage was accomplished. To provide additional verification of the protein isoform, the chemical formula of the protein (along with any proposed modifications) was used to create a theoretical isotope envelope with Thermo Qual Browser (part of the Thermo Fisher Scientific XCalibur® 2.0.7 Software package) and compared with the highly resolved deconvoluted (with Thermo Xtract®) isotope envelope of the observed protein isoform. These efforts allowed us to characterize modifications of these human ribosomal proteins, especially the novel isoforms.

Chapter 3: Results

This research project successfully characterized the modifications in altered ribosomal proteins of drug resistant and drug susceptible precursor MCF7 breast cancer cell lines. As will be discussed, the altered proteins characterized by this study play integral roles in the ribosome in protein synthesis. We suggest based on our results and the research of other scientists that the plasticity observed in our cancer ribosomes/ribosomal proteins is possibly a natural mechanism in eukaryotic cell development and survival that is utilized by our ribosomes to survive exposure to chemotherapeutic drugs by altering their structure during translation.

Ribosome Isolation

Improvements in the protein yield observed from ribosome isolation and protein extraction from the whole ribosome was accomplished through modifications of protocols previously established in our laboratory [1]. Modifications to our original protocol occurred in the final centrifugation step where a swinging bucket rotor was observed to improve the quality and size of the ribosomal pellet when compared with that of the fixed angle rotor previously used (See Table 3.1).

For each harvest of the cell lines, flasks were seeded at the same time and their growth monitored. Harvests for MXR and MXS which were compared 1:1 with each other (either by gel array or HPLC) and were completed as close together as possible (in 3 of 4 cases they were harvested one day apart). Protein concentrations for both cell lines in a set were determined using a Bio-Rad RC/DC protein assay. Chicken egg lysozyme was substituted as the standard protein in the protein assay for

the commonly used bovine serum albumin (BSA). This allowed for the standard protein in the assay to more closely mimic the amino acid composition of the ribosomal proteins. In general, a higher number of cells were harvested from the MXS cell line with each harvest when compared with the MXR cell line (see Fig 3.1). To compensate for this discrepancy, a larger number of flasks of the resistant cell line were seeded and harvested to ensure enough material from each was collected.

	Average Cell Pellet Weight (g)	Average Protein Concentration (mg/mL)
Fixed Angle Rotor	1.406	0.46
Swinging Bucket Rotor	1.543	1.52

Table 3.1 Average cell pellet weight (wet weight in g) and MXR protein concentration before and after change in protocol illustrates that after the protocol change, protein yield improved

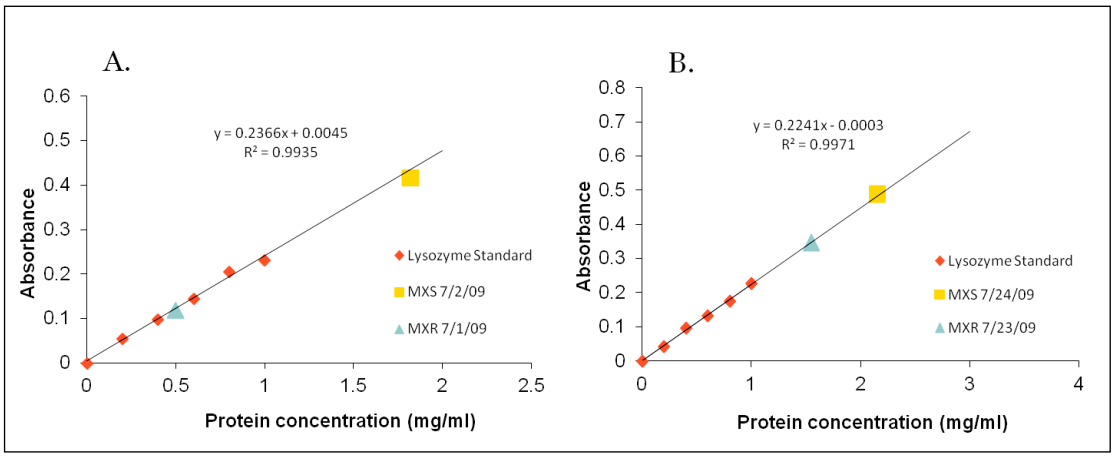

Figure 3.1 Protein concentrations of MXR and MXS as determined by Bio Rad RC/DC protein Assay.
The cell pellet weight (shown in Table 3.1) and protein yield in rehydration buffer (for gel arrays) of the MXR cell line was initially low (as seen in A) and compensated for by an increase in the number of flasks seeded with MXR per harvest (B) and the use of a swinging bucket rotor for ribosomal pellet isolation

HPLC fractionation of ribosomal proteins

Consulting previous studies that involved the HPLC fractionation of ribosomal proteins [72; 73; 91; 105; 106; 107; 108] it was determined most prudent to test both C-4 and C-18 columns for their efficacy of resolving the MCF7 ribosomal proteins. In addition to different resins, acetonitrile gradients were tested with the standard protein mixture and yeast ribosomal proteins for their effectiveness. An assessment of time of the gradient run versus the cost/benefit of protein resolution was considered. Twelve gradients in total were tested before the following gradient was decided upon; (1) 0 – 83 min; 15 – 95% ACN, (2) 83 – 93 min; 95% ACN, (3) 93 – 94 min; 95 – 15% ACN, (4) 94 – 113 min; 15% ACN. Approximately 10.5nmol of ribosomal proteins were injected manually with a flow rate of 0.55mL/min.

Fractions were manually collected every minute beginning at 5 minutes until 60 minutes. At 60 minutes, fractions were subsequently collected every 5 minutes until 80 minutes at which point one fraction was collected every ten minutes. Typical UV chromatograms for both standard proteins and ribosomal proteins are seen in Figures 3.2 and 3.3 respectively. Each fraction was aliquoted for bottom-up analysis and molecular mass analysis.

Inspection of the whole proteins in each fraction using the MALDI-TOF revealed which fractions contained proteins. Fractions containing proteins were digested in solution with trypsin to identify protein components and molecular mass measurements were made with ESI-LTQ-Orbitrap on the whole protein fraction aliquots to determine the masses of individual proteins.

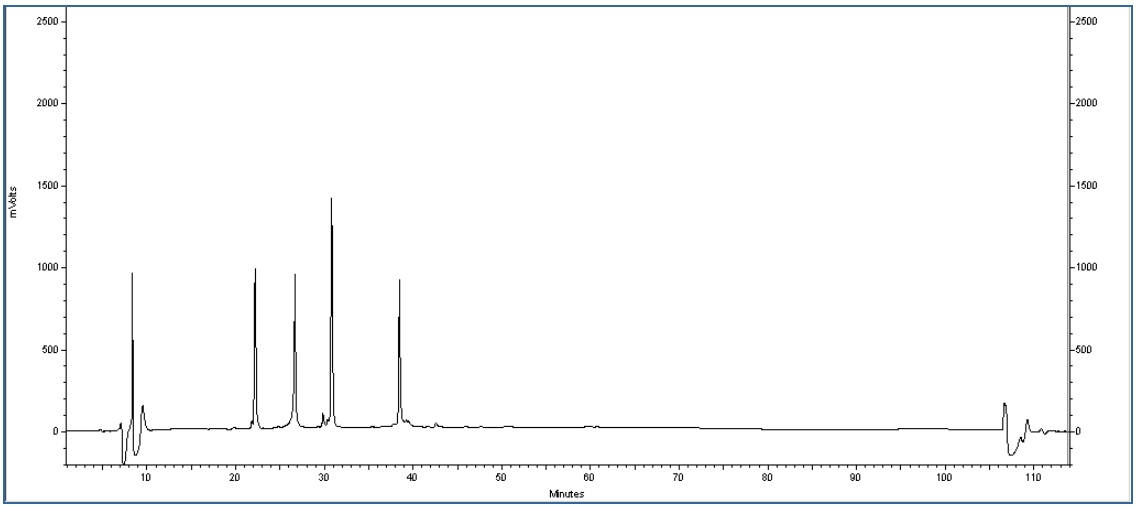

Figure 3.2 UV chromatogram of the mixture of four standard proteins
The mixture consisted of lysozyme, ribonuclease A, cytochrome C and myoglobin eluting in that order out of the C-18 column.

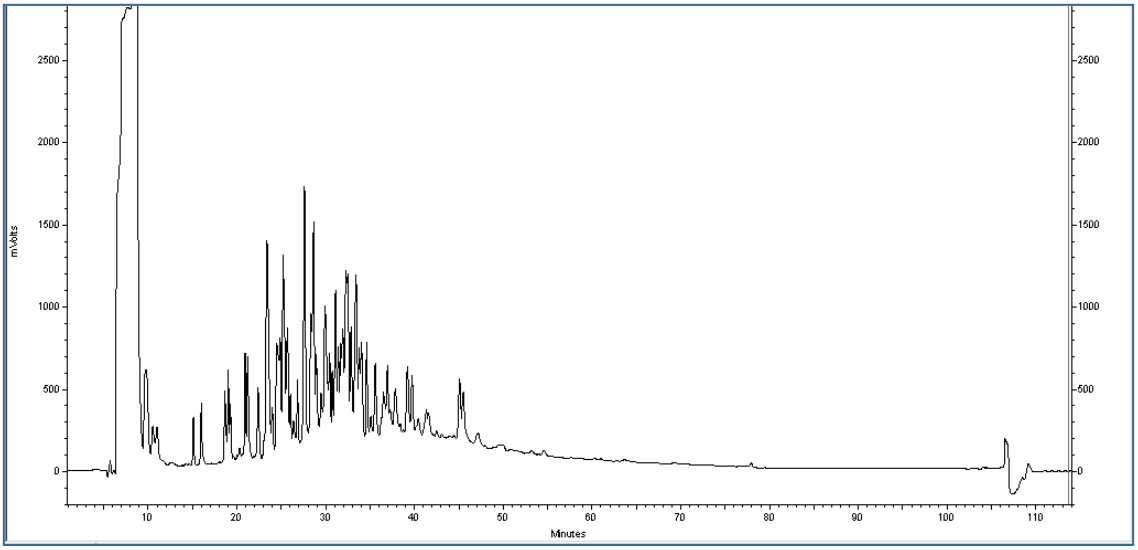

Figure 3.3 UV chromatogram of the MXR ribosomal protein mixture

Detection of Proteins in HPLC fractions

HPLC fractions were first evaluated with the Shimadzu-Axima CFR + MALDI-TOF instrument to ensure the presence of protein. Yeast ribosomal proteins were first assessed to determine the sample application methods and instrument parameters best suited to fractions from different regions of the acetonitrile gradient. The m/z range initially examined by MALDI was 7k – 70k and then later changed to 10k - 70k to avoid signal interference from low molecular weight contaminants. As was to be expected, samples that eluted late in the gradient and which thus had a high acetonitrile concentration produced better spectra for many proteins when the MALDI matrix was mixed with 5% Triton-X 100 (Fig 3.4 – Fig 3.5). When these methods were applied to the human ribosomal protein HPLC fractions, of the approximately 65 fractions analyzed, 55 of these had protein(s) at detectable levels.

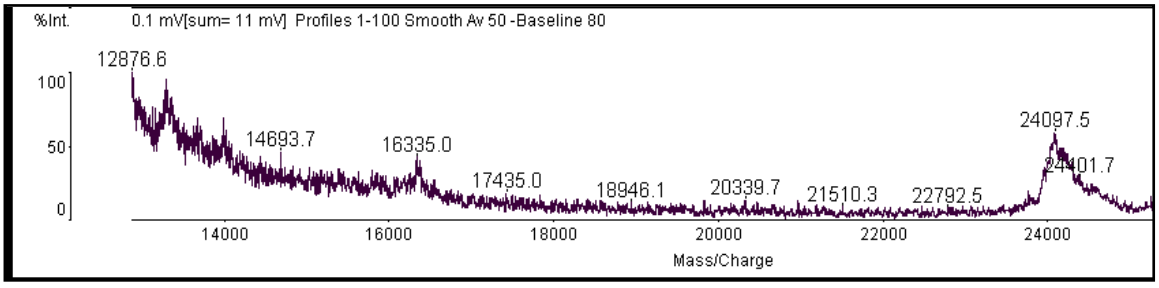

Fig 3.4 MALDI spectra without detergent
MALDI spectra of MXR ribosomal protein 36 minute HPLC fraction spotted on MALDI plate with MALDI matrix (10mg/mL sinapinic acid in 50% acetonitrile/1% trifluoroacetic acid) without 5% Triton-X 100.

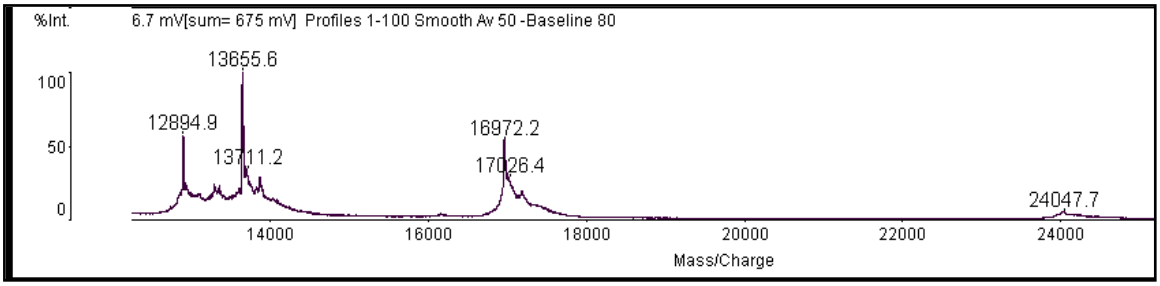

Fig 3.5 MALDI spectra with detergent
MALDI spectra of MXR ribosomal protein 36 minute HPLC fraction spotted on MALDI plate with MALDI matrix (10mg/mL sinapinic acid in 50% acetonitrile/1% trifluoroacetic acid) with 5% Triton-X 100.

ESI-MS analysis of HPLC fractions

Trypsin digestion

Peptides generated from an in-solution digestion of the individual HPLC fractions were fractionated and analyzed using the LC-LTQ-Orbitrap. The 5 most abundant ions were selected from each precursor scan for tandem MS analysis and MS/MS spectra were searched against the IPI human database as described in the Experimental section. Initial runs showed extended elution profiles for certain peptide hits. Alterations in the sample injection volume, gradient and dynamic exclusion parameters were used to minimize this along with exclusion lists.

Ribosomal proteins identified in these HPLC fraction digests from the database search were compared with their theoretical isoelectric point (http://expasy.org/tools/) and mapped over time in the protein gradient. Frequently, multiple proteins were identified by their peptides (a minimum of 2) in a fraction.

In the first LC-MS/MS analysis of the ribosomal protein HPLC fractions, peptides from 71 of the 79 ribosomal proteins were observed in the 36 fractions (not observed; RPS27, RPL26, RPL32, RPL37, RPL37A, RPL39, RPL41) however the sequence coverage for each protein was low, frequently with only 2 peptides confidently identified for an individual protein (data not shown). The second LC-MS/MS analysis utilized a slightly different gradient (shorter; changed from 95min to 48min) and a total of 404 (346 non-overlapping) ribosomal peptides were confidently identified in the 56 fractions, corresponding to 57 ribosomal proteins. An example of a total ion chromatogram from one of the HPLC fractions is shown in Figure 3.6. Sequence coverage of these proteins averaged 36.2% and ranged from 14% to 95% (See Table 3.2). In order to improve sequence coverage and discovery of post translational modifications, additional digestion methods were investigated.

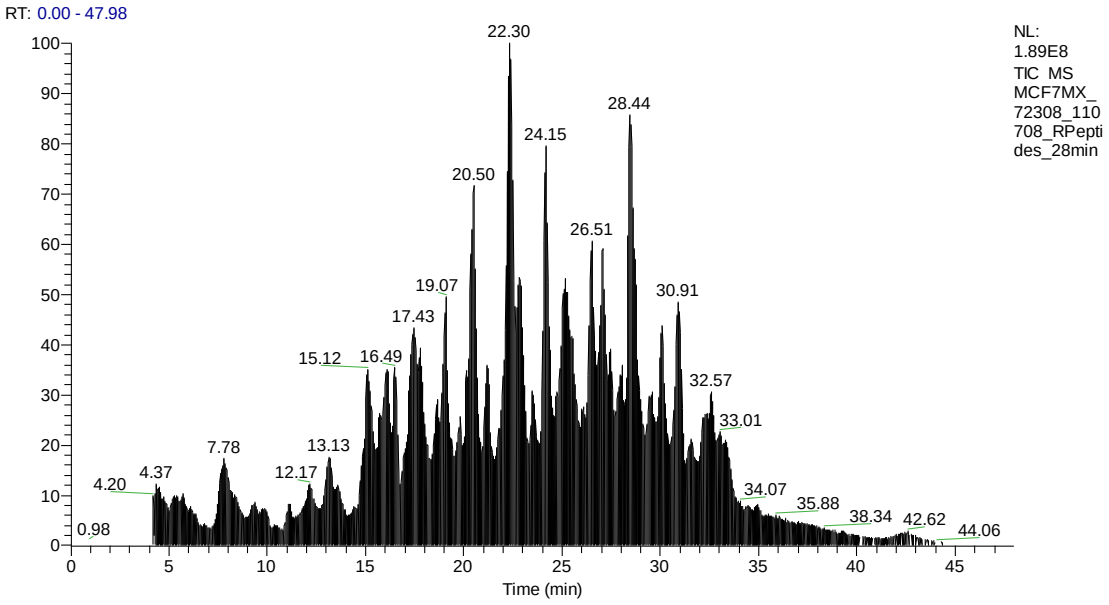

Figure 3.6 The Total Ion Chromatogram from the 28 minute HPLC fraction that confidently identified a protein of interest, RPL23A

Protein	Total # unique peptides	% sequence coverage
RPLP0	**7**	**42.35%**
RPLP1	**4(3 non-overlaps)**	**56.14%**
RPLP2	**8 (5 non-overlaps)**	**94.78%**
RPL3	**3**	**16.50%**
RPL4	**15 (12 non-overlaps)**	**37.24%**
RPL5	**4**	**17.17%**
RPL6	**16 (14 non-overlaps)**	**55.21%**
RPL7	**13 (10 non-overlaps)**	**43.15%**
RPL7A	**12 (9 non-overlaps)**	**37.59%**
RPL8	**8 (7 non-overlaps)**	**31.52%**
RPL9	**4**	**22.92%**
RPL10	**5**	**16.52%**
RPL10A	**11 (8 non-overlaps)**	**40.09%**
RPL11	**5 (4 non-overlaps)**	**23.16%**
RPL12	**7 (6 non-overlaps)**	**58.79%**
RPL13	**11 (10 non-overlaps)**	**40.28%**
RPL13A	**12 (11 non-overlaps)**	**50.25%**
RPL14	**6 (5 non-overlaps)**	**37.90%**
RPL18	**7 (6 non-overlaps)**	**31.91%**
RPL18A	**6**	**29.55%**
RPL19	**4**	**20.41%**
RPL23	**2**	**24.56%**
RPL23A	**9 (7 non-overlaps)**	**48.73%**
RPL26	**8 (7 non-overlaps)**	**34.64%**
RPL27	**6**	**34.56%**
RPL27A	**4**	**31.08%**
RPL28	**6**	**32.12%**
RPL30	**4**	**33.91%**
RPL32	**2**	**14.81%**
RPL35	**6 (5 non-overlaps)**	**43.90%**
RPL36A	**2**	**15.09%**
RPL37A	**2**	**29.35%**
RPL38	**4 (3 non-overlaps)**	**52.86%**
RPS2	**10 (9 non-overlaps)**	**33.79%**
RPS3	**12 (8 non-overlaps)**	**46.09%**
RPS3A	**10 (8 non-overlaps)**	**32.95%**
RPS4X	**7 (6 non-overlaps)**	**19.39%**
RPS5	**8 (6 non-overlaps)**	**40.20%**
RPS6	**9 (7 non-overlaps)**	**26.91%**
RPS7	**12 (7 non-overlaps)**	**62.37%**
RPS8	**3 (2 non-overlaps)**	**13.94%**
RPS9	**15**	**64.95%**

RPS10	**6 (5 non-overlaps)**	**40.61%**
RPS11	**8**	**45.57%**
RPS12	**4**	**34.09%**
RPS13	**12 (10 non-overlaps)**	**68.87%**
RPS16	**8 (7 non-overlaps)**	**48.63%**
RPS18	**11 (9 non-overlaps)**	**51.97%**
RPS19	**9 (7 non-overlaps)**	**47.59%**
RPS20	**3**	**28.57%**
RPS23	**4**	**24.48%**
RPS25	**10**	**66.40%**
RPS28	**2**	**33.33%**
RPSA	**9**	**41.67%**

Table 3.2 Ribosomal proteins identified by 2 or more peptides in a *single* LC-MS/MS analysis of the HPLC fractions based on PepArML

Whole Ribosomal Proteome

Acid Digestion

An alternative approach to digesting the ribosome and improving sequence coverage was utilized by the microwave accelerated acid digestion of the whole human ribosomal proteome. Both ESI and NSI (NSI-Orbitrap analysis will be discussed later) were utilized in the automated peptide analysis of acid digested ribosomal proteins as described in the Experimental section. For ESI-Orbitrap analysis, 5 injections of acid digested MXR ribosomal peptides were combined to confidently identify 217 distinct peptides corresponding to 63 of the 79 ribosomal proteins. Seventy one peptides were identified as having masses above 3kDa, with charges as high as 11+. An example of one such peptide, in this case from RPSA, part of the RPS2 family (known as RPS0 in yeast) is found in Fig 3.7 and Fig 3.8.

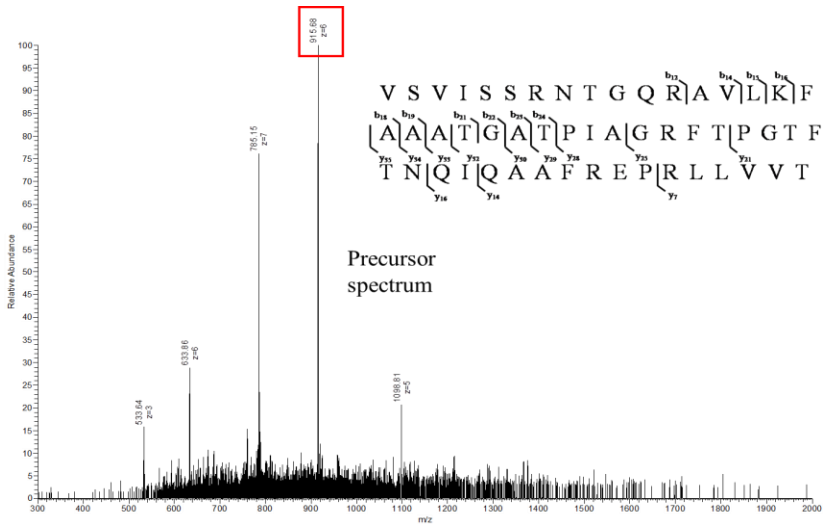

Figure 3.7 Precursor spectrum of the 5604Da peptide confidently identified by ProSightPC 2.0 with an E-value of 6.93E-24.

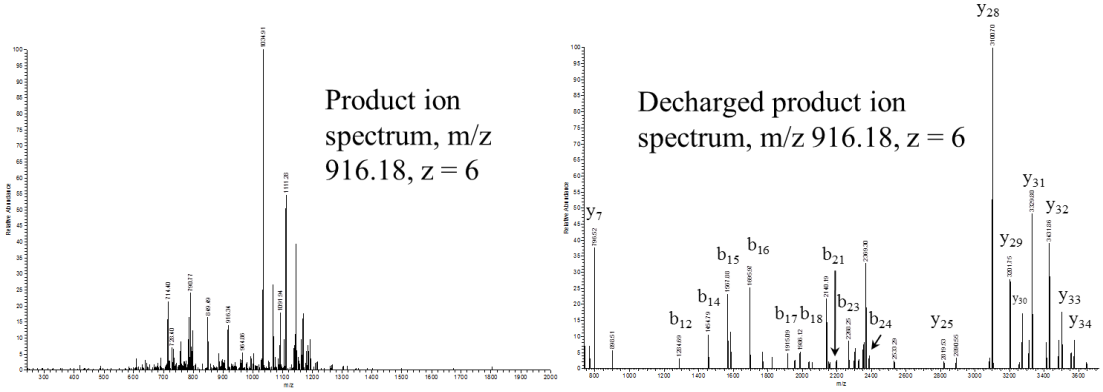

Figure 3.8 Product ion spectrum and decharged product ion spectrum of the precursor ion shown in Figure 3.7

Intact Mass Measurements of Human Ribosomal Proteins

Molecular ion measurements of intact ribosomal proteins were obtained using both the whole ribosomal proteome mixture as well as collected HPLC fractions of

ribosomal proteins as described in the Experimental section. Molecular mass measurements of the most abundant protein components in the HPLC fractions were matched with proteins confidently identified by in-solution digestion of the corresponding fraction (as determined by both PepArML and/or Scaffold). Abundant molecular ions which did not match proteins identified in the in-solution digestion (with 2 or more peptides) were compared with those proteins identified in the fraction with one confidently identified peptide; if the molecular ion exactly *matched* the theoretical mass of the protein, it was accepted as identification. For example; RPS29 was identified in the 23 minute fraction by one peptide and the theoretical molecular mass of RPS29 is 6544.27 which is nearly identical to the observed mass of 6544.28.

This technique allowed for the identification of 37 proteins not including isoforms of individual proteins. Protein modifications were observed that have been confirmed by the investigations conducted in this lab, by the bottom-up data from the LC fractions and/or the in-gel digestions (see Table 3.3), or by other published research on human ribosomal proteins. For example, a study of these MXR proteins conducted previously in this lab[1] made the observation of a 47Da ± 5Da increase of RPL11 found in spot #25. The bottom up research conducted in the current investigation of these ribosomes has revealed that this mass change can be attributed to an N-terminal acetylation. Another protein, RPL31 was found in the previous study to lose several hundred Daltons from its theoretical mass. This same protein was observed with the same loss of mass in an HPLC fraction of this investigation with 48.8% sequence coverage and 11 unique peptides at the 99.9% cut off criteria for Scaffold. Neither the in-gel digestion from the previous investigation in this

laboratory[1] or the current in solution digestion data were able to confidently detect the N-terminal end of this protein. This suggests the possibility of an N-terminal truncation. If the first 3 amino acid residues on the N-terminal end were truncated, a protein that would originally be expected to weigh 14459.95Da is now theoretically expected to weigh 14160.82Da. The deconvoluted experimental mass observed in the LC fraction was 14162.78Da. The gel extracted mass observed previously in this lab by MALDI-MS[1] was 14183 ± 5 Da (Refer to Table 3.3). The difference between the MALDI-MS and the experimental mass determined by LC-ESI-Orbitrap is evidence of a sodium adduct (+21), a common artifact observed with MALDI-MS.

Protein isoforms were also noted in the HPLC fractions. For example, RPS25 was observed in two different fractions, once with its N-terminal methionine and once without. Most of the protein isoforms of interest were identified in the HPLC fractions as well as the gel extracted protein. For example, one of the RPS10 isoforms was identified via the trypsin digestion in the fraction collected at 31 minutes (See Fig 3.9, 3.10, and 3.11) and the molecular mass observed corresponded with the RPS10 isoform found in spot 29.

{MATRIX SCIENCE} Mascot Search Results

```
User           : Karen Lohnes
Email          : klohnes@umd.edu
Search title   : 110708_MCFMX31min
MS data file   : MCF7MX_72308_110708_RPeptides_31min.mgf
Database       : IPI_human human_20080530 (71884 sequences; 30064464 residues)
Timestamp      : 20 Nov 2008 at 19:28:29 GMT
Protein hits   : IPI00008438 Tax_Id=9606 Gene_Symbol=RPS10 40S ribosomal protein S10
                 IPI00329389 Tax_Id=9606 Gene_Symbol=RPL6 60S ribosomal protein L6
                 IPI00024933 Tax_Id=9606 Gene_Symbol=RPL12 60S ribosomal protein L12
                 IPI00216587 Tax_Id=9606 Gene_Symbol=RPS8 40S ribosomal protein S8
                 IPI00816063 Tax_Id=9606 Gene_Symbol=RPL12 RPL12 protein
                 IPI00472119 Tax_Id=9606 Gene_Symbol=- Uncharacterized protein ENSP00000343748
                 IPI00477279 Tax_Id=9606 Gene_Symbol=LOC646875 similar to ribosomal protein L12
                 IPI00025512 Tax_Id=9606 Gene_Symbol=HSPB1 Heat shock protein beta-1
                 IPI00221089 Tax_Id=9606 Gene_Symbol=RPS13 40S ribosomal protein S13
                 IPI00186712 Tax_Id=9606 Gene_Symbol=RPS26 Novel protein similar to ribosomal protein 26 RPS26
                 IPI00013415 Tax_Id=9606 Gene_Symbol=RPS7 40S ribosomal protein S7
                 IPI00003918 Tax_Id=9606 Gene_Symbol=RPL4 60S ribosomal protein L4
```

Figure 3.9 Mascot search results for the fraction collected at 31 minutes when searched against the IPI human database. RPS10 was found to be the most abundant protein in the fraction (see Fig 3.10 and 3.11)

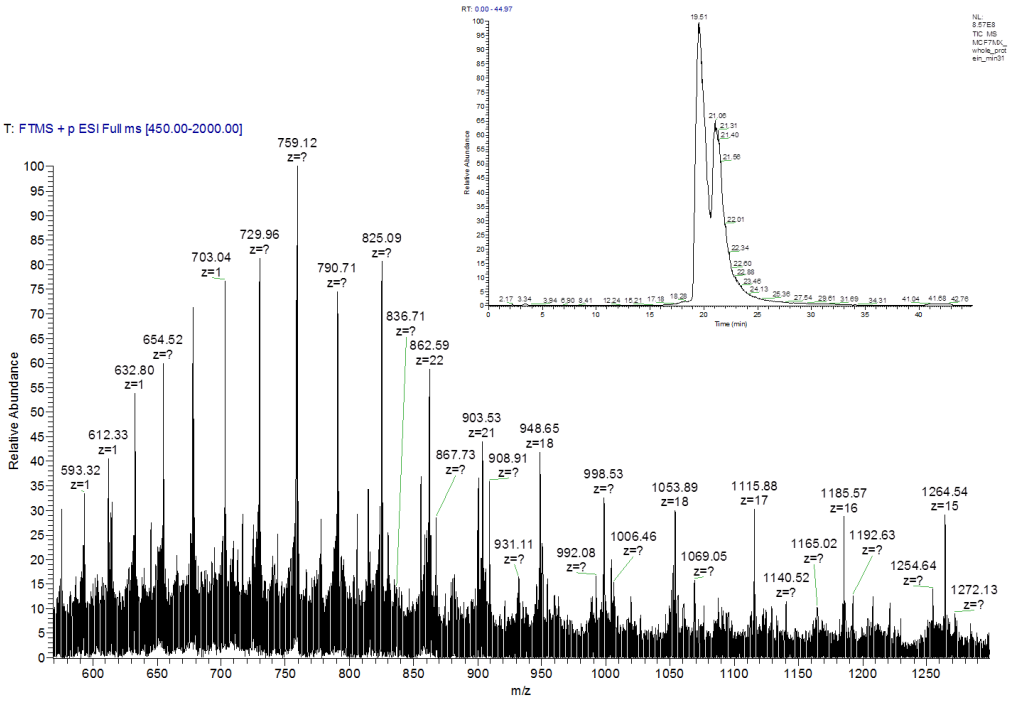

Figure 3.10 RPS10 containing fraction TIC (inset) and ESI-Orbitrap MS spectrum of the most abundant peak in the RPLC fraction

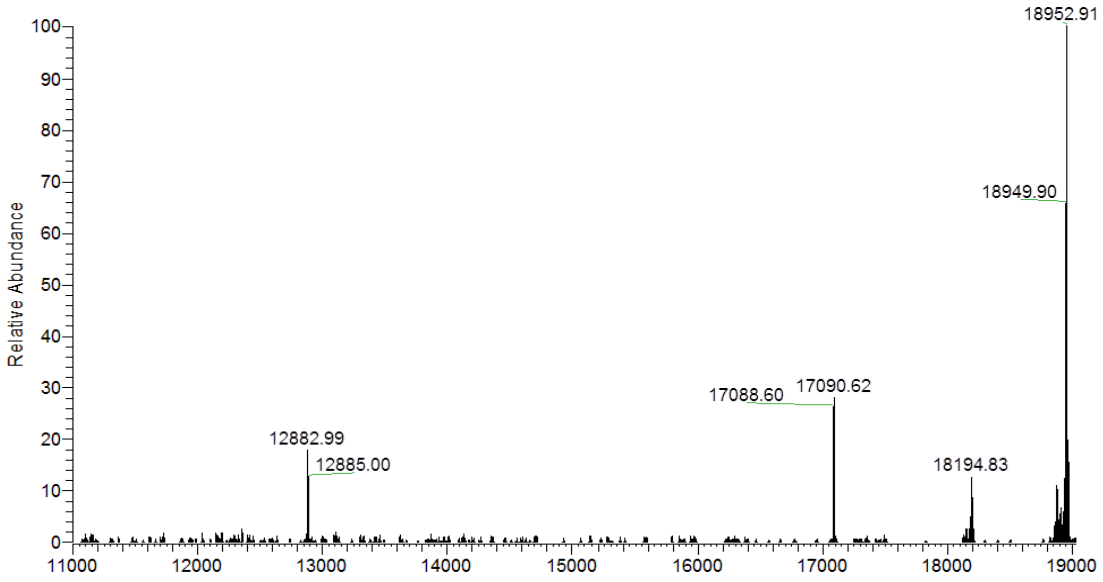

Figure 3.11 Deconvoluted mass spectrum of the HPLC fraction collected at 31 minutes. Notice the abundant mass shown at 18952.91 which is one of the RPS10 isoforms found in the gel extracted experiment (spot 29; Refer to Figs 3.23, 3.30 and 3.54)

Fraction(s) observed (min)	Protein match	Theoretical Mass	Observed Mass	% Sequence coverage
18	RPL36A/RPL36a mass +oxidation (Met loss)	12306.79/ 12322.78	12305.72/ 12321.75	52.83
20	RPL26	17255.56	17256.53	65.52
21	RPL24	17775.92	17777.85	59.24
22	RPL27 (Met loss)	15663.73	15665.67	43.38
22, 23	RPL37a (Met loss)	10141.45	10142.41	38.04
22	RPS29 (Met loss)	6544.27	6542.23	35.71
22, 23	RPL38 (Met loss)	8084.7	8085.69	60
23	RPL28 (Met loss + acetyl)	15655.66	15657.59	53.28
24	RPL35A (Met loss + acetyl)	12445.76	12447.71	45.45
24	RPSII (Met loss + acetyl)	18337.99	18339.94	56.33
24	RPL31 (loss of 1st 3 N-terminal AA)	14160.82	14162.78	48.8
25	RPS24 (+ acetyl)	15461.46	15464.33	30
25	RPS23 (+ oxidation)	15689.69	15690.55	55.94
26	RPL35 (Met loss)	14417.54	14419.40	36.59

26	RPL27A (Met loss + oxidation)	16443.02	16444.87	56
26	RPL18A	20758.93	20760.72	48.86
26,27	RPS19 (Met loss)	15926.52	15927.46	51.03
27	RPL17 (Met loss)	21263.32	21265.24	45.11
27	RPL23 (Met loss + acetyl)	14773.05	14775.01	58.57
27	RPL22 (Met loss)	14653.79	14655.75	42.19
27	RPL10 (Met loss)	24468.86	24470.82	28.26
28	RPL23A (Met loss + acetyl)	17603.13	17604.10	46.84
28	RPL23A (Met loss + acetyl + phosphorylation +2 oxidation)	17717.1	17719.10	46.84
28	RPS25 (Met loss + *formyl*)	13636.67	13638.67	38.4
28	RPS27 (Met loss + acetyl)	9369.82	9370.83	40.48
28	RPL30 (Met loss)	12649.72	12651.69	59.13
29	RPS16 (Met loss)	16311.02	16314.00	47.26
29	RPL11 (Met loss + acetyl)	20162.61	20163.60	64.97
29	RPS15 (Met loss + acetyl)	16948.17	16949.08	18.62
30	RPL32 (Met loss)	15725.79	15727.72	52.29
30	RPL18 (Met loss)	21501.06	21500.95	44.68
31	RPS10 (2 dimethyl)	18952.97	18952.91	40
31	RPS13	17089.64	17090.62	45.7
33	RPL13A (Met loss + acetyl)	23485.68	23486.30	38.92
34	RPS12 (Met loss + Acetyl)	14422.47	14423.45	65.91
35	RPL9	21860.83	21862.80	60.42
35	RPS25 (2 oxidations)	13771.71	13774.61	40
36	RPSA (Met loss + Acetyl)	32760.45	32764.43	77.39
37	RPS3 (Met loss + acetyl)	26598.46	26598.42	69.14
37	RPS3 (Met loss + acetyl + phospho+oxidation)	26694.44	26698.46	69.14
37	RPS3 (Met loss)	26556.44	26.00	69.14

Table 3.3 Proteins observed in HPLC fractions and their corresponding sequence coverage (99.9% peptide confidence Scaffold)

Fragmentation of Intact Ribosomal Proteins

The examination of the whole ribosomal proteome also included additional analyses of high resolution mass measurements with fragmentation products of the proteins. As described previously, ProSightPC 2.0 was used to decharge the spectra of the precursor and fragment ions and to search the MS/MS spectra. A custom database of the 79 human ribosomal proteins was used for this search however the data was also searched against the IPI human protein database to ensure confidence in the ribosomal protein identifications. One analysis of each cell line was conducted in this fashion and from these two sample runs, a total of 18 proteins were identified excluding isoforms of the individual proteins. Variations of the observed proteins included loss of N-terminal methionine, acetylation and oxidation. For example, RPS23 was observed with and without oxidation (with oxidation in the MXR cell line). It should be noted that the analysis of rat fibroblast small subunit ribosomal proteins in the Ahn lab also indicated two forms of RPS23 which differed in mass by 15.9 mass units (proposed as containing either an internal methylation or hydroxylation; oxidation considered less likely since corresponding sequence lacked methionine). A table of the observed proteins is found in Table 3.4. An example of one of the proteins identified with an N-terminal Met loss and an acetylation (N-acetyl-L-ala), RPS11, is shown in Figures 3.12, 3.13, 3.14, 3.15 and 3.16.

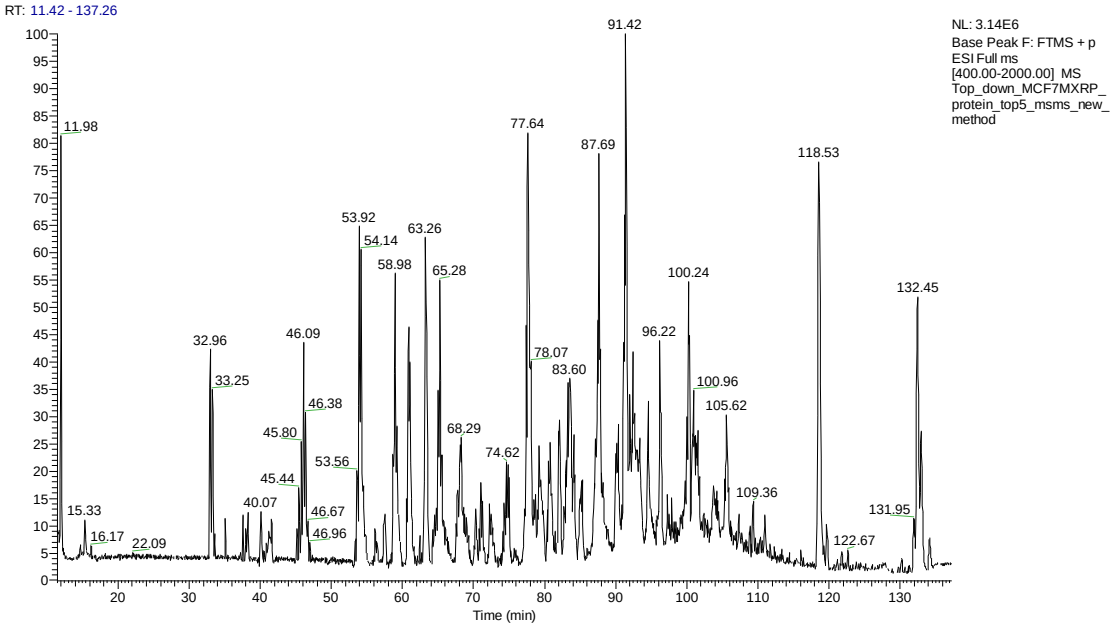

Figure 3.12 Base peak chromatogram of the top-down analysis of the human MXR ribosomal proteome. The RPS11 protein was observed in the peak seen at 65.28 minutes

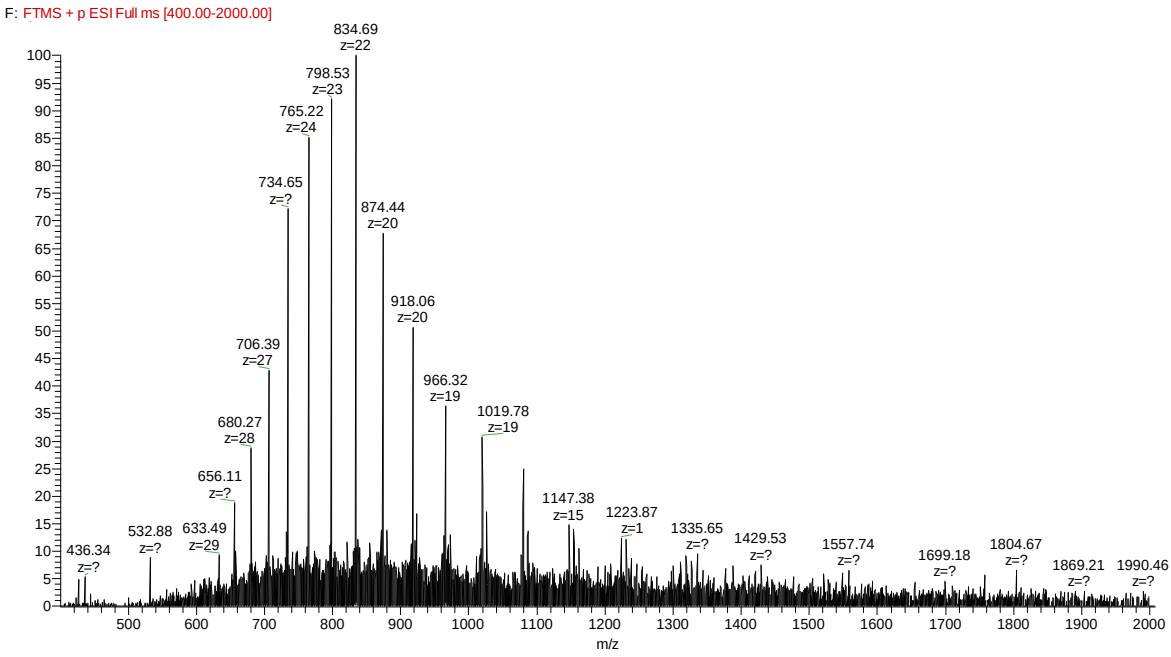

Figure 3.13 Precursor spectrum of the confidently identified ribosomal protein RPS11. The most abundant precursor ion was selected for fragmentation (see Figure 3.14)

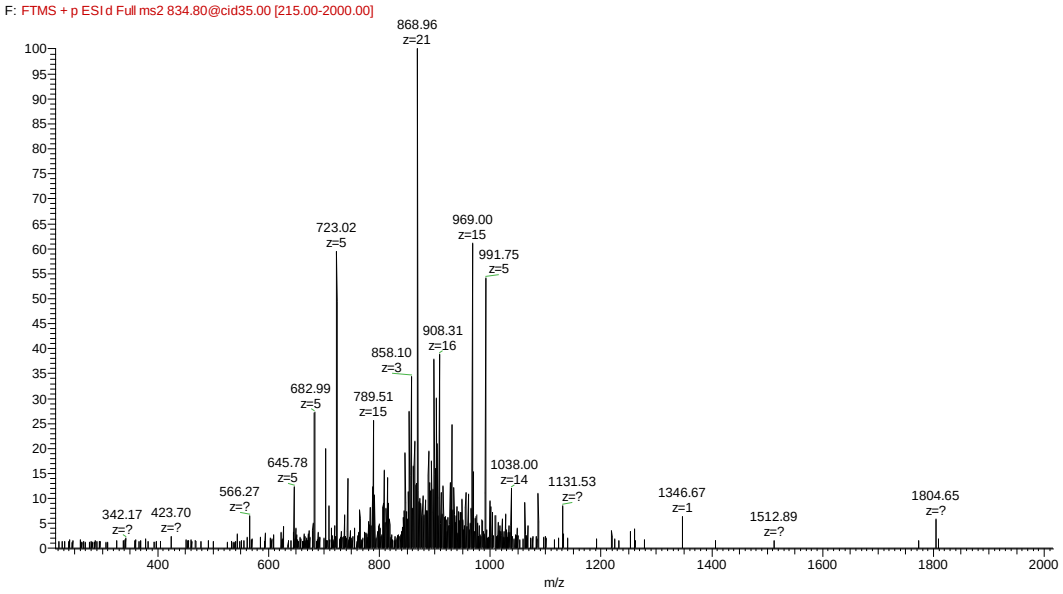

Figure 13.14 MS/MS spectrum of the precursor ion at *m/z* 834.80 (22+ charge state, intact mass = 18341.01 Da)

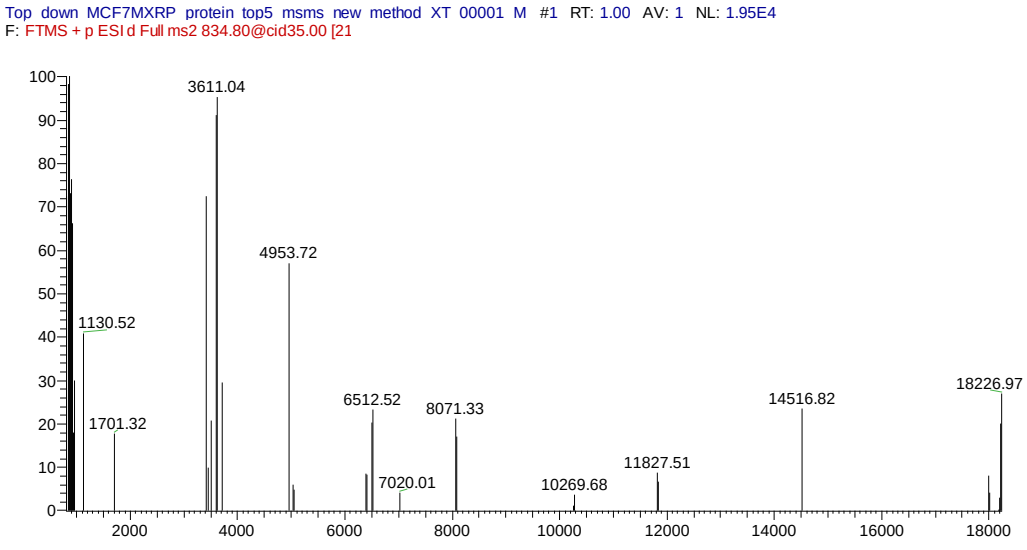

Figure 13.15 Decharged MS/MS spectrum of Fig 3.14. ProSightPC 2.0 confidently assigned 15 fragment ions in this spectrum to RPS11 with E-value = 8.36E-10

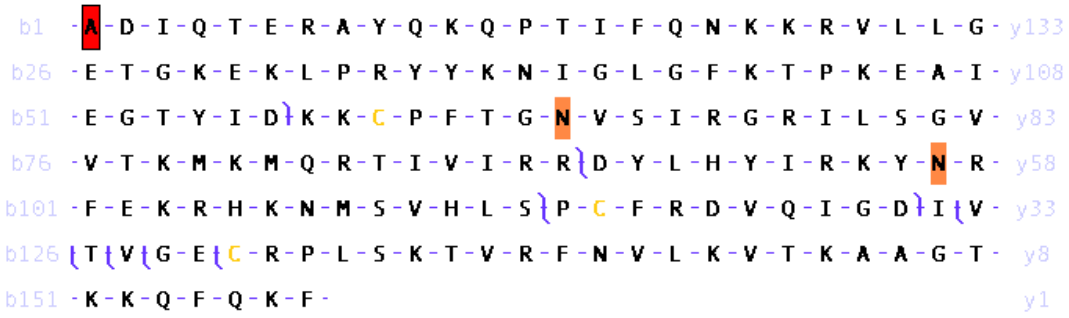

Figure 13.16 Protein sequence of RPS11 showing the 15 fragmentation sites assigned by ProSightPC 2.0. The highlighted red alanine residue indicates the acetylation

Ribosomal Protein	E-Value	Modification	Theoretical Mass	Observed Mass
RPS11	8.36E-10	Met loss, N-acetyl-L-Ala	18337.99	18341.01
RPS15A	1.30E-14	Met loss	14707.95	14708.00
RPS16	4.00E-05	Met loss	16311.02	16313.04
RPS19	7.60E-09	Met loss	15926.52	15928.48
RPS20	3.00E-04	Met loss, N-acetyl-L-Ala	13281.30	13281.28
RPS21	1.00E-04	N-acetyl-L-Met	9151.59	9154.59
RPS23	4.90E-05	Met loss, *oxidation*	15689.69	15691.65
RPS23	1.00E-04	Met loss	15673.70	15675.68
RPS24	4.80E-07	*Formylation*	15449.46	15447.45
RPS24	0.0009	N-acetyl-L-Met	15462.47	15464.43
RPS27	7.60E-05	Met loss, *Acetylation*	9369.82	9371.80
RPS28	1.10E-23	N-acetyl-L-Met	7882.22	7882.25
RPL27	2.70E-12	Met loss	15663.73	15665.79
RPL28	3.20E-05	Met loss, N-acetyl-L-Ser	15655.66	15656.66
RPL30	3.50E-11	Met loss	12649.72	12652.69
RPL32	1.10E-07	Met loss	15725.79	15728.72
RPL32	2.00E-05	Met loss, -14.12Da	15725.79	15711.67

RPL35A	3.50E-08	Met loss, N-acetyl-L-Ser	12445.76	12447.68
RPL35A	1.70E-08	Met loss, N-acetyl-L-Ser, + 18.08Da	12445.76	12463.83
RPL37A	1.10E-05	Met loss	10141.45	10140.44
RPL38	3.50E-21	Met loss	8084.70	8086.71
RPLP2	1.80E-07	+162.06Da	11662.86	11824.92

Table 3.4 Proteins observed with top-down analysis of both cell lines (1 sample injection per cell line)

nLC-Orbitrap Analysis of ribosomal proteins

Acid digested Ribosomal Proteins

As mentioned previously, automated peptide analyses of the acid digestion products of the ribosomal proteome were made using both ESI-MS and NSI-MS. The nLC-Orbitrap analysis included 4 injections of the MXR peptides and 3 injections of the MXS peptides. Data was processed using ProSightPC 2.0 and peptide identification accepted with an E-value of less than 1.0E-03. There were similar retention times for shared peptides between the two cell lines however the MXS cell line contained approximately half as many matched spectra as the MXR cell line. The data was searched against a human ribosomal protein database as well as the IPI human database. As there were no discernible differences in the non-ribosomal proteins discovered between the two cell lines, the difference in matched spectra is attributed to a probable error in determination of protein concentration in the MXS cell line.

Ribosomal protein identifications overlapped between the cell lines with the distinguishing characteristics of MXR and MXS ribosomal proteins based on peptides

unique to that cell line. Across the seven injections, 366 distinct peptides were confidently identified including redundancies (where identified peptides differed only in the addition or removal of an N- or C-terminal Asp). After removing redundant peptides, 276 peptide identifications remained. Sixty five of the 79 human ribosomal proteins were identified by two or more peptides and 70 by one confidently identified peptide (See Appendix for a list of peptides). Average sequence coverage was 46.2% for these 70 proteins. Forty four percent of these peptides were identified as having masses above 3kDa, with charges as high as 12+. An example of one such peptide, in this case from ribosomal protein RPL24, is found in Fig 3.17, 3.18 and 3.19. This peptide was confidently identified by ProSightPC 2.0 with 15 fragments and an E-value = 1.29E-21.

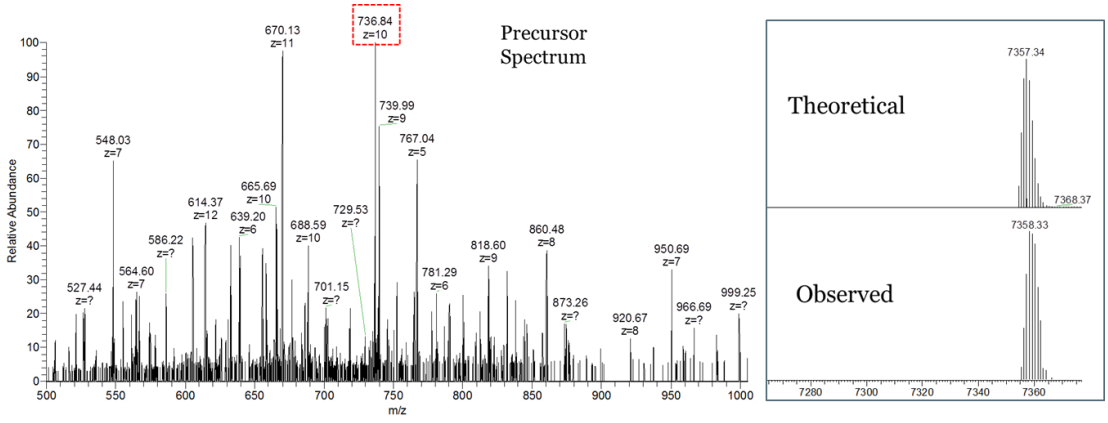

Figure 3.17 Precursor spectrum of RPL24 peptide observed during nLC-Orbitrap analysis of MXR ribosomal protein acid digest and theoretical vs. observed decharged mass of that peptide

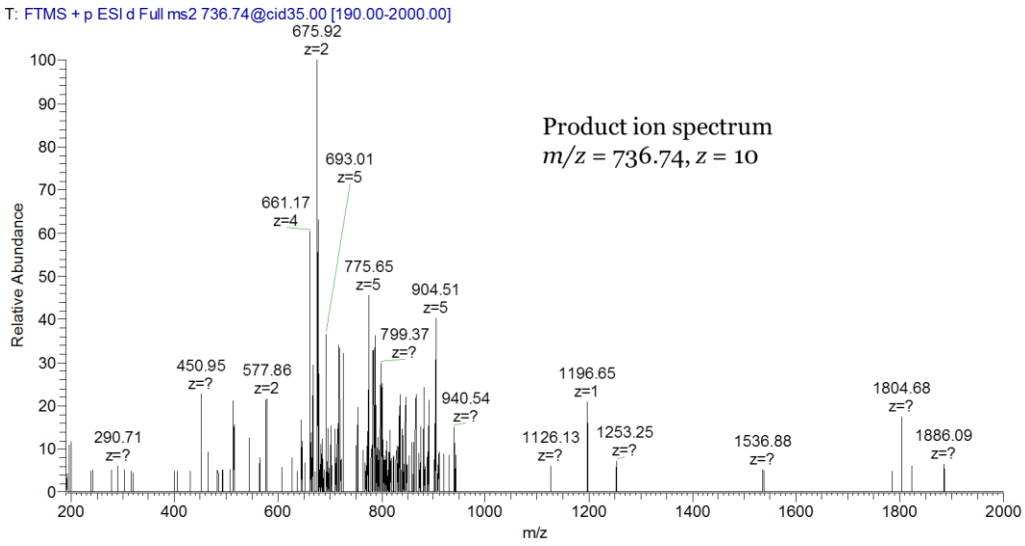

Figure 3.18 Product ion spectrum of precursor ion shown in Fig 3.17. The precursor ion chosen for this peptide was m/z = 736.74 and as a result, the mass difference (Da) between theoretical and observed in this case is - 0.01

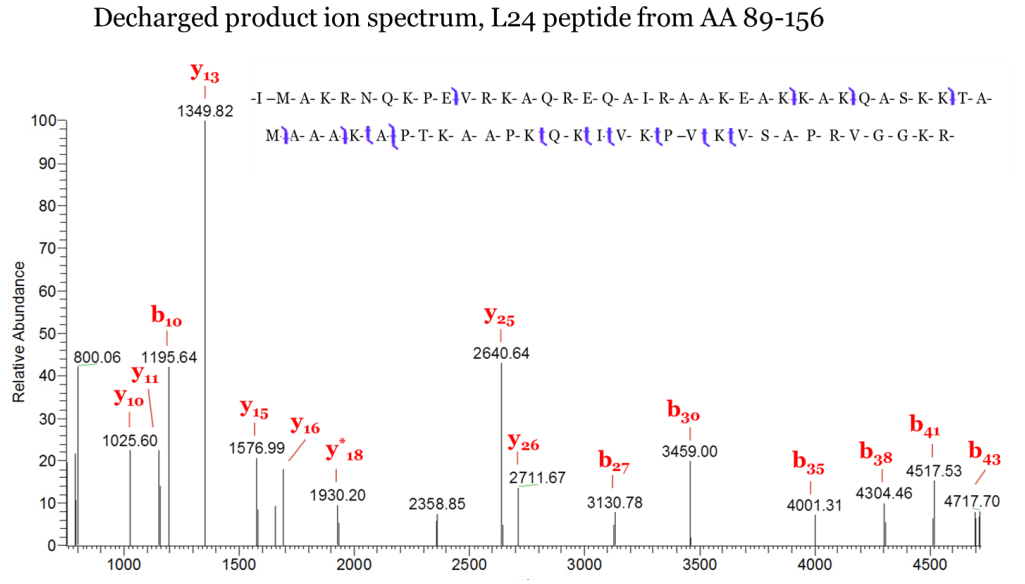

Figure 3.19 Deconvoluted product ion spectrum of RPL24 peptide from AA 89-156. This peptide was confidently identified by ProSightPC 2.0 with 15 fragments and assigned an E-value = 1.29E-21

Gel Extracted Proteins

Proteins were extracted from duplicate gel arrays and combined for LC-Orbitrap analysis as described in the Experimental section. HPLC columns packed with three different resins were tested for mass measurement/top-down analysis of the gel extracted proteins; C-18, C-4 and PLRP-S. In addition to type of column, two sources and instrument parameters were tested. The C-18 and C-4 columns were both manufactured by Agilent (0.1 x 150mm 5µ, 300Å pore size). These columns were fitted to the Thermo NSI source with the dynamic NSI probe and an emitter purchased from New Objective, Inc. The instrument was externally calibrated using the standard Thermo mixture (caffeine, MRFA, Ultramark 1600) and tuned with cytochrome C prior to sample analysis.

The most effective combination which provided the data discussed henceforth used the Advance CaptiveSpray Plug-and-Play source. Samples were loaded onto a 0.3 x 2.5 mm^2 TARGA Piccolo 5µ C-18 precolumn at 5% solvent B for 10 minutes for desalting. Proteins were then eluted into a PLRP-S capillary column. Not only was it found that this source/column and trap combination were an effective configuration but the specifications of the flow-rate and back-pressure for the loading of the sample onto the trap proved critical. This was achieved by having the fused silica that loaded the trap (for desalting) of a larger inner diameter (0.025µm) than the fused silica on the other side of the trap (0.015µm). In doing this, another advantage was that a small change in solvent concentration when the valve switched out of the trapping position (an intermittent increase in solvent B) would have less impact on the trapped proteins.

A limitation of the Advance CaptiveSpray Plug-and-Play source is the fact that the capillary specific to that source is not to be used above 200˚C. However with the ability to use the sheath gas (sheath gas flow rate set at 2.00 arbitrary units) and the tube lens voltage increased (175V), the capillary temperature was not an issue. The verification of this workflow came with the observation of identical masses of proteins seen in both the gel extracted proteins and the HPLC fraction molecular mass measurement of the same proteins. An example of one such case is seen by comparing the spectra in Figure 3.20.

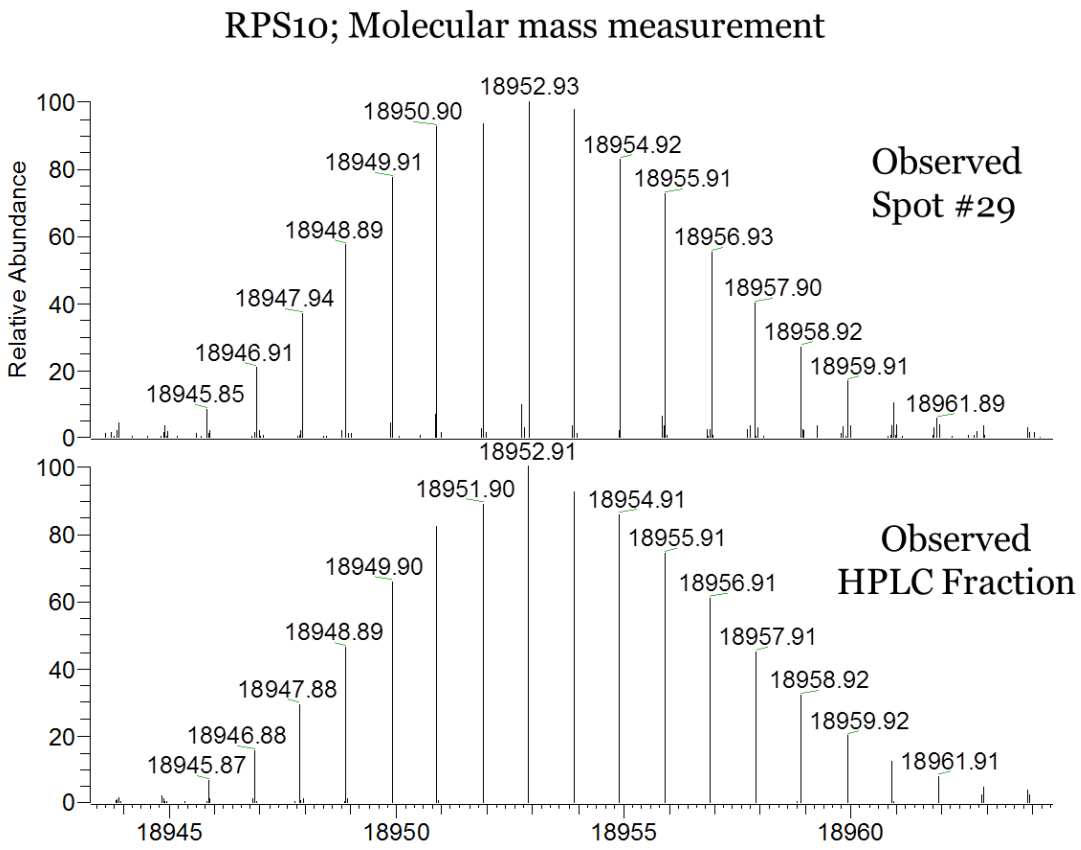

Figure 3.20 Deconvoluted spectrum of the protein peak observed from the gel extracted protein identified as RPS10 (top) compared with the deconvoluted spectrum of this protein identified in the HPLC fraction (bottom)

Two-Dimensional Gel Electrophoresis

Sample Loading Methods

Four different methods/conditions were tested for the loading of samples in the first dimension of the gel arrays. All methods except the test of rehydration with Destreak™ used an electrode wick dipped in 15mM DTT at the cathode end to provide a continuous influx of DTT which is a weak acid and as a result, migrates out of the basic part of the IEF gel. The first method had previously been used in this lab for ribosomal protein visualization with gel array which involved in-gel rehydration of the sample (both passive and active loading) in standard rehydration buffer (7M urea, 2M thiourea, 2% CHAPS, 0.5% IPG buffer, 50mM DTT). Both active and passive rehydration of the sample led to inconsistencies in the sample loading and streaking in the final gel image as seen in Fig 3.21. The rehydration buffer was altered in subsequent gels to contain 15% IPA and 2.5% glycerol which further optimized the gel appearance but still resulted in some sample streaking (refer to Fig 3.21). A more traditional approach to working with basic proteins in IEF was then tested with cup loading of the sample at the cathode end of the IPG strip. This proved to be the most effective method of sample loading with reproducible patterns seen in the gel arrays of the same cell line as seen in Fig 3.22. All gels used in this comparative analysis were produced using cup-loading at the anode end with DTT, 15% IPA and 2.5% glycerol in the rehydration buffer.

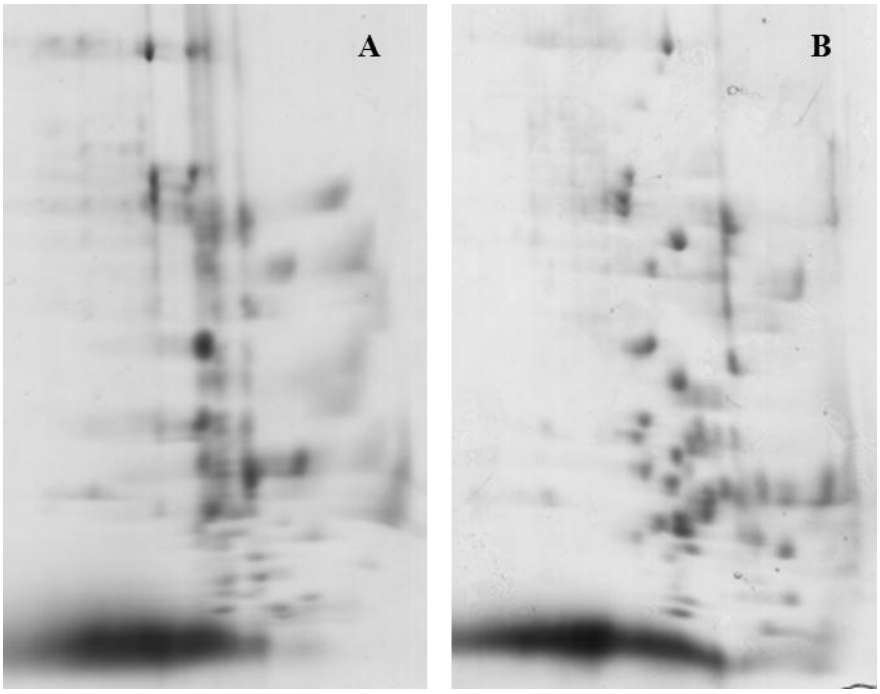

Figure 3.21 Image A shows an example of a gel produced with rehydration loading of the sample in standard rehydration buffer which led to streaking (both horizontal and vertical streaking were observed though vertical streaking is attributed to the second dimension). Image B shows an example of a gel produced with rehydration loading of the sample with modified rehydration buffer containing 15% IPA and 2.5% glycerol

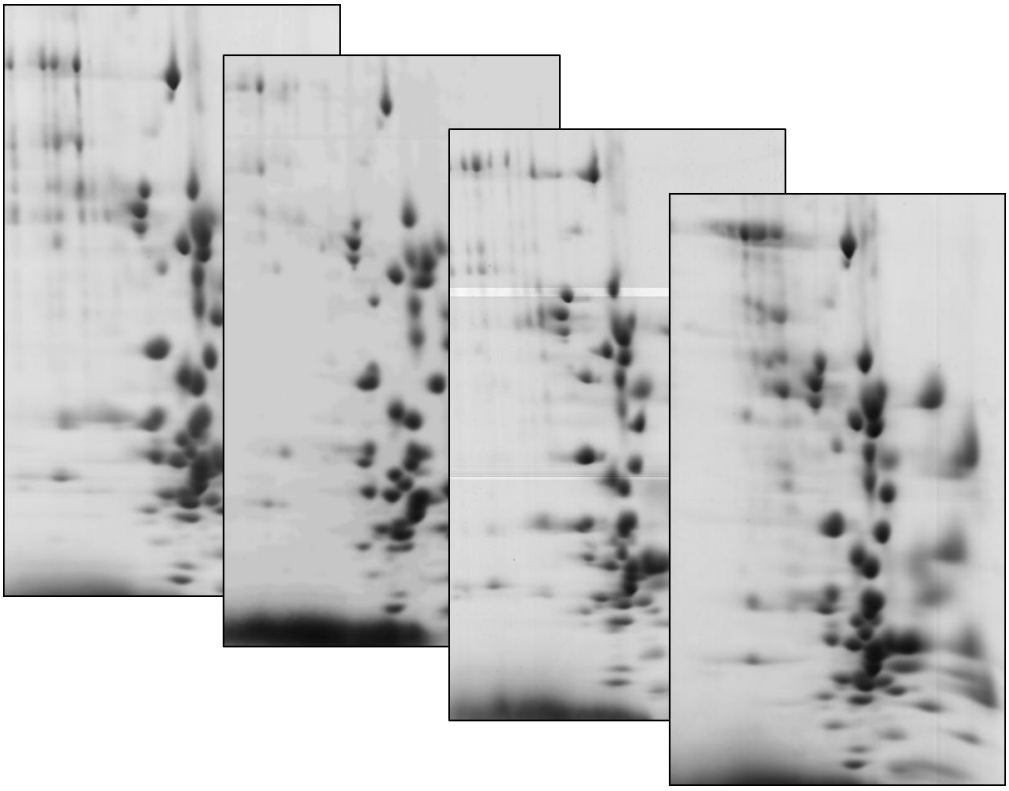

Figure 3.22 Reproducibility of spot patterns in gel arrays of the same cell line (MXR in this case) when samples were loaded using cup-loading at the anode end of the IPG strip in rehydration buffer that also contained 15% IPA and 2.5% glycerol

In-Gel Digestion and Protein Sequence Coverage

Gel arrays of MXR ribosomes with proteins spots that had been annotated and identified previously in this lab[1] were used as a reference for comparison to aid in the identification of protein spots (See Figure 3.23). In-gel digestion was used to confirm the identities of the protein spots observed in the gels and to determine the modifications of the proteins in the spots of interest. Spots chosen for in-gel digestion were not only those found in differential abundance between the two cell

lines (as determined by imaging software discussed in the next section) but also reference spots to ensure that the protein spots identified were homologous between gels.

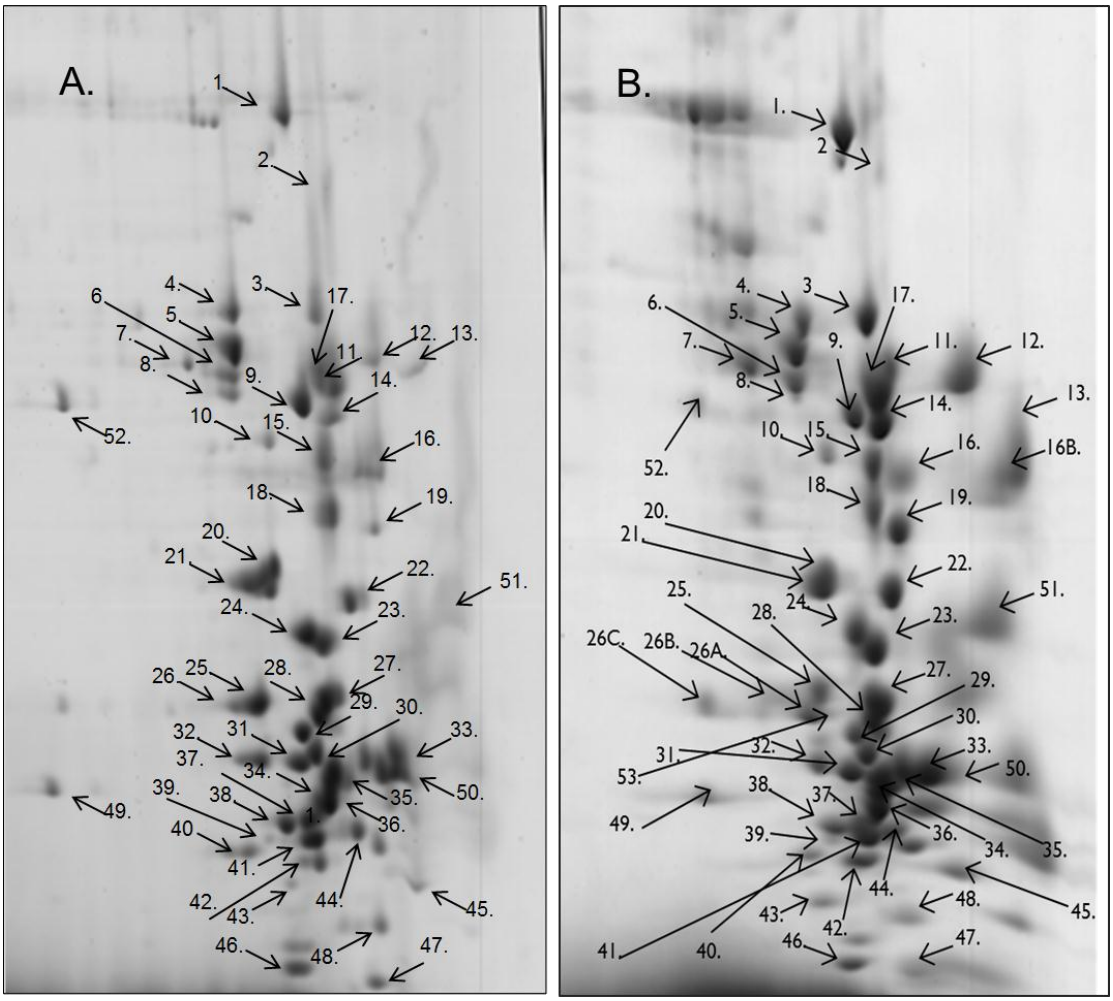

Figure 3.23 Annotated MXR gel of previous research[1] compared with an annotated MXR gel from the current study.
Gel (A.) on the left shows *previously published work* from this lab[1] with spot annotation while gel (B.) on the right shows the current work using the same spot annotation

Twenty three spots including spots found in differential abundance or unique to a given cell line were chosen for in-gel digestion. One of these spots contained

peptides from more than one protein (spot 20; refer to Fig 3.23). Identifications of proteins in nearby spots were considered to rule out the possibility of overlap from a neighboring spot. When a neighboring spot could not assist in restricting the protein composition of a spot of interest, molecular masses of the extracted gel proteins were used in determining which protein was most abundant in the selected spot.

Protein sequence coverage was ultimately determined by a combination of the information from the in-gel digestion, intact molecular mass of extracted proteins, and acid digestion data. Modifications of the proteins were assigned exclusively by confidently identified peptides from the in-gel digestion of protein spots except for the non-acetylated form of the N-terminal end of RPS3 which was identified in the acid digestion data and confirmed by the intact molecular mass. The twenty two spots have a combined digestion (acid and trypsin) sequence coverage which ranged from 50% to 99% shown in Table 3.5. Peptides that were confidently identified by trypsin digestion only are shown in red, those confidently identified by acid digestion only are shown in blue and those confidently identified by both acid and trypsin digestion are identified by red with black outline text. For the nine spots of interest (including one unchanging spot; spot #7) the sequence coverage ranged from 65% to 98%. The sequence coverage of these proteins specifically is summarized in Table 3.6.

#	Protein	Sequence Coverage	% Sequence Coverage
3	RPL6	MAGEKVEKPD TKEKKPEAKK VDAGGKVKKG NLKAKKPKKG KPHCSRNPVL VRGIGRYSRS AMYSRKAMYK RKYSAAKSKV EKKKKEKVLA TVTKPVGGDK NGGTRVVKLR KMPRYYPTED VPRKLLSHGK KPFSQHVRKL RASITPGTIL IILTGRHRGK RVVFLKQLAS GLLLVTGPLV LNRVPLRRTH QKFVIATSTK IDISNVKIPK HLTDAYFKKK KLRKPRHQEG EIFDTEKEKY EITEQRKIDQ KAVDSQILPK IKAIPQLQGY LRSVFALTNG IYPHKLVF	83
5	RPS3A	MAVGKNKRLT KGGKKGAKKK VVDPFSKKDW YDVKAPAMFN IRNIGKTLVT RTQGTKIASD GLKGRVFEVS LADLQNDEVA FRKFKLITED VQGKNCLTNF HGMDLTRDKM CSMVKKWQTM IEAHVDVKTT DGYLLRLFCV GFTKKRNNQI RKTSYAQHQQ VRQIRKKMME IMTREVQTND LKEVVNKLIP DSIGKDIEKA CQSIYPLHDV FVRKVKMLKK PKFELGKLME LHGEGSSSGK ATGDETGAKV ERADGYEPPV QESV	74
6	RPS3	MAVQISKKRK FVADGIFKAE LNEFLTRELA EDGYSGVEVR VTPTRTEIII LATRTQNVLG EKGRRIRELT AVVQKRFGFP EGSVELYAEK VATRGLCAIA QAESLRYKLL GGLAVRRACY GVLRFIMESG AKGCEVVVSG KLRGQRAKSM KFVDGLMIHS GDPVNYYVDT AVRHVLLRQG VLGIKVKIML PWDPTGKIGP KKPLPDHVSI VEPKDEILPT TPISEQKGGK PEPPAMPQPV PTA	98

7	RPS3	MAVQISKKRK FVADGIFKAE LNEFLTRELA EDGYSGVEVR VTPTRTEIII LATRTQNVLG EKGRRIRELT AVVQKRFGFP EGSVELYAEK VATRGLCAIA QAESLRYKLL GGLAVRRACY GVLRFIMESG AKGCEVVVSG KLRGQRAKSM KFVDGLMIHS GDPVNYYVDT AVRHVLLRQG VLGIKVKIML PWDPTGKIGP KKPLPDHVSI VEPKDEILPT TPISEQKGGK PEPPAMPQPV PTA	85
8	RPS3	MAVQISKKRK FVADGIFKAE LNEFLTRELA EDGYSGVEVR VTPTRTEIII LATRTQNVLG EKGRRIRELT AVVQKRFGFP EGSVELYAEK VATRGLCAIA QAESLRYKLL GGLAVRRACY GVLRFIMESG AKGCEVVVSG KLRGQRAKSM KFVDGLMIHS GDPVNYYVDT AVRHVLLRQG VLGIKVKIML PWDPTGKIGP KKPLPDHVSI VEPKDEILPT TPISEQKGGK PEPPAMPQPV PTA	89
13	RPL8	MGRVIRGQRK GAGSVFRAHV KHRKGAARLR AVDFAERHGY IKGIVKDIIH DPGRGAPLAK VVFRDPYRFK KRTELFIAAE GIHTGQFVYC GKKAQLNIGN VLPVGTMPEG TIVCCLEEKP GDRGKLARAS GNYATVISHN PETKKTRVKL PSGSKKVISS ANRAVVGVVA GGGRIDKPIL KAGRAYHKYK AKRNCWPRVR GVAMNPVEHP FGGGNHQHIG KPSTIRRDAP AGRKVGLIAA RRTGRLRGTK TVQEKEN	84
15	RPS8	MGISRDNWHK RRKTGGKRKP YHKKRKYELG RPAANTKIGP RRIHTVRVRG GNKKYRALRL DVGNFSWGSE CCTRKTRIID VVYNASNNEL VRTKTLVKNC IVLIDSTPYR QWYESHYALP LGRKKGAKLT PEEEEILNKK RSKKIQKKYD ERKKNAKISS LLEEQFQQGK LLACIASRPG QCGRADGYVL EGKELEFYLR KIKARKGK	72

16B	RPL13	MAPSRNGMVL KPHFHKDWQR RVATWFNQPA RKIRRRKARQ AKARRIAPRP ASGPIRPIVR CPTVRYHTKV RAGRGFSLEE LRVAGIHKKV ARTIGISVDP RRRNKSTESL QANVQRLKEY RSKLILFPRK PSAPKKGDSS AEELKLATQL TGPVMPVRNV YKKEKARVIT EEEKNFKAFA SLRMARANAR LFGIRAKRAK EAAEQDVEKK K	77
18	RPL10	MGRRPARCYR YCKNKPYPKS RFCRGVPDAK IRIFDLGRKK AKVDEFPLCG HMVSDEYEQL SSEALEAARI CANKYMVKSC GKDGFHIRVR LHPFHVIRIN KMLSCAGADR LQTGMRGAFG KPQGTVARVH IGQVIMSIRT KLQNKEHVIE ALRRAKFKFP GRQKIHISKK WGFTKFNADE FEDMVAEKRL IPDGCGVKYI PNRGPLDKWR ALHS	72
19	RPL13A	MAEVQVLVLD GRGHLLGRLA AIVAKQVLLG RKVVVVRCEG INISGNFYRN KLKYLAFLRK RMNTNPSRGP YHFRAPSRIF WRTVRGMLPH KTKRGQAALD RLKVFDGIPP PYDKKKRMVV PAALKVVRLK PTRKFAYLGR LAHEVGWKYQ AVTATLEEKR KEKAKIHYRK KKQLMRLRKQ AEKNVEKKID KYTEVLKTHG LLV	71
20	RPL9	MKTILSNQTV DIPENVDITL KGRTVIVKGP RGTLRRDFNH INVELSLLGK KKKRLRVDKW WGNRKELATV RTICSHVQNM IKGVTLGFRY KMRSVYAHFP INVVIQENGS LVEIRNFLGE KYIRRVRMRP GVACSVSQAQ KDELILEGND IELVSNSAAL IQQATTVKNK DIRKFLDGIY VSEKGTVQQA DE	89
20	RPS7	MFSSSAKIVK PNGEKPDEFE SGISQALLEL EMNSDLKAQL RELNITAAKE IEVGGGRKAI IIFVPVPQLK SFQKIQVRLV RELEKKFSGK HVVFIAQRRI LPKPTRKSRT KNKQKRPRSR TLTAVHDAIL EDLVFPSEIV GKRIRVKLDG SRLIKVHLDK AQQNNVEHKV ETFSGVYKKL TGKDVNFEFP EFQL	81

24	RPL17	MVRYSLDPEN PTKSCKSRGS NLRVHFKNTR ETAQAIKGMH IRKATKYLKD VTLQKQCVPF RRYNGGVGRC AQAKQWGWTQ GRWPKKSAEF LLHMLKNAES NAELKGLDVD SLVIEHIQVN KAPKMRRRTY RAHGRINPYM SSPCHIEMIL TEKEQIVPKP EEEVAQKKKI SQKKLKKQKL MARE	91
25	RPL11	MAQDQGEKEN PMRELRIRKL CLNICVGESG DRLTRAAKVL EQLTGQTPVF SKARYTVRSF GIRRNEKIAV HCTVRGAKAE EILEKGLKVR EYELRKNNFS DTGNFGFGIQ EHIDLGIKYD PSIGIYGLDF YVVLGRPGFS IADKKRRTGC IGAKHRISKE EAMRWFQQKY DGIILPGK	92
26A	RPL23A	MAPKAKKEAP APPKAEAKAK ALKAKKAVLK GVHSHKKKKI RTSPTFRRPK TLRLRRQPKY PRKSAPRRNK LDHYAIIKFP LTTESAMKKI EDNNTLVFIV DVKANKHQIK QAVKKLYDID VAKVNTLIRP DGEKKAYVRL APDYDALDVA NKIGII	65
26B	RPL11	MADQGEKEN PMRELRIRKL CLNICVGESG DRLTRAAKVL EQLTGQTPVF SKARYTVRSF GIRRNEKIAV HCTVRGAKAE EILEKGLKVR EYELRKNNFS DTGNFGFGIQ EHIDLGIKYD PSIGIYGLDF YVVLGRPGFS IADKKRRTGC IGAKHRISKE EAMRWFQQKY DGIILPGK	81
26C	RPL23A	MAPKAKKEAP APPKAEAKAK ALKAKKAVLK GVHSHKKKKI RTSPTFRRPK TLRLRRQPKY PRKSAPRRNK LDHYAIIKFP LTTESAMKKI EDNNTLVFIV DVKANKHQIK QAVKKLYDID VAKVNTLIRP DGEKKAYVRL APDYDALDVA NKIGII	72

29	RPS10	MLMPKKNRIA IYELLFKEGV MVAKKDVHMP KHPELADKNV PNLHVMKAMQ SLKSRGYVKE QFAWRHFYWY LTNEGIQYLR DYLHLPPEIV PATLRRSRPE TGRPRPKGLE GERPARLTRG EADRDTYRRS AVPPGADKKA EAGAGSATEF QFRGGFGRGR GQPPQ	93
39	RPS10	MLMPKKNRIA IYELLFKEGV MVAKKDVHMP KHPELADKNV PNLHVMKAMQ SLKSRGYVKE QFAWRHFYWY LTNEGIQYLR DYLHLPPEIV PATLRRSRPE TGRPRPKGLE GERPARLTRG EADRDTYRRS AVPPGADKKA EAGAGSATEF QFRGGFGRGR GQPPQ	83
43	RPS15A	MVRMNVLADA LKSINNAEKR GKRQVLIRPC SKVIVRFLTV MMKHGYIGEF EIIDDHRAGK IVVNLTGRLN KCGVISPRFD VQLKDLEKWQ NNLLPSRQFG FIVLTTSAGI MDHEEARRKH TGGKILGFFF	98
44	RPL23	MSKRGRGGSS GAKFRISLGL PVGAVINCAD NTGAKNLYII SVKGIKGRLN RLPAAGVGDM VMATVKKGKP ELRKKVHPAV VIRQRKSYRR KDGVFLYFED NAGVIVNNKG EMKGSAITGP VAKECADLWP RIASNAGSIA	99
45	RPL35A	MSGRLWSKAI FAGYKRGLRN QREHTALLKI EGVYARDETE FYLGKRCAYV YKAKNNTVTP GGKPNKTRVI WGKVTRAHGN SGMVRAKFRS NLPAKAIGHR IRVMLYPSRI	50
46	RPL38	MPRKIEEIKD FLLTARRKDA KSVKIKKNKD NVKFKVRCSR YLYTLVITDK EKAEKLKQSL PPGLAVKELK	83

53	RPL12	MPPKFDPNEI KVVYLRCTGG EVGATSALAP KIGPLGLSPK KVGDDIAKAT GDWKGLRITV KLTIQNRQAQ IEVVPSASAL IIKALKEPPR DRKKQKNIKH SGNITFDEIV NIARQMRHRS LARELSGTIK EILGTAQSVG CNVDGRHPHD IIDDINSGAV ECPAS	77

Table 3.5 Protein sequence coverage of twenty ribosomal proteins selected for in-gel digestion in the MXR gel arrays is indicated above. Peptides that were confidently identified by trypsin digestion only are shown in red, those confidently identified by acid digestion only are shown in blue and those confidently identified by both acid and trypsin digestion are identified by green text

Spot #	Protein Name	Percent Sequence Coverage
6	RPS3	98
7	RPS3	85
8	RPS3	89
25	RPL11	92
26A	RPL23A	65
26B	RPL11	81
26C	RPL23A	67
29	RPS10	93
39	RPS10	83

Table 3.6 Sequence coverage of the ribosomal proteins found in altered abundance between MXR and MXS using comparative densitometry

Comparative Densitometry between MXR and MXS Cell Lines

To provide confidence in the identification of altered proteins between the two cell lines, two imaging software programs were used concurrently to analyze the protein abundance profiles in the gel arrays, PDQuest™ from Bio-Rad and CompugenZ3™. Both imaging programs controlled for normalization of the spot intensity between images to take into account differences between sample loadings. In addition, both programs allowed for the alignment of the gel arrays and the creation of a master image/registered image of averaged gels for each cell line. Four gels of each cell line were used in order to create these master images for comparative densitometry. Composite gel arrays were evaluated in this way between the cell lines indicating homologous spots that were increased or decreased in abundance greater than 2-fold or only observed in one cell line (on/off). PDQuest™ also granted the additional option of a T-test comparison between the master images of the cell lines. Relative quantitation of the proteins was determined using the Compugen™ software. Figure 3.24 shows the composite 2-D gel maps of the MXS (left) and MXR (right) cell lines with Compugen™ spot assignment based on user defined parameters; Minimum spot area = 50, Minimum spot contrast = 20.

All of these analyses were taken into account and both PDQuest™ and CompugenZ3™ agreed on 8 protein spots whose protein abundance was altered. Figures 3.25 and 3.26 each show the protein spots chosen as differentially abundant by CompugenZ3™ and PDQuest™ respectively. Yellow circles in the Compugen™ gel comparison Figure indicate those proteins found in lower abundance in the MXR

cell line while blue circles indicate proteins found in higher abundance. Red circles indicate protein spots which were only found in the MXR cell line. Since PDQuest™ separated gel spots by on/off, >2-fold abundance and T-test results into different gel maps, one figure was created to show only those 8 protein spots shared with the Compugen™ analysis. PDQuest™ spot notation in Fig 3.26 is as follows; yellow circles with no fill indicate proteins found in higher abundance in the MXR cell line while yellow circles with black fill indicate proteins found in lower abundance in the MXR cell line. Red circles were added to indicate protein spots determined by PDQuest™ to be found only in the MXR gel arrays (those shared in common with CompugenZ3™ analysis).

In-gel digestion allowed for the confident identification of these proteins as isoforms of RPS3, RPS10, RPL11 and RPL23A. In-gel digestion also allowed the addition of a ninth spot, an additional isoform of RPS3, which served as a reference spot as it remained unchanged between the two cell lines. In the case of the Compugen™ analysis, one of the differentially abundant RPS10 isoforms (spot #39) was considered to be unique to the MXR cell line while in the PDQuest™ analysis, this spot was classified as being greater than 2-fold more abundant in the MXR cell line. As a result, this isoform was considered to be present in both cell lines however in significantly greater abundance in the MXR cell line. A zoomed in image of each of these changes between the two cell lines is seen in Figures 3.27, 3.28, 3.29, and 3.30. Spots in these Figures are labeled following the color scheme used in the Compugen™ program. Relative quantitation of the protein spots was determined by the Compugen™ program and is seen in Table 3.7.

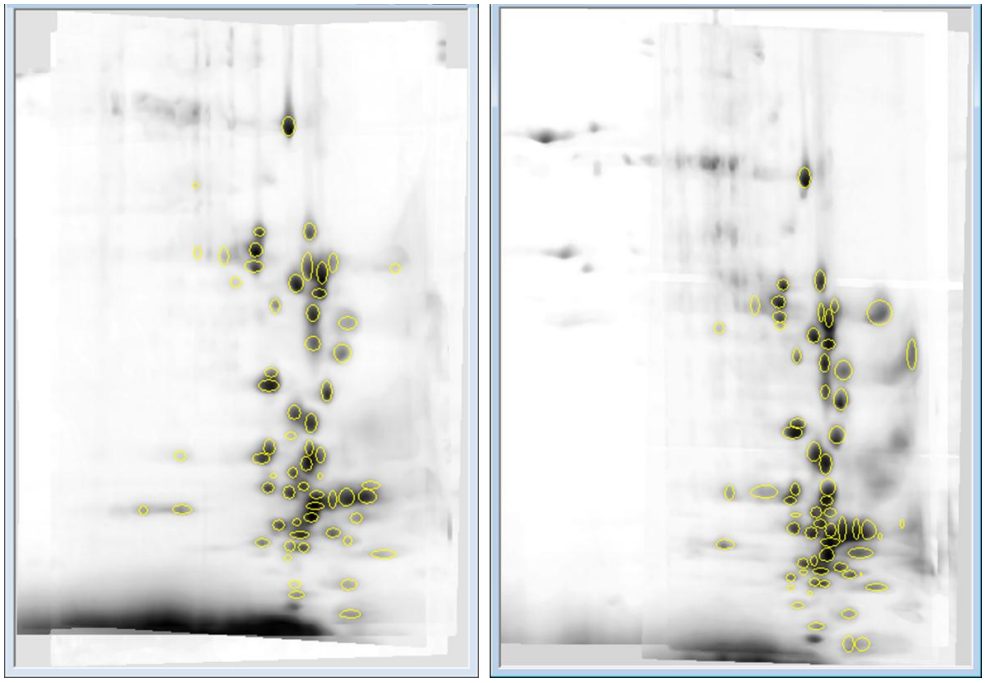

Figure 3.24 Composite gel maps of the MXS (left) and the MXR (right) gel arrays with Compugen™ assigned spots (maps were also manually inspected and spots verified)

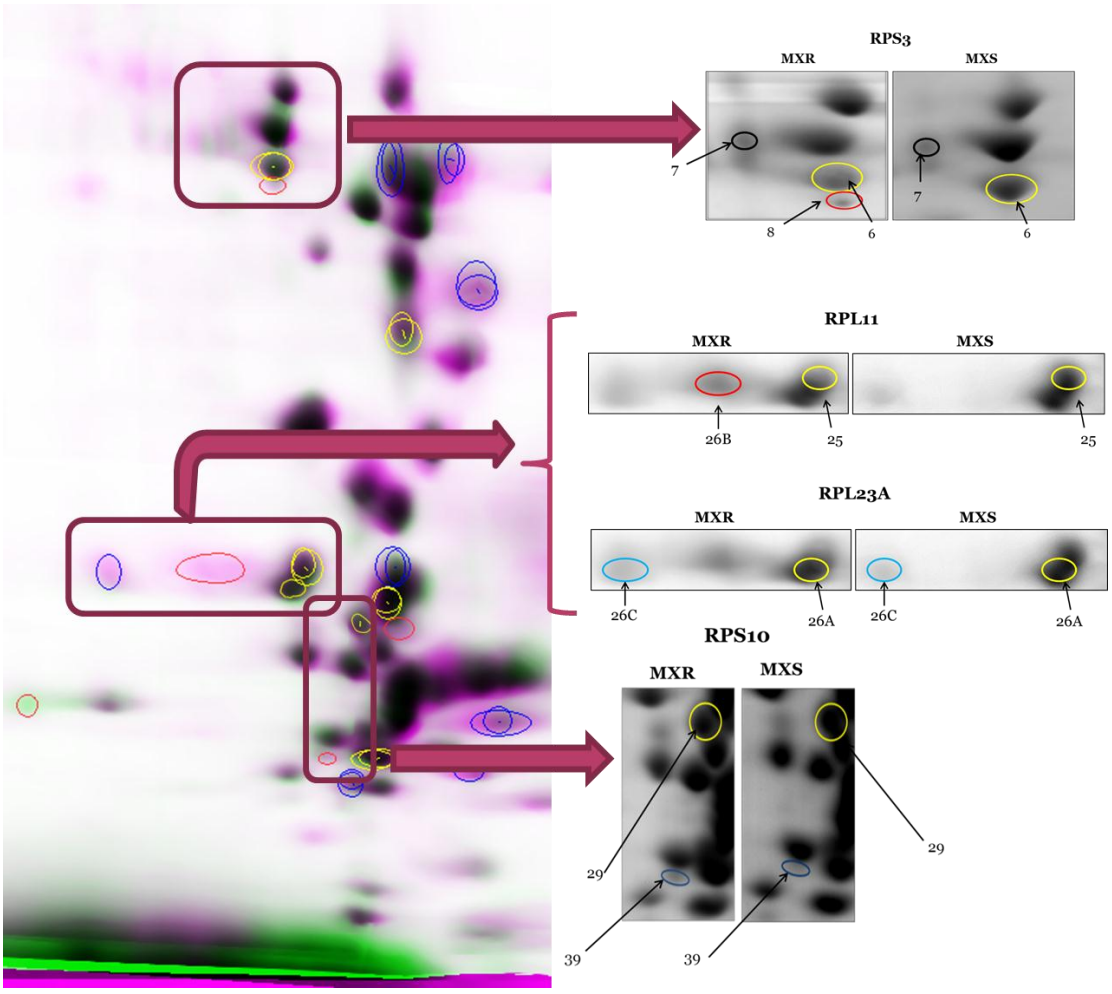

Figure 3.25 Overlaid gel image for comparative densitometry conducted by CompugenZ3™ and the sets of spots assigned which corresponded with proteins also found differentially abundant by PDQuest™ (Fig 3.26). Blue circles represent proteins found in higher abundance in the MXR cell line, yellow circles represent those proteins found in lower abundance in the MXR cell line and red circles represent spots unique to the MXR cell line

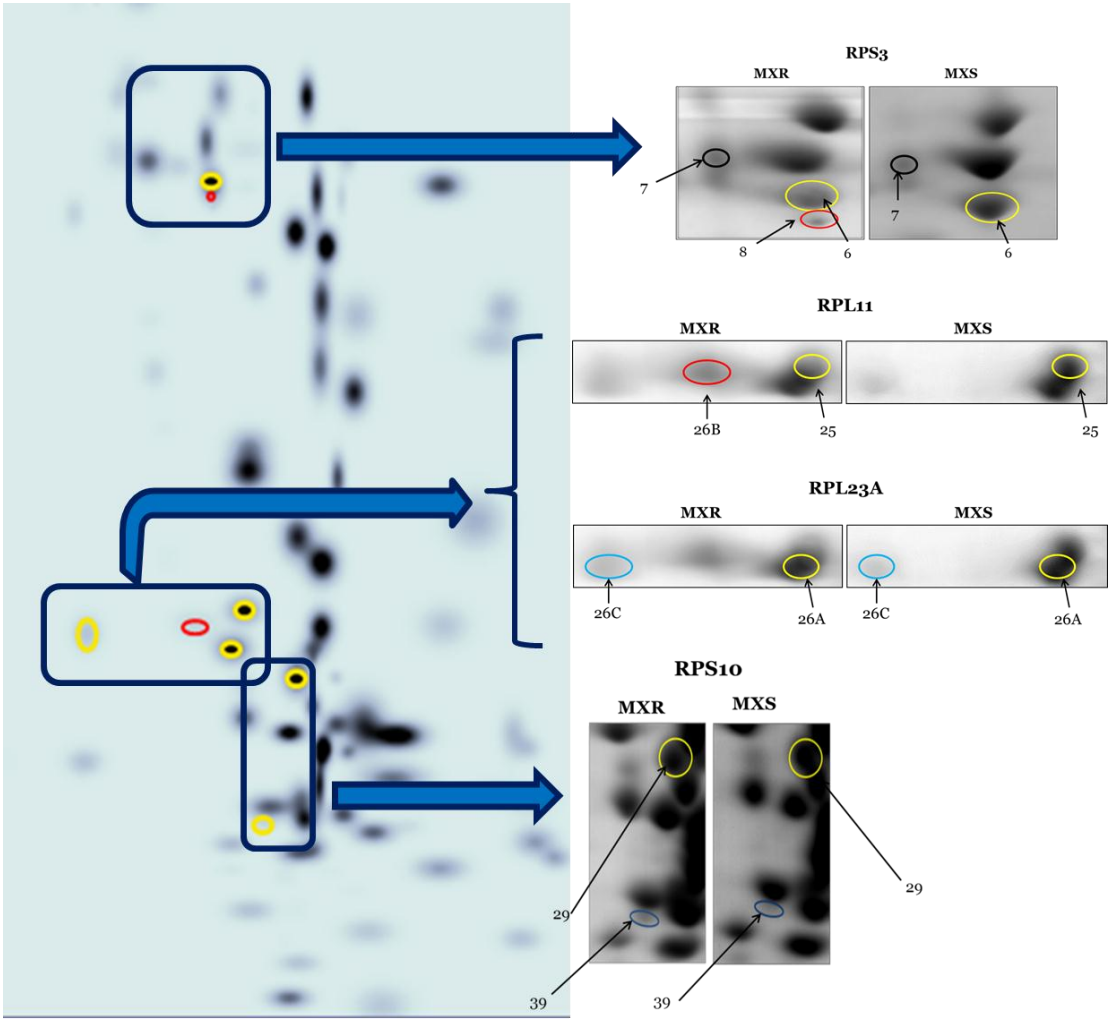

Figure 3.26 Overlayed composite of MXR and MXS gels for comparison conducted by PDQuest™. Proteins that were in lower abundance in the MXR cell line are represented by yellow filled in circles, while those proteins in higher abundance are represented by yellow circles without fill. Proteins found only in the MXR cell line are represented here by red circles. The gel images on the right adopts the color scheme used by CompugenZ3™ (refer to Fig 3.25)

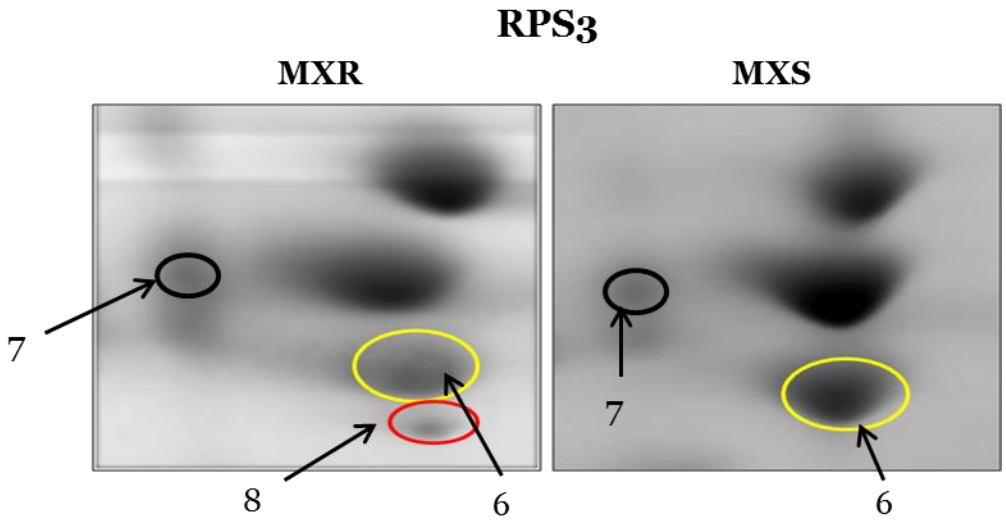

Figure 3.27 RPS3 protein abundance changes between the MXR and MXS cell lines.

The yellow circle indicates a protein isoform that is found in lower abundance in the MXR cell line while the red circle indicates a spot which is novel to the MXR cell line (unmatched). Spot number 7 indicated with the black circle is an RPS3 isoform that remained unchanged between the two cell lines and served as a reference spot

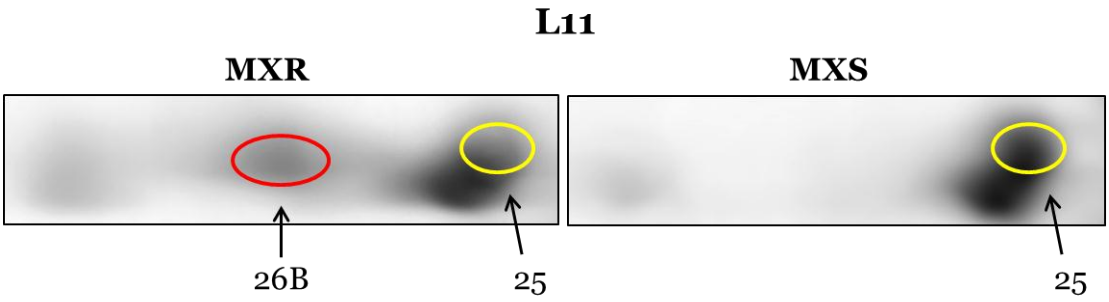

Figure 3.28 RPL11 protein abundance changes between the MXR and MXS cell lines.

The yellow circle indicates a protein spot found in higher abundance in the MXS cell line while the red circle indicates a protein which was unique to the MXR cell line (unmatched)

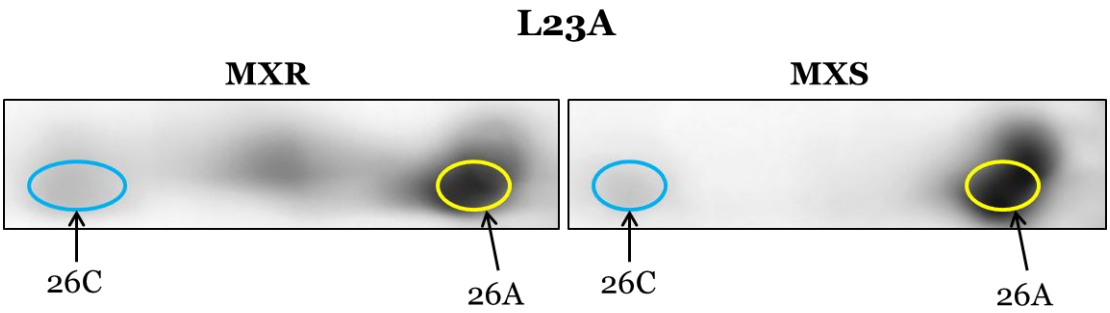

Figure 3.29 RPL23A protein abundance changes between the MXR and MXS cell lines.

The yellow circle indicates a protein spot found in higher abundance in the MXS cell line while the blue circle indicates a protein which is found in higher abundance in the MXR cell line

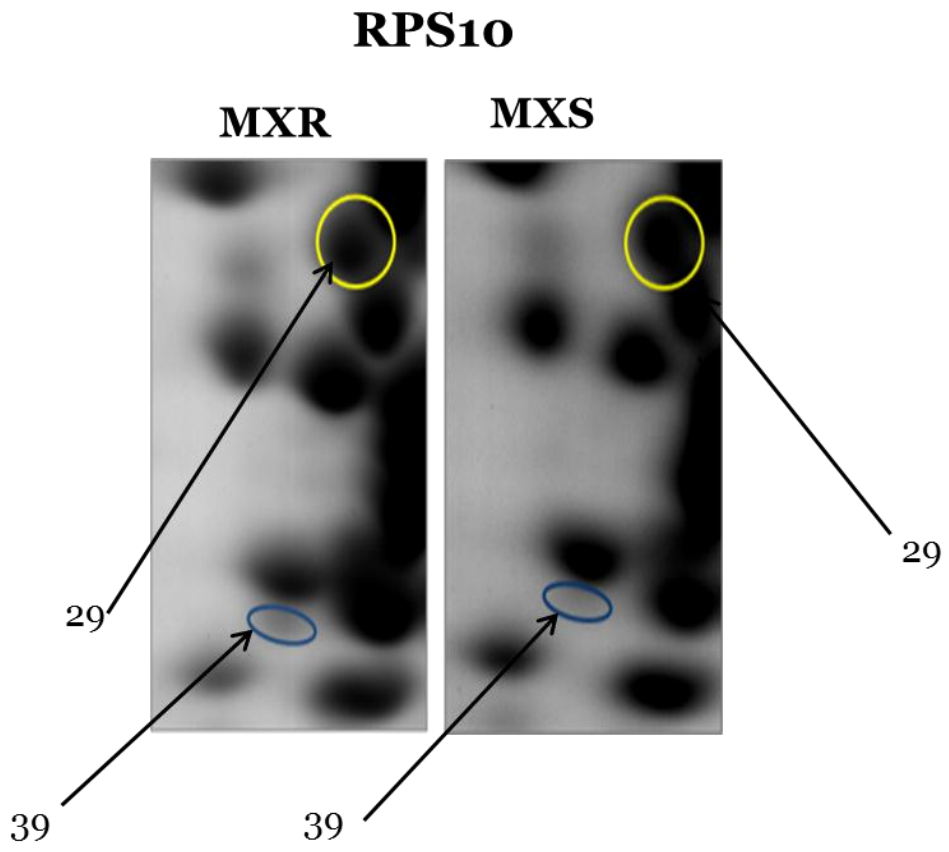

Figure 3.30 RPS10 protein abundance changes between the MXR and MXS cell lines.

The yellow circle indicates a protein isoform which was found in lower abundance in the MXR cell line while the blue circle indicates a protein isoform which was found in higher abundance in the MXR cell line. Spot 39 was considered to be unique (unmatched) to the MXR cell line by the CompugenZ3™ software program while it was considered to be in greater than 2-fold abundance in the MXR cell line based on the PDQuest™ software package

Spot Name	Protein Isoform	MXS Cell Line	MXR Cell Line
6	RPS3	1	0.45 ± 0.02
7	RPS3	1	1
8	RPS3	Not Present	Present
25	RPL11	1	0.23 ± 0.02
26B	RPL11	Not Present	Present
26A	RPL23A	1	0.51 ± 0.02
26C	RPL23A	1	2.57 ± 0.10
29	RPS10	1	0.53 ± 0.03
39	RPS10	Not Present	Present (> 2-fold PDQuest™)

Table 3.7 CompugenZ3™ determined relative quantitation of protein isoforms with altered abundance between the two cell lines. Spot #39 was considered unique (unmatched) to the MXR cell line by CompugenZ3™ but labeled as a spot found in greater than 2-fold abundance by PDQuest™

Protein Isoform Characterization

Results of the gel-extracted molecular masses were integrated with in-gel digestion data, MXR acid digestion data, HPLC fraction molecular masses and HPLC fraction trypsin digestion data in order to fully characterize the protein isoforms found in the MXR cell line. The initial step in the protein isoform characterization was determination of an *approximate* molecular mass with MALDI (this process also verified the presence of protein in the gel extraction solution). An aliquot of the gel extraction solution consisted of MALDI matrix solvent (50%ACN/0.1% TFA) which was mixed 1:1 with MALDI matrix (10mg/mL sinapinic acid) in matrix solvent with 5% Triton-100X (final concentration 2.5%) as described in the Experimental section and spotted on the MALDI plate using the sandwich method. As was the case with the in-gel digestions, additional protein spots were selected for protein extraction for gel validation purposes. Twenty three spots in all were examined with MALDI. The change in mass from the theoretical mass was initially intended to be used as a rough guide for possible post translational modifications however it served as an early indication of what was later verified with the in-gel digestion analysis to be *incomplete alkylation* (with IAA) of many of the proteins. Instrument settings were as described in the Experimental. A table of the molecular masses observed with MALDI and the corresponding spot name and protein identification is found in Table 3.8.

#	Protein Name	# of Cys	Theoretical mass (incl. IAA alkylation)	MALDI Mass	Delta Mass (Da)
3	RPL6	1	32654	32792	138
5	RPS3A	4	30042	28367	-1675
6	RPS3	3	26728	26828	100
7	RPS3	3	26728	26898	170
8	RPS3	3	26728	26704	-24
13	RPL8	4	28121	28095	-26
15	RPS8	5	24359	24201	-158
16B	RPL13	1	24187	24335	148
18	RPL10	8	24929	24954	25
19	RPL13A	1	23503	23711	208
20	RPS7	0	22126	22188	62
20	RPL9	2	21977	21900	77
24	RPL17	4	21494	21296	-198
25	RPL11 iso1	4	20349	20385	36
26A	RPL23A	0	17563	17562	1
26B	RPL11 iso2	4	20352(with Met) 20221(without Met)	20392	40 (with Met) 171 (without Met)
26C	RPL23A	0	17563	17671	108
29	RPS10	0	18897	18991	93
39	RPS10	0	18897	15922	-2975
43	RPS15A	2	14822	14793	-29

44	RPL23	2	14979	14954	-25
45	RPL35A	1	12595	12559	-36
46	RPL38	1	8144	8145	1
53	RPL12	3	17990	18021	31

Table 3.8 Whole proteins which were extracted from the gels were first evaluated with MALDI to verify the presence of sample. The delta mass values were the first indication that there was incomplete alkylation of the proteins in many cases as well as oxidation of the proteins as a result of sample handling and storage

Following MALDI analysis of the gel protein extracts, protein isoforms were characterized by comparison of their nLC-Orbitrap molecular mass measurements with information determined from the in-gel digestion data. If it was determined that further information was required to determine the modifications of the protein in question, information gathered from the MXR acid digestion as well as digestion data and molecular masses observed during HPLC fraction analysis were also considered.

One major issue which proved problematic with the molecular mass interpretation was the incomplete alkylation with IAA first observed with MALDI in cysteine containing proteins. In RPS3 for example, more than one version of each protein isoform was observed containing between 0 – 3 carbamidomethylations. Depending on the modification(s) and number of artifactual methionine oxidations (from sample handling and storage), overlap between the protein isotopic envelopes on more than one occasion was observed. Digestion data was found to support the incomplete alkylation of the proteins. To simplify data interpretation, the **most abundant** molecular ion in the molecular mass data of a protein isoform was selected

under the condition that this molecular mass **agreed with earlier observations** made regarding the protein in question (in-gel digestion, acid digestion, etc.). As described previously, once a protein isoform with modifications was determined based on mass and digestion data, the high resolution deconvoluted isotope peaks of each protein was compared with the theoretical isotope envelope expected for the average mass of the amino acid composition of the given protein (along with any detected modifications; PTM or artifactual).

In the instance of RPS3, three protein spots identified as numbers 6, 7 and 8 were noted in the gel arrays corresponding to each protein isoform (See Figures 3.23 and 3.37). The sequence annotation of RPS3 has noted the loss of the initiator methionine and was observed to be the case with the digestion data analysis. For spot #6, the most abundant experimental mass was 26598.45Da (the unalkylated form of the protein; supported by digestion data and HPLC fraction molecular mass), shown in the bottom of Figure 3.31. The mass difference between the theoretical mass of RPS3 and the observed mass of spot #6 corresponded exactly with the addition of an acetylation/trimethylation. Inspection of the digestion data revealed that the N-terminal acetylated peptide of the protein had been detected in both the MXR acid digestion data and the trypsin digestion data as seen in Figures 3.32 and 3.33. The theoretical spectrum corresponding to this protein with an acetylation is shown in the top panel of Fig 3.31. The most abundant mass of the isoform of RPS3 observed in spot #7 was measured as 26897.52Da as seen in the bottom panel of Figure 3.35. The presence of additional protein species in the sample resulted in some issues with deconvolution. The in-gel digestion verified the presence of three methionine

oxidations (the result of sample handling and storage) as shown in Table 3.9, however, the most abundant mass contained three oxidations. In-gel digestion data verified that the mass difference between the experimental mass and the theoretical mass was attributable to the addition of an N-terminal acetylation and one phosphorylation found on T221. A comparison of the theoretical spectrum of all three oxidized species with the observed species is seen in Figure 3.34. Figure 3.36 illustrates the phosphopeptide observed for this protein. The theoretical spectrum for the protein with the proposed modifications is shown in the top panel of Fig 3.35. Another isoform of RPS3 was detected in spot #8 at a mass of 26667.46Da as shown in the bottom panel of Figure 3.37. This corresponded with the mass of the protein without any modifications except for 2 carbamidomethylations (without acetylation). This was further supported by examination of the acid digestion data which found the N-terminal peptide of RPS3 without an acetylation. The theoretical spectrum for the protein with the proposed modifications is shown in the top panel of Fig 3.37. The sequence coverage for all three isoforms is shown aligned in Figure 3.38. A list of all the modified peptides observed for RPS3 is included in Table 3.10.

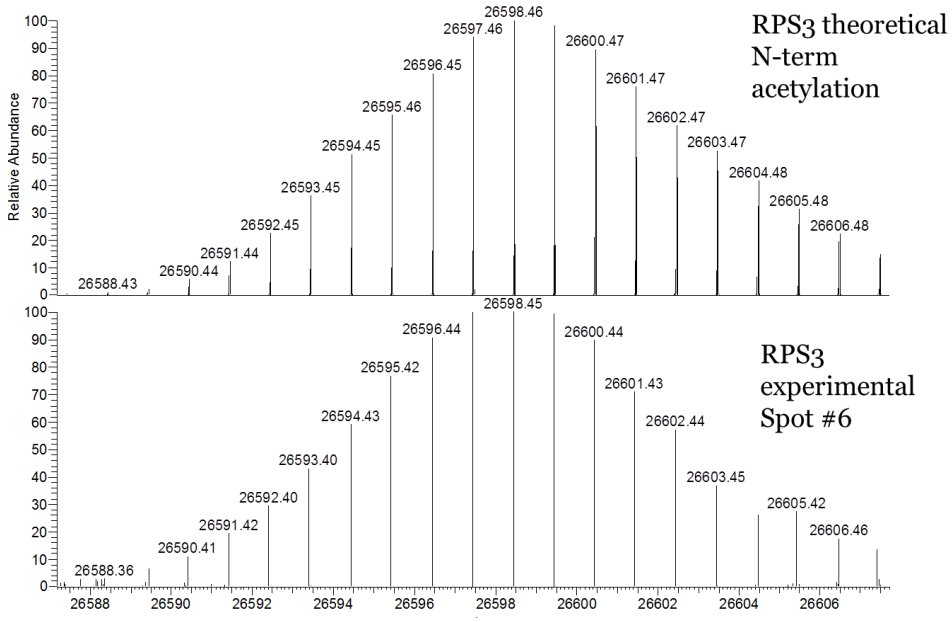

Figure 3.31 Experimental spectrum of extracted RPS3 protein found in spot #6 (bottom panel) compared with theoretical spectrum with the corresponding modification (top panel)

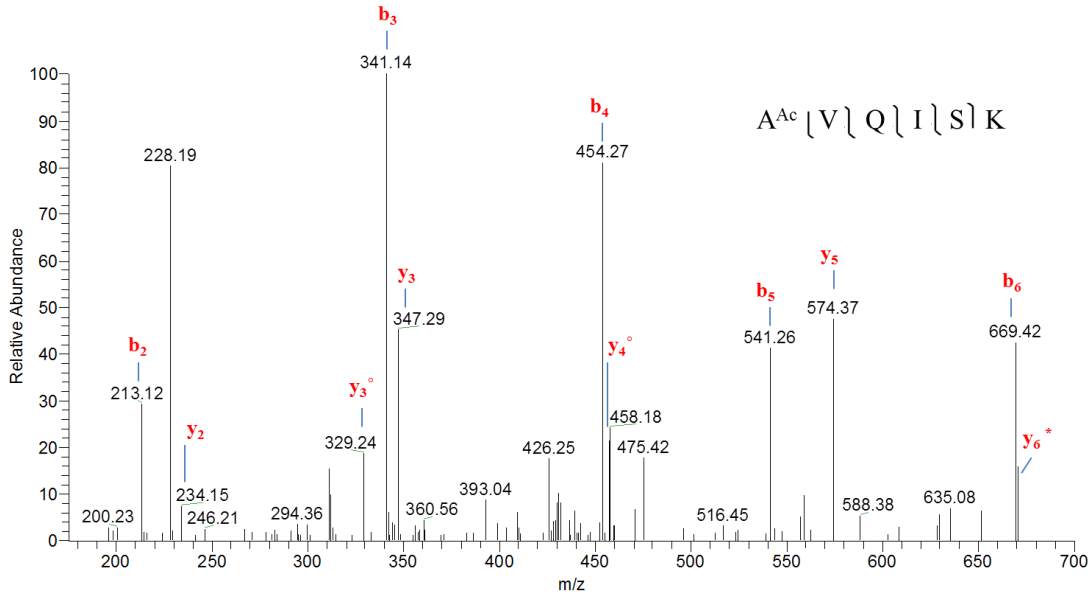

Figure 3.32 MS/MS spectrum of N-terminal acetylation of RPS3 identified from in-gel digestion. Neutral water loss was noted in both this and the acid digestion fragment ions (refer to Fig 3.33)

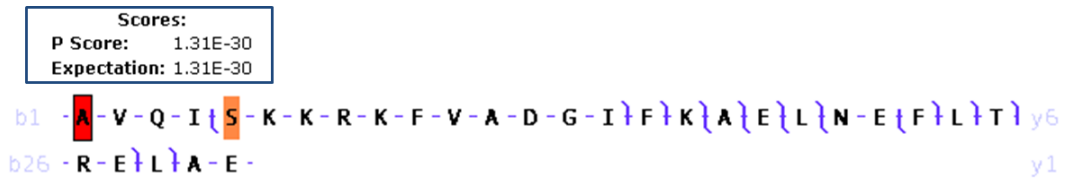

Figure 3.33 Fragment ions identified during nLC-Orbitrap analysis of the MXR acid digest to confidently identify the N-terminal acetylation of RPS3 in spot 6 and spot 7. The A residue highlighted in red signifies the acetylation while the S residue highlighted in orange signifies that that fragment was dehydrated (neutral water loss)

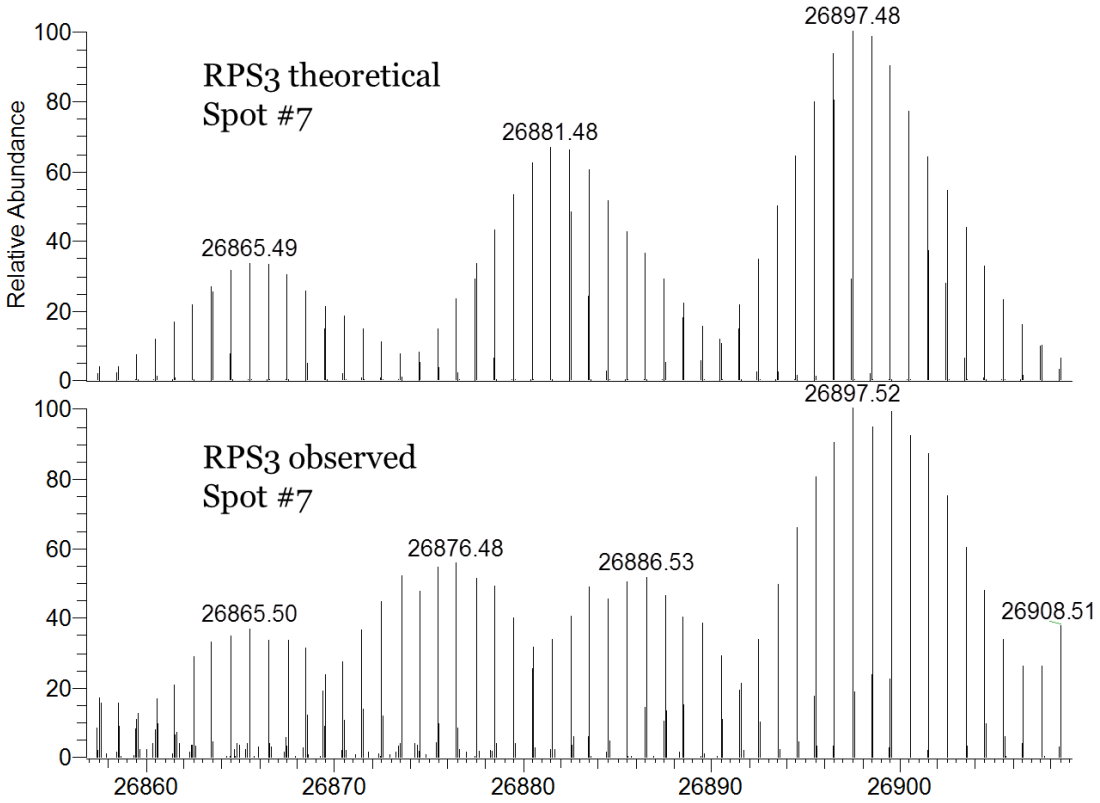

Figure 3.34 Theoretical (top panel) and observed (bottom panel) mass spectrum for RPS3 isoform found in spot #7 with 1, 2 and 3 methionine oxidations. The lack of agreement is attributed to unidentified contaminating protein species which interfered with the deconvolution

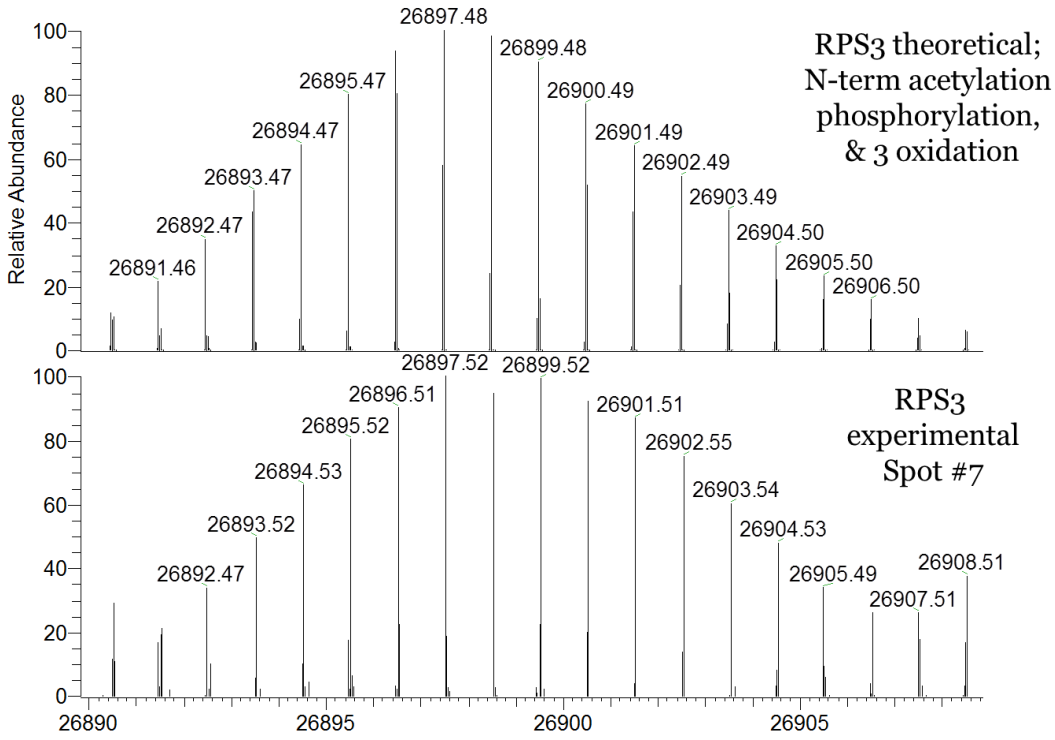

Figure 3.35 Theoretical (top panel) and observed (bottom panel) mass spectrum for RPS3 isoform found in spot #7. Digestion data showed this isoform to contain an N-terminal acetylation (as seen in Figure 3.31 for spot 6), 3 oxidations (refer to Table 3.9), and a phosphorylation on T221 (refer to Fig 3.36)

Number of Oxidations	Molecular Mass Observed	Theoretical Mass	Supporting peptides observed in bottom-up analyses w/ and without oxidation
3	26897.52	26897.48	K.-FVDGLMIHSGDPVNYYVDTAVR.-H, K.-IMLPWDPTGK.-I, K.-GGKPEPPAMPQPVPTA-
2	26883.53	26881.48	K.-FVDGLMIHSGDPVNYYVDTAVR.-H, K.-IMLPWDPTGK.-I, K.-GGKPEPPAMPQPVPTA-
1	26865.50	26865.49	K.-FVDGLMIHSGDPVNYYVDTAVR.-H, K.-IMLPWDPTGK.-I, K.-GGKPEPPAMPQPVPTA-

Table 3.9 As with many other gel extracted proteins, the molecular mass of spot #7 was observed with multiple oxidations. The most abundant molecular mass for this RPS3 isoform contained 3 methionine oxidations. These oxidations were also observed in the in-gel digestion data and attributed to sample handling

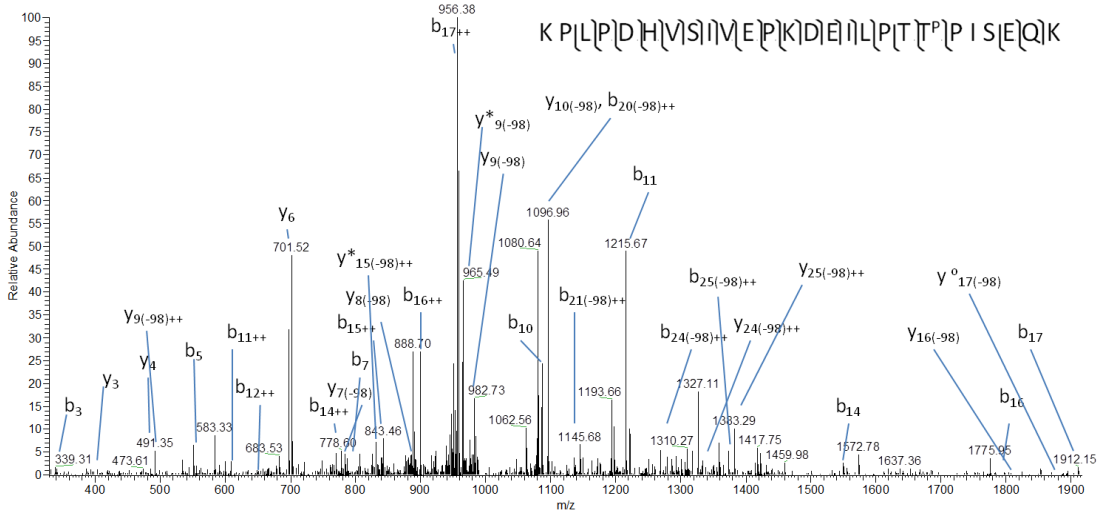

Figure 3.36 MS/MS spectrum of the phosphopeptide found on T221 in the RPS3 isoform found in spot 7

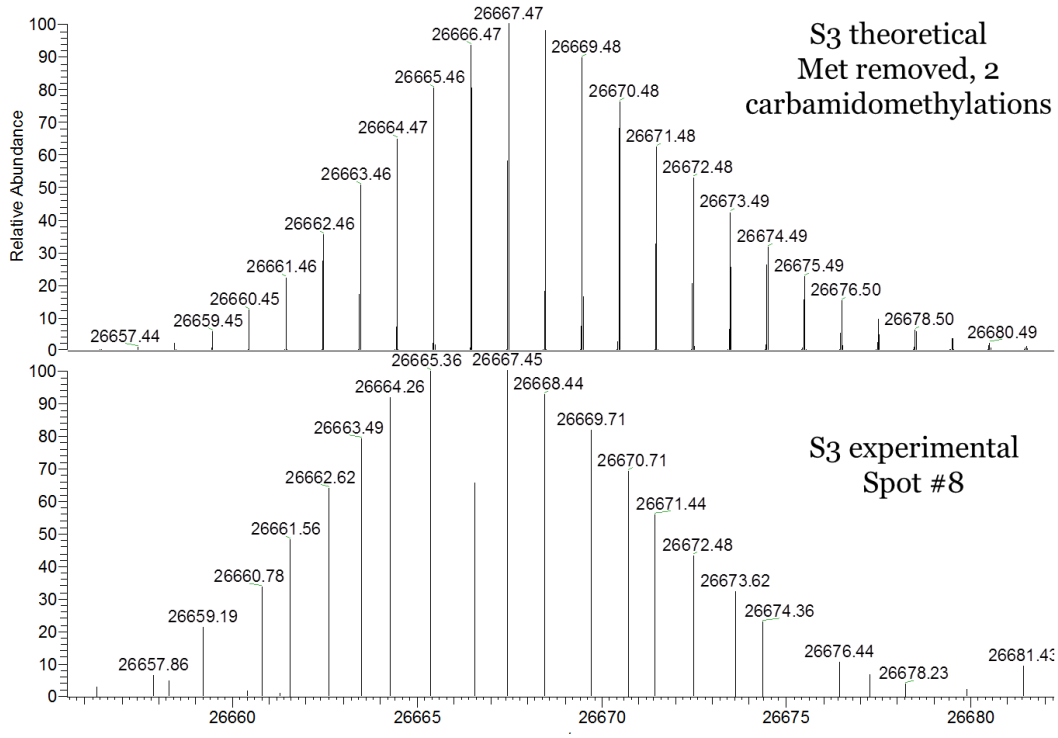

Figure 3.37 Theoretical (top panel) and experimental (bottom panel) mass spectrum for RPS3 isoform found in spot 8

Figure 3.38 Sequence coverage observed for all three RPS3 isoforms. Underlined residues indicate a modification was observed

Two spots in the gel arrays were identified as isoforms of RPL11 and labeled as spot #25 and #26B (see Fig 3.23 and 3.28). There are two spliced variants of RPL11 noted in the currently accepted human ribosomal protein database (UniProtKB/Swiss-Prot; P62913-1, P62913-2) which differ in sequence by one amino acid (Q in position 3 in isoform 1 or the loss of Q in isoform 2). Both of these variants of RPL11 were observed in the in-gel digestion data and the MS/MS spectra of the identifying peptides are shown in Figures 3.40 and 3.42. The initiator methionine is noted in the database sequence annotation as being removed and was observed as such in the digestion data. Observations of a mass difference between the most abundant nLC-Orbitrap experimental molecular mass and the theoretical mass of RPL11 isoform 1 corresponded with the addition of a single

acetylation/trimethylation and were noted in spot #25 with a mass of 20161.61Da seen in Figure 3.39. This mass change was supported with the in-gel digestion data by the discovery of the N-terminal end of the protein with an N-acetyl-L-alanine (shown in Figure 3.40). Cysteine containing peptides were observed in the digestion data with and without carbamidomethylation (4 in total). The other isoform of RPL11 was labeled as spot #26B. The most abundant mass was observed as 20410.62Da and is shown in Figure 3.41. Inspection of the in-gel digestion data revealed that this variant of RPL11 is isoform 2 contained 4 acetylations, 1 deamidation and 1 oxidation. The deamidated peptide ion precursor was manually inspected to determine the validity of the assignment of this modification. Although a possibility exists that the deamidation could be misassigned, the precursor ion was found to align more closely with the expected ion of the deamidated peptide versus the unmodified peptide. This can be seen in Figure 3.42. MS/MS spectra of two of these peptides are shown in Figures 3.43 and 3.44. Sequence coverage of the two RPL11 isoforms aligned is shown in Figure 3.45. All modified peptides of RPL11 are listed in Table 3.10.

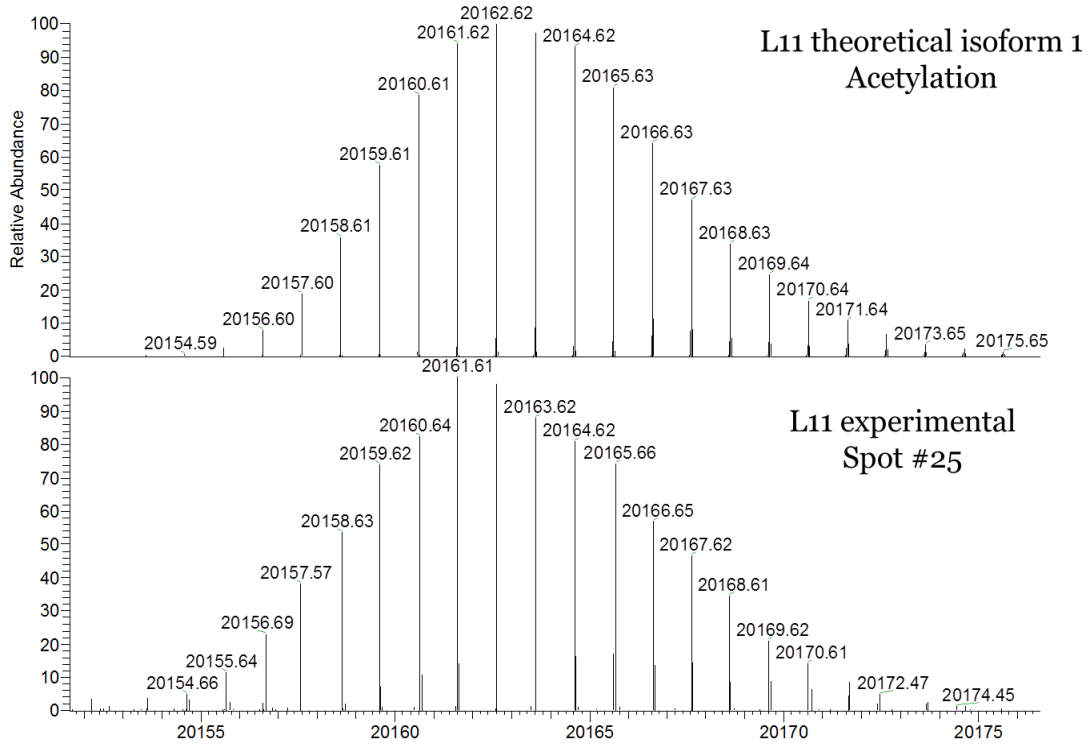

Figure 3.39 Theoretical (top panel) and experimental (bottom panel) mass spectrum of isoform 1 of RPL11 observed in spot 25 with an acetylation (refer to Fig 3.40). This protein isoform was not fully alkylated and was also observed in the HPLC fraction analysis

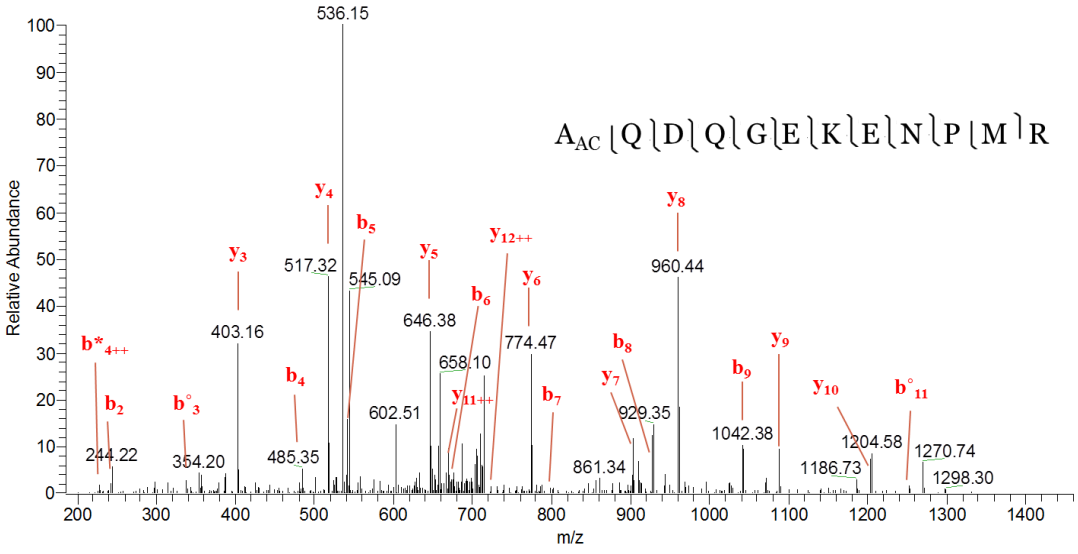

Figure 3.40 MS/MS spectrum identified as the acetylated N-terminal end of isoform 1 of RPL11 found in spot 25

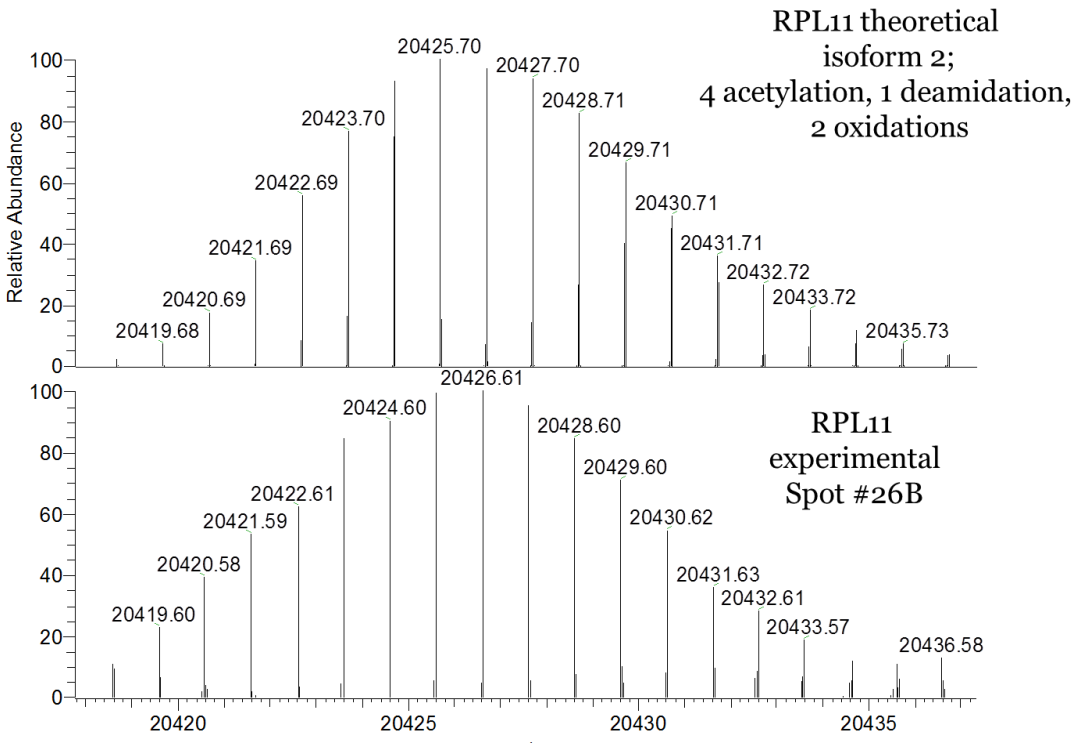

Figure 3.41 Theoretical (top panel) and experimental (bottom panel) mass spectrum of RPL11 isoform 2 and the corresponding PTMs identified by bottom-up analysis some of which are shown in Figures 3.43 and 3.44

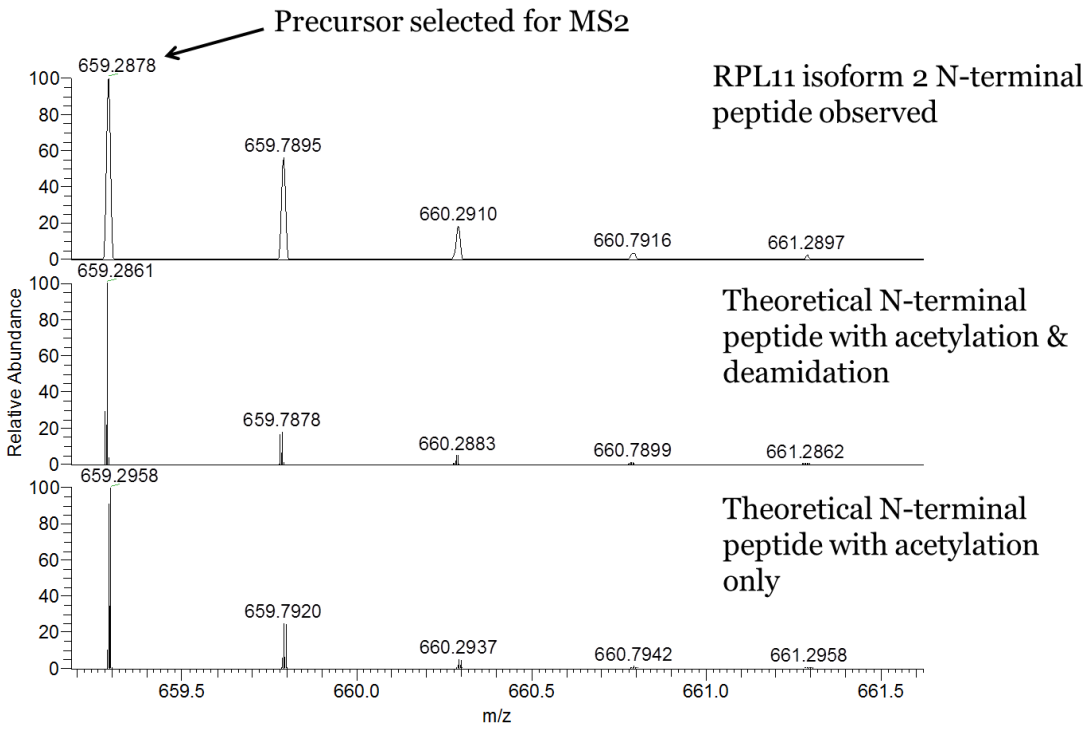

Figure 3.42 Precursor ion (z =2+) spectrum identified as the acetylated N-terminal end of RPL11 isoform 2 found in spot 26B compared with the theoretical precursor ions with and without deamidation

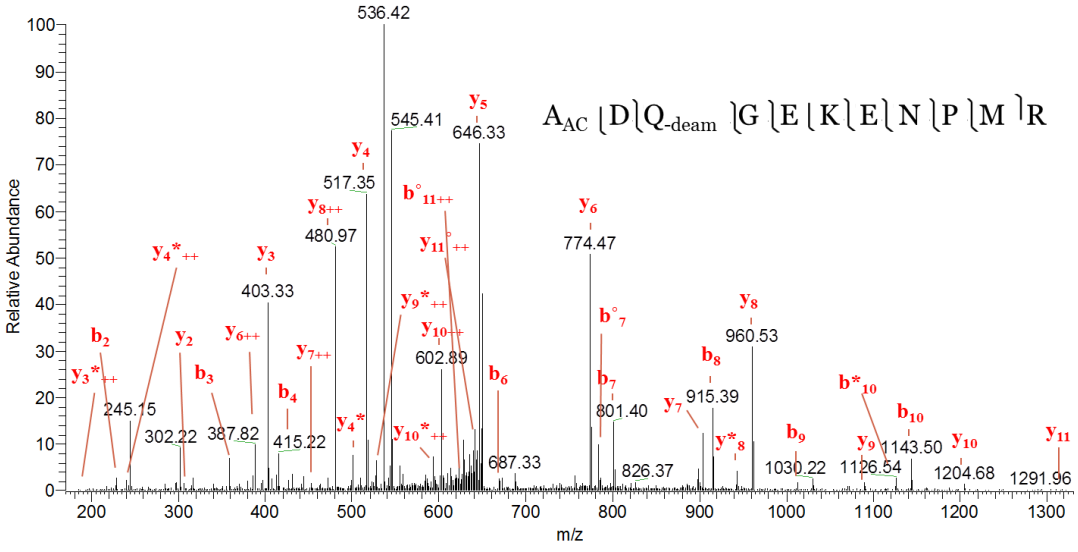

Figure 3.43 MS/MS spectrum identified as the acetylated N-terminal end of RPL11 isoform 2 found in spot 26B. This peptide was only observed with a deamidation on Q3 and was not observed without the deamidation

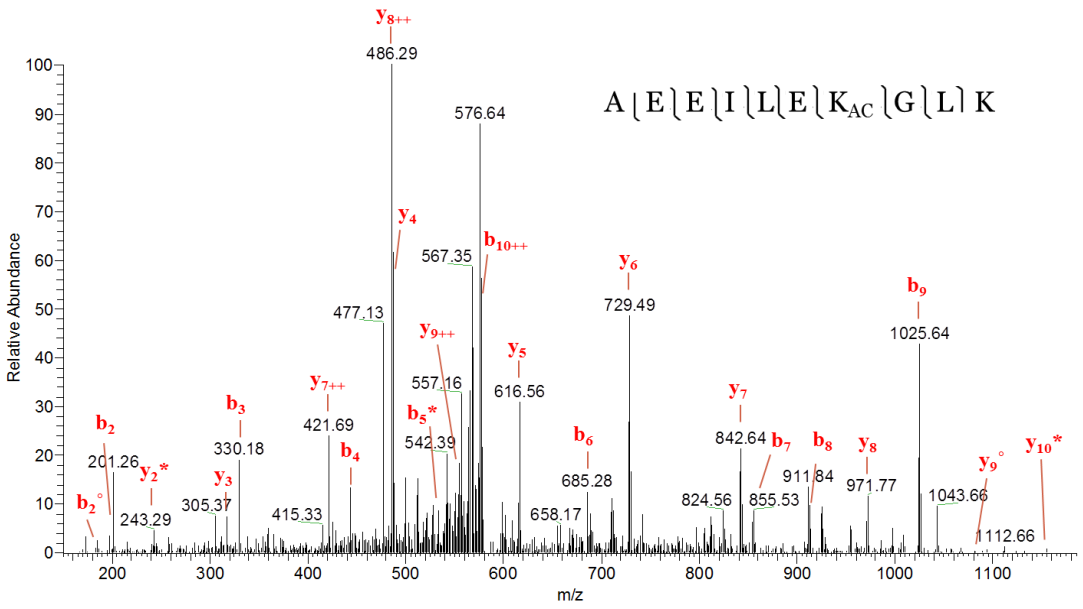

Figure 3.44 MS/MS spectrum identified as lysine acetylated peptide of RPL11 found in spot 26B

```
Spot 25    MAQDQGEKEN  PMRELRIRKL  CLNICVGESG  DRLTRAAKVL  EQLTGQTPVF
Spot 26B   MADQGEKEN   PMRELRIRKL  CLNICVGESG  DRLTRAAKVL  EQLTGQTPVF

Spot 25    SKARYTVRSF  GIRRNEKIAV  HCTVRGAKAE  EILEKGLKVR  EYELRKNNFS
Spot 26B   SKARYTVRSF  GIRRNEKIAV  HCTVRGAKAE  EILEKGLKVR  EYELRKNNFS

Spot 25    DTGNFGFGIQ  EHIDLGIKYD  PSIGIYGLDF  YVVLGRPGFS  IADKKRRTGC
Spot 26B   DTGNFGFGIQ  EHIDLGIKYD  PSIGIYGLDF  YVVLGRPGFS  IADKKRRTGC

Spot 25    IGAKHRISKE  EAMRWFQQKY  DGIILPGK
Spot 26B   IGAKHRISKE  EAMRWFQQKY  DGIILPGK
```

Acetylation: N-terminal end (A2); Spot 25, 26B
K-acetylation: K51, K67, K85; Spot 26B (all noted in literature)

Legend
X = obs. w trypsin
X = obs. w acid digest
X = obs. w both acid/trypsin
X = not obs.

Figure 3.45 Sequence coverage of RPL11 isoforms aligned. Underlined residues indicate a modification was observed. Both isoforms of RPL11 were observed (isoform 1 in spot 25 and isoform 2 in spot 26B; distinguished by residue 3 in sequence). Observed acetylations have been noted in the literature[109; 110; 111] previously (A2, K52, K67 & K85) except N-terminal acetylation of isoform 2. Oxidations were observed in in-gel digestion peptide data on methionine (M11 & M162)

Two altered spots were identified in the gel arrays as isoforms of RPL23A and labeled as spots #26A and #26C (see Fig 3.23 and 3.29). Sequence annotation of this protein notes loss of the initiator methionine and this was consistent with molecular mass measurements observed (the peptide corresponding with the N-terminal end however was not detected). The most abundant molecular mass noted for spot #26A is shown in Figure 3.46 and was 17638.14Da. Based on the theoretical mass of the protein and observations made with the in-gel digestion, this experimental mass corresponds with the addition of an acetylation (or trimethylation) and 2 methionine oxidations, all of which were discovered in the digestion data. The MS/MS spectrum corresponding to the acetylation noted for this protein in spot#26A (also observed in #26C) is shown in Figure 3.47. The MXR HPLC fraction mass measurement containing RPL23A was also found to contain this isoform (without the 2 oxidations). The theoretical and experimental mass of the HPLC fraction isoform corresponding to the protein found in spot 26A is shown in Figure 3.48. The isoform of RPL23A discovered in spot#26C was found to be most abundant experimentally at a molecular mass of 17718.18Da and is shown in Figure 3.49. Modifications that support this mass change from the theoretical mass include the acetylation determined to be present in both isoforms (see Fig 3.47) in addition to a phosphorylation detected in the MXR HPLC fraction digestion and two artifactual oxidations (1 methionine and 1 histidine) detected in the digestion data (detected by both in-gel and HPLC fraction trypsin digestion). As discussed with spot 7 of RPS3, these oxidations are attributed to sample handling. The MS/MS spectrum for the phosphorylation is shown in

Figure 3.50. The sequence coverage for both RPL23A isoforms is shown in Figure 3.51. A list of all the modified peptides found in RPL23A is included in Table 3.10.

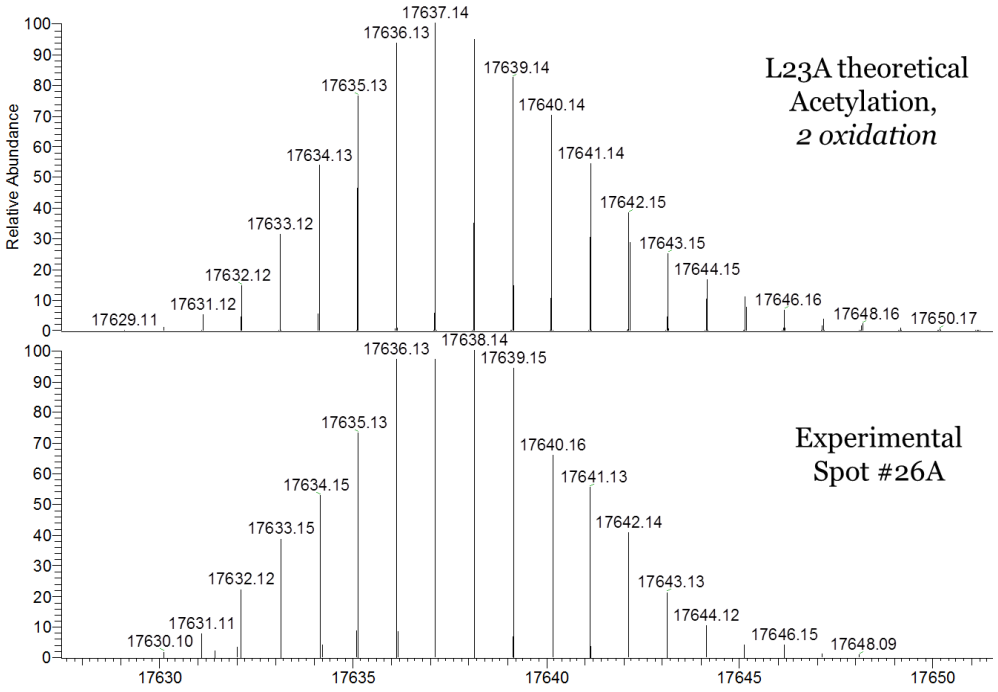

Figure 3.46 Theoretical (top panel) and experimental (bottom panel) MS of protein observed in spot 26A identified as RPL23A and determined by bottom-up analysis to contain an acetylation (refer to Fig 3.47)

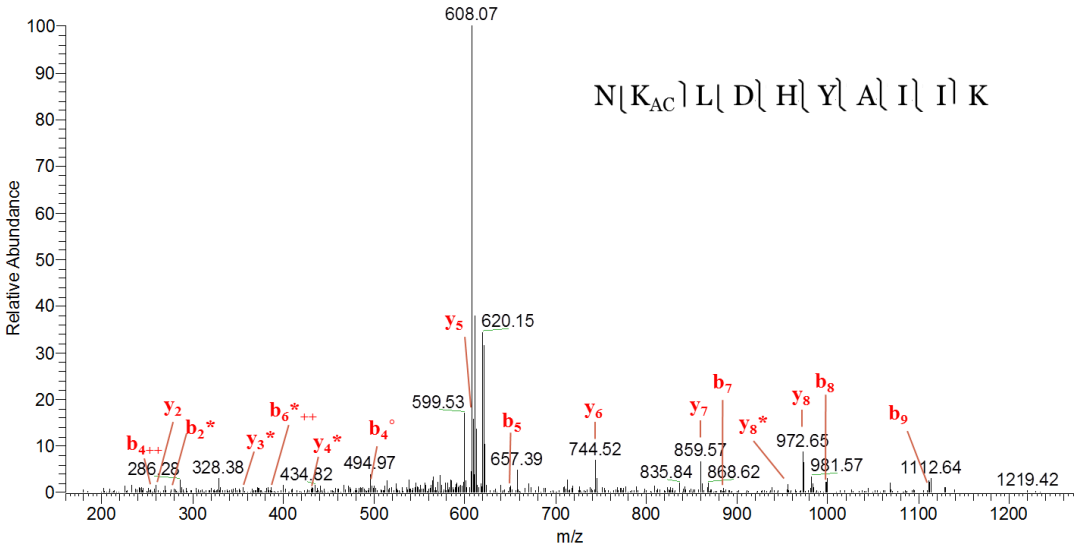

Figure 3.47 MS/MS spectrum of acetylated peptide from RPL23A identified in both spot 26A and 26C

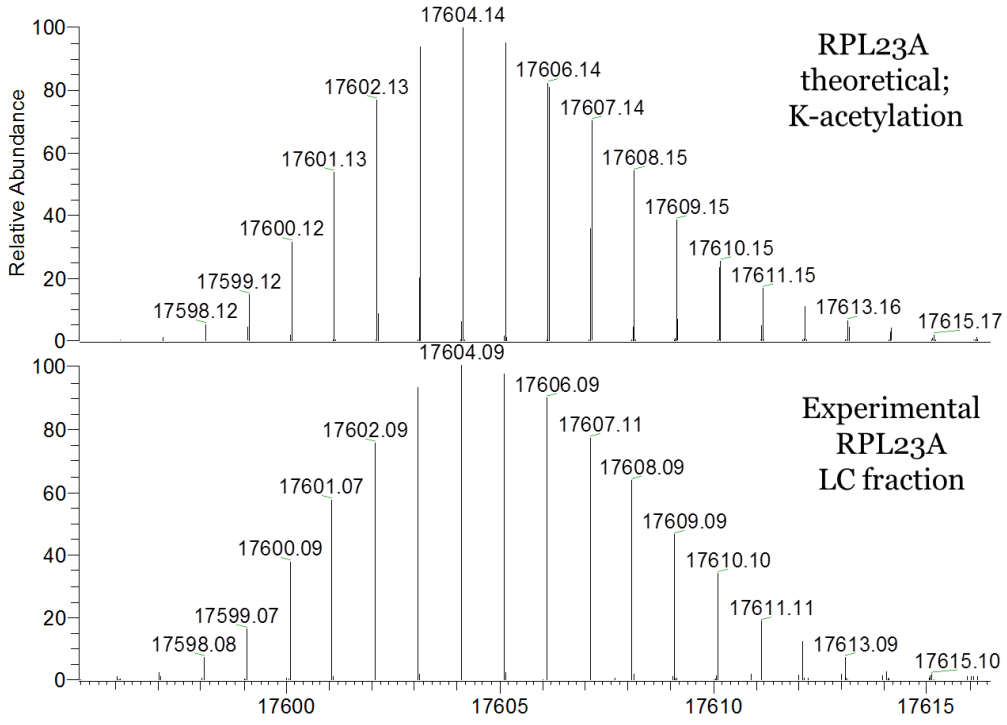

Figure 3.48 Spot #26A RPL23A protein isoform was also observed in the LC analysis (both fraction and top-down whole ribosomal proteome) with acetylation and without oxidation (therefore oxidation was attributed to sample handling and storage)

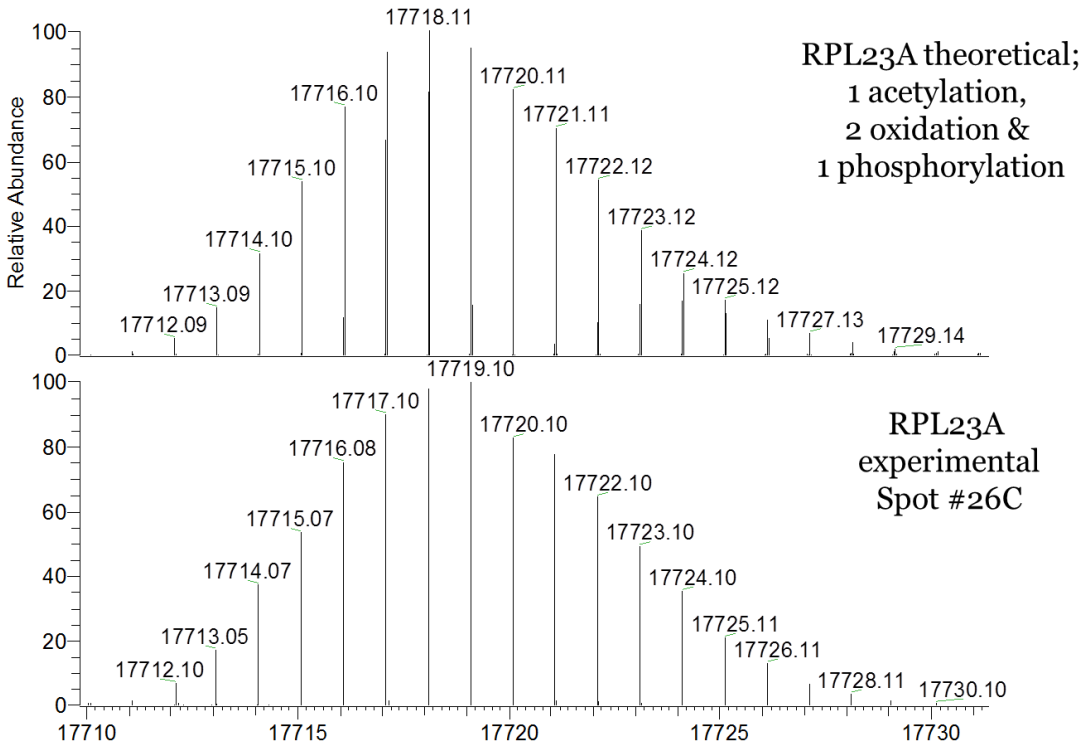

Figure 3.49 Theoretical (top panel) and experimental (bottom panel) molecular mass observed for spot 26C identified as RPL23A by bottom-up analysis with 1 acetylation, 2 oxidation and a phosphorylation

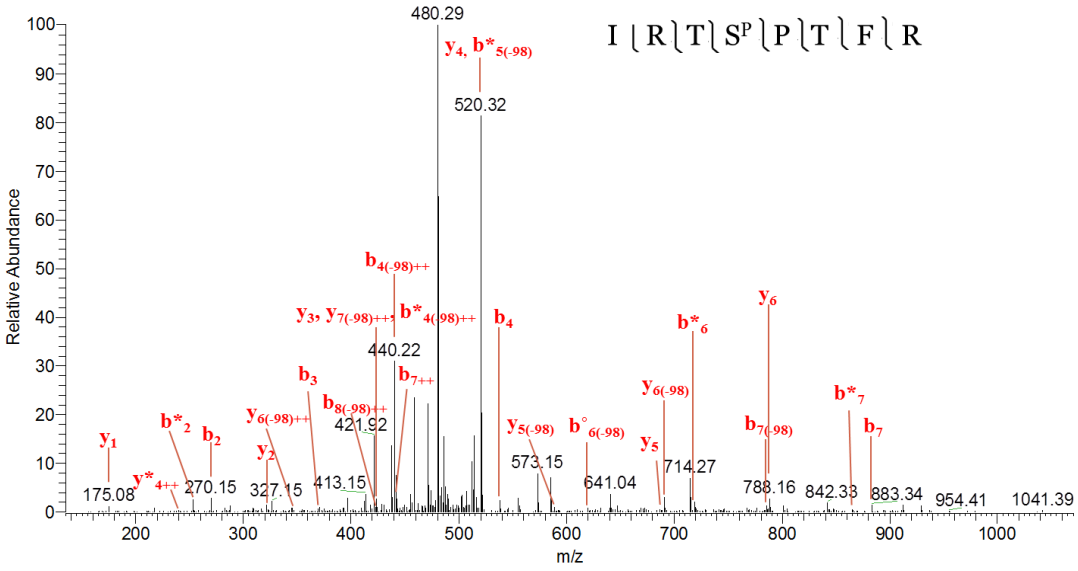

Figure 3.50 MS/MS spectrum of the phosphopeptide observed in RPL23A in spot 26C

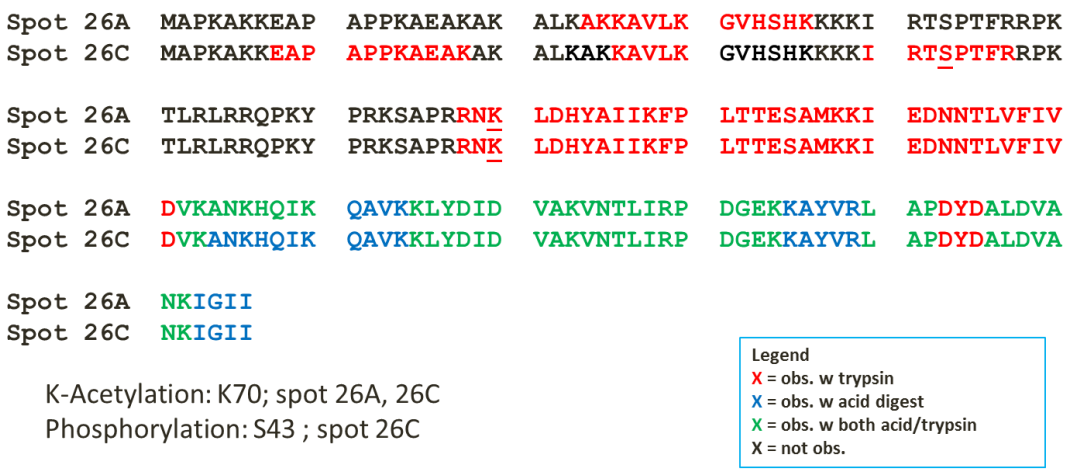

Figure 3.51 RPL23A isoform sequence coverage; underlined residues indicate a modification was observed. Both the S43 phosphorylation and the K70 acetylation are noted in the Uniprot database as common modifications. Artifactual oxidations (resulting from sample handling) were observed on H73 and M87

The two remaining altered protein spots which were investigated in this study were found to be isoforms of RPS10 and labeled as #29 and #39 (see Fig 3.23 and 3.30). The most abundant molecular mass observed for spot #29 is seen in Figure 3.52 and was 18952.91Da. Examination of the digestion data for RPS10 revealed that aside from a 12 residue peptide found on the C-terminal end of the protein, there was complete sequence coverage. A protein modification consistent with the mass change from the theoretical mass of the protein was not observed in the digestion data. Noted in the sequence annotation for RPS10 is dimethylation of two arginine residues on the C-terminal end (R158 and R160), a modification found as a step in the incorporation of RPS10 into the ribosome and for ribosome function as a whole. It is noted in the literature that without the addition of this modification, poor subunit association and enhanced proteasomal degradation of RPS10 was reported[112]. The mass change

associated with the addition of these two dimethylations matches the observed mass change versus the theoretical mass of RPS10 based on the amino acid sequence. The theoretical mass spectrum of the modified RPS10 is shown with that observed for spot #29 in Figure 3.52. The molecular mass measurement of the MXR HPLC fraction containing RPS10 (as determined by trypsin digestion) revealed a highly abundant deconvoluted molecular mass which corresponded with the isoform observed in spot 29. The experimental molecular mass observed for this HPLC fraction sample is shown in Figure 3.53. The molecular mass observed for spot #39 was 16019.49Da and is shown in Figure 3.54. Bottom-up analysis of this protein spot revealed no modifications aside from an oxidation (attributed to sample handling). There was a marked lack of observation of peptides found near the C-terminal end. Given that the observed mass was notably lower than the theoretical mass, the logical first assumption was that the protein was truncated on the C-terminal end. Incorporating those peptides observed during the MXR acid digestion with the tryptic digestion data for spot #39 indicated that all but the last 28 residues of the protein had been observed. If the protein were truncated at that point, the theoretical mass would be 16019.45Da which is only 0.04Da from the observed molecular mass. An oxidized version of this protein was also observed and is shown in Figure 3.55. The sequence coverage of both isoforms is found in Figure 3.56.

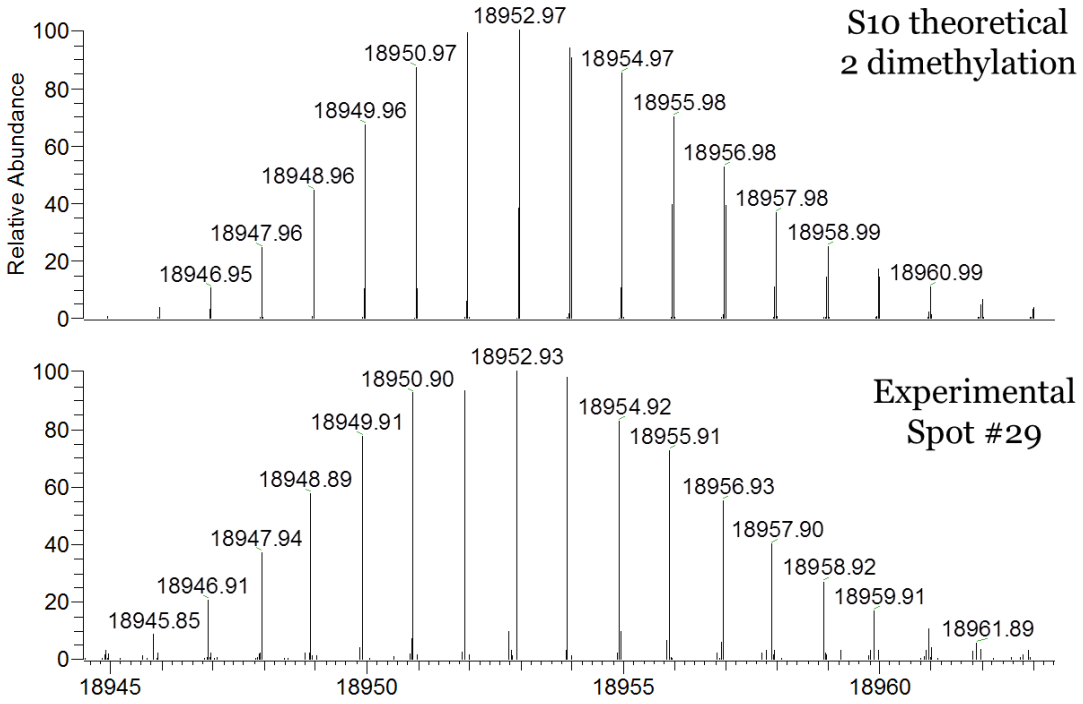

Figure 3.52 Theoretical (top panel) and experimental (bottom panel) mass measurement for spot 29 identified as RPS10 and believed to contain 2 arginine dimethylations

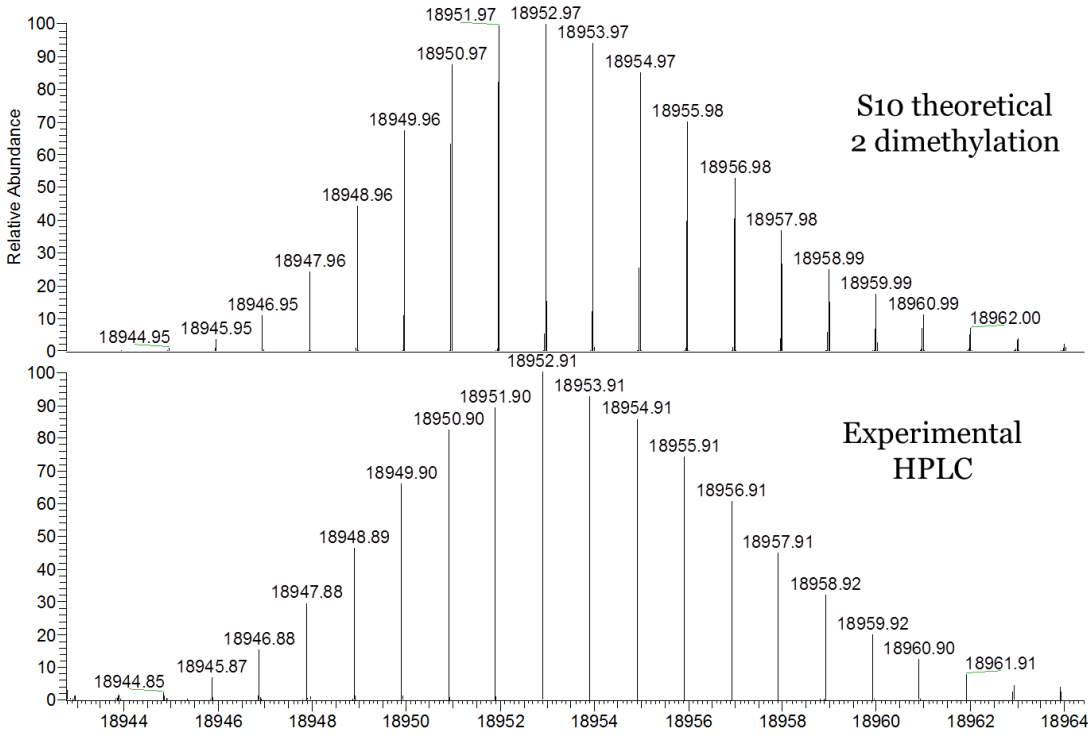

Figure 3.53 Theoretical (top panel) and experimental (bottom panel) molecular mass observed in the MXR HPLC fraction identified as containing RPS10 which corresponded with the isoform of RPS10 found in spot 29

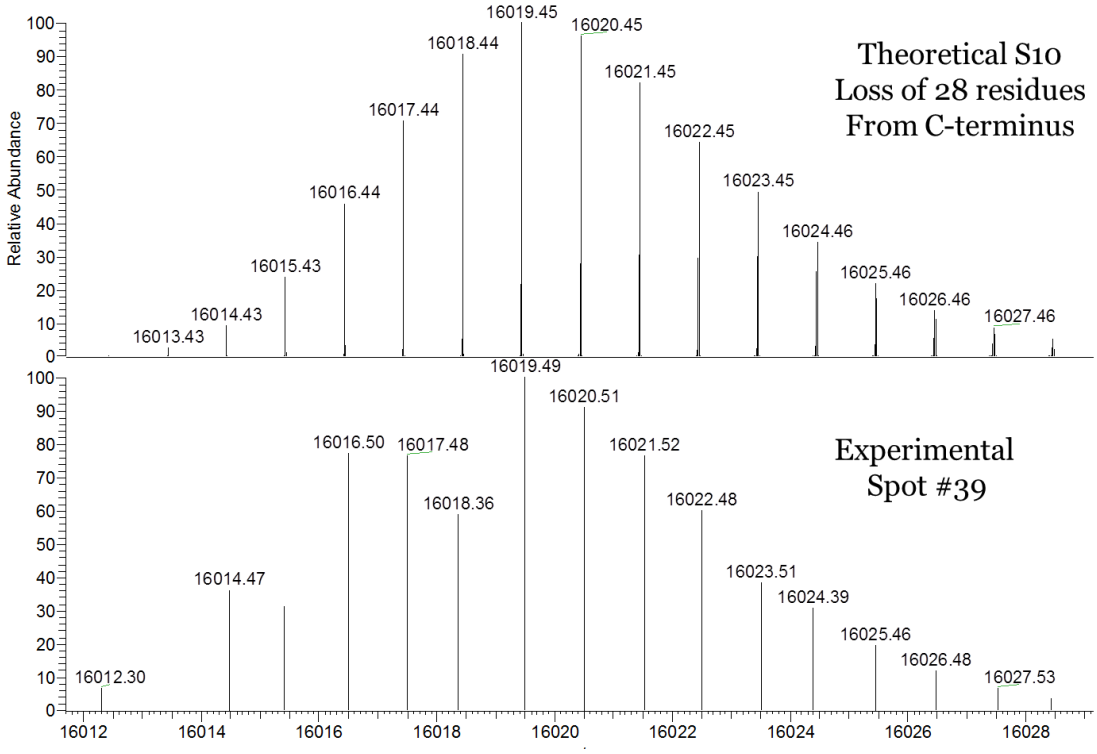

Figure 3.54 Theoretical (top panel) and experimental (bottom panel) mass measurement observed for spot 39 believed to have a C-terminal truncation of 28 residues based on in-gel digestion data

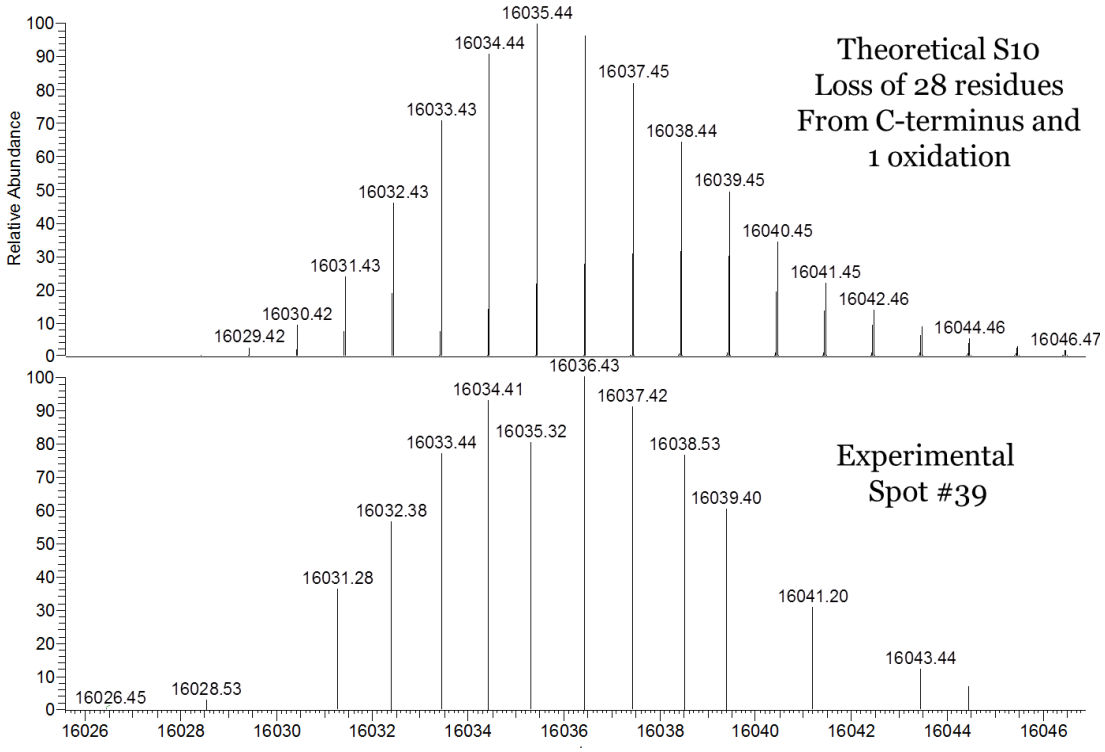

Figure 3.55 Theoretical and additional experimental mass measurement observed for spot 39 showing the proposed C-terminal truncation based on the digestion data and 1 oxidation also seen in the digestion data

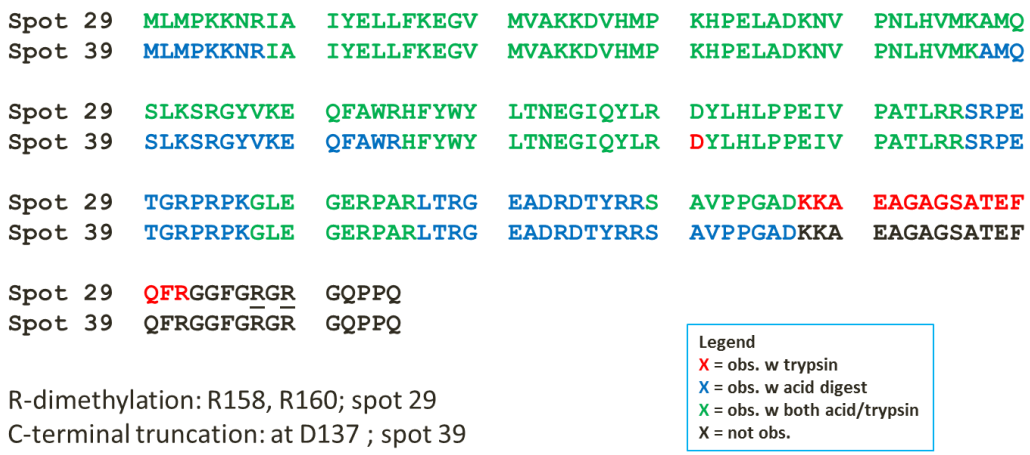

Figure 3.56 Sequence coverage for the two isoforms of RPS10 is shown above with legend. Underlined residue indicates modification (M oxidation)

A summary of the molecular masses of the gel extracted proteins of interest along with their modifications are listed in Table 3.11. As will be discussed later, all the post translational modifications (acetylation, phosphorylation, dimethylation) which were observed (or proposed in the case of dimethylation) have been noted previously in the literature. The molecular mass of all but two of the altered protein isoforms (spot 26B and spot 39) were confirmed in MXR HPLC fraction molecular mass analysis and/or top-down analysis of the MXR ribosomal proteome. It is believed that these variants were not observed in the LC-ESI-Orbitrap analysis due to their abundance in the case of spot 39 or combination of abundance and size in the case of spot 26B.

Gene Name	Spot #	Peptide	Modification
RPS3	6, 7	M.A^{Ac}VISK.K	N-terminal acetylation
RPS3	7	K.DEILPTT^{p}PISEQK.P	Phosphorylation (T)
RPS3	7	K.KPLPDHVSIVEPKDEILPTT^{p}PISEQK.P	Phosphorylation (T)
RPS3	7	K.FVDGLMIHSGDPVNYYDTAVR.H	Oxidation (M)
RPS3	7	K.IMLPWDPTGK.I	Oxidation (M)
RPS3	7	K.GGKPEPPAMPQPVPTA.-	Oxidation (M)
RPL11	25	M.A^{Ac}QDQGEKENPMR.E	N-terminal acetylation
RPL11	26B	M.A^{Ac}DQGEKENPMR.E	N-terminal acetylation, deamidation (Q)
RPL11	26B	K.VLEQLTGQTPVFSK^{Ac}AR.Y	Acetylation (K)
RPL11	26B	R.NEK^{Ac}IAVHCTVR.G	Acetylation (K)
RPL11	26B	K.AEEILEK^{Ac}GLK.V	Acetylation (K)
RPL11	26B	R.ISKEAMR.W	Oxidation (M)
RPL23A	26C	K.IRTS^{P}PTFR.R	Phosphorylation (S)
RPL23A	26A, 26C	R.RNK^{Ac}LDHYAIIK.F	Acetylation (K), oxidation (H)
RPL23A	26A, 26C	K.FPLTTESAMK.K	Oxidation (M)

RPS10	39	K.NVPNLHVMK.A	Oxidation (M)

Table 3.10 Modified peptides found in the altered protein isoforms including oxidations (oxidations were the result of sample handling and are listed here only as a result of the molecular mass being the oxidized version of the protein)

#	Gene Name	Theoretic. MW	PTM's	Theoretical MW w/ PTM's	Experimental MW (most abundant)	Delta mass
6	RPS3	26553.43	N-term acetylation	26598.46	26598.45	0.01
7	RPS3 + alkylation with IAA	26724.47	N-term acetylation, phosphorylation, 3 oxidations (sample handling)	26897.48	26897.52	0.04
8	RPS3 + alkylation with IAA	26724.49	Incomplete alkylation	26667.47	26667.45	0.02
25	RPL11 isoform 1	20117.63	N-terminal acetylation	20162.62	20161.61	1.01
26B	RPL11 isoform 2 + alkylation with IAA	20217.65	N-terminal acetylation, 3 lysine acetylation, 1 deamidation, 2 oxidations (sample handling)	20425.70	20425.61	0.09
26A	RPL23A	17563.15	Lysine acetylation, 2 oxidation (sample handling)	17637.14	17638.14	1.00
26C	RPL23A	17563.15	Lysine acetylation, 2 oxidation (sample handling), phosphorylation	17718.11	17719.10	0.99
29	RPS10	18896.87	2 Arg dimethylations	18952.97	18952.93	0.04
39	RPS10	18896.87	28 residue C-terminal truncation	16019.45	16019.49	0.04

Table 3.11 Molecular masses of protein isoforms found in altered abundance between the MXR and MXS cell lines as indicated by comparative densitometry (spot 7 was used as a reference spot). Incomplete alkylation (by IAA) was observed in many of the gel spots. The most abundant molecular mass was used and modifications considered only if relevant to other experimental findings

Chapter 4: Discussion

The ribosome is not a static organelle. Alteration in the structure of rRNA and ribosomal protein composition have been noted in relation to everything from inherent traits such as cell type and developmental state (embryo vs. adult) to external conditions such as physiological stress on the cell or organism[77; 82; 87; 88; 89; 113; 114; 115]. This ribosome heterogeneity has been proposed to impact ribosome function. As previously discussed, the ribosome filter hypothesis proposes that mechanisms involving differential mRNA capture allow for the ribosomal subunits to affect the translation of particular mRNAs. A principle behind the ribosome filter hypothesis is that heterogeneity in the structure of a ribosomal protein or proteins may impact accessibility of binding sites within the ribosome (i.e. the mRNA exit channel)[82; 83]. Kondrashov *et al* found that a mutation in one ribosomal protein (RPL38) influenced the recruitment of HOX mRNA (critical in skeletal development) to the ribosome during embryonic development[89]. Mutations in other select ribosomal proteins (RPS19, RPS20, RPL29, and RPL24) did not have the same effect on the HOX mRNA. It was concluded from functional and biochemical studies that RPL38 was able to exert a well-defined role in translational control of HOX mRNA as a constituent of the ribosome. The location of RPL38 (near expansion segment 27) in the eukaryotic ribosome was proposed to exert a conformational change in the ribosome where it might influence accessibility to subsets of mRNA. The specialized role RPL38 was found to play in translational control led to Kondrashov and colleagues to refer to these ribosomes as "specialized ribosomes"[82; 83; 89]. An

investigation into the role that post translational modifications play in "specialized ribosomes"/heterogeneous ribosomes has not yet been published.

Modifications of ribosomal proteins such as those noted in this study (truncation, phosphorylation, acetylation, etc.) have been noted as distinguishing features of ribosomal proteins in disease and studies of the cell cycle [25; 73; 74; 75; 79; 87; 90]. Given the role the ribosome plays in the regulation of the cell cycle and the link it has to some forms of drug resistance in prokaryotic organisms, the goal of this investigation was to compare the ribosomal proteome in chemotherapeutic resistant human cell lines with a drug susceptible cell line. A combination of bottom-up, middle-down and top-down approaches was used to characterize the ribosomal proteins found in the MXR and MXS cell lines.

Comparison of methods

Number of ribosomal proteins

Bottom-up vs. middle-down characterization of the MXR ribosome

In a bottom-up proteomic strategy, the proteins of interest are first digested using enzymatic or chemical cleavage. These peptides are then analyzed by one of several mass spectrometric platforms. One of the most common approaches traditionally first involves the separation of the protein mixture (in this case the ribosomal proteome) by 2-D gel electrophoresis followed by individual gel spots being subjected to tryptic digestion and analyzed by tandem mass spectrometry. Another common workflow requires digestion of the entire protein mixture, fractionation of the peptides via multi-step chromatography and analysis using

tandem mass spectrometry. Each of these methods produces MS/MS spectra, which can then be searched against a given database[52; 57]. Both of these approaches were used; the first, as described, was used to compare the two cell lines to identify proteins with altered abundances. The second method was used, with a slight variation, in that the ribosomal proteome was first fractionated with RP-HPLC prior to digestion with trypsin of each fraction and tandem mass spectrometry.

The 2-D gel arrays produced in this study were optimized through the use of cup-loading in the first dimension. Cup-loading is widely recommended for basic proteins in 2DGE[93; 99; 100]. Previous research in this laboratory involved the 2DGE of the MXR ribosomal proteome and the annotated gels from that study were used for comparison and spot identification[1]. All of the spots selected from the gel in this study; those identified as being in altered abundance and the spots used for reference and method validation, were cross-referenced with the gels produced from the previous study (refer to Fig 3.23). Proteins identified by digestion in the 23 gel spots evaluated in the current investigation all correlated with identifications made in the previous study with the exception of one protein spot, spot 43. Given that more than one protein can be present in a protein spot, it is very likely that the identification made in the earlier study is valid. Two additional proteins were identified with the bottom-up analyses of these 23 gel spots; RPL23A and RPS7. In total, the combined 2-D gel array studies allowed for the annotation of 50 ribosomal proteins and 2 ribosomal associated proteins listed in Table 4.1.

Spot#	*Protein*	*Average MW*	*Theoretical pI*
1	Polyadenylate Binding Protein	70671	9.52
2	RPL3	45978	10.19
3	RPL6	32597	10.59
4	RPL5	34231	9.73
5	RPS3A	29814	9.75
6	RPS3	26557	9.68
7	RPS3	26557	9.68
8	RPS3	26557	9.68
9	RPS4X	29467	10.16
10	RPL10A	24700	9.94
	RPL13	24130	11.65
11	RPL7A	29864	10.61
12	RPS6	28681	10.85
13	RPL8	27893	11.04
14	RPL7	29226	10.66
15	RPS8	24074	10.32
16	RPL14	23301	10.94
17	RPS2	31193	10.25
18	RPL10	24473	10.11
19	RPL13A	23446	10.94
20	RPL9	21863	9.96
	RPS7	22127	10.09
21	RPS5	22876	9.73
22	RPS9	22460	10.66
23	RPL21	18434	10.49
24	RPL17	21266	10.18
25	RPL11 iso1	20121	9.64
26A	RPL23A	17564	10.44
26B	RPL11 iso2	20121	9.64
26C	RPL23A	17564	10.44
27	RPL26	17258	10.55
28	RPS11	18300	10.31
29	RPS10	18898	10.15
30	RPS15	16909	10.39
	RPS16	16314	10.21
	RPL27	15667	10.56
	RPS25	13742	10.12
	RPS15A	14708	10.14
31	RPS25	13742	10.12
32	RPS17	15419	9.85
33	RPS18	17587	10.99
34	RPS13	17091	10.53
35	RPL27	15667	10.56
36	RPL31	14463	10.54
	RPS14	16142	10.08
37	RPS16	16314	10.21
38	RPS20 iso2	16006	9.40

39	RPS10	18898	10.15
40	RPL30	12653	9.65
41	RPS19	15929	10.31
	RPS15A	14708	10.14
42	RPS15A	14708	10.14
43*	RPS15A*	14708	10.14
	RPS10*	18898	10.15
44	RPL23	14865	10.51
45	RPL35A	12538	11.07
46	RPL38	8087	10.10
47	RPS29 iso1	6546	10.17
48	RPL37A	10144	10.44
49	RPL22	14656	9.22
50	RPL36A	12310	10.56
51	RPL24	17779	11.26
52	FKBP3 (FK506)	25046	9.29
53	RPL12	17819	9.48

Table 4.1 Proteins identified by gel array in the current study and previous study[1] using the bottom-up approach. There was agreement between the two studies on all but one protein spot *43; identified as RPS10 in the previous study and RPS15A in the current investigation

Bottom-up analysis using fractionation at the protein level by RP-HPLC followed by trypsin digestion and tandem mass spectrometry via LC-ESI-LTQ-Orbitrap of the fractions identified a larger number of proteins than the 2D-gel arrays. Of the 79 human ribosomal proteins, 62 proteins were confidently identified by two or more unique peptides and 71 were confidently identified by one peptide. There was however differences between the proteins identified in these two methods as shown in Figure 4.1. A comparison of the proteins identified by bottom-up analysis using 2-D gel arrays and in-gel digestion with the RP-HPLC fractionation of the ribosomal protein mixture followed by digestion of the fractions revealed that 4 unique proteins were identified by bottom-up analysis in the 2-D gel arrays and 25 unique proteins were identified in the HPLC fractions when proteins identified by one peptide were included. When only those proteins identified by two or more peptides

were examined, there were 18 unique proteins identified by the HPLC fraction digestion analysis versus 6 identified by the gel. These differences could possibly be attributed to the limitations set on the proteins in the first dimension of the 2-D gel electrophoresis as they are enriched for a selected pI range; IPG strips in the range of 7-11 were used.

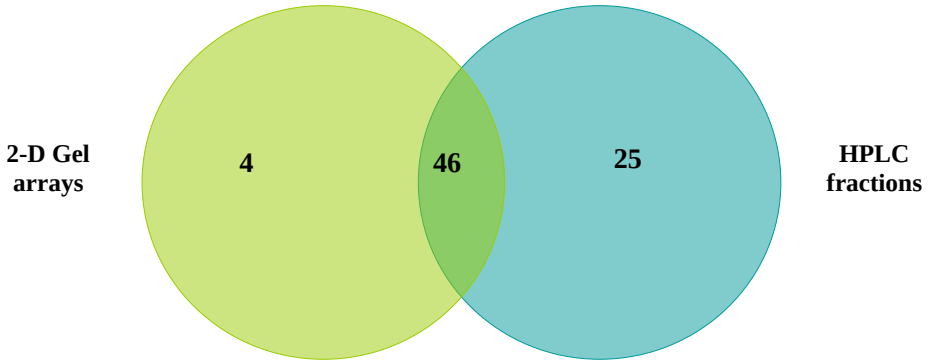

Figure 4.1 Venn diagram comparing the protein identifications made with bottom-up analysis by 2-D gel array and in-gel digestion vs. RP-HPLC fractionation of the proteins followed by in-solution trypsin digestion of the fractions

Middle-down analysis was another approach used to identify the proteins in the MXR ribosomal proteome. The middle-down approach typically takes advantage of enzymatic or chemical cleavage with selectivity for a single amino acid residue. The resulting proteolytic products produced from the proteins of interest are large, with polypeptides typically observed between 3kDa to 10kDa in size. Larger

polypeptides have been observed to fractionate with improved resolution by HPLC. These peptides can be analyzed using a variety of mass spectrometric platforms, often through a combination of methods used in top-down and bottom-up proteomics based on the nature and complexity of the sample being investigated. The approach used in this investigation involved microwave accelerated acid digestion of the MXR ribosomal proteome which cleaves on either side of aspartic acid residues as discussed previously by this laboratory[64]. This was followed by fractionation of the peptides via nanoLC interfaced with a Thermo LTQ-Orbitrap, for high resolution (survey scans acquired at 30K resolution) tandem mass spectrometry of both precursor and product ions. Large polypeptides carry a higher number of charges when electrosprayed which enhances CID. The middle-down approach is especially favored in the investigation of proteins containing multiple PTMs or proteomes which contain multiple protein isoforms such as the current study[2; 62; 63; 64; 65; 66].

Middle-down analysis with the acid digestion products of the MXR ribosomal proteome allowed for the confident identification of 62 of the 79 human ribosomal proteins by two or more unique peptides (70 proteins were confidently identified by one peptide). When compared with the bottom-up methods, the proteins identified again differ between the methods. This is demonstrated in the Venn diagrams seen in Figure 4.2 and 4.3. Based on these results, it would appear that an equal number of proteins could be identified with samples collected by HPLC fractionation and trypsin digestion versus a sample of the entire MXR ribosomal proteome cleaved with microwave accelerated acid digestion. A combination of 2-D gel array in-gel digestion, acid digestion and the samples from HPLC fractionation allowed for the

identification of a total of 76 of the 79 ribosomal proteins. As can be seen in the Venn diagrams and Table 4.1, if only one of these methods could be employed, the number of proteins identified by the middle-down approach is very comparable to the HPLC fractionation- in-solution trypsin digestion and provides a more rapid sample preparation time. The main disadvantage of using the middle-down approach as will be discussed with regards to sequence coverage is the fact that protein isoforms are more difficult to distinguish using this method.

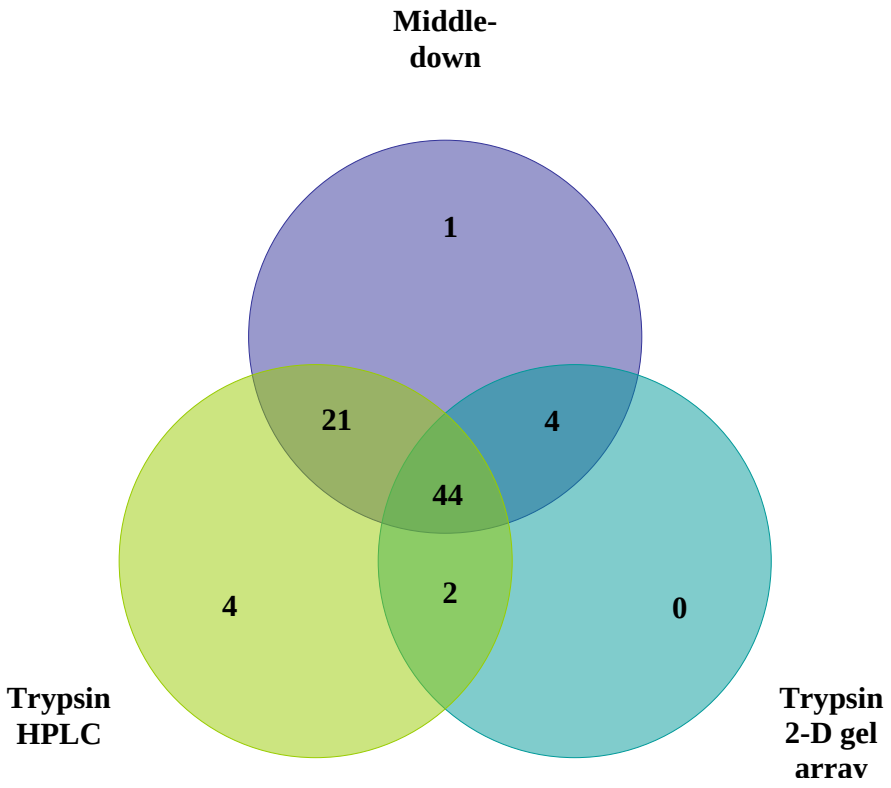

Figure 4.2 Venn diagram illustrating overlap/differences between the protein identifications made by bottom-up methods and a middle-down analysis of microwave accelerated acid digestion products. These are protein identifications based on *1 confidently identified peptide*

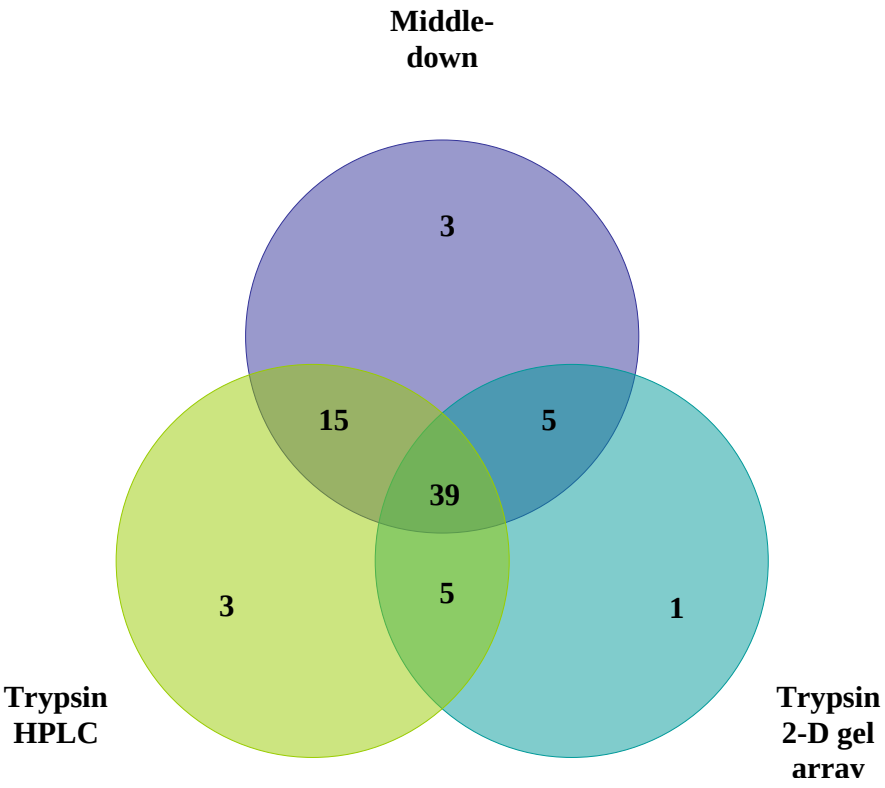

Figure 4.3 Venn diagram illustrating overlap/differences between the protein identifications made by bottom-up methods and a middle-down analysis of microwave accelerated acid digestion products. These are protein identifications based on *2 or more confidently identified peptides*

Sequence coverage of the ribosomal proteins and their modifications

Optimal sequence coverage of the ribosomal proteins was achieved in this investigation by a combination of the results from three digestion methods; in-solution trypsin digestion of HPLC fractions, in-gel digestion of the MXR protein isoforms in the 2-D gel arrays and microwave accelerated acid digestion of the entire MXR proteome. For bottom-up analysis based around HPLC fractionation of the ribosomal proteome, the average sequence coverage for the 71 proteins identified was 33% with a range for peptide coverage of individual proteins from 6% to 95%. The

23 protein spots which were chosen for this investigation from the 2-D gel arrays had average sequence coverage (solely from the trypsin digestion) of 62% with a range of peptide coverage for individual proteins from 25% to 92%. When the sequence coverage observed from the previous study of the in-gel digestion of ribosomal proteins is taken into account (50 ribosomal proteins identified in total), the average sequence coverage becomes 65% with a peptide coverage range for individual proteins that remains from 25% to 92%[1]. The average sequence coverage of the 70 ribosomal proteins observed using acid digestion was 52% with a peptide coverage range for individual proteins from 6% to 99%[2]. A comparison of the sequence coverage observed for each of the ribosomal proteins based on these three methods and based on the isoelectric point of the ribosomal protein is found in Figures 4.4, 4.5 and 4.6. Under ideal circumstances, the peptide products from these methods would be complimentary (non-redundant). By combining the peptides identified by the different methods for a given protein the average individual ribosomal protein sequence coverage was increased to 75% of the 76 ribosomal proteins identified by a minimum of 1 confident peptide identification and 78% for the 72 ribosomal proteins with two or more confident peptide identifications. The peptide sequence coverage range for individual proteins was from 8% to 100% for the 76 proteins confidently identified by 1 peptide and from 19% to 100% for the 72 proteins confidently identified by a minimum of 2 peptides as can be seen in Table 4.2.

When the digestion methods are examined on their own, the cost and benefit of each relates to the goal of the investigation and the desire to localize post translational modifications. Based on the results of this investigation, the location of

a given post translational modification is also of significance. Acid digestion provided the majority of protein sequence coverage information on the N-terminal of a given protein (perhaps due in part to this area being more readily accessible for chemical cleavage). Many ribosomal proteins lose their N-terminal methionine and are subsequently acetylated on the new N-terminal residue[115; 116] and the use of acid digestion allowed for the identification of these N-terminal peptides. As can be seen in Table 4.2, digestion with these three workflows allowed for the identification of the N-terminus of 52 proteins (not including protein variants). The N-terminal fragments of an additional 7 proteins were exclusively identified with top-down fragmentation of intact proteins (18 proteins total) which will be discussed in a subsequent section. Identifications are noted in Table 4.2 by a letter; G signifies in-gel digestion, A signifies acid digestion, H signifies HPLC fraction trypsin digestion and T signifies top-down fragmentation. Acid digestion allowed for the identification of 39 N-terminal peptides with 18 of these peptides being exclusively observed by acid digestion. By comparison, 19 N-terminal peptides were identified by the in-gel digestion workflow (not counting N-terminal peptide variants) with only 5 observations unique to that method. Similarly, 17 N-terminal peptides were identified with the HPLC fractionation and digestion workflow with 4 exclusive to that workflow. Of the 52 N-terminal peptides identified by these three digestion methods, 40 were observed to have lost their N-terminal methionine and 19 were acetylated (4 of which were N-acetyl-L-methionine). Twelve of the 19 N-terminal acetylations were observed with acid digestion with 6 of those being unique to acid digestion. Although the acid digestion was more effective at producing peptides

found on the N-terminal end of the protein sequence, the in-gel digestion also allowed for the identification of variants of a ribosomal protein produced by alternative processing of the N-terminal end of a protein. This was noted for acid digestion in the case of RPS3 where the N-terminal peptide was observed with and without acetylation. For in-gel digestion, two forms of RPS5 were noted, one with the N-terminal methionine and the second without the N-terminal methionine (the version without N-terminal methionine is acetylated). This variable processing of the RPS5 N-terminus has been noted in several references listed in the UniProt knowledge database entry for this protein[117].

Post translational modifications found within the protein (as opposed to on the N-terminal end) were best identified using the 2-D gel array and in-gel digestion workflow. A somewhat less superior alternative was in-solution digestion of the HPLC fractions. The modifications identified (listed in Table 4.2) included phosphorylation, lysine acetylation and a deamidation (methionine oxidation was also identified but was not listed as it is considered the result of sample handling). In-gel digestion identified a phosphorylation site on RPL23 at T64. This has been noted previously by Rigbolt and colleagues in an investigation of human embryonic stem cells[118; 119]. The intact molecular mass noted for RPL23 in the HPLC fraction collected at 27 minutes did not account for a phosphorylation leading to the conclusion that the phosphorylated form of the protein was less abundant or found in another HPLC fraction. Another example of a modification observed by in-gel digestion is the phosphorylation of S54 in RPL10. This modification has been noted as having been detected by mass spectrometry in a curated modification database

(PhosphoSitePlus®) in 4 different human samples which included the Jurkat cell line (T-cell leukemia) and the K562 cell line (chronic myelogenous leukemia)[119]. Four search engines supported evidence of a phosphorylation found on T123 of RPS7. This phosphorylation has been noted in the PhosphoSitePlus® database for 5 human samples and 7 mouse samples[119]. Human samples in this study with this modification were listed as HELA cells (cervical carcinoma), NCI-H3255 (non-small cell lung cancer) and MKN-45 (gastric carcinoma)[120].

Despite these findings and the fact that the in-gel digestion provided the highest average peptide sequence coverage for individual proteins, the preparation of the 2-D gel arrays was labor intensive and contamination of the protein spots (inadvertently from sample handling, carry-over between LC-MS/MS runs, and the presence of more than one protein in a spot) posed concerns. Additionally, the number of proteins identified by the 2-D gel array workflow was impacted in part by the restrictions placed on the pI range of the proteins investigated in the first dimension (See Fig 4.5). The PTMs confidently identified using in-solution digestion of the HPLC fractions was not as numerous as can be seen in Table 4.2 which may be attributed to the average sequence coverage of these proteins being much lower than the other two digestion methods (33%). This would be expected to be improved with more concentrated samples and repeated sample injections[121].

In the context of this investigation, the largest amount of information about the MXR protein isoforms was obtained using 2-D gel arrays and in-gel digestion. The contribution of the MXR proteome acid digestion cannot be understated, particularly concerning the regions of these proteins which lacked sequence coverage

such as the N-terminal regions. In most cases, the proteins where information was lacking from the in-gel digestion regarding the N-terminal end involved proteins with a basic residue near the N-terminal end. An alternative approach to sequencing these proteins could have been to use in-gel digestion with another endoproteinase such as Lys-C.

ID	Average Molecular Weight	Swiss-prot ID	pI	% sequence coverage HPLC fractions	% sequence coverage acid digest	% sequence coverage in-gel digest	Total % sequence coverage
RPL10	24473	P27635	10.11	17%	24%	54%	72%
RPL10a	24700	P62906	9.94	40%	47%	72%	84%
RPL11	20121	P62913	9.64	23%	19%	92%	92%
RPL12	17819	P30050	9.48	59%	23%	74%	81%
RPL13	24130	P26373	11.65	40%	58%	42%	79%
RPL13A	23446	P40429	10.94	50%	6%	71%	77%
RPL14	23301	P50914	10.94	28%	39%	43%	67%
RPL15	24015	P61313	11.62	14%	47%	0%	52%
RPL17	21266	P18621	10.18	25%	64%	57%	91%
RPL18	21503	Q07020	11.73	32%	18%	0%	43%
RPL18A	20762	Q02543	10.72	30%	43%	0%	53%
RPL19	23466	P84098	11.48	20%	39%	0%	47%
RPL21	18434	P46778	10.49	22%	61%	77%	94%
RPL22	14656	P35268	9.22	10%	66%	50%	91%
RPL23	14865	P62829	10.51	25%	92%	75%	100%
RPL23A	17564	P62750	10.44	44%	26%	59%	78%
RPL24	17779	P83731	11.26	0%	59%	66%	87%
RPL26	17258	P61254	10.55	30%	69%	59%	92%
RPL27	15667	P61353	10.56	35%	95%	67%	100%
RPL27A	16430	P46776	11	31%	68%	0%	69%
RPL28	15616	P46779	12.02	32%	63%	0%	76%
RPL29	17621	P47914	11.66	32%	37%	0%	54%
RPL3	45978	P39023	10.19	6%	44%	62%	76%
RPL30	12653	P62888	9.65	34%	11%	84%	87%
RPL31	14463	P62899	10.54	32%	44%	66%	77%
RPL32	15729	P62910	11.32	15%	30%	0%	45%

RPL34	13162	P49207	11.48	7%	24%	0%	24%
RPL35	14420	P42766	11.04	44%	97%	0%	100%
RPL35A	12538	P18077	11.07	6%	33%	41%	50%
RPL36	12123	Q9Y3U8	11.59	6%	98%	0%	99%
RPL36A	12310	P83881	10.56	15%	47%	49%	80%
RPL37	10947	P61927	11.74	-	-	-	-
RPL37A	10144	P61513	10.44	29%	0%	82%	82%
RPL38	8087	P63173	10.1	53%	83%	54%	83%
RPL39	6275	P62891	12.55	20%	0%	0%	20%
RPL4	47566	P36578	11.07	37%	42%	0%	58%
RPL40	6181	P62987	10.32	-	-	-	-
RPL41	3456	P62945	12.96	-	-	-	-
RPL5	34231	P46777	9.73	17%	49%	68%	79%
RPL6	32597	Q02878	10.59	55%	64%	25%	83%
RPL7	29226	P18124	10.66	43%	24%	72%	90%
RPL7a	29864	P62424	10.61	38%	85%	59%	97%
RPL8	27893	P62917	11.04	32%	68%	57%	87%
RPL9	21863	P32969	9.96	23%	49%	72%	90%
RPLP0	34274	P05388	5.7	42%	27%	0%	42%
RPLP1	11383	P05386	4.21	56%	0%	0%	56%
RPLP2	11665	P05387	4.38	96%	74%	0%	100%
RPS10	18898	P46783	10.15	41%	73%	78%	93%
RPS11	18300	P62280	10.31	46%	78%	76%	96%
RPS12	14384	P25398	7.01	34%	62%	0%	72%
RPS13	17091	P62277	10.53	69%	98%	83%	99%
RPS14	16142	P62263	10.08	0%	50%	46%	68%
RPS15	16909	P62841	10.39	12%	58%	64%	76%
RPS15A	14708	P62244	10.14	0%	88%	62%	98%
RPS16	16314	P62249	10.21	49%	84%	60%	97%
RPS17	15419	P08708	9.85	0%	44%	85%	97%
RPS18	17587	P62269	10.99	52%	39%	69%	76%
RPS19	15929	P39019	10.31	48%	52%	65%	82%
RPS2	31193	P15880	10.25	34%	33%	55%	79%
RPS20	16006	P60866	9.4	29%	40%	88%	79%
RPS21	9111	P63220	8.68	8%	52%	0%	60%
RPS23	15676	P62266	10.5	24%	38%	0%	55%
RPS24	15423	P62847	10.79	9%	55%	0%	55%
RPS25	13742	P62851	10.12	66%	58%	52%	78%
RPS26	12884	P62854	11.01	21%	70%	0%	72%
RPS27	9330	P42677	9.58	0%	32%	0%	32%
RPS27A	9418	P62979	9.86	8%	0%	0%	8%

RPS28	7841	P62857	10.7	33%	99%	0%	100%
RPS29	6546	P62273	10.17	34%	0%	87%	91%
RPS3	26557	P23396	9.68	46%	30%	91%	98%
RPS30	6648	P62861	12.15	19%	0%	0%	19%
RPS3A	29814	P61247	9.75	33%	53%	47%	76%
RPS4X	29467	P62701	10.16	19%	73%	72%	94%
RPS5	22876	P46782	9.73	40%	12%	70%	92%
RPS6	28681	P62753	10.85	27%	16%	60%	61%
RPS7	22127	P62081	10.09	62%	52%	65%	81%
RPS8	24074	P62241	10.32	14%	44%	59%	72%
RPS9	22460	P46781	10.66	65%	43%	87%	94%
RPSA	32723	P08865	4.79	42%	68%	0%	80%

Table 4.2 A list of the 79 human ribosomal proteins along with their average molecular weight (as determined by Expasy), their theoretical pI (also determined by Expasy) and sequence coverage observed for each of three methods; HPLC-fractionation and in-solution trypsin digestion, microwave accelerated acid digestion[2] and in-gel digestion of 2-D gel arrays and the sequence coverage observed when these observations were combined. Those in-gel digestion values reported in RED were determined from in-gel digestion analysis previously reported from this laboratory[1; 2]. Those values underlined were determined by a single peptide

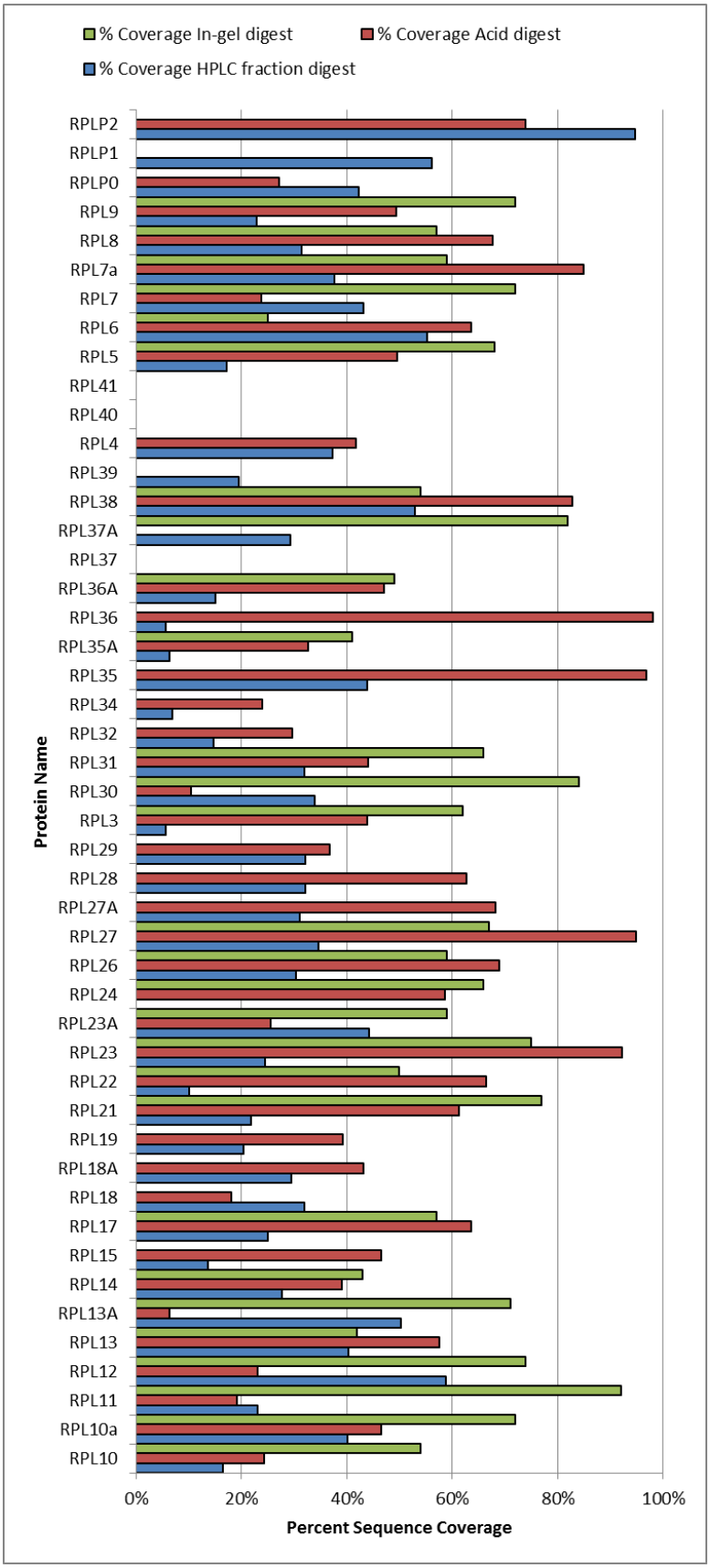

Figure 4.4 Sequence coverage of the ribosomal proteins found in the large subunit using 3 methods

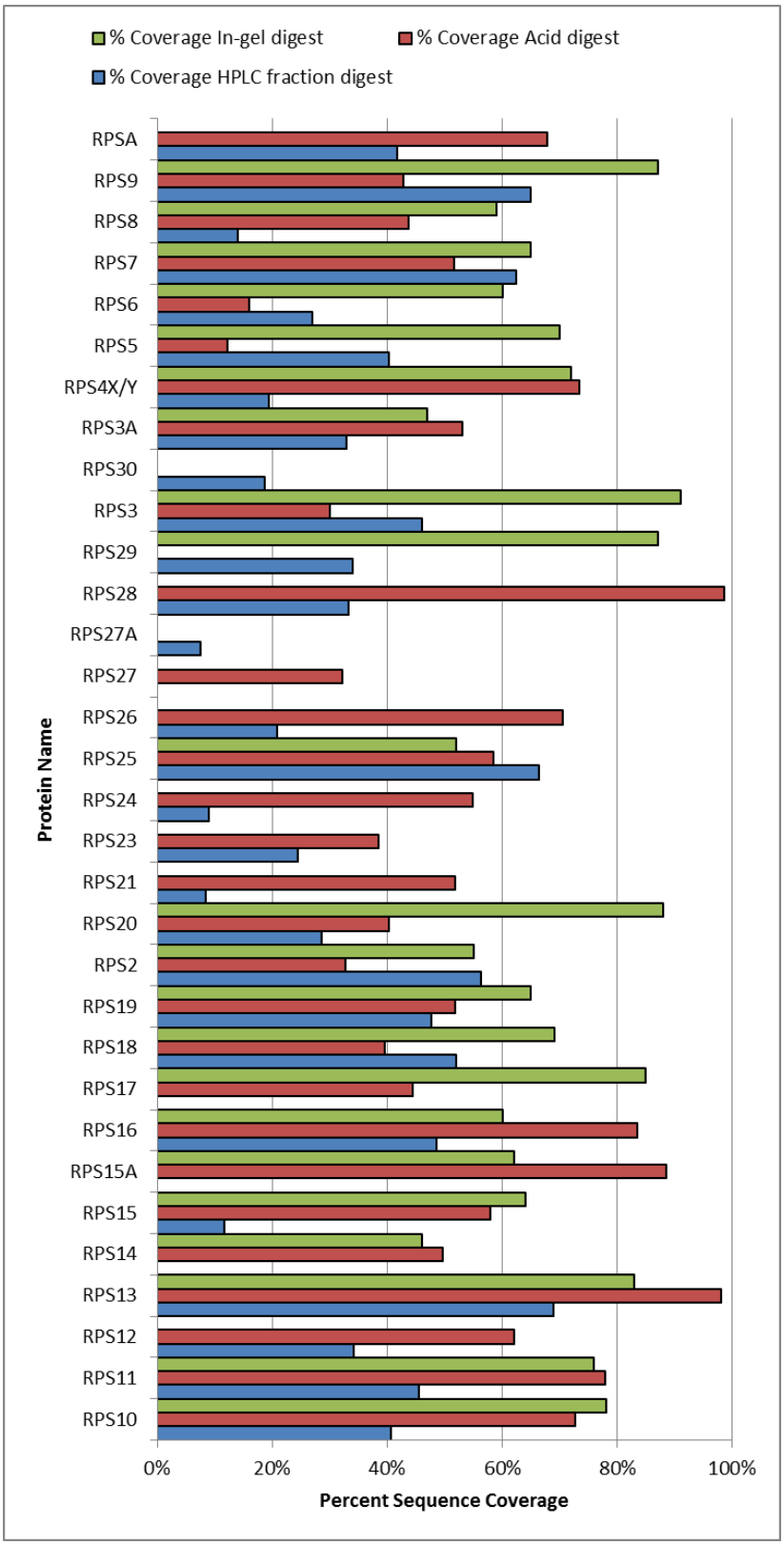

Figure 4.5 Sequence coverage of the ribosomal proteins found in the small subunit using 3 methods

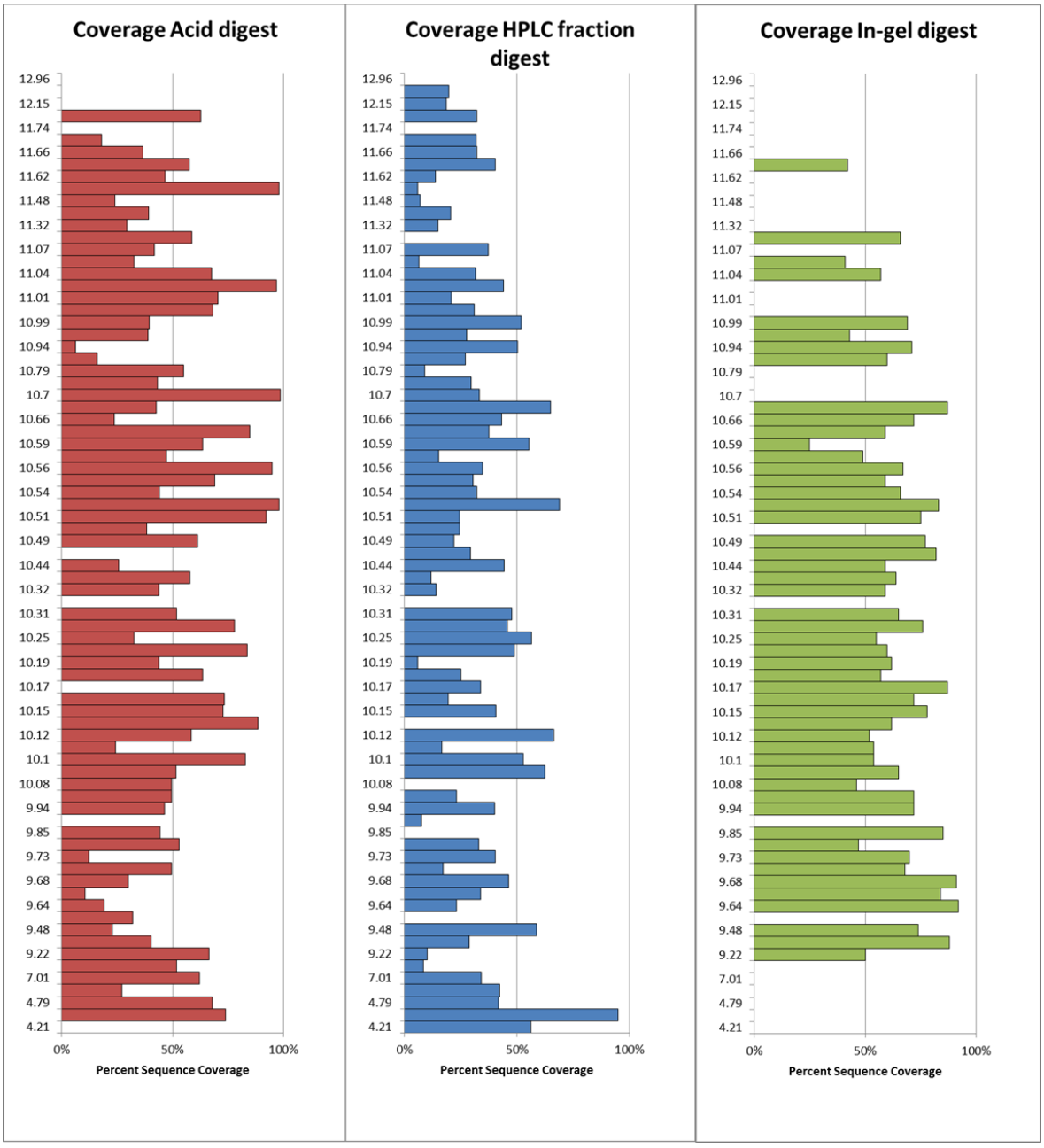

Figure 4.6 Sequence coverage of the ribosomal proteins as determined by their isoelectric point using each of the 3 methods

Protein	N-term obs.	- Met	N-term Acetyl	+ Phos	Deam	K Acetyl
RPL10	A; G	A; G		G		
RPL10a	G	G				H
RPL11	G; H	G; H	G; H		G	G; H
RPL13	A	A				
RPL13A	G	G	G			
RPL14	A	A				
RPL15	A	A				
RPL17	G	G				
RPL19	A	A	A			
RPL21	G; A					
RPL22	A	A				
RPL23	A	A	A	G		
RPL23A	-			G		G
RPL24	G; A					
RPL26	A					
RPL27	A; T	A; T				
RPL28	A; T	A; T	A; T			
RPL3	A; G	A; G				
RPL30	T	T				
RPL32	A; T	A; T				
RPL35	A	A				
RPL35A	A; T	A; T	A; T			
RPL36	A	A				
RPL36A	A	A				
RPL37A	T	T				
RPL38	A; T	A; T				
RPL4	H; A	H; A	H; A			
RPL5	A	A				
RPL6	H; G	H; G				
RPL7	G					
RPL7a	A	A				
RPL8	A	A				
RPL9	A					
RPLP0	H		H			
RPLP1	H	H	H			
RPLP2	A; T					H
RPS10	G; A					

RPS11	T; A; G; H	T; A; G; H	T; A; G; H			
RPS13	A; G	A; G				
RPS15	A; H	A; H	A; H			
RPS15A	T	T				
RPS16	T; A; G; H	T; A; G; H				
RPS17	A	A				
RPS18	A; H	A; H	A; H			
RPS19	T; A; H	T; A; H				
RPS2	H					
RPS20	T; A; G	T; A; G	T; A; G			
RPS21	T; H		T; H			
RPS23	T	T				
RPS24	T		T			
RPS27	T	T	T			
RPS28	A; T		A; T			
RPS29	G; H	G; H				
RPS3	A; H; G	A; H; G	A; H; G	G		
RPS3A	G	G				
RPS4X/Y	A	A				
RPS5	H; G	H; G	H; G			
RPS6	A; H			H		
RPS7	A		A	G		
RPSA	H	H	H			

Table 4.3 Protein name and modification identified (by method); G = in-gel digestion, A = acid digestion, H = HPLC fraction in-solution digestion, T = top-down fragmentation

Intact mass measurements and top-down characterization of the MXR ribosome

Intact mass measurements and top-down mass spectrometry was used to characterize the MXR ribosomal proteome. The top-down approach uses mass spectrometry to weigh intact protein ions and multi-stage tandem mass spectrometry to produce sequence tags from large protein fragments. The entire MXR ribosomal proteome was injected for LC-ESI-LTQ-Orbitrap MS/MS analysis. Of one MXR sample injection, 18 ribosomal proteins were identified using this method. An

alternative approach involves only the acquisition of the molecular ion using a high resolution instrument (the LTQ-Orbitrap again in this case) to determine the intact protein mass. Molecular mass measurements were acquired with aliquots of the HPLC protein fractions a well as gel extracted ribosomal proteins.

The ribosomal proteins identified with top-down fragmentation are listed in Table 3.3. Additional structural information such as the retention /loss of the N-terminal methionine and acetylation of N-terminus was obtained on all of these proteins except for three proteins; RPL32 (loss of 14.12Da), RPL35A (gain of 18.08Da) and RPLP2 (gain of 162.16Da). According to the Uniprot database alone, RPLP2 contains as many as 6 phosphorylation sites and two acetylation sites. The lysine acetylation site which was observed on K98 of RPLP2 in the HPLC fraction digest is noted as a frequent site of ubiquitination in both multiple myeloma cells and mouse embryonic stem cells. A link between sites of lysine acetylation and ubiquitination has been proposed by numerous researchers[122; 123; 124].

Molecular mass measurements are not stand-alone measurements. In other words, additional data is required in order to draw conclusions regarding the identity of the protein in question and any modifications that protein may possess. Without detailed information relating the protein being weighed and PTMs found on the protein in question, the wrong conclusions can be made about a particular protein. For example the protein isoform found in spot 39 was identified as being RPS10 and because of the information regarding the sequence coverage observed, the C-terminal truncation found in that isoform was able to be correctly identified. Looking as the molecular mass on its own, however, 16019.49Da, the assumption that a different

ribosomal protein with a lower theoretical mass was present in that spot could have easily been made.

Despite the shortcomings of molecular mass measurements described in the paragraph above, bottom-up analysis on its own can sometimes provide information about a modification which could be attributed to either an artifact of sample processing or a significant biological event. In such circumstances, it is often *only* with molecular mass values that conclusions can be drawn regarding modifications or changes to a protein. As a result, this analysis is best considered as complementary to bottom-up/middle-down.

Proteomic applications

Rapid assessment of the ribosomal proteome

The most efficient stand-alone method to rapidly assess the ribosomal proteome would be to employ a middle-down approach such as the microwave accelerated acid digestion used in this investigation. A more informative approach however would be to use this technique in conjunction with a top-down approach. As discussed above, a drawback of the middle-down approach is the inability to detect certain PTMs and to easily distinguish protein isoforms (without matching top-down measurements). To offset this loss, the results of this study would suggest a good complementary/alternative method would be to visualize the ribosomal proteome using 2-D gel electrophoresis with in-gel digestion using either trypsin or Lys-C to determine proteins with altered abundance prior to a top-down/molecular mass measurement analysis. In a comparative study such as this investigation, the ability

to visually compare and evaluate the proteome in a gel format allows for selectivity of proteins of interest and molecular mass measurements of extracted gel proteins allows for confidence in the identification of PTMs/alteration in protein primary structure. This provided a larger number of peptide identifications with a very small time commitment. Digesting the entire ribosomal proteome though fast and able to be conducted with relative ease, did not allow for peptides to be used to discern between protein isoforms. In addition, certain post translational modifications such as phosphorylation are easily hydrolyzed during microwave accelerated acid digestion, thus losing the ability to identify this PTM with confidence.

Comparison with previous proteomic studies of ribosomal proteins

Every approach applied in this investigation (top-down, bottom-up and middle-down) had been previously utilized in proteomic studies to characterize ribosomal proteins. A common conclusion of many proteomic investigations attempting to characterize a ribosomal proteome is that ribosomal proteins are as the laboratory of David Lubman stated "*notoriously difficult to assess*"[125]. Despite the assumption that all ribosomal proteins should be in equimolar concentration and thus the dynamic range of the investigation should be 1, there exist to my knowledge no investigations that have successfully characterized all the ribosomal proteins expected for a given sample. The most elaborate characterizations of ribosomal proteomes conducted by the Reilly laboratory and Carroll and colleagues detected all but one of the expected ribosomal proteins in their respective investigations (53 of 54 in the case of Reilly and 79 of 80 in the case of Carroll)[72; 107]. The unique chemical

characteristics of these proteins (their basic isoelectric point) are considered to contribute to this problem. For example, one of the 3 proteins that were not detected using any of the methods employed in this investigation is RPL41 (refer to Table 4.2). This is the most basic of the ribosomal proteins with a theoretical pI of 12.96 in the human ribosome (calculated at over 13 in the rat ribosome). In addition it is one of the smallest of the ribosomal proteins with a molecular mass expected around 3456Da. The ribosomal protein isolation techniques utilized for the 2DGE aspect of this investigation has been proposed to cause the loss of smaller ribosomal proteins[126].

The average sequence coverage observed from a combination of the three approaches (bottom-up, middle-down, and top-down) was found to be 76% for the 76 ribosomal proteins detected. There was 75% sequence coverage for the 44 ribosomal proteins observed from the large subunit and 76% for the 32 ribosomal proteins of the small subunit. This surpasses the sequence coverage of many of the other reported proteomic studies of the ribosomal proteome. The study by the Natalie Ahn laboratory of the rat-1 fibroblast small subunit ribosomal proteins observed an average sequence coverage of 59%[73]. The tryptic peptide mass maps of the 53 (of 54 total) *Caulobacter crescentus* ribosomal proteins detected by the Reilly laboratory were reported with an average sequence coverage of 62%[72]. The Leary laboratory reported observing 31 of the 32 small subunit proteins with 97% coverage however the sequence coverage of each protein was not reported in the publication for comparison to our results[74].

Effect on ribosome function

Earlier investigations in this laboratory on the capacity of these MXR ribosomes to function looked at two properties of the MXR cells and ribosomes; the number of ribosomes in the cells as compared with the MXS cells and the translational efficiency of each ribosomal type using methionine incorporation study.

Met incorporation study

Given that there were differences detected in the ribosomal protein composition between the 2 cell lines, an investigation was conducted by Alexey Petrov and Jaclyn Wolff in the laboratory of Professor Jonathan Dinman to determine if differences existed between the rates at which these 2 types of ribosomes processed mRNA. The results of this experiment can be seen in Figure 4.7. Over the 9 hours period, the methionine incorporation measurements demonstrated that the MXR ribosomal activity was decreased by about 25%.

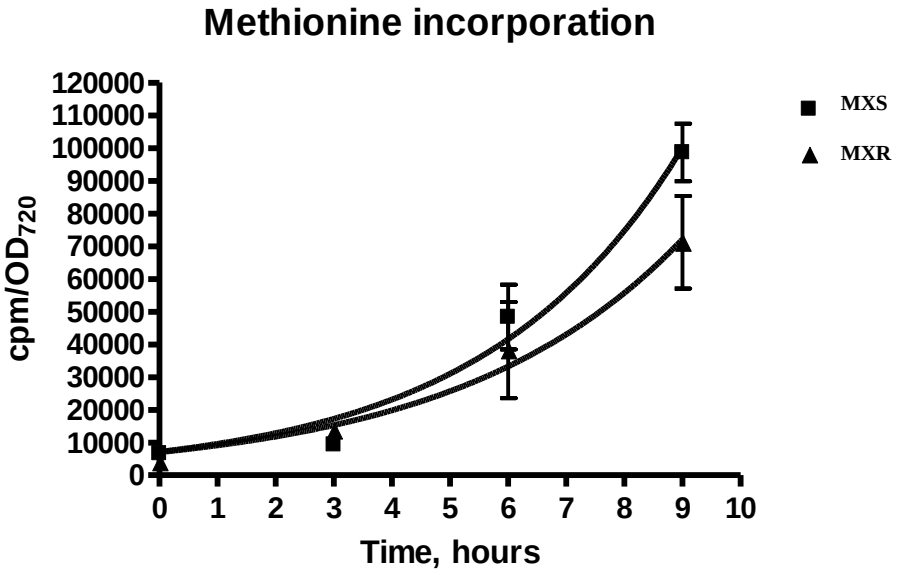

Figure 4.7 The result of the methionine incorporation study illustrated that ribosomal activity in the MXR ribosomes decreased by about 25% over 9 hours when compared with the MXS ribosomes

Ribosome abundance

Given that differences were observed in the ribosomal activity between the two cell lines, it is reasonable to assume that there may be differences in the number of ribosomes in the cells causing this discrepancy. An experiment comparing the number of ribosomes in the cells between the two cell lines was conducted in an earlier investigation as described[1]. This investigation illustrated that in fact there is *no significant difference* in the number of ribosomes between the MXR and MXS cell lines as seen in Figure 4.8. Based on this information, it is clear that a physical feature or features of the ribosome in the MXR cell line was leading to the change in translational efficiency.

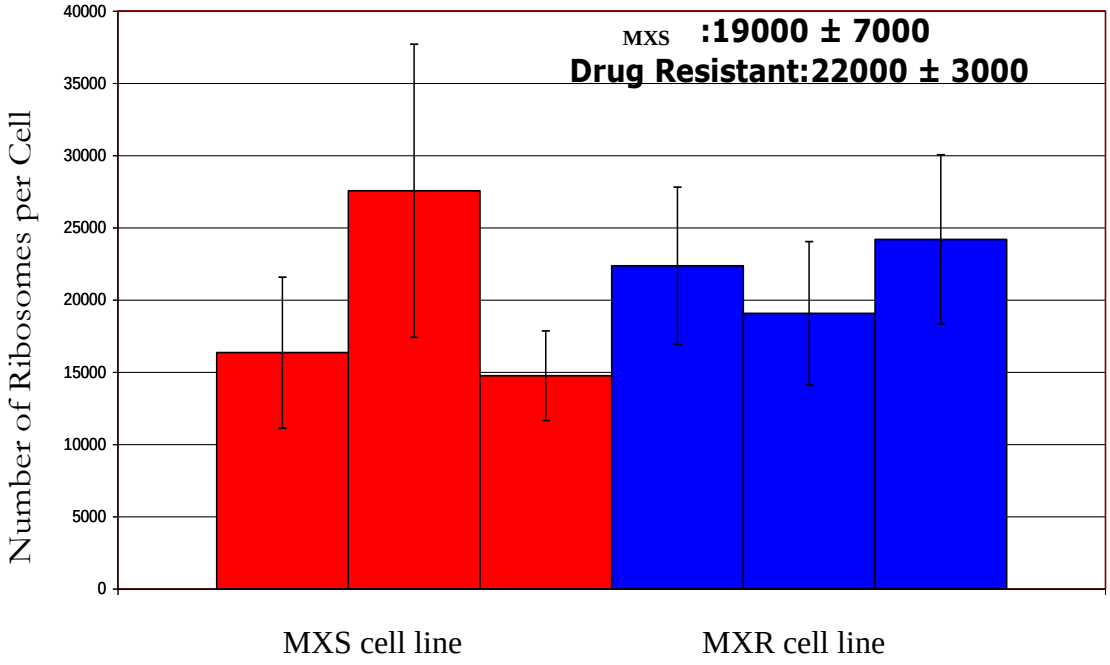

Figure 4.8 Over replicate harvests, an earlier investigation[1] has shown no significant differences in the number of ribosomes between the two cell lines Technical replicates; N = 8 for ribosome count, N = 2 for cell count

Implications

The implications of these two experiments is that there is a physical difference between the ribosome found in the MXR cell line versus the ribosome found in the MXS cell line, the result of which leads to a decrease in translational efficiency in the MXR cell line. As was surmised based on the proteomic evaluation of these ribosomes, the primary protein composition is altered between the two cell types. The investigations illustrated in Figures 4.7 and 4.8 suggest that the alteration in the primary structure of these ribosomal proteins is leading to or contributing to the change in translational efficiency.

The connection between treatment with a chemotherapeutic agent and the ribosome

The effect that treatment with chemotherapeutic drugs including mitoxantrone has on the ribosome is rooted in ribosome biogenesis. In a study of 36 chemotherapeutic drugs on 2fTGH (a human fibrosarcoma cell line) using *in vivo* labeling of RNA and rRNA analysis and western blotting, Burger and colleagues determined that clinically relevant concentrations of most of these drugs led to a loss of nucleolar integrity and interruption of rRNA synthesis. In the case of mitoxantrone, exposure of the cells to a concentration below that used for clinical treatment led to a fast decrease in the appearance of 47S (35S in yeast) rRNA. This was interpreted as a causal relationship between treatment with mitoxantrone and an interruption in the transcription of 47S rRNA[36].

A recent publication from Lee *et al* demonstrated that mitoxantrone binds to nucleolar and coiled body phosphoprotein 1 (NOLC1) also known as Nopp140 or hNopp140 (human Nopp140)[127]. The transcription of 47S rRNA is likely affected by the interaction that mitoxantrone has with the C-terminal end of Nopp140 (NOLC1). Nopp140 has been shown to bind to RNA polymerase I in a coimmunoprecipitation study and a double immunofluorescence investigation illustrated that it colocalizes with RNA polymerase I at the rDNA (rRNA genes) transcription active foci in the nucleolus. Cells which were transfected with only the amino terminal portion of hNopp140 or induced for overexpression of hNopp140 resulted in mislocalization of RNA polymerase I and altered nucleolar phenotypes such as that observed when cells are treated with actinomycin D, a known inhibitor of rRNA synthesis[36; 128]. The mitoxantrone interaction with Nopp140 was found to enhance the interaction of

Nopp140 with protein kinase CK2 (casein kinase 2). Protein kinase CK2 plays a role in the control of cell growth, the regulation of rDNA transcription and apoptosis. The interaction CK2 has with Nopp140, which is controlled by the phosphorylation state of Nopp140, suppresses the catalytic activity of CK2[127]. As mentioned above, mitoxantrone acts as a positive effector of the interaction between these two proteins thus ensuring the suppression of the catalytic activity of CK2. Since elevated CK2 activity has been linked to many cancers, the suppression of this activity by the interactions CK2 has with Nopp140 and mitoxantrone might assist in cancer treatment[127].

The cause of the interruption in transcription of 47S rRNA is likely multifactorial given the fact that there is more than one binding partner of mitoxantrone and they participate in multiple functions in the cell. Gopinath *et al* published a study in 2005 that showed that mitoxantrone binds to specific vault RNAs (vRNAs) known as hvg-1 and hvg-2, components of vault cytoplasmic ribonucleoprotein particles (eukaryotic organelle three times the size of ribosomes found in higher eukaryotes whose function is still not well understood). It was suggested from this work that this may contribute to multidrug resistance due to the fact that vRNAs are (1) observed to be overexpressed in cells treated with cytotoxic compounds such as mitoxantrone, (2) that they are involved in nucleo-cytoplasmic transport and (3) that the compound that the cell is exposed to is found *within* the vRNA complex allowing for the export of the toxic compounds for intracellular detoxification. This interaction though likely relevant to some of the mitoxantrone

interactions in the cell, does not explain the relationship between the effect of mitoxantrone treatment and alterations found in the ribosomal proteins[38].

Factors linking acquired mitoxantrone resistance to altered ribosomal proteins

This investigation has focused on 4 proteins determined to be in altered abundance between the MXR and MXS ribosome: RPL11, RPL23A, RPS3, and RPS10. On further examination, the interactions that these proteins have in the cell; in conditions of ribosomal stress, in the fully functional ribosome of a healthy cell and in ribosome biogenesis help explain why they would be altered in a drug resistant cell.

Ribosomal protein RPL11 has numerous interacting partners, among them p53, MDM2, and c-myc. Numerous studies have shown that under circumstances of cellular stress, RPL11 can act in concert with other ribosomal proteins (RPL5 and RPL23) and 5S rRNA to inhibit the activity of MDM2 (an E3 ubiquitin ligase that targets p53 for degradation) thus resulting in the activation of the p53 pathway[129]. The Myc protein is reputed to bind to and hypothetically control the transcription of *at least 15% of the eukaryotic genome*. Myc is known to enhance RNA polymerase I and III rRNA catalyzed synthesis. It also participates in coordinating the processing of genes of proteins that contribute to rRNA processing, ribosome assembly as well as nuclear-cytoplasmic transport of mature ribosomal subunits. Myc participates in the RNA polymerase II-dependent transcription of ribosomal protein genes and increases the transcription of many of these genes (RPL11 is an example of a transcription target)[76]. An investigation into the ribosomal response after treatment of

mammalian cells with a drug that inhibits transcription of 47S rRNA (actinomycin D; a member of the same class of drugs as mitoxantrone) has shown that cytoplasmic RPL11 acts as a first responder to the stress placed on the cell, recruiting miR-24 loaded miRISC Ago2 (microRNA silencing complex argonaute 2) to the c-myc mRNA for mRNA decay and silencing[130]. In this respect, RPL11 and Myc participate in a negative feedback loop which under normal circumstances would prevent cell growth and ribosome biogenesis during times of ribosomal stress. The fact that we observe altered abundance and a new protein isoform of RPL11 in a chemotherapeutic resistant suggests a possible connection to the role it plays in the cell. These pathways have been altered in the resistant cell line in order for the cells to survive continued exposure to mitoxantrone. This could imply one of three things. Either (1) these RPL11 proteins are altered as an effect of a previous interaction with their binding partners (bearing in mind that most of the communication involved in these pathways involves the presence of certain PTMs), (2) the altered isoforms are a result of changes in these cellular pathways that did not directly involve the protein itself or (3) the protein becomes altered in the context of the mature ribosome while interacting with either an mRNA or a protein involved in these pathways. Options 2 and 3 are not exclusive of one another. Only additional investigations could further clarify which of these possibilities has occurred.

The ribosomal protein RPL23A (referred to as RPL25 in yeast) also interacts with proteins that contribute to ribosomal biogenesis and cell signaling. Unlike RPL11, there is no question as to whether the contribution/interaction of RPL23A is in the context of a fully mature cytoplasmic ribosome. An investigation by Oh *et al*

found that RPL23A binds to mTORC2 to promote the co-translational stabilization and phosphorylation of nascent Akt polypeptide[131]. A study by Zinzalla *et al* found that ribosomes exposed to protein translation inhibitors are capable of facilitating mTORC2 signaling independent of protein synthesis, suggesting that it is the physical structure of the 80S ribosome and not the activity that mediate mTORC2 activity. Their investigation found that the 80S ribosome was in fact required for mTORC2 activity[132]. The mTOR protein, mammalian target of rapamycin, is a serine threonine kinase known to be correlated with cell growth, cell survival, *rRNA transcription, and protein synthesis*. It is associated with two different protein complexes; mTORC1 and mTORC2. The complex studied by Oh and colleagues which associates with RPL23A in actively translating ribosomes consists of mTOR, rictor, SIN1and mLST8 and is known for its role in cell survival and actin cytoskeletal reorganization. As part of the role it plays in promoting cell survival, mTORC2 has been found to mediate the phosphorylation of several members of the AGC antiapoptotic kinase family of proteins such as Akt, PKB and SGK which in turn activate these kinases towards substrates such as Fox03a and NDRG1. The phosphorylation event on the carboxyterminal tail turn motif associated with the Akt nascent polypeptide is required for proper Akt folding and maturation. Without co-translational phosphorylation, ubiquitination of Akt occurs during translation leading to its destruction[131; 133]. Phosphorylated Akt is known to feed into the pathway that signals to Myc to transcribe rRNA[134]. A connection between acquired chemotherapeutic drug resistance in attached cells treated with an inhibitor of mTOR and PI3K showed an adaptive response on the part of the drug exposed cells by their up-regulation of

pro-survival proteins. The proteins whose expression was increased varied among cell lines (often times the apoptosis regulating protein Bcl-2 was observed) but the end result was always the same; increased cap-independent translation and FOXO transcription[135]. Changes in the structure of the ribosomal protein, RPL23A, that interacts with a key component of this process, mTORC2 could possibly either influence the MXR cells ability to maintain an interaction with the mTORC2 complex (even in the face of protein signals that would drive the cell towards apoptosis) or if the interaction with mTORC2 was altered, drive the pathways that would lead to an increase in cap-independent translation and FOXO transcription.

The ribosomal protein RPS3 is not only an important participant in the ribosome in translation but it is also known to play a role in DNA repair, apoptosis, ribosome biogenesis and cell cycle regulation. It is a substrate of multiple protein kinases (PKCδ, Akt, ERK, Cdk1) resulting in cell signaling phosphorylation which many argue determine the localization and function of the protein in the cell at that time[136; 137; 138; 139]. An investigation of yeast ribosome biogenesis revealed that the phosphorylation of RPS3 and subsequent dephosphorylation was required for the RPS3 to be stably integrated into the pre-40S ribosome for proper 40S formation and export from the nucleus[140]. It has been suggested that under certain conditions, RPS3 may be capable of leaving the ribosome in the same manner as RPL11[136]. The environment of the cell and stage of the cell cycle determine whether RPS3 will be phosphorylated, what kinase will phosphorylate it and on which residue(s) it will be phosphorylated. It should be noted that in our investigation, both cell lines were found to contain the T221 phosphorylated RPS3 isoform observed in spot number 7

with equal abundance. This is the only isoform of RPS3 we observed with a phosphorylation. Yoon *et al* in 2011 has shown that T221 is phosphorylated by Cdk1 (if *both* S6 and T221 are phosphorylated, this implicates the PKCδ kinase). Experiments with cell cycle inhibitors suggested that phosphorylation of RPS3 on T221 by Cdk1 is for the purposes of targeting RPS3 for nuclear transport during the cell cycle (particularly at the G2/M phase of the cell cycle)[141]. It is not entirely clear the function served by T221 phosphorylation in RPS3 in both cell lines in the current investigation. The relation of this phosphorylation event to cell cycle events is undeniable. Whether the phosphorylation event precedes the departure of RPS3 from the ribosome or signals some other function that RPS3 plays in the ribosomal machine is unclear at this time however due to the fact that it is detected in equal abundance in both cell lines, it does not directly factor as a structural difference between the MXR and MXS ribosomes.

Perhaps even more crucial to the primary structure of RPS3 is its interaction with the protein RACK1 (Receptor for Activated C Kinase 1) in mature ribosomes. RACK1 is often times referred to as a ribosomal protein of the small subunit since it is commonly found in association with the 40S ribosome. In fact, the chemical associations that RACK1 has with the ribosomal proteins and rRNA are so strong that after studying the crystal structure of the eukaryotic 40S subunit it was concluded that free RACK1 would be observed as a result of RACK1 up-regulation as opposed to its dissociation/release from the ribosome. RACK1 has been described as the "central cellular signaling hub" of the ribosome[142]. It directly interacts with several ribosomal proteins on the head region of the 40S subunit, specifically RPS3, RPS16 and RPS17.

The C-terminal end of RPS3 in particular interacts with RACK1. RACK1 is known to promote translation by recruiting PKC and eIF6 to the small subunit for the PKC-driven phosphorylation of eIF6 to stimulate subunit association. After 80S formation, RACK1 is known to recruit signaling molecules to the ribosome. It has been proposed that in addition to the recruitment of signaling molecules, RACK1 recruits ribosomes to different cellular sites and to stimulate the translation of specific mRNAs[142; 143]. RPS3 not only interacts directly with RACK1 but also with many of the neighbors of RACK1 in addition to a direct interaction with RPS17. This is illustrated in Figure 4.9. It goes without saying that any alteration in the structure of RPS3 could have a domino effect on RACK1 and the interactions it has with the 40S and 80S ribosome as well as any proteins it recruits. Any alterations in these pathways could provide an explanation for the change in the MXR ribosomal proteins however how this is advantageous to drug resistant cells would require further investigation.

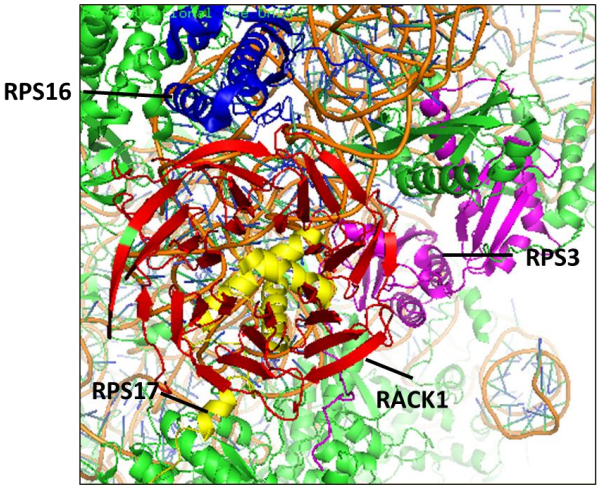

Figure 4.9 RACK1 shown in red, interacts with RPS16, RPS17 and the C-terminal end of RPS3 (shown in magenta). Figure created using Pymol[144]

Ribosomal protein RPS10 is an example of one of the proteins that interacts with RPS3 in the small subunit. As a result, a change in the structure of RPS10, particularly a C-terminal truncation such as that observed in the protein isoform found in spot 39 would be expected to affect ribosomal structure. The effect of RPS10 on ribosomal structure goes beyond its interaction with RPS3. Ren *et al* found that RPS10 is a substrate for protein arginine methyltransferase 5 (PRMT5) which results in the dimethylation of R158 and R160. Without dimethylation at both of these sites they found that the contribution of RPS10 to ribosome biogenesis both in terms of being incorporated into the ribosome and contributing to 18S rRNA maturation was less than optimal leading to an imbalance in the number of 40S subunits to 60S subunits. In addition, the RPS10 which was not dimethylated was not as effectively incorporated into the ribosome and free RPS10 was subject to proteosomal degradation[112]. A recent study in the laboratory of Marc Wilkins found that arginine dimethylation can in fact be related to the age of a protein. When a protein is found to be dimethylated, the lifespan of the protein is significantly longer[145]. The implication exists that the metabolism of the ribosomal proteins in the MXR cells is possibly different. The truncated form of RPS10 observed in this investigation would alter the interactions in these pathways and would no longer (if it ever did) possess the dimethylarginines found in the C-terminal end of the protein. Given this information, it is hard to predict what advantage or purpose this might serve in the MXR ribosome. It is possible that the truncation is a product of proteosomal degradation of the RPS10 lacking the dimethylations. If that is the case, the protein

isoform is not likely behind the cause of the resistance but more the result of an effect of a pathway that has been disturbed.

Protein isoforms and their location in the ribosome

This investigation has established that at least 4 proteins have altered abundance between the MXR and MXS ribosome; RPS3, RPS10, RPL23A, and RPL11. The modifications we have identified on these ribosomal proteins as well as the other ribosomal proteins/ribosomal associated proteins that they interact with in the context of the ribosomal machine may play a role in the function of these ribosomes. The crystal structure of the eukaryotic ribosome was published in the last year with the positions of all but one ribosomal protein determined (human RPL10A)[146; 147].

Ribosomal protein RPL11

There is ribosomal movement during the translocation of mRNA and tRNA frequently referred to as ratcheting. The coordination of the small and large subunit in *Saccharomyces cerevisiae* was published by Ben-Shem *et al*[147]. Though protein sequences are not identical between yeast and humans, the interactions between most proteins and overall ribosomal structure is expected to be very similar. A key protein involved in the interactions between the 60S and 40S subunit is RPL11, referred to as the central protuberance protein. Coordinated with RPL11 are the small subunit proteins RPS18 and RPS15 (refer to Figure 4.10). Based on the 3.0 Angstrom resolution crystal structure of 80S ribosome from yeast, at one point in the subunit interaction, the RPL11 K85 residue is coordinated with RPS15 F42 residue[146; 147]. It

should be noted again that the yeast sequence for RPL11 is not identical to the human sequence (a Blast search reveals that fission yeast and human RPL11 share 74% identity[117]), If the interactions are the same in humans, RPL11 K85 is acetylated in the MXR cell line, changing the chemistry of the interaction between the two residues. RPL11 was described along with the other proteins in the central protuberance, to undergo considerable rearrangements and a shift in position of all regions of the RPL11 sequence except for the N-terminal end. It was suggested that this structural rearrangement is the possible reason for a separate5S rRNA which interacts with these proteins. Based on its role in subunit interaction and the observed PTMs found in RPL11 in the MXR cell line, it is not surprising that alterations in PTMs might affect the movement of mRNA through the mRNA tunnel. Based on investigations of the mRNA tunnel with a structural translation inhibitor called Stm1 (a homologous protein in humans is unknown), the path through the mRNA tunnel of the Stm1 protein is believed to contact nine different ribosomal proteins indicating the presence of these proteins in the tunnel opening. RPL11 is the only 60S subunit protein that is believed to come into direct contact with Stm1 suggesting RPL11 interacts with mRNA as it passes through the processing tunnel. This would make any alterations in the primary structure of RPL11 relevant to the accessibility of the tunnel in the ribosome for mRNA translation[146; 147].

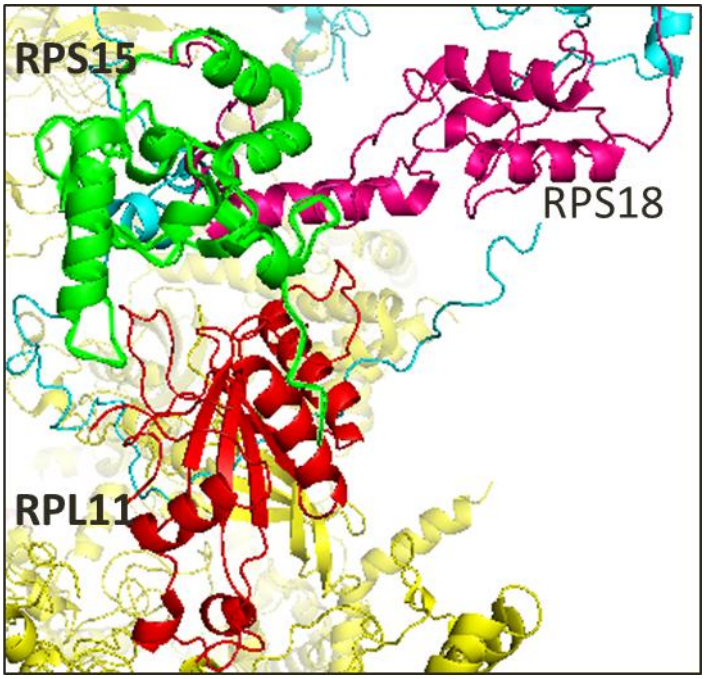

Figure 4.10 The ratcheted eukaryotic ribosome illustrating the interaction between RPL11, the 40S subunit with the small subunit proteins RPS15 and RPS18. Figure created using Pymol[144]

Ribosomal protein RPL23A

RPL23A is found near the exit tunnel in the eukaryotic ribosome (refer to Figure 4.11). As previously mentioned, it is capable of interacting with protein factors in actively translating ribosomes such as the mTORC2 complex discussed above. One of the most important interactions RPL23A has is as a member of the nascent polypeptide associated complex, otherwise known as the NAC, and the signal recognition particle, otherwise known as the SRP. The nascent chain of an emerging polypeptide must come into contact with a large variety of factors in the eukaryotic cell including protein chaperones, methionine aminopeptidases, N-acetyltransferases, the NAC, the ribosome associated membrane proteinERj1p, the SRP, the translocon Sec61 and SR. NAC is a heterodimeric cytosolic protein complex with a α-subunit

and β-subunit found in abundance in higher eukaryotes. Altered intracellular levels of NAC subunits have been linked with numerous disease processes including Alzheimer's disease, Down syndrome and malignant brain tumors[148]. Although NAC deletion only causes minor growth defects in yeast it has been linked with embryonic lethality in mice, flies and nematodes. With limited space around the exit tunnel of the ribosome, certain ribosomal proteins are used in a functional capacity for the NAC. Numerous ribosomal proteins at the site of the exit tunnel including RPL23A have been identified as the docking site for both the NAC and the SRP with the binding being determined by protein sequence. In the case of hydrophobic stretches of residues, recent studies have suggested that even before they emerge from the ribosome, the SRP is recruited by changing conformation of the ribosome. It has also been suggested that the NAC is a modulator for the specificity of the SRP. Regardless, both of these so called CLIPS (chaperones linked to protein synthesis), have been found to directly interact with RPL23A[148; 149]. It is interesting to note that the interactions that RPL23A has with mTORC2 in monitoring that Akt is modified correctly so that it may fold in the proper conformation and avoid ubiquitination is not unlike the role that it plays in the NAC and the SRP. Since ribosomal conformations are noted as changing to acquiesce to the need of the NAC and SRP for access to certain nascent polypeptides, if structural changes interfere with this ability, this would influence which proteins are successfully folded and modified for their role in the cell. The SRP has been shown to display particular specificity towards hydrophobic sequences and transmembrane regions of proteins with the majority of proteins associated with the SRP destined for the Golgi apparatus,

endoplasmic reticulum or the plasma membrane. The SRP has been noted for altering the configuration of the ribosome upon binding due to these hydrophobic stretches of sequence so it is possible that these proteins would be more affected were the structure of the ribosome to be altered [149]. Regardless, if the structure of RPL23A is altered as in the case of the MXR cells, then it is reasonable to expect that this would influence its interactions with the NAC and SRP.

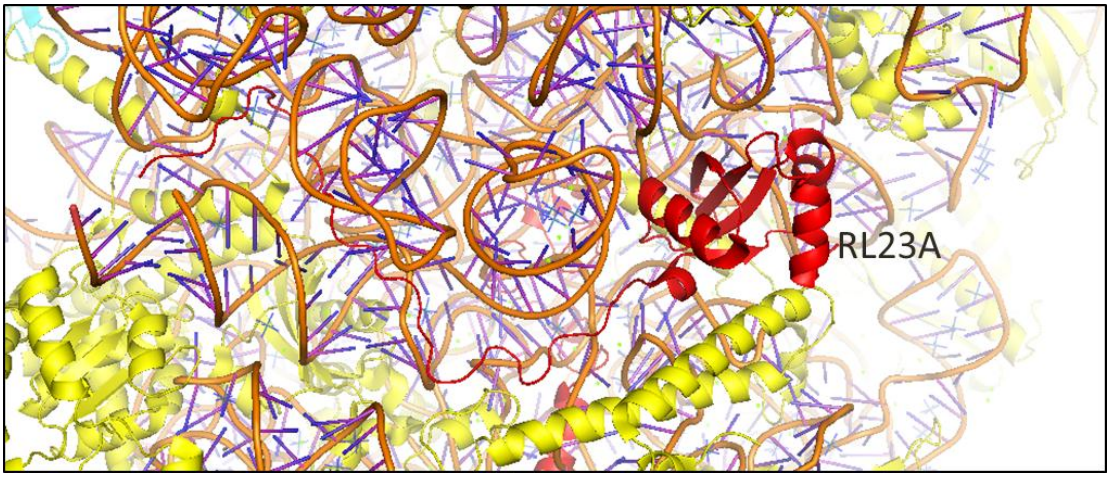

Figure 4.11 RPL23A found with extension running through the exit tunnel of the ribosome. Figure created using Pymol[144]

Ribosomal protein RPS3

The interaction of ribosomal protein RPS3 with RACK1 is critical not only to cell signaling events but also to the structure of the 80S ribosome. The C-terminal end of RPS3 is involved in at least two salt bridges with RACK1 according to structural studies of the 40S ribosome conducted by Rabl *et al*. Their studies indicated that RPS3 spans much of the distance of the mRNA tunnel and interacts with RACK1, RPS17, RPS10, RPS20, RPS29 and RPS2 in the small subunit alone[142]. An illustration of RPS3 in the mRNA tunnel is shown in Figure 4.12. Deviations from the structure of RPS3 needed at a given time (in the cell cycle or otherwise)

would impact much of the interactions between the proteins in the ribosome. The structural change between the RPS3 isoform observed in spot 6 and spot 8 is minimal with the N-terminal acetylation being the only difference. The significance of this difference could either come from the chemistry of the change (interactions and size) or it could be rooted in the message that this acetylation gives the cell; preservation of the protein and destination in the cell[150]. It is interesting to note that although the phosphorylation was common feature of an isoform seen in both cell lines, the location of that phosphorylation is on the C-terminal end which would affect whether or how RACK1 can interact with it and if RACK1 is not able to interact/continue interacting with the ribosome[142]. Under these circumstances this could be the possible reason why certain proteins signal that RPS3 should relocate to the nucleus[139]. One could even posit further that the proteins that then interact with the ribosome and set proteins like RPS3 on its course are determined by the state of the cell (DNA repair, pathogen response, undergoing a change in the cell cycle, etc.).

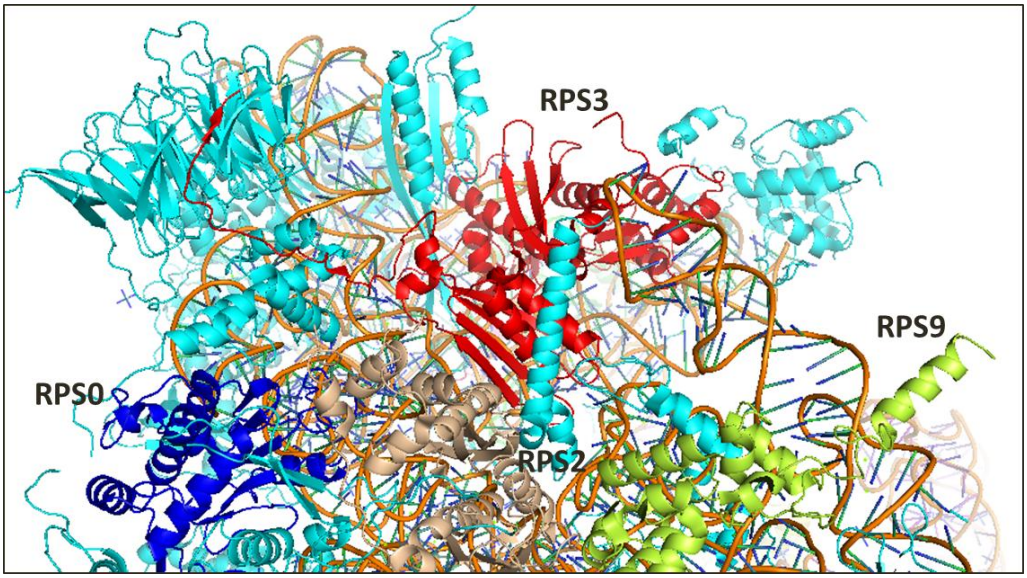

Figure 4.12 RPS3 with the mRNA entry and exit sites. Figure created using Pymol[144]

Ribosomal protein RPS10

In the small subunit ribosomal protein RPS10 interacts with RPS3 as well as RPS12, RPS20 and RPS29. The incorporation of RPS10 and successful interaction in the ribosome requires the full length protein with dimethyl arginine modifications found on R158 and R160. The MXR ribosome contained a novel RPS10 isoform with a truncation of the last 28 residues from the C-terminal end. With this truncation (from K137 to Q165), the dimethylarginine modification would obviously not be present and the interactions of this protein with other ribosomal proteins would be dramatically altered[112]. It is suspected that this isoform is the result of proteosomal degradation, either of the RPS10 prior to its incorporation into the ribosome or as a result of the C-terminal end being vulnerable to proteases in the cell even in the context of the ribosomal machine. In either case, the shift in ribosomal structure would impact the efficiency with which these ribosomes translate mRNA and could at least provide an explanation for the 25% decrease in ribosomal activity as determined by the methionine incorporation study.

Summary and Prospectus

This investigation has found that the acquired resistance after treatment of MCF7 breast cancer cells with mitoxantrone, a known inhibitor of ribosome biogenesis[36] and binding partner of hNopp140[127] (a phosphoprotein necessary for the localization of RNA polymerase I and binding partner of protein kinase CK2), leads to changes in the primary structures of the ribosomal proteins. Specifically, the 80S

ribosomes of the drug resistant cells were found to have altered abundance between the MXR and MXS ribosomes of isoforms of RPS3, RPL11, RPL23A and RPS10 and novel isoforms of RPS3and RPL11. Investigations of the methionine incorporation of the drug resistant ribosomes found that it was 25% lower than the activity found in their drug susceptible counterparts. This could not be explained by a change in the number of ribosomes between the two cell lines.

Based on our findings and knowledge of the interacting partners of the different ribosomal protein isoforms, we would propose that the impact of the altered ribosomal proteins within the ribosomal machine extends to the access to ribosomal mRNA binding sites. By altering the access the mRNA binding sites, this could allow for differential selective translational activity of the MXR ribosome. Looking at the possible impact that exposure to mitoxantrone may have on the cell (increased levels of ROS in the cellular environment), a need for cell survival proteins exists. The concept behind differential selective translation is that cell survival proteins would be produced during times of cellular stress, such as that which occurs during exposure to a chemotherapeutic drug. Hand in hand with the concept of differential selective translation is the ability of the ribosome to utilize cap-independent translation (IRES)[151]. This is supported by the interacting partners of RPL23A, namely mTOR.

Further characterization of these ribosomes is needed in terms of the modifications such as phosphorylation (RPS3 T221), in the proteins linked with their role in the ribosome (RACK1, Myc, mTOR, NAC, etc.) and in additional proteins whose abundance was found altered (for example RPS15A, RPS3A, RPS6, RPL12,

RPL10, RPL5). A proteomic approach is most practical. In certain cases these proteins can be investigated using immunological techniques as evidenced by the large number of antibodies available for modifications like phosphorylation and for specific ribosomal proteins such as RPS3.

Appendices

ID	Left Amino Acid	Sequence	Right Amino Acid	E Value
RPSA	D	VLQMKEEDVLKFLAAGTHLGGTNL	D	1.56E-24
RPSA	D	VLKFLAAGTHLGGTNL	D	1.33E-26
RPSA	D	VLKFLAAGTHLGGTNLDFQMEQYIYKRKS	D	2.12E-08
RPSA	D	FQMEQYIYKRKS	D	2.78E-06
RPSA	D	VSVISSRNTGQRAVLKFAAATGATPIAGRFTPGTFTNQIQAAFREPRLLVVT	D	1.28E-18
RPSA	D	PRADHQPLTEASYVNLPTIALCNT	D	2.89E-13
RPSA	D	IAIPCNNKGAHSVGLMWWMLAREVLRMRGTISREHPWEVMP	D	0.0000242
RPSA	D	IAIPCNNKGAHSVGLMWWMLAREVLRMRGTISREHPWEVMPDLYFYR	D	0.000193
RPSA	D	PEEIEKEEQAAAEKAVTKEEFQGEWTAPAPEFTATQPEVA	D	3.26E-16
RPS2	D	KEWMPVTKLGRLVKD	M	6.79E-07
RPS2	D	MKIKSLEEIYLFSLPIKESEIID	F	0.000351
RPS2	D	FFLGASLK	D	9.66E-12
RPS2	D	FFLGASLKDEVLKIMPVQKQTRAGQRTRFKAFVAIG	D	8.54E-16
RPS2	D	HLVKTHTRVSVQRTQAPAVATT	-	8.42E-09
RPS3	M	AVQISKKRKFVA	D	7.77E-20
RPS3	D	GIFKAELNEFLTRELAE	D	5.37E-22
RPS3	D	TAVRHVLLRQGVLGIKVKIMLPW	D	9.57E-11
RPS3	D	PTGKIGPKKPLPDHVSIVEPK	D	6.36E-11

RPS3A	D	VKAPAMFNIRNIGKTLVTRTQGTKIAS	D	2.53E-19
RPS3A	D	GLKGRVFEVSLA	D	1.52E-08
RPS3A	D	LQNDEVAFRKFKLITE	D	3.33E-09
RPS3A	D	EVAFRKFKLITE	D	0.0000233
RPS3A	D	VQGKNCLTNFHGM	D	2.2E-14
RPS3A	D	LKEVVNKLIPD	S	5.14E-10
RPS3A	D	LKEVVNKLIPDSIGK	D	1.54E-14
RPS3A	D	SIGKDIEKACQSIYPLH	D	6.74E-25
RPS3A	D	IEKACQSIYPLH	D	1.26E-10
RPS3A	D	VFVRKVKMLKKPKFELGKLMELHGEGSSSGKATG	D	2.2E-22
RPS3A	D	VFVRKVKMLKKPKFELGKLMELHGEGSSSGKATGDETGAKVERA	D	6.63E-23
RPS4X	M	ARGPKKHLKRVAAPKHWMLD	K	1.59E-10
RPS4X	D	KLTGVFAPRPSTGPHKLRECLPLIIFLRNRLKYALTG	D	3.9E-09
RPS4X	D	KTGENFRLIY	D	9.42E-07
RPS4X	D	TKGRFAVHRITPEEAKYKLCKVRKIFVGTKGIPHLVTH	D	1.21E-33
RPS4X	D	LETGKITDFIKF	D	0.0000432
RPS4X	D	TGNLCMVTGGANLGRIGVITNRERHPGSFD	V	9.11E-06
RPS4X	D	TGNLCMVTGGANLGRIGVITNRERHPGSFDVVHVK	D	2.83E-14
RPS4X	D	ANGNSFATRLSNIFVIGKGNKPWISLPRGKGIRLTIAEER	D	1.71E-10
RPS5	D	IKLFGKWST	D	2.91E-18

RPS5	D	STRIGRAGTVRRQAVD	V	0.000178
RPS6	-	MKLNISFPATGCQKLIEV	D	9.85E-13
RPS6	D	ANLSVLNLVIVKKGEKD	I	5.19E-13
RPS6	D	ANLSVLNLVIVKKGEKDIPGLT	D	4.78E-07
RPS7	-	MFSSSAKIVKPNGEKPD	E	1.01E-09
RPS7	-	MFSSSAKIVKPNGEKPDEFESGISQALLELEMNS	D	3.39E-20
RPS7	D	AILEDLVFPSEIVGKRIRVKL	D	3.29E-16
RPS7	D	LVFPSEIVGKRIRVKLD	G	0.000515
RPS7	D	GSRLIKVHL	D	7.25E-10
RPS7	D	KAQQNNVEHKVETFSGVYKKLTGKD	V	1.42E-12
RPS7	D	KAQQNNVEHKVETFSGVYKKLTGKDVNFEFPEFQL	-	1.74E-10
RPS7	D	VNFEFPEFQL	-	1.23E-15
RPS8	D	VVYNASNNELVRTKTLVKNCIVLID	S	4.55E-09
RPS8	D	STPYRQWYESHYALPLGRKKGAKLTPEEEEILNKKRSKKIQKKY	D	8.52E-06
RPS8	D	GYVLEGKELEFYLRKIKARKGK	-	3.27E-20
RPS9	D	PRRLFEGNALLRRLVRIGVLD	E	0.0000902
RPS9	D	PRRLFEGNALLRRLVRIGVLDEGKMKL	D	0.000307
RPS9	D	YILGLKIE	D	3.15E-18
RPS9	D	YILGLKIEDFLERRLQTQVFKLGLAKSIHHARVLIRQRHIRVRKQVVNIPSFIVRL	D	1.37E-06
RPS9	D	FLERRLQTQVFKLGLAKSIHHARVLIRQRHIRVRKQVVNIPSFIVRL	D	9.15E-07
RPS10	-	MLMPKKNRIAIYELLFKEGVMVAKKD	V	1.1E-30

RPS10	-	MLMPKKNRIAIYELLFKEGVMVAKKDVHMPKHPELA	D	1.83E-19
RPS10	D	KNVPNLHVMKAMQSLKSRGYVKEQFAWRHFYWYLTNEGIQYLR	D	3.78E-16
RPS10	D	YLHLPPEIVPATLRRSRPETGRPRPKGLEGERPARLTRGEA	D	3.6E-12
RPS11	M	ADIQTERAYQKQPTIFQNKKRVLLGETGKEKLPRYYKNIGLGFKTPKEAIEGTYID	K	1.14E-06
RPS11	D	IQTERAYQKQPTIFQNKKRVLLGETGKEKLPRYYKNIGLGFKTPKEAIEGTYID	K	3.03E-14
RPS11	D	YLHYIRKYNRFEKRHKNMSVHLSPCFRD	V	0.0000146
RPS11	D	YLHYIRKYNRFEKRHKNMSVHLSPCFRDVQIG	D	4.75E-06
RPS11	D	VQIGDIVTVGECRPLSKTVRFNVLKVTKAAGTKKQFQKF	-	2.38E-21
RPS11	D	IVTVGECRPLSKTVRFNVLKVTKAAGTKKQFQKF	-	1.77E-25
RPS12	D	VNTALQEVLKTALIH	D	1.85E-26
RPS12	D	GLARGIREAAKAL	D	3.2E-07
RPS12	D	EPMYVKLVEALCAEHQINLIKV	D	0.0000115
RPS12	D	NKKLGEWVGLCKI	D	1.62E-08
RPS12	D	YGKESQAKDVIEEYFKCKK	-	7.55E-16
RPS12	D	VIEEYFKCKK	-	0.0000153
RPS13	M	GRMHAPGKGLSQSALPYRRSVPTWLKLTSD	D	3.3E-09
RPS13	D	DVKEQIYKLAKKGLTPSQIGVILR	D	4.03E-24
RPS13	D	SHGVAQVRFVTGNKILRILKSKGLAPD	L	1.24E-11

RPS13	D	LPEDLYHLIKKAVAVRKHLERNRKD	K	2.1E-07
RPS13	D	LYHLIKKAVAVRKHLERNRKD	K	4.55E-13
RPS13	D	AKFRLILIESRIHRLARYYKTKRVLPPNWKYESSTA SALVA	-	3.76E-08
RPS14	D	TFVHVTDLSGKETICRVTGGMKVKA	D	9.98E-15
RPS14	D	LSGKETICRVTGGMKVKA	D	1.84E-21
RPS14	D	VAQRCKELGITALHIKLRATGGNRTKTPGPGAQSA LRALARSGMKIGRIE	D	1.62E-17
RPS15	M	AEVEQKKKRTFRKFTYRGVD	L	1.15E-06
RPS15	M	AEVEQKKKRTFRKFTYRGVDL	D	3.62E-09
RPS15	D	MIILPEMVGSMVGVYNGKTFNQVEIKPEMIGHYLG EFSITYKPVKHGRPGIGATHSSRFIPLK	-	2.55E-11
RPS15A	D	ALKSINNAEKRGKRQVLIRPCSKVIVRFLTVMMKH GYIGEFEIID	D	3.28E-07
RPS15A	D	DHRAGKIVVNLTGRLNKCGVISPRF	D	0.000188
RPS15A	D	LEKWQNNLLPSRQFGFIVLTTSAGIMDHEEARRKH TGGKILGFFF	-	1.36E-17
RPS15A	D	HEEARRKHTGGKILGFFF	-	0.0000106
RPS16	M	PSKGPLQSVQVFGRKKTATAVAHCKRGNGLIKVN GRPLEMIEPRTLQYKLLEPVLLLGKERFAGV	D	3.96E-14
RPS16	D	IRVRVKGGGHVAQIYAIRQSISKALVAYYQKYVD	E	6.74E-12
RPS16	D	PRRCESKKFGGPGARARYQKSYR	-	1.04E-10
RPS17	M	GRVRTKTVKKAARVIIEKYYTRLGND	F	1.83E-08
RPS17	D	TKEMLKLLD	F	2.38E-09
RPS17	D	TKEMLKLLDFGSLSNLQVTQPTVGMNFKTPRGPV	-	3.58E-16
RPS17	D	FGSLSNLQVTQPTVGMNFKTPRGPV	-	1.02E-21

RPS18	M	SLVIPEKFQHILRVLNTNID	G	4.88E-14
RPS18	D	GRRKIAFAITAIKGVGRRYAHVVLRKAD	I	1.04E-06
RPS18	D	IDLTKRAGELTE	D	1.31E-06
RPS19	M	PGVTVKDVNQQEFVRALAAFLKKSGKLKVPEWVD	T	1.48E-07
RPS19	D	VNQQEFVRALAAFLKKSGKLKVPEWVD	T	1.02E-17
RPS19	D	VNQQEFVRALAAFLKKSGKLKVPEWVDTVKLAKHKELAPY	D	1.64E-11
RPS19	D	TVKLAKHKELAPYD	E	4.65E-10
RPS19	D	GGRKLTPQGQRDLDRIAGQVAAANKKH	-	0.0000094
RPS19	D	LDRIAGQVAAANKKH	-	6.7E-09
RPS20	M	AFKDTGKTPVEPEVAIHRIRITLTSRNVKSLEKVCA	D	3.24E-06
RPS20	D	RFQMRIHKRLID	L	8.67E-06
RPS21	D	LYVPRKCSASNRIIGAKD	H	3.95E-06
RPS21	D	KVTGRFNGQFKTYAICGAIRRMGES	D	7.38E-07
RPS23	D	GCLNFIEENDEVLVAGFGRKGHAVG	D	1.95E-19
RPS23	D	EVLVAGFGRKGHAVGD	I	0.000421
RPS23	D	IPGVRFKVVKVANVSLLALYKGKKERPRS	-	8.38E-37
RPS24	D	TVTIRTRKFMTNRLLQRKQMVID	V	1.17E-06
RPS24	D	VLHPGKATVPKTEIREKLAKMYKTTPD	V	4.36E-18
RPS24	D	VLHPGKATVPKTEIREKLAKMYKTTPDVIFVFGFRTHFGGGKTTGFGMIY	D	0.000264
RPS24	D	VIFVFGFRTHFGGGKTTGFGMIY	D	7.51E-16
RPS25	D	KLNNLVLF	D	1.64E-10
RPS25	D	KATYDKLCKEVPNYKLITPAVVSERLKIRGSLARAALQELLSKGLIKLVSKHRAQVIYTRNTKGG	D	8.32E-11

RPS25	D	KLCKEVPNYKLITPAVVSERLKIRGSLARAALQELL SKGLIKLVSKHRAQVIYTRNTKGG	D	6.54E-28
RPS26	D	KAIKKFVIRNIVEAAAVRD	I	5.23E-29
RPS26	D	KAIKKFVIRNIVEAAAVRDISEASVF	D	8.09E-09
RPS26	D	AYVLPKLYVKLHYCVSCAIHSKVVRNRSREARKD	R	0.0000363
RPS26	D	RTPPPRFRPAGAAPRPPPKPM	-	1.47E-11
RPS27	D	LLHPSPEEEKRKHKKKRLVQSPNSYFM	D	0.000218
RPS28	-	MDTSRVQPIKLARVTKVLGRTGSQGQCTQVRVEF M	D	5.4E-08
RPS28	D	TSRVQPIKLARVTKVLGRTGSQGQCTQVRVEFM	D	1.58E-14
RPS28	D	TSRSIIRNVKGPVREGD	V	2.26E-07
RPS28	D	TSRSIIRNVKGPVREGDVLTLLESEREARRLR	-	4.22E-27
RPS28	D	VLTLLESEREARRLR	-	1.52E-17
RPL3	M	SHRKFSAPRHGSLGFLPRKRSSRHRGKVKSFPK	D	3.74E-14
RPL3	D	PSKPVHLTAFLGYKAGMTHIVREVD	R	6.25E-29
RPL3	D	FSSMKKYCQVIRVIAHTQMRLLPLRQKKAHLMEIQ VNGGTVAEKL	D	1.6E-11
RPL3	D	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKI D	L	1.73E-11
RPL3	D	FVMLKGCVVGTKKRVLTLRKSLLVQTKRRALEKI DLKFID	T	2.81E-06
RPL3	D	TTSKFGHGRFQTMEEKKAFMGPLKK	D	3.09E-24
RPL3	D	TTSKFGHGRFQTMEEKKAFMGPLKKDRIAKEEGA	-	1.96E-20
RPL4	M	ACARPLISVYSEKGESSGKNVTLPAVFKAPIRPD	I	9.28E-26
RPL4	D	KVEGYKKTKEAVLLLKKLKAWND	I	5.74E-11

RPL4	D	NGIIKAFRNIPGITLLNVSKLNILKLAPGGHVGRFCIWTESAFRKL	D	2.9E-23
RPL4	D	KAAAAAAALQAKSDEKAAVAGKKPVVGKKGKKAAVGVKKQKKPLVGKKAAATKKPAPEKKPAEKKPTTEEKKPAA	-	2.45E-09
RPL4	D	EKAAVAGKKPVVGKKGKKAAVGVKKQKKPLVGKKAAATKKPAPEKKPAEKKPTTEEKKPAA	-	0.0000118
RPL5	M	GFVKVVKNKAYFKRYQVKFRRRREGKTD	Y	1.57E-08
RPL5	D	KNKYNTPKYRMIVRVTNR	D	2.68E-06
RPL5	D	AGLARTTTGNKVFGALKGAVD	G	3.55E-22
RPL5	D	GGLSIPHSTKRFPGY	D	1.97E-10
RPL5	D	SESKEFNAEVHRKHIMGQNVA	D	0.000132
RPL5	D	YMRYLMEE	D	9.83E-10
RPL5	D	AYKKQFSQYIKNSVTPD	M	0.00023
RPL5	D	RVAQKKASFLRAQERAAES	-	3.47E-12
RPL6	D	KNGGTRVVKLRKMPRYYPTE	D	3.65E-07
RPL6	D	VPRKLLSHGKKPFSQHVRKLRASITPGTILIILTGRHRGKRVVFLKQLASGLLLVTGPLVLNRVPLRRTHQKFVIATSTKID	I	8.22E-06
RPL6	D	ISNVKIPKHLTD	A	9.97E-20
RPL6	D	ISNVKIPKHLTDAYFKKKKLRKPRHQEGEIF	D	4.23E-13
RPL6	D	AYFKKKKLRKPRHQEGEIF	D	6.23E-13
RPL6	D	AYFKKKKLRKPRHQEGEIFDTEKEKYEITEQRKI	D	2.97E-24
RPL6	D	TEKEKYEITEQRKID	Q	1.04E-11
RPL6	D	SQILPKIKAIPQLQGYLRSVFALTNGIYPHKLVF	-	7.04E-23
RPL7	D	NALIARSLGKYGIICME	D	2.57E-23

RPL7	D	LIHEIYTVGKRFKEANNFLWPFKLSSPRGGMKKKT THFVEGG	D	4.15E-24
RPL7A	M	PKGKKAKGKKVAPAPAVVKKQEAKKVVNPLFEK RPKNFGIGQD	I	3.1E-14
RPL7A	M	PKGKKAKGKKVAPAPAVVKKQEAKKVVNPLFEK RPKNFGIGQDIQPKR	D	7.69E-06
RPL7A	D	LTRFVKWPRYIRLQRQRAILYKRLKVPPAINQFTQ AL	D	0.0000492
RPL7A	D	RQTATQLLKLAHKYRPETKQEKKQRLLARAEKKA AGKGD	V	3.41E-18
RPL7A	D	VPTKRPPVLRAGVNTVTTLVENKKAQLVVIAHD	V	8.6E-28
RPL7A	D	VPTKRPPVLRAGVNTVTTLVENKKAQLVVIAHDV	D	1.01E-14
RPL7A	D	PIELVVFLPALCRKMGVPYCIIKGKARLGRLVHRK TCTTVAFTQVNSED	K	0.000181
RPL7A	D	KGALAKLVEAIRTNYND	R	8.99E-09
RPL7A	D	KGALAKLVEAIRTNYNDRY	D	1.38E-12
RPL8	M	GRVIRGQRKGAGSVFRAHVKHRKGAARLRAVD	F	1.35E-13
RPL8	M	GRVIRGQRKGAGSVFRAHVKHRKGAARLRAVDFA ERHGYIKGIVK	D	3.68E-06
RPL8	D	FAERHGYIKGIVKD	I	6.02E-12
RPL8	D	FAERHGYIKGIVKDIIHD	P	2.58E-24
RPL8	D	PGRGAPLAKVVFRD	P	3.05E-22
RPL8	D	PYRFKKRTELFIAAEGIHTGQFVYCGKKAQLNIGN VLPVGTMPEGTIVCCLEEKPG	D	7.64E-12
RPL8	D	RGKLARASGNYATVISHNPETKKTRVKLPSGSKKV ISSANRAVVGVVAGGGRID	K	3.11E-29
RPL9	-	MKTILSNQTV	D	3.91E-27

RPL9	D	IPENVDITLKGRTVIVKGPRGTLRR	D	0.0000176
RPL9	D	ITLKGRTVIVKGPRGTLRR	D	9.72E-15
RPL9	D	FNHINVELSLLGKKKKRLRVD	K	3.74E-07
RPL9	D	IELVSNSAALIQQATTVKNK	D	7.6E-15
RPL9	D	IELVSNSAALIQQATTVKNKDIRKFL	D	0.000081
RPL9	D	GIYVSEKGTVQQA	D	1.63E-10
RPL10	M	GRRPARCYRYCKNKPYPKSRFCRGVPD	A	0.000539
RPL10	D	AKIRIFDLGRKKAKVD	E	3.79E-15
RPL10	D	MVAEKRLIP	D	2.35E-06
RPL10A	D	TLYEAVREVLHGNQRKRRKFLETVELQISLKNY	D	1.07E-12
RPL10A	D	IPHMDIEALKKLNKNKKLVKKLAKKY	D	3.4E-12
RPL10A	D	IEALKKLNKNKKLVKKLAKKYD	A	3.01E-22
RPL10A	D	AFLASESLIKQIPRILGPGLNKAGKFPSLLTHNENMVAKVD	E	3.12E-13
RPL11	D	TGNFGFGIQEHID	L	5.91E-16
RPL11	D	PSIGIYGL	D	6.28E-10
RPL11	D	FYVVLGRPGFSIA	D	2.03E-15
RPL12	D	PNEIKVVYLRCTGGEVGATSALAPKIGPLGLSPKKVGD	D	1.77E-07
RPL13	M	APSRNGMVLKPHFHKD	W	0.0000316
RPL13	D	PRRRNKSTESLQANVQRLKEYRSKLILFPRKPSAPKKGD	S	1.64E-09
RPL13	D	SSAEELKLATQLTGPVMPVRNVYKKEKARVITEEEKNFKAFASLRMARANARLFGIRAKRAKEAAEQ	D	2.01E-17
RPL13A	D	KYTEVLKTHGLLV	-	3.84E-20

RPL14	M	VFRRFVEVGRVAYVSFGPHAGKLVAIVD	V	1.62E-09
RPL14	M	VFRRFVEVGRVAYVSFGPHAGKLVAIVDVI	D	1.68E-09
RPL14	M	VFRRFVEVGRVAYVSFGPHAGKLVAIVDVIDQNRALV	D	2.88E-16
RPL14	D	FILKFPHSAHQKYVRQAWQKAD	I	0.0000285
RPL14	D	INTKWAATRWAKKIEARERKAKMTD	F	0.0000477
RPL15	M	GAYKYIQELWRKKQSD	V	1.38E-08
RPL15	D	STYKFFEVILI	D	4.77E-13
RPL15	D	PFHKAIRRNPD	T	0.000212
RPL15	D	TQWITKPVHKHREMRGLTSAGRKSRGLGKGHKFHHTIGGSRRAAWRRRNTLQLHRYR	-	0.0000918
RPL17	D	PENPTKSCKSRGSNLRVHFKNTRETAQAIKGMHIRKATKYLKD	V	1.83E-11
RPL17	D	SLVIEHIQVNKAPKMRRRTYRAHGRINPYMSSPCHIEMILTEKEQIVPKPEEEVAQKKKISQKKLKKQKLMARE	-	0.000451
RPL18	D	VRVQEVPKLKVCALRVTSRARSRILRAGGKILTF	D	1.9E-07
RPL18A	D	LTTAGAVTQCYR	D	2.37E-08
RPL18A	D	MGARHRARAHSIQIMKVEEIAASKCRRPAVKQFHD	S	4.12E-17
RPL18A	D	SKIKFPLPHRVLRRQHKPRFTTKRPNTFF	-	2.71E-12
RPL19	M	SMLRLQKRLASSVLRCGKKKVWLD	P	1.69E-07
RPL19	D	PNETNEIANANSRQQIRKLIKD	G	9.52E-07
RPL19	D	RHMYHSLYLKVKGNVFKNKRILMEHIHKLKA	D	0.000529

RPL21	-	MTNTKGKRRGTRYMFSRPFRKHGVVPLATYMRIYKKG	D	0.000786
RPL21	D	IKGMGTVQKGMPHKCYHGKTGRVYNVTQHAVGIVVNKQVKGKILAKRINVRIEHIKHSKSR	D	3.64E-13
RPL22	M	APVKKLVVKGGKKKKQVLKFTLD	C	6.46E-11
RPL22	D	AANFEQFLQERIKVNGKAGNLGGGVVTIERSKSKITVTSEVPFSKRYLKYLTKKYLKKNNLR	D	0.000626
RPL23	M	SKRGRGGSSGAKFRISLGLPVGAVINCA	D	7.77E-14
RPL23	D	NTGAKNLYIISVKGIKGRLNRLPAAGVG	D	7.59E-21
RPL23	D	MVMATVKKGKPELRKKVHPAVVIRQRKSYRRKD	G	1.41E-09
RPL23	D	NAGVIVNNKGEMKGSAITGPVAKECA	D	1.64E-27
RPL23	D	LWPRIASNAGSIA	-	2.88E-10
RPL23A	D	VKANKHQIKQAVKKLY	D	8.2E-16
RPL23A	D	VKANKHQIKQAVKKLYDI	D	0.00068
RPL23A	D	IDVAKVNTLIRP	D	0.0000125
RPL23A	D	ALDVANKIGII	-	2.52E-21
RPL23A	D	VANKIGII	-	2.76E-20
RPL24	-	MKVELCSFSGYKIYPGHGRRYART	D	7.33E-15
RPL24	D	IMAKRNQKPEVRKAQREQAIRAAKEAKKAKQASKKTAMAAAKAPTKAAPKQKIVKPVKVSAPRVGGKR	-	1.46E-21
RPL26	-	MKFNPFVTSD	R	2.47E-14
RPL26	D	DEVQVVRGHYKGQQIGKVVQVYRKKYVIYIERVQREKANGTTVHVGIHPSKVVITRLKL	D	0.0000417
RPL26	D	RKKILERKAKSRQVGKEKGKYKEETIEKMQE	-	8.23E-17
RPL27	M	GKFMKPGKVVLVLAGRYSGRKAVIVKNID	D	4.45E-17
RPL27	D	RPYSHALVAGI	D	7.16E-09

RPL27	D	RPYSHALVAGIDRYPRKVTAAMGKKKIAKRSKIKSFVKVYNYNHLMPTRYSV	D	0.000213
RPL27	D	RYPRKVTAAMGKKKIAKRSKIKSFVKVYNYNHLMPTRYSVD	I	1.01E-10
RPL27	D	IPLDKTVVNKDVFR	D	3.18E-19
RPL27	D	PALKRKARREAKVKFEERYKTGKNKWFFQKLRF	-	4.25E-21
RPL27A	D	KYHPGYFGKVGMKHYHLKRNQSFCPTVNL	D	5.55E-09
RPL27A	D	KLWTLVSEQTRVNAAKNKTGAAPIID	V	0.0000229
RPL27A	D	VVRSGYYKVLGKGKLPKQPVIVKAKFFSRRAEEKIKSVGGACVLVA	-	3.43E-09
RPL28	M	SAHLQWMVVRNCSSFLIKRNKQTYSTEPNNLKARNSFRYNGLIHRKTVGVEPAA	D	1.08E-09
RPL28	D	LRMAAIRRASAILRSQKPVMVKRKRTRPTKSS	-	0.0000799
RPL29	D	PKFLRNMRFAKKHNKKGLKKMQANNAKAMSARAEAIKALVKPKEVKPKIPKGVSRKLD	R	9.01E-15
RPL30	D	IIRSMPEQTGEK	-	8.84E-16
RPL31	D	TRLNKAVWAKGIRNVPYRIRVRLSRKRNE	D	0.0000804
RPL31	D	SPNKLYTLVTYVPVTTFKNLQTVNVD	E	4.47E-15
RPL32	M	AALRPLVKPKIVKKRTKKFIRHQSD	R	2.19E-18
RPL32	D	RYVKIKRNWRKPRGI	D	0.000455
RPL34	D	RIKRAFLIEEQKIVVKVLKAQAQSQKAK	-	2.17E-38
RPL35	M	AKIKARDLRGKKKEELLKQL	D	2.69E-11
RPL35	D	LRGKKKEELLKQL	D	1.06E-27

RPL35	D	LKVELSQLRVAKVTGGAASKLSKIRVVRKSIARVLTVINQTQKENLRKFYKGKKYKPL	D	5.15E-09
RPL35	D	LRPKKTRAMRRRLNKHEENLKTKKQQRKERLYPLRKYAVKA	-	8.42E-11
RPL35A	M	SGRLWSKAIFAGYKRGLRNQREHTALLKIEGVYARD	E	2.35E-08
RPL36	M	ALRYPMAVGLNKGHKVTKNVSKPRHSRRRGRLTKHTKFVRD	M	3.52E-17
RPL36	D	MIREVCGFAPYERRAMELLKVSKD	K	0.0000072
RPL36	D	KRALKFIKKRVGTHIRAKRKREELSNVLAAMRKAAAKK	D	4.58E-13
RPL36A	M	VNVPKTRRTFCKKCGKHQPHKVTQYKKGKD	S	6.53E-17
RPL36A	M	VNVPKTRRTFCKKCGKHQPHKVTQYKKGKDSLYAQGKRRY	D	1.96E-15
RPL36A	D	KKRKGQVIQF	-	0.000113
RPL38	M	PRKIEEIKDFLLTARRKD	A	8.29E-13
RPL38	D	NVKFKVRCSRYLYTLVITDKEKAEKLKQSLPPGLAVKELK	-	2.47E-07
RPL38	D	KEKAEKLKQSLPPGLAVKELK	-	1.01E-23
RPLP0	D	RATWKSNYFLKIIQLL	D	5.81E-11
RPLP0	D	MLLANKVPAAARAGAIAPCEVTVPAQNTGLGPEKTSFFQALGITTKISRGTIEILS	D	1.47E-10
RPLP0	D	YTFPLAEKVKAFLA	D	3.32E-25
RPLP2	-	MRYVASYLLAALGGNSSPSAK	D	2.44E-25
RPLP2	-	MRYVASYLLAALGGNSSPSAKDIKKIL	D	3.99E-14
RPLP2	D	RLNKVISELNGKNIED	V	3.09E-06

RPLP2	D	VIAQGIGKLASVPAGGAVAVSAAPGSAAPAAGSAP AAAEEKK	D	2.4E-20

Appendix Table 1: Acid Digestion peptide identifications and their corresponding ProSightPC 2.0 E-Values

Bibliography

1. Hays, F. A. (2006). Mass Spectrometric Analysis of Cytoplasmic Ribosomal Proteins in Drug Resistant and Drug Susceptible Human Cell Lines. Doctor of Philosophy, University of Maryland.
2. Cannon, J., Lohnes, K., Wynne, C., Wang, Y., Edwards, N. & Fenselau, C. (2010). High-Throughput Middle-Down Analysis Using an Orbitrap. *Journal of Proteome Research* **9**, 3886-3890.
3. Nobelprize.org. (2009). The Nobel Prize in Chemistry 2009.
4. Spirin, A. S. (2002). Ribosome as a Molecular Machine. *FEBS Letters Protein Biosynthesis* **514**, 2-10.
5. Perry, R. (2005). The Architecture of Mammalian Ribosomal Protein Promoters. *BMC Evolutionary Biology* **5**, 15.
6. Perry, R. P. (2007). Balanced Production of Ribosomal Proteins. *Gene* **401**, 1-3.
7. Warner, J. R. (1977). In the Absence of Ribosomal RNA Synthesis, the Ribosomal Proteins of HeLa Cells are Synthesized Normally and Degraded Rapidly. *Journal of Molecular Biology* **115**, 315-333.
8. Lam, Y. W., Lamond, A. I., Mann, M. & Andersen, J. S. (2007). Analysis of Nucleolar Protein Dynamics Reveals the Nuclear Degradation of Ribosomal Proteins. *Current biology : CB* **17**, 749-760.
9. Cech, T. R. (2000). The Ribosome Is a Ribozyme. *Science* **289**, 878-879.
10. Matragkou, C. N., Eleni T. Papachristou, Sotirios S. Tezias, Asterios S. Tsiftsoglou, Theodora Choli-Papadopoulou, Ioannis S. Vizirianakis,. (2008). The Potential Role of Ribosomal Protein S5 on Cell Cycle Arrest and Initiation of Murine Erythroleukemia Cell Differentiation. *Journal of Cellular Biochemistry* **104**, 1477-1490.
11. Worbs, M., Wahl, M. C., Lindahl, L. & Zengel, J. M. (2002). Comparative Anatomy of a Regulatory Ribosomal Protein. *Biochimie* **84**, 731-743.
12. Lafontaine, D. L. J. & Tollervey, D. (2001). The Function and Synthesis of Ribosomes. *Nat Rev Mol Cell Biol* **2**, 514-520.
13. Steitz, T. A. (2008). A Structural Understanding of the Dynamic Ribosome Machine. *Nat Rev Mol Cell Biol* **9**, 242-253.
14. Chandramouli, P., Topf, M., Ménétret, J.-F., Eswar, N., Cannone, J. J., Gutell, Robin R., Sali, A. & Akey, C. W. (2008). Structure of the Mammalian 80S Ribosome at 8.7 Å Resolution. *Structure* **16**, 535-548.
15. Kenmochi, N., Kawaguchi, T., Rozen, S., Davis, E., Goodman, N., Hudson, T. J., Tanaka, T. & Page, D. C. (1998). A Map of 75 Human Ribosomal Protein Genes. *Genome Res.* **8**, 509-523.
16. Lewis, J. D. & Tollervey, D. (2000). Like Attracts Like: Getting RNA Processing Together in the Nucleus. *Science* **288**, 1385-1389.
17. Lecompte, O., Ripp, R., Thierry, J.-C., Moras, D. & Poch, O. (2002). Comparative Analysis of Ribosomal Proteins in Complete Genomes: An

Example of Reductive Evolution at the Domain Scale. *Nucl. Acids Res.* **30**, 5382-5390.
18. Wilson, D. N. & Nierhaus, K. H. (2005). Ribosomal Proteins in the Spotlight. *Critical Reviews in Biochemistry and Molecular Biology* **40**, 243 - 267.
19. Henras, A., Soudet, J., GÃcrus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. & Henry, Y. (2008). The Post-transcriptional Steps of Eukaryotic Ribosome Biogenesis. *Cellular and Molecular Life Sciences (CMLS)* **65**, 2334-2359.
20. Nakao, A., Yoshihama, M. & Kenmochi, N. (2004). RPG: the Ribosomal Protein Gene database. *Nucl. Acids Res.* **32**, D168-170.
21. Nazar, R. N. (1980). A 5.8 S rRNA-like Sequence in Prokaryotic 23 S rRNA. *FEBS Letters* **119**, 212-214.
22. Wool, I. G., Yuen-Ling Chan, and Anton Glück. (1995). Structure and Evolution of Mammalian Ribosomal Proteins. *Biochem. Cell Biol.* **73**, 933-947.
23. Xaplanteri, M. A., Papadopoulos, G., Leontiadou, F., Choli-Papadopoulou, T. & Kalpaxis, D. L. (2007). The Contribution of the Zinc-Finger Motif to the Function of Thermus thermophilus Ribosomal Protein S14. *Journal of Molecular Biology* **369**, 489-497.
24. Dresios, J., Chan, Y.-L. & Wool, I. G. (2002). The Role of the Zinc Finger Motif and of the Residues at the Amino Terminus in the Function of Yeast Ribosomal Protein YL37a. *Journal of Molecular Biology* **316**, 475-488.
25. Mazumder, B., Sampath, P., Seshadri, V., Maitra, R. K., DiCorleto, P. E. & Fox, P. L. (2003). Regulated Release of L13a from the 60S Ribosomal Subunit as A Mechanism of Transcript-Specific Translational Control. *Cell* **115**, 187-198.
26. Hemmerich, P., Stefan Bosbach, Anna Mikecz, Ulrich Krawinkel,. (1997). Human Ribosomal Protein L7 Binds RNA with an alpha-Helical Arginine-Rich and Lysine-Rich Domain. *European Journal of Biochemistry* **245**, 549-556.
27. Baxter-Roshek, J. L., Petrov, A. N. & Dinman, J. D. (2007). Optimization of Ribosome Structure and Function by rRNA Base Modification. *PLoS ONE* **2**, e174.
28. Jack, K., Bellodi, C., Landry, Dori M., Niederer, Rachel O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, Alison M., Thompson, Sunnie R., Ruggero, D. & Dinman, Jonathan D. (2011). rRNA Pseudouridylation Defects Affect Ribosomal Ligand Binding and Translational Fidelity from Yeast to Human Cells. *Molecular Cell* **44**, 660-666.
29. Acker, M. G. & Lorsch, J. R. (2008). Mechanism of Ribosomal Subunit Joining During Eukaryotic Translation Initiation. *Biochemical Society Transactions* **036**, 653-657.
30. Gebauer, F. & Hentze, M. W. (2004). Molecular Mechanisms of Translational Control. *Nat Rev Mol Cell Biol* **5**, 827-835.
31. Kapp, L. D. & Lorsch, J. R. (2004). The Molecular Mechanics of Eukaryotic Translation. *Annual Review of Biochemistry* **73**, 657-704.

32. Lee, T.-H., Blanchard, S. C., Kim, H. D., Puglisi, J. D. & Chu, S. (2007). The Role of Fluctuations in tRNA Selection by the Ribosome. *Proceedings of the National Academy of Sciences* **104**, 13661-13665.
33. Budkevich, T. V., El'skaya, A. V. & Nierhaus, K. H. (2008). Features of 80S Mammalian Ribosome and its Subunits. *Nucl. Acids Res.* **36**, 4736-4744.
34. Pisarev, A. V., Hellen, C. U. T. & Pestova, T. V. (2007). Recycling of Eukaryotic Posttermination Ribosomal Complexes. *Cell* **131**, 286-299.
35. Nakagawa, M., Schneider, E., Dixon, K. H., Horton, J., Kelley, K., Morrow, C. & Cowan, K. H. (1992). Reduced Intracellular Drug Accumulation in the Absence of P-glycoprotein (mdr1) Overexpression in Mitoxantrone-Resistant Human MCF-7 Breast Cancer Cells. *Cancer Res* **52**, 6175-81.
36. Burger, K., Mühl, B., Harasim, T., Rohrmoser, M., Malamoussi, A., Orban, M., Kellner, M., Gruber-Eber, A., Kremmer, E., Hölzel, M. & Eick, D. (2010). Chemotherapeutic Drugs Inhibit Ribosome Biogenesis at Various Levels. *Journal of Biological Chemistry* **285**, 12416-12425.
37. Mewes, K., Blanz, J., Ehninger, G., Gebhardt, R. & Zeller, K.-P. (1993). Cytochrome P-450-induced Cytotoxicity of Mitoxantrone by Formation of Electrophilic Intermediates. *Cancer Research* **53**, 5135-5142.
38. Gopinath, S. C. B., Matsugami, A., Katahira, M. & Kumar, P. K. R. (2005). Human Vault-Associated Non-Coding RNAs Bind to Mitoxantrone, a Chemotherapeutic Compound. *Nucleic Acids Research* **33**, 4874-4881.
39. Gopinath, S. C. B., Wadhwa, R. & Kumar, P. K. R. (2010). Expression of Noncoding Vault RNA in Human Malignant Cells and Its Importance in Mitoxantrone Resistance. *Molecular Cancer Research* **8**, 1536-1546.
40. Meskauskas, A. & Dinman, J. D. (2007). Ribosomal Protein L3: Gatekeeper to the A Site. *Molecular Cell* **25**, 877-888.
41. Meskauskas, A., Russ, J. R. & Dinman, J. D. (2008). Structure/Function Analysis of Yeast Ribosomal Protein L2. *Nucl. Acids Res.* **36**, 1826-1835.
42. Kawai, S., Murao, S., Mochizuki, M., Shibuya, I., Yano, K. & Takagi, M. (1992). Drastic Alteration of Cycloheximide Sensitivity by Substitution of One Amino Acid in the L41 Ribosomal Protein of Yeasts. *J. Bacteriol.* **174**, 254-262.
43. Wilcox, S. K., Cavey, G. S. & Pearson, J. D. (2001). Single Ribosomal Protein Mutations in Antibiotic-Resistant Bacteria Analyzed by Mass Spectrometry. *Antimicrob. Agents Chemother.* **45**, 3046-3055.
44. Carr, J. F., Gregory, S. T. & Dahlberg, A. E. (2005). Severity of the Streptomycin Resistance and Streptomycin Dependence Phenotypes of Ribosomal Protein S12 of Thermus thermophilus Depends on the Identity of Highly Conserved Amino Acid Residues. *Journal of Bacteriology* **187**, 3548-3550.
45. Carr, J. F., Hamburg, D.-M., Gregory, S. T., Limbach, P. A. & Dahlberg, A. E. (2006). Effects of Streptomycin Resistance Mutations on Posttranslational Modification of Ribosomal Protein S12. *J. Bacteriol.* **188**, 2020-2023.
46. Warner, J. R. & McIntosh, K. B. (2009). How Common Are Extraribosomal Functions of Ribosomal Proteins? *Molecular cell* **34**, 3-11.

47. Shi, Y., Zhai, H., Wang, X., Han, Z., Liu, C., Lan, M., Du, J., Guo, C., Zhang, Y., Wu, K. & Fan, D. (2004). Ribosomal Proteins S13 and L23 Promote Multidrug Resistance in Gastric Cancer Cells by Suppressing Drug-Induced Apoptosis. *Experimental Cell Research* **296**, 337-346.
48. Gottesman, M. M., Fojo, T. & Bates, S. E. (2002). Multidrug Resistance in Cancer: Role of ATP-Dependent Transporters. *Nat Rev Cancer* **2**, 48-58.
49. Wasinger, V. C., Stuart J. Cordwell, Anne Cerpa-Poljak, Jun X. Yan, Andrew A. Gooley, Marc R. Wilkins, Mark W. Duncan, Ray Harris, Keith L. Williams, Ian Humphery-Smith. (1995). Progress with Gene-Product Mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **16**, 1090-1094.
50. Wilkins, M. R. S., Jean-Charles; Gooley, Andrew A.; Appel, Ron D.; Humphery-Smith, Ian; Hochstrasser, Denis F.; Williams, Keith L. (1996). Progress with Proteome Projects: Why all Proteins Expressed by a Genome Should be Identified and How to Do It. *Biotechnology & Genetic Engineering Reviews* **13**, 19-50.
51. Anderson, N. L., Norman G. Anderson. (1998). Proteome and Proteomics: New Technologies, New Concepts, and New Words. *Electrophoresis* **19**, 1853-1861.
52. Dass, C. (2007). *Fundamentals of Contemporary Mass Spectrometry*, John Wiley and Sons Ltd.
53. Consortium, T. U. (2011). Ongoing and Future Developments At the Universal Protein Resource. *Nucleic acids research* **39**, D214-D219.
54. Hayden, E. C. (2011). Cells May Stray From 'Central Dogma'; The Ability to Edit RNA to Produce 'New' Protein-Coding Sequences Could be Widespread in Human Cells. *Nature* doi: doi:10.1038/news.2011.304
55. Lei, T., He, Q.-Y., Wang, Y.-L., Si, L.-S. & Chiu, J.-F. (2008). Heparin Chromatography to Deplete High-Abundance Proteins for Serum Proteomics. *Clinica Chimica Acta* **388**, 173-178.
56. Simpson, R. J. (2003). *Proteins and Proteomics : a Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
57. Eng, J. K., McCormack, A. L. & Yates, J. R. (1994). An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *Journal of the American Society for Mass Spectrometry* **5**, 976-989.
58. Siuti, N. & Kelleher, N. L. (2007). Decoding Protein Modifications Using Top-Down Mass Spectrometry. *Nature Methods* **4**, 817-821.
59. Pesavento, J. J., Mizzen, C. A. & Kelleher, N. L. (2006). Quantitative Analysis of Modified Proteins and Their Positional Isomers by Tandem Mass Spectrometry: Human Histone H4. *Anal. Chem.* **78**, 4271-4280.
60. Pesavento, J. J., Kim, Y.-B., Taylor, G. K. & Kelleher, N. L. (2004). Shotgun Annotation of Histone Modifications: A New Approach for Streamlined Characterization of Proteins by Top Down Mass Spectrometry. *J. Am. Chem. Soc.* **126**, 3386-3387.
61. Doerr, A. (2008). Top-Down Mass Spectrometry. *Nature Methods* **5**, 24.
62. Zee BM, Young NL & BA, G. (2011). Quantitative Proteomic Approaches to Studying Histone Modifications. *Curr Chem Genomics* **5**, 106-14.

63. Mann, M. & Kelleher, N. L. (2008). Precision Proteomics: The Case for High Resolution and High Mass Accuracy. *Proceedings of the National Academy of Sciences* **105**, 18132-18138.
64. Swatkoski, S., Gutierrez, P., Wynne, C., Petrov, A., Dinman, J., Edwards, N. & Fenselau, C. (2008). Evaluation of Microwave-Accelerated Residue-Specific Acid Cleavage for Proteomic Applications. *J. Proteome Res.* **7**, 579-586.
65. Hauser, N. J., Han, H., McLuckey, S. A. & Basile, F. (2008). Electron Transfer Dissociation of Peptides Generated by Microwave D-Cleavage Digestion of Proteins. *Journal of Proteome Research* **7**, 1867-1872.
66. Wu, S.-L., Hühmer, A. F. R., Hao, Z. & Karger, B. L. (2007). On-Line LC−MS Approach Combining Collision-Induced Dissociation (CID), Electron-Transfer Dissociation (ETD), and CID of an Isolated Charge-Reduced Species for the Trace-Level Characterization of Proteins with Post-Translational Modifications. *Journal of Proteome Research* **6**, 4230-4244.
67. Lee, S.-W., Berger, S. J., Martinović, S., Paša-Tolić, L., Anderson, G. A., Shen, Y., Zhao, R. & Smith, R. D. (2002). Direct Mass Spectrometric Analysis of Intact Proteins of the Yeast Large Ribosomal Subunit Using Capillary LC/FTICR. *Proceedings of the National Academy of Sciences* **99**, 5942-5947.
68. Vladimirov, S. N., Anton V. Ivanov, Galina G. Karpova, Alekxander K. Musolyamov, Tsezi A. Egorov, Bernd Thiede, Brigitte Wittmann-Liebold & Albrecht Otto. (1996). Characterization of the Human Small-Ribosomal-Subunit Proteins by N-Terminal and Internal Sequencing, and Mass Spectrometry. *European Journal of Biochemistry* **239**, 144-149.
69. Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M. & Yates, J. R. (1999). Direct Analysis of Protein Complexes Using Mass Spectrometry. *Nat Biotech* **17**, 676-682.
70. Forbes, A. J., Patrie, S. M., Taylor, G. K., Kim, Y.-B., Jiang, L. & Kelleher, N. L. (2004). Targeted Analysis and Discovery of Posttranslational Modifications in Proteins from Methanogenic Archaea by Top-Down MS. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 2678-2683.
71. Strader, M. B., VerBerkmoes, N. C., Tabb, D. L., Connelly, H. M., Barton, J. W., Bruce, B. D., Pelletier, D. A., Davison, B. H., Hettich, R. L., Larimer, F. W. & Hurst, G. B. (2004). Characterization of the 70S Ribosome from Rhodopseudomonas palustris Using an Integrated "Top-Down" and "Bottom-Up" Mass Spectrometric Approach. *J. Proteome Res.* **3**, 965-978.
72. Running, W. E., Ravipaty, S., Karty, J. A. & Reilly, J. P. (2007). A Top-Down/Bottom-Up Study of the Ribosomal Proteins of Caulobacter crescentus. *J. Proteome Res.* **6**, 337-347.
73. Louie, D. F., Resing, K. A., Lewis, T. S. & Ahn, N. G. (1996). Mass Spectrometric Analysis of 40 S Ribosomal Proteins from Rat-1 Fibroblasts. *J. Biol. Chem.* **271**, 28189-28198.

74. Yu, Y., Ji, H., Doudna, J. A. & Leary, J. A. (2005). Mass Spectrometric Analysis of the Human 40S Ribosomal Subunit: Native and HCV IRES-Bound Complexes. *Protein Sci* **14**, 1438-1446.
75. Carroll, A. J., Heazlewood, J. L., Ito, J. & Millar, A. H. (2008). Analysis of the Arabidopsis Cytosolic Ribosome Proteome Provides Detailed Insights into Its Components and Their Post-translational Modification. *Mol Cell Proteomics* **7**, 347-369.
76. van Riggelen, J., Yetil, A. & Felsher, D. W. (2010). MYC as a Regulator of Ribosome Biogenesis and Protein Synthesis. *Nat Rev Cancer* **10**, 301-309.
77. Bykhovskaya, Y., Mengesha, E. & Fischel-Ghodsian, N. (2009). Phenotypic Expression of Maternally Inherited Deafness is Affected by RNA Modification and Cytoplasmic Ribosomal Proteins. *Molecular Genetics and Metabolism* **97**, 297-304.
78. Mao-De, L. & Jing, X. (2007). Ribosomal Proteins and Colorectal Cancer. *Current Genomics* **8**, 43-49.
79. Dai, M.-S. & Lu, H. (2008). Crosstalk Between c-Myc and Ribosome in Ribosomal Biogenesis and Cancer. *Journal of Cellular Biochemistry* **105**, 670-677.
80. Ellis, S. R. & Massey, A. T. (2006). Diamond Blackfan Anemia: A Paradigm for a Ribosome-Based Disease. *Medical hypotheses* **66**, 643-648.
81. Zhang, Y. & Lu, H. (2009). Signaling to p53: Ribosomal Proteins Find Their Way. *Cancer Cell* **16**, 369-377.
82. Mauro, V. P. & Edelman, G. M. (2007). The Ribosome Filter Redux. *Cell Cycle* **6**, 2246-2251.
83. Mauro, V. P. & Edelman, G. M. (2002). The Ribosome Filter Hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12031-12036.
84. Haselbacher, G. K., Humbel, R. E. & Thomas, G. (1979). Insulin-Like Growth Factor: Insulin or Serum Increase Phosphorylation of Ribosomal Protein S6 During Transition of Stationary Chick Embryo Fibroblasts into Early G1 Phase of the Cell Cycle. *FEBS Letters* **100**, 185-190.
85. Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M. & Finley, D. (2000). Cell Cycle-Regulated Modification of the Ribosome by a Variant Multiubiquitin Chain. *Cell* **102**, 67-76.
86. Porras-Yakushi, T., Whitelegge, J. & Clarke, S. (2006). A Novel SET Domain Methyltransferase in Yeast: Rkm2-Dependent Trimethylation of Ribosomal Protein L12ab at Lysine 10. *The Journal of biological chemistry* **281**, 35835 - 35845.
87. Komili, S., Farny, N. G., Roth, F. P. & Silver, P. A. (2007). Functional Specificity among Ribosomal Proteins Regulates Gene Expression. *Cell* **131**, 557-571.
88. Bellodi, C., Krasnykh, O., Haynes, N., Theodoropoulou, M., Peng, G., Montanaro, L. & Ruggero, D. (2010). Loss of Function of the Tumor Suppressor DKC1 Perturbs p27 Translation Control and Contributes to Pituitary Tumorigenesis. *Cancer Research* **70**, 6026-6035.

89. Kondrashov, N., Pusic, A., Stumpf, C. R., Shimizu, K., Hsieh, Andrew C., Xue, S., Ishijima, J., Shiroishi, T. & Barna, M. (2011). Ribosome-Mediated Specificity in Hox mRNA Translation and Vertebrate Tissue Patterning. *Cell* **145**, 383-397.
90. Kasai, H., Nadano, D., Hidaka, E., Higuchi, K., Kawakubo, M., Sato, T.-A. & Nakayama, J. (2003). Differential Expression of Ribosomal Proteins in Human Normal and Neoplastic Colorectum. *J. Histochem. Cytochem.* **51**, 567-574.
91. Hardy SJ, K. C., Voynow P, Mora G. (1969). The Ribosomal Proteins of Escherichia coli. I. Purification of the 30S Ribosomal Proteins. *Biochemistry.* **8**, 2897-905.
92. Barritault, D., Expert-BezanÇOn, A., GuÉRin, M.-F. & Hayes, D. (1976). The Use of Acetone Precipitation in the Isolation of Ribosomal Proteins. *European Journal of Biochemistry* **63**, 131-135.
93. Görg, A., Obermaier, C., Boguth, G., Csordas, A., Diaz, J.-J. & Madjar, J.-J. (1997). Very Alkaline Immobilized pH Gradients for Two-Dimensional Electrophoresis of Ribosomal and Nuclear Proteins. *Electrophoresis* **18**, 328-337.
94. Görg, A., Weiss, W. & Dunn, M. J. (2004). Current Two-Dimensional Electrophoresis Technology for Proteomics. *Proteomics* **4**, 3665-3685.
95. Kane, L. A., Yung, C. K., Agnetti, G., Neverova, I. & Van Eyk, J. E. (2006). Optimization of Paper Bridge Loading for 2-DE Analysis in the Basic pH Region: Application to the Mitochondrial Subproteome. *Proteomics* **6**, 5683-5687.
96. Lamberti, C., Pessione, E., Giuffrida, M. G., Mazzoli, R., Barello, C., Conti, A. & Giunta, C. (2007). Combined cup loading, bis(2-hydroxyethyl) disulfide, and protein precipitation protocols to improve the alkaline proteome of Lactobacillus hilgardii. *Electrophoresis* **28**, 1633-1638.
97. Görg, A., Drews, O., Lück, C., Weiland, F. & Weiss, W. (2009). 2-DE with IPGs. *Electrophoresis* **30**, S122-S132.
98. Bartkowiak, K., Wieczorek, M., Buck, F., Harder, S. n., Moldenhauer, J., Effenberger, K. E., Pantel, K., Peter-Katalinic, J. & Brandt, B. H. (2009). Two-Dimensional Differential Gel Electrophoresis of a Cell Line Derived from a Breast Cancer Micrometastasis Revealed a Stem/Progenitor Cell Protein Profile. *Journal of Proteome Research* **8**, 2004-2014.
99. Gorg, A., Klaus, A., Luck, C., Welland, F. & Weiss, W. (2007). Two-dimensional Electrophoresis with Immobilized pH Gradients for Proteome Analysis. In *A Laboratory Manual* 3 edit., pp. 166. Technical University of Munich, Freising-Weihenstephan, Germany.
100. Li, X.-M., Patel, B. B., Blagoi, E. L., Patterson, M. D., Seeholzer, S. H., Zhang, T., Damle, S., Gao, Z., Boman, B. & Yeung, A. T. (2004). Analyzing Alkaline Proteins in Human Colon Crypt Proteome. *Journal of Proteome Research* **3**, 821-833.
101. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. & Mann, M. (2007). In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. *Nat. Protocols* **1**, 2856-2860.

102. Mirza, U. A., Liu, Y.-H., Tang, J. T., Porter, F., Bondoc, L., Chen, G., Pramanik, B. N. & Nagabhushan, T. L. (2000). Extraction and Characterization of Adenovirus Proteins from Sodium Dodecylsulfate Polyacrylamide Gel Electrophoresis by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **11**, 356-361.
103. Edwards, N., Wu, X. & Tseng, C.-W. (2009). An Unsupervised, Model-Free, Machine-Learning Combiner for Peptide Identifications from Tandem Mass Spectra. *Clinical Proteomics* **5**, 23-36.
104. Wynne, C., Fenselau, C., Demirev, P. A. & Edwards, N. (2009). Top-Down Identification of Protein Biomarkers in Bacteria with Unsequenced Genomes. *Analytical Chemistry* **81**, 9633-9642.
105. Kamp, R. M. & Wittmann-Liebold, B. (1984). Purification of Escherichia coli 50 S Ribosomal Proteins by High Performance Liquid Chromatography. *FEBS Letters* **167**, 59-63.
106. Cooperman, B. S., Weitzmann, C. J. & Buck, M. A. (1988). Reversed-Phase High-Performance Liquid Chromatography of Ribosomal Proteins. In *Methods in Enzymology* (Harry F. Noller, J. K. M., ed.), Vol. Volume 164, pp. 523-532. Academic Press.
107. Carroll, A., Heazlewood, J., Ito, J. & Millar, A. (2008). Analysis of the Arabidopsis Cytosolic Ribosome Proteome Provides Detailed Insights Into Its Components and Their Post-Translational Modification. *Mol Cell Proteomics* **7**, 347 - 369.
108. Gilar, M., Bouvier, E. S. P. & Compton, B. J. (2001). Advances in Sample Preparation in Electromigration, Chromatographic and Mass Spectrometric Separation Methods. *Journal of Chromatography A* **909**, 111-135.
109. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V. & Mann, M. (2009). Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* **325**, 834-840.
110. Schwer, B., Eckersdorff, M., Li, Y., Silva, J. C., Fermin, D., Kurtev, M. V., Giallourakis, C., Comb, M. J., Alt, F. W. & Lombard, D. B. (2009). Calorie Restriction Alters Mitochondrial Protein Acetylation. *Aging Cell* **8**, 604-606.
111. Odintsova, T. I., Müller, E.-C., Ivanov, A. V., Egorov, T. A., Bienert, R., Vladimirov, S. N., Kostka, S., Otto, A., Wittmann-Liebold, B. & Karpova, G. G. (2003). Characterization and Analysis of Posttranslational Modifications of the Human Large Cytoplasmic Ribosomal Subunit Proteins by Mass Spectrometry and Edman Sequencing. *Journal of Protein Chemistry* **22**, 249-258.
112. Ren, J., Wang, Y., Liang, Y., Zhang, Y., Bao, S. & Xu, Z. (2010). Methylation of Ribosomal Protein S10 by Protein-arginine Methyltransferase 5 Regulates Ribosome Biogenesis. *Journal of Biological Chemistry* **285**, 12695-12705.
113. Dez, C. & Tollervey, D. (2004). Ribosome Synthesis Meets the Cell Cycle. *Current Opinion in Microbiology* **7**, 631-637.

114. Di, R., Blechl, A., Dill-Macky, R., Tortora, A. & Tumer, N. E. (2010). Expression of a Truncated Form of Yeast Ribosomal Protein L3 in Transgenic Wheat Improves Resistance to Fusarium Head Blight. *Plant Science* **178**, 374-380.
115. Kamita, M., Kimura, Y., Ino, Y., Kamp, R. M., Polevoda, B., Sherman, F. & Hirano, H. (2011). Nα-Acetylation of Yeast Ribosomal Proteins and its Effect on Protein Synthesis. *Journal of Proteomics* **74**, 431-441.
116. Bommer, U.-A. & Stahl, J. (2001). Ribosomal Proteins in Eukaryotes. In *eLS*. John Wiley & Sons, Ltd.
117. Consortium, T. U. (2012). Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic acids research* **40**, D71-D75.
118. Rigbolt, K. T. G., Prokhorova, T. A., Akimov, V., Henningsen, J., Johansen, P. T., Kratchmarova, I., Kassem, M., Mann, M., Olsen, J. V. & Blagoev, B. (2011). System-Wide Temporal Characterization of the Proteome and Phosphoproteome of Human Embryonic Stem Cell Differentiation. *Sci. Signal.* **4**, rs3-.
119. Hornbeck, P. V., Chabra, I., Kornhauser, J. M., Skrzypek, E. & Zhang, B. (2004). PhosphoSite: A Bioinformatics Resource Dedicated to Physiological Protein Phosphorylation. *Proteomics* **4**, 1551-1561.
120. Moritz, A., Li, Y., Guo, A., Villen, J., Wang, Y., MacNeill, J., Kornhauser, J., Sprott, K., Zhou, J., Possemato, A., Ren, J. M., Hornbeck, P., Cantley, L. C., Gygi, S. P., Rush, J. & Comb, M. J. (2010). Akt-RSK-S6 Kinase Signaling Networks Activated by Oncogenic Receptor Tyrosine Kinases. *Sci. Signal.* **3**, ra64-.
121. Liu, H., Sadygov, R. G. & Yates, J. R. (2004). A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. *Analytical Chemistry* **76**, 4193-4201.
122. Sadoul, K., Boyault, C., Pabion, M. & Khochbin, S. (2008). Regulation of Protein Turnover by Acetyltransferases and Deacetylases. *Biochimie* **90**, 306-312.
123. Caron, C., Boyault, C. & Khochbin, S. (2005). Regulatory Cross-Talk Between Lysine Acetylation and Ubiquitination: Role in the Control of Protein Stability. *BioEssays* **27**, 408-415.
124. Yang, X.-J. & Seto, E. (2008). Lysine Acetylation: Codified Crosstalk with Other Posttranslational Modifications. *Molecular Cell* **31**, 449-461.
125. Kreunin, P., Yoo, C., Urquidi, V., Lubman, D. M. & Goodison, S. (2007). Differential Expression of Ribosomal Proteins in a Human Metastasis Model Identified by Coupling 2-D Liquid Chromatography and Mass Spectrometry. *Cancer Genomics - Proteomics* **4**, 329-339.
126. Suh, M.-J. (2004). Investigation of methods suitable for the matrix-assisted laser desorption/ionization mass spectrometric analysis of proteins from ribonucleoprotein complexes. *European mass spectrometry* **10**, 89-99.
127. Lee, W.-K., Lee, S.-Y., Na, J.-H., Jang, S., Park, C. R., Kim, S.-Y., Lee, S.-H., Han, K.-H. & Yu, Y. G. (2012). Mitoxantrone Enhances the Interaction Between Nopp140 and CK2 Mitoxantrone Binds to Nopp140, an Intrinsically

Unstructured Protein, and Modulate its Interaction with Protein Kinase CK2. *Bulletin of the Korean Chemical Society* **33**, 2005-2011.
128. Chen, H.-K., Pai, C.-Y., Huang, J.-Y. & Yeh, N.-H. (1999). Human Nopp140, Which Interacts with RNA Polymerase I: Implications for rRNA Gene Transcription and Nucleolar Structural Organization. *Molecular and cellular biology* **19**, 8536-8546.
129. Deisenroth, C. & Zhang, Y. (2010). Ribosome Biogenesis Surveillance: Probing the Ribosomal Protein-Mdm2-p53 Pathway. *Oncogene* **29**, 4253-4260.
130. Challagundla, K. B., Sun, X.-X., Zhang, X., DeVine, T., Zhang, Q., Sears, R. C. & Dai, M.-S. (2011). Ribosomal Protein L11 Recruits miR-24/miRISC To Repress c-Myc Expression in Response to Ribosomal Stress. *Molecular and cellular biology* **31**, 4007-4021.
131. Oh, W. J., Wu, C. c., Kim, S. J., Facchinetti, V., Julien, L. A., Finlan, M., Roux, P. P., Su, B. & Jacinto, E. (2010). mTORC2 Can Associate with Ribosomes to Promote Cotranslational Phosphorylation and Stability of Nascent Akt Polypeptide. *EMBO J* **29**, 3939-3951.
132. Zinzalla, V., Stracka, D., Oppliger, W. & Hall, Michael N. (2011). Activation of mTORC2 by Association with the Ribosome. *Cell* **144**, 757-768.
133. Mayer, C. & Grummt, I. (2006). Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**, 6384-6391.
134. Montanaro, L., Treré, D. & Derenzini, M. (2012). Changes in Ribosome Biogenesis May Induce Cancer by Down-Regulating the Cell Tumor Suppressor Potential. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1825**, 101-110.
135. Muranen, T., Selfors, Laura M., Worster, Devin T., Iwanicki, Marcin P., Song, L., Morales, Fabiana C., Gao, S., Mills, Gordon B. & Brugge, Joan S. (2012). Inhibition of PI3K/mTOR Leads to Adaptive Resistance in Matrix-Attached Cancer Cells. *Cancer Cell* **21**, 227-239.
136. Kim, T.-S., Kim, H. D. & Kim, J. (2009). PKCδ-dependent functional switch of rpS3 between translation and DNA repair. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1793**, 395-405.
137. Kim, T.-S., Kim, H. D., Shin, H.-S. & Kim, J. (2009). Phosphorylation Status of Nuclear Ribosomal Protein S3 Is Reciprocally Regulated by Protein Kinase Cδ and Protein Phosphatase 2A. *Journal of Biological Chemistry* **284**, 21201-21208.
138. Kim, H. D., Lee, J. Y. & Kim, J. (2005). Erk Phosphorylates Threonine 42 Residue of Ribosomal Protein S3. *Biochemical and Biophysical Research Communications* **333**, 110-115.
139. Gao, X. & Hardwidge, P. R. (2011). Ribosomal Protein S3: A Multifunctional Target of Attaching/Effacing Bacterial Pathogens. *Fronteirs in Microbiology* **2011**, 1-6.
140. Schäfer, T., Maco, B., Petfalski, E., Tollervey, D., Böttcher, B., Aebi, U. & Hurt, E. (2006). Hrr25-Dependent Phosphorylation State Regulates Organization of the Pre-40S Subunit. *Nature* **441**, 651-655.

141. Yoon, I.-S., Chung, J. H., Hahm, S.-H., Park, M. J., Lee, Y. R., Ko, S. I., Kang, L.-W., Kim, T.-S., Kim, J. & Han, Y. S. (2011). Ribosomal Protein S3 is Phosphorylated by Cdk1/cdc2 During G2/M Phase. *BMB Reports* **44**, 529-534.
142. Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A. & Ban, N. (2011). Crystal Structure of the Eukaryotic 40S Ribosomal Subunit in Complex with Initiation Factor 1. *Science* **331**, 730-736.
143. Sengupta, J., Nilsson, J., Gursky, R., Spahn, C. M. T., Nissen, P. & Frank, J. (2004). Identification of the Versatile Scaffold Protein RACK1 on the Eukaryotic Ribosome by Cryo-EM. *Nat Struct Mol Biol* **11**, 957-962.
144. Schrodinger, LLC. (2010). The PyMOL Molecular Graphics System, Version 1.3r1.
145. Pang, C., Gasteiger, E. & Wilkins, M. (2010). Identification of Arginine- and Lysine-Methylation in the Proteome of Saccharomyces cerevisiae and Its Functional Implications. *BMC Genomics* **11**, 92.
146. Ben-Shem, A., Jenner, L., Yusupova, G. & Yusupov, M. (2010). Crystal Structure of the Eukaryotic Ribosome. *Science* **330**, 1203-1209.
147. Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. & Yusupov, M. (2011). The Structure of the Eukaryotic Ribosome at 3.0 Å Resolution. *Science* **334**, 1524-1529.
148. Pech, M., Spreter, T., Beckmann, R. & Beatrix, B. (2010). Dual Binding Mode of the Nascent Polypeptide-associated Complex Reveals a Novel Universal Adapter Site on the Ribosome. *Journal of Biological Chemistry* **285**, 19679-19687.
149. Alamo, M. d., Hogan, D. J., Pechmann, S., Albanese, V., Brown, P. O. & Frydman, J. (2011). Defining the Specificity of Cotranslationally Acting Chaperones by Systematic Analysis of mRNAs Associated with Ribosome-Nascent Chain Complexes. *PLoS Biol* **9**, e1001100.
150. Arnesen, T. (2011). Towards a Functional Understanding of Protein N-Terminal Acetylation. *PLoS Biol* **9**, e1001074.
151. Holcik, M. & Sonenberg, N. (2005). Translational Control in Stress and Apoptosis. *Nat Rev Mol Cell Biol* **6**, 318-327.