Quantitative cross-linking mass spectrometry for protein structural stability studies

Martin Mathay

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

James E. Bruce, Chair

Shao-En Ong

Devin Schweppe

Program Authorized to Offer Degree:
Genome Sciences

University of Washington

Abstract

Quantitative cross-linking mass spectrometry for

protein structural stability studies

Martin Mathay

Chair of the Supervisory Committee:
James E. Bruce
Genome Sciences

Mass spectrometry-based methods for protein-ligand identification have expanded classical techniques for the bioanalytical characterization of small molecule target engagement and their modes of action. In the last decade, a series of techniques have coupled mass spectrometry readout, structure-function framework, and thermodynamic stability to expand the suite of proteomics techniques for protein-ligand interactions. Although these methods have proven powerful, due to the complex nature of these large-scale studies, having multiple avenues of assessment is critical for the proper evaluation of clinical value. In this work, the interfacing of these protein-denaturation experimental designs with cross-linking mass spectrometry sample workflows is investigated to better understand the protein topologies in these protein-ligand large-scale analyses. The developed method, protein-denaturation and quantitative cross-linking mass spectrometry, offers another strategy in the unbiased assessment of protein target engagement studies. Additionally, from a basic science perspective, this method also provides data in understanding the molecular principles of protein folding in structure-(dys)function studies. First, I validated a proof-of-concept of protein-denaturation with quantitative cross-linking mass spectrometry in a standard protein and known ligand. Then, I adapted and assessed the viability of this method on the proteome-level scale. Although this method has much room for optimization for tackling large-scale studies, its data provides promise with smaller complex proteomes. Overall, quantitative cross-linking mass spectrometry during protein unfolding is a reliable assay that can be used alone or provide complementary information to the current generation of protein-denaturation mass-spectrometry methods for generating target-engagement atlases.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACN	acetonitrile
BDP-NHP	biotin aspartate proline N-hydroxyphthalimide
DMSO	dimethyl sulfoxide
DTT	dithiothreitol
FA	formic acid
FPLC	fast protein liquid chromatography
HPLC	high-pressure liquid chromatography
HSP	heat shock protein
IAA	iodoacetamide
MS	mass spectrometry
m/z	mass-to-charge ratio
PPI	protein–protein interactions
PAO	*Pseudomonas aeruginosa*
PP	Pulse Proteolysis
SPROX	Stability of Proteins from Rates of Oxidation
TCA	trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
TPP	Thermal Proteome Profiling
XL-MS	cross-linking mass spectrometry

ACKNOWLEDGEMENTS

There are many people I am deeply grateful to for helping make this dissertation possible, me the scientist, and me the person I am today. Graduate school continually stressing the "re" in research was both challenging and humbling; and there is no doubt that my communities supported me through this time. I sincerely thank you all.

I am fortunate enough to have had support prior to coming to graduate school. My parents always emphasized the importance of education. They did well in imbuing that message and going against their own mental map in braving my own ventures in electing an education in a newly founded high school and liberal arts college. There I learned how to perform bench work, read, and generate metaphors from the guidance of Dr. Julie Nowicki, Dr. Scott Gilbert, Dr. Gwynn Kessler, Dr. Alison Holiday, and Dr. Amy Vollmer. The decision for one more academic venture of graduate school on the west coast was the only logical next step.

I am grateful to the department and specifically my thesis committee Dr. Colin Manoil, Dr. Shao-En Ong, Dr. Devin Schweppe, and Dr. Pradeep Singh for their patience, engagement, and guidance throughout my PhD. I am especially grateful to Colin sharing expertise, providing opportunities to share my ideas with an excited audience, and having his lab arrange for the multiple bacterial strains I have requested throughout the years. Additionally, I would like to thank the magical administrative team of the Department, particularly Brian Giebel, Casey Hudlow, Maureen Larsen, Serena Newhall, for making this process a very smooth one.

Certainly, I would like to thank James E Bruce for his role in chaperoning my scientific state during graduate school. Jim's mentorship has taken many forms over the last few years. The freedom he has allowed me to make the good and bad mistakes, many failures, and occasional

accomplishments. Many individuals from the lab have enhanced the completion of my training. Specifically, Dr. Juan Chavez and Dr. Xuefei Zhong for their patience and guidance especially in my first few years with the lab. Dr. Michael Cammarata provided much collaborative conversations during the last two years. Fellow graduate students Anna Bakhtina, Arianne Caudal, and Jared Mohr supplied a great deal of time, knowledge, and friendship at a shared stage of life. More broadly, all members of the Bruce Lab, past and present, have influenced the way I approach problems and contributed to the success I have found. I have enjoyed watching the Bruce Lab find successes throughout my time in the lab and I am excited to see what they discover and develop next.

Over time span of a dissertation, there is movement outside of the lab as much as inside. Tyler Fragie, Katie Lobanova, Lindsay Taylor have made my interactions within the University of Washington entertaining in the midst of graduate work. Christopher Chalaka, Mandy Gao, Sarah Gonzales, Eva Li, and Raven Weber have expanded that enjoyment in and out Seattle. These out of lab connections, together with previous ones with Ariel Gewirtz, Paul Bierman, Molly Feldman, Rishi Jaggernauth, and Tina Zhu proved to be particularly important during the unprecedented time during the latter half of my graduate career. And, my wonderful orange cat, Arlo, and black cat, June, proved to be the best 2020 decisions any graduate student could ask for.

Chapter 1. INTRODUCTION

1.1 PROTEIN FOLDING, STRUCTURE, AND STABILITY

Biological function requires the performance and regulation of various biochemical reactions, which are performed by proteins. Proteins are essential biomolecules that perform a variety of functions in the living organism, such as providing cell structure, catalyzing metabolic reactions, and transporting signaling molecules. Proteins are synthesized as linear chains of amino acids and need to fold into precise three-dimensional architecture to perform their specific biological function.[1] This process is known as protein folding and must be achieved in a crowded intracellular milieu.[2]

The simplest model for protein folding conceives only two states, the unfolded and folded state. However, the solution phase structures of proteins are not static. Proteins repeatedly unfold and refold during their lifetime and cellular health depends on this success.[3] The conformational changes that proteins undergo can be as dramatic as a loss of all higher order structure on a global scale or as subtle as a specific element of secondary structure in a more localized unfolding event (**Figure 1.1**). The precarious folding landscapes pose a fundamental problem for phenotypic stability. The dysregulation of protein states favors the transition from a healthy to disease state, consequently driving the cellular system to threaten the whole organism.[4] Thus, knowledge about stabilizing structures and interactions involved in the conformational changes that proteins undergo is crucial for understanding fundamental biological processes, diseased states, and protein-targeting therapeutic modes of actions.

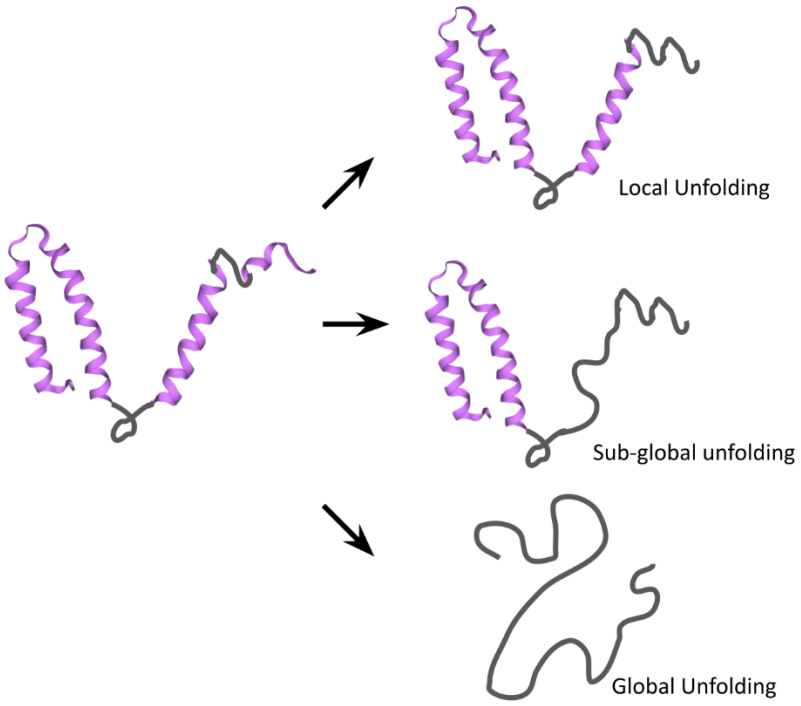


Figure 1.1. Schematic representation of the range of conformational changes that proteins can undergo in solution. [5]

Historically, the classical biochemistry and analytical chemistry techniques employed for and characterizing the intricacies of protein structure, protein folding, and protein-ligand engagement has been slow and tedious. These techniques, such as X-ray crystallography, fluorescence resonance energy transfer, and isothermal titration calorimetry, require copious amounts of starting material, extensive sample processing, and *a priori* knowledge of interacting partners. Moreover, these are often not suitable for high-throughput studies. Nevertheless, these experiments have led to an astonishing body of important discoveries and advances in our understanding of protein dysregulation and the cellular diseased state.

With the recent emergence of high throughput "omics" technologies, our ability to detect, identify, and characterize proteins from their primary sequence, three-dimensional structure, and variety of conformational changes and chemical modifications have been radically changed.[6] The proteomics field, the large-scale study of proteins produced or modified by an organism or living

system, has been largely driven by technological and methodological advances.[7] For example, the development of matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), two soft ionization techniques, enabled the generation of intact gas phase ions of large biological molecules allowing for the direct mass spectrometric (MS) detection of proteins. The high sensitivity of the mass spectrometer has enabled detection and accurate mass analysis at low abundance not possible with classical techniques.

Traditionally, quantitative proteomic studies of biological states have focused on determination of protein abundance level changes to gain new functional insight.[8] The presence or absence of certain proteins by differential comparison to the entire or subset of a source's ensemble generates direct information on cellular function. However, this approach is hindered as the link between a protein's altered abundance level and altered function is not always clear. A protein's functional characterization in the cellular milieu requires an understanding of its dynamic structural changes, processing modifications, and microenvironment.

Drug discovery approaches largely fall under target-based drug discovery or phenotypic drug discovery. Target-based discovery, also referred to as reversed chemical genetics, starts with a target relevant to a disease. Then, once the target has been validated, assays are developed for a high-throughput screening of chemical libraries to identify small molecules that bind with the target. In contrast, phenotypic drug development has small molecules tested directly for their impact on a diseased phenotype. Because pathways and targets are not considered prior, this method provides the opportunity of discovering new therapeutic targets. Candidate molecules are then assayed for target identification, mode-of-action elucidation, and target safety risks.[9]

Numerous MS-based methods have been developed to assess and validate ligand-target interactions.[10-12] For example, affinity purification (AP-MS) remains the most well-known

strategy for target identification. Here, proteins of interest are allowed to reversibly bind to the immobilized drug on the stationary phase. After a wash step, eluted binding proteins are identified and quantified by MS. Although straightforward, this method requires the immobilization of a drug to a solid matrix or modification of drug with an enrichment handle, either which could potentially change the biological or binding properties. Furthermore, AP-MS suffers from high false discovery for the ease of purifying the large amount of abundant proteins with weak binding interactions.[13]

Recently, a new generation of MS-based techniques have been developed to address limitations of AP-MS. Many of these approaches have expanded previous designs while utilizing the fundamental link between protein folding, biological function, and thermodynamic stability to exploit the principle that ligand binding shifts the protein folding equilibrium.[14,15] Conventional techniques of protein thermodynamics have induced a stress to generate an unfolded protein population to compare to the folded state.[16] The use of denaturant, lowering pH, and increasing temperature have facilitated the unfolding process and enable comparison of the folded/unfolded equilibrium between two biological states. To screen for ligands that bind to target proteins *in vitro*, the shift in the protein's thermodynamic stability can be monitored in the absence or presence of compounds.[17] For example, Cellular Thermal Shift Assay (CETSA) screens for ligands that bind to a target protein *in vitro* by the shift in a protein's thermal stability in the absence or presence of compounds.[18] In the CETSA protocol, cells or lysates are treated with a drug or vehicle and are then heated at a range of different temperatures. Following heating, the remaining soluble protein is detected by, for example, a Western Blot. Utilizing the loss of detected, quantified protein across the thermal denaturation curves can generate a "melting" temperature (T_m) at which the native and denatured proteins are the same. Ligand-bound proteins, however, are stabilized by their

interacting partner and will therefore melt at a higher temperature when exposed to the same heat gradient, resulting in a T_m shift. However, CETSA is limited in requiring previous knowledge and antibody availability in its required Western Blot detection.

These MS-based protein stability profiling methods have the potential to identify more functionally relevant protein targets than AP-MS and protein expression level-based methods. While each technique has unique advantages and disadvantages, the combined effort of multiple strategies provides a more complete understanding of the drug-disease state. Therefore, the design and validation of new techniques will expand the protein-denaturation MS-based toolkit of ligand-target experiments. To consider new strategies, this introduction will cover a few of these experimental techniques.

1.2 MS-BASED PROTEOMIC APPROACHES IN STRUCTURAL STABILITY

Several mass spectrometry-based (MS) proteomic approaches have been developed to characterize ligand-induced changes in the structural changes of proteins and do not require any chemical modification of the ligand. For example, limited proteolysis and Drug Affinity Responsive Target Stability (DARTS) assay protein structure based on protease susceptibility. More recently, MS-based approaches have incorporated protein-denaturant sample workflows to induce thermodynamic stability for ligand-induced changes. These approaches have included Stability of Proteins from Rates of Oxidation (SPROX), Pulse Proteolysis (PP), and Thermal Proteome Profiling (TPP).[19-21] These approaches involve application of increasing denaturant followed by a chemical modification, enzymatic digestion, or protein precipitation. The presence of a ligand (or comparison to mutant phenotype) shifts the modification, enzymatic, or precipitation profile at these sites as a function of denaturant. The general workflows of these methods are summarized in **Figure 1.2**.

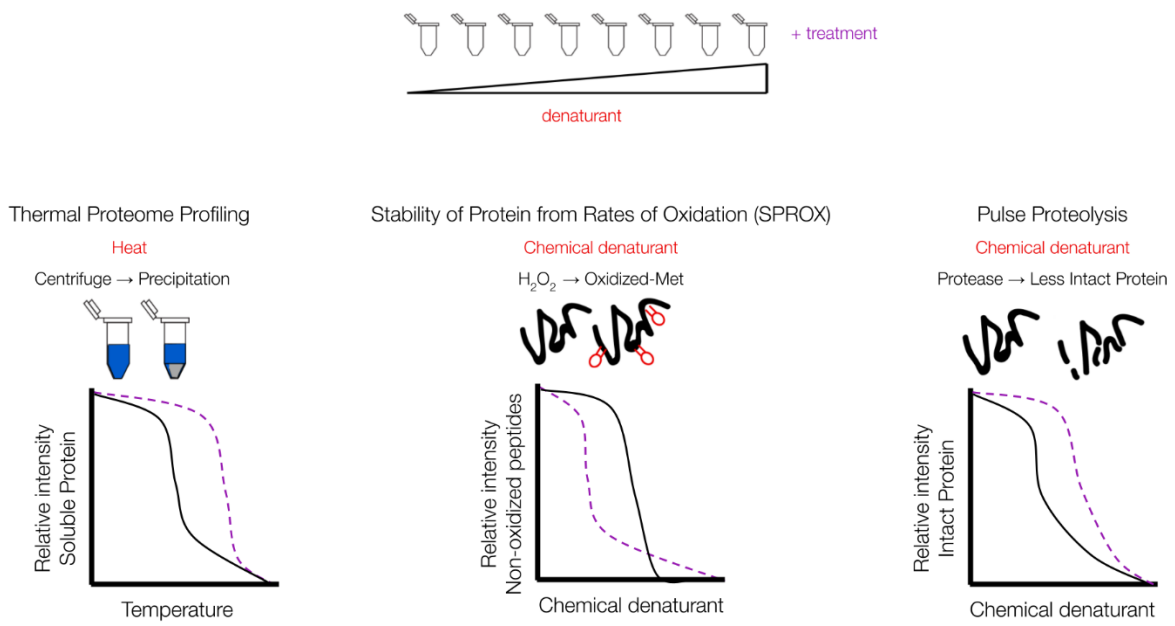


Figure 1.2. Schematic representation of current protein-denaturation MS-based techniques for protein-ligand structural stability measurement. [22]

Thermal proteome profiling is the current "gold standard" of protein structural stability profiling. In the TPP method, temperature dependent protein aggregation is used to report on the thermal denaturation properties of a protein. By adapting CETSA with an MS-based output, TPP generates an unbiased method for drug-target engagement with use of isobaric quantitative mass tags. TPP has been applied to identify protein-drug interactions, protein-protein interactions, and subcellular localization changes.[20,23] Although a powerful technique, TPP relies on thermal stabilization of proteins upon interaction and some proteins are not susceptible to thermal (de)stabilization and, thus are not characterized by this approach.[24] Moreover, the soluble protein-level readout of TPP requires global unfolding and cannot provide information on ligand binding at the residue or regional level.

SPROX and PP both provide a peptide-level readout by utilizing chemical denaturation. The PP protocol involves incubating a protein sample with increasing concentrations of urea followed by treatment with a non-specific protease. This "pulse" of a protease generates nonspecific proteolytic digestion that is used to evaluate the protein's unfolding reaction across the urea concentrations. Originally, the relative intensities of protein bands generated by gel electrophoresis separations could be compared with and without ligand at each denaturant concentration assessed. The method has been adapted with stable isotope labeling with amino acids in cell culture (SILAC) and an MS-readout to show viability on the proteome-wide-scale.[25] In this modified protocol, the ligand is incubated in either the light or heavy lysate before the sample is aliquoted into a series of denaturant-containing buffers. Afterwards, samples subjected to LC-MS/MS analysis and light/heavy ratios are assessed at each denaturant concentration. The change in ratios suggests protein-ligand interaction.

Similar to PP, SPROX follows increased urea or guanidine concentrations with hydrogen peroxide treatment. The conditions of SPROX have been optimized under which the primary site of oxidation is the thioester group in the side chain of methionine residue. In theory, the increasing denaturation of a protein causes a methionine to become more solvent-exposed and susceptible to this oxidation reaction. In the first demonstration of SPROX, model protein systems were directly analyzed by MALDI-time-of-flight MS.[19] The feasibility of this methods to quantify protein-ligand engagement on the proteome-scale have been proved with use of isobaric iTRAQ tags in an Hsp90 inhibitor and ATP activation.[26,27] SPROX for large-scale quantitation has also been interfaced with SILAC and TMT.[28,29] An alternative strategy for proteome-wide SPROX using heavy-labeled hydrogen peroxide ($H_2{}^{18}O_2$) has also been reported.[30] Despite these successful examples, SPROX has seen limited applicability to proteins with methionine residues.

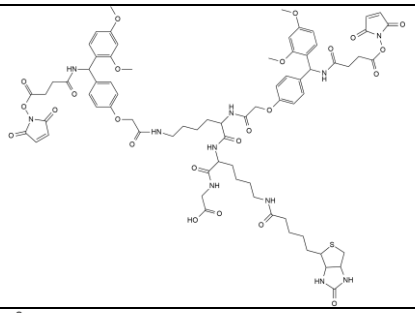
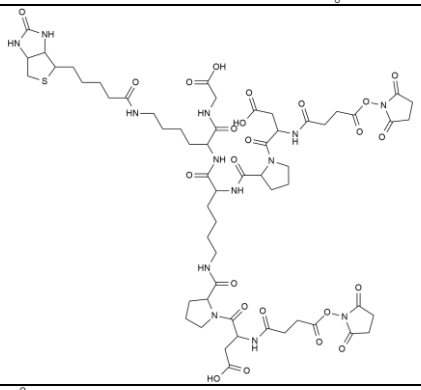
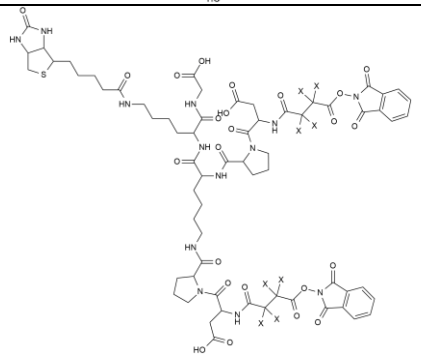
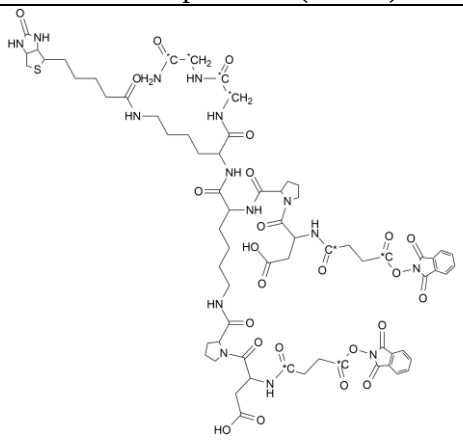
The work described in this dissertation involves the application of protein-denaturation mass-spectrometry workflow with quantitative chemical cross-linking for a cross-linked peptide-pair readout to uniquely reveal residue-level protein conformational changes during denaturant-induced unfolding. Moreover, the workflow provides similar proteome-wide profiling to ligand binding and mutant analysis. The underlying principles of XL-MS is reviewed below.

1.3 APPLICATION OF CROSS-LINKING MASS SPECTROMETRY

Cross-linking mass spectrometry is a chemical modification-based MS-technique that covalently links accessible, reactive amino acid residues that are proximal to one another. The identified cross-links serve as "molecular rulers" that impose distance restraints and serve as a basis for subsequent computational modeling of protein three-dimensional structure. The covalent linkage itself is strong enough to survive stringent sampling steps despite the label's step potential to capture weak and transient interactions. With recent advances in MS instrumentation, MS-cleavable cross-linkers, and informatics tools, XL-MS has been increasingly applied to large-scale protein structural biology in a variety of lysates, cell types, and tissues.[31-33]

The design of a cross-linker inherently affects the amino acids it can react with as well as the maximum cross-linkable distance. The general design of a cross-linker contains two or more reactive groups connected via a spacer chain. Recently, homo-reactive N-hydroxysuccinimide (NHS) esters as both reactive groups have become the most widely used. The NHS ester shows high reactivity at physiological pH with primary amines, thus linking protein N-termini or lysine residues. The N-hydroxyphthalimide (NHP) ester, an analog of the NHS ester, was reported to have a faster reaction rate to the NHS ester and has been utilized in the majority of cross-linkers in Protein Interaction Reporter (PIR) technology.[34,35]

Table 1.1. Evolution of PIR cross-linkers

PIR cross-linker	Structure	Reference
B-Rink-NHS		Zhang et al.[36]
BDP-NHS		Zheng at al.[37]
BDP-NHP	 X denotes H or D isotope for d0(X = H) or d8(X = D)	Zhong et al.[38]
iqPIR (2-plex)	 *denotes the position can be isotope-coded with a C13	Chavez et al.[39]

PIR for XL-MS has gone through various changes to achieve quantitation of identified cross-linked peptide pairs. Some iterations of PIR cross-linkers have been detailed in **Table 1.1**. The original design of PIR, B-RINK-NHS, utilized the NHS ester and was applied to identify *in vivo* protein-protein interactions in *Shewanella oneidensis*. Quantitation of PIR cross-links first utilized label-free, MS1-based methods to monitor binding of calmodulin.[40] Later, PIR was applied with stable isotope labelling by amino acids in cell culture for the comparison of cancer cell lines.[41] The use of PIR with SILAC provided *in vivo* quantitative measurements in protein topologies have been extended to compare conformational changes upon drug treatment.[42,43] However, SILAC has a notable limitation in requiring in-cell amino acid labeling, and not all biological systems are affordably compatible with such requirement. This prevents SILAC in clinical isolates as well as cells that can endogenously synthesize the amino acid label. As such, a PIR design incorporated deuterated cross-linker pairs (d_0-BDP-NHP/ d_8-BDP-NHP) for the quantitative analysis through comparison of peak areas of d_0 and d_8 cross-linked peptide pair precursors.[38] However, like other MS1-based quantification approaches for proteome-scale data-dependent experiments, challenges come from accurate quantification from low abundance cross-linked species as well as the co-elution of cross-linked and non-cross-linked peptide pairs.

The most recent generation of PIR cross-linkers incorporates stable heavy isotopes into the cross-linker design to allow for MS2 based quantitation. Taking advantage of the cleavable cross-linker design, isobaric quantitative PIR (iqPIR) allows for cross-linked peptide pairs original from different samples to have exactly the same mass in MS1 measurements. However, once fragmented in MS2 space, distinct isotope signatures can be quantified. Like other isobaric mass tags, such as ITRAQ and TMT, iqPIR sees increased signal-to-noise by not splitting a light-heavy pair into two signals, generation of multiple fragment ions carrying quantitative information, and

avoidance of chromatographic alignment that carries the burden of retention time shifts.[39] The quantitative performance of the iqPIR strategy has been demonstrated in cardiac diseased phenotypes and drug treatment.[44,45]

1.4 THE RIBOSOME AS AN INTEREST OF PROTEIN BIOLOGY

Ribosomes are complex molecular machines that synthesize proteins as translated from the genetic information on messenger RNAs, and therefore exists at critical junction of the genotype-phenotype and genes-proteins relation in all species. Given the fundamental nature of protein synthesis, ribosomes are a main target for antibiotics. Antibiotics that inhibit protein synthesis by targeting the ribosome have targeted specific functional. Structural studies of the ribosome and of its complexes with translation factors and inhibitors have provided valuable insights on the mechanism of protein synthesis and antibiotic modes of action. Meanwhile, the rise of antimicrobial resistant infections have become increasingly difficult, creating a global public health threat.[46] This has spurred renewed interest in structural studies of protein synthesis inhibitors.[47]

Fully assembled ribosomes have a large and small subunit; the bacterial 70S ribosome is comprised of the 50S and 30S subunits. In *Escherichia coli*, the 30S subunit is composed of 21 ribosomal proteins and a single 16S ribosomal RNA whereas the 50S subunit contains 33 ribosomal proteins and two rRNAs (5S, 23S).[47] Through highly regulated and choreographed steps, translation factors guide the ribosome during the protein synthesis cycle, which can be divided into four main steps: initiation, elongation, termination, and ribosome recycling (**Figure 1.3**). During the initiation phase, the small subunit binds to mRNA and decodes the three-nucleotide codon; the large subunit harbors the peptidyl transferase center and main catalytic function for the formation of the growing amino acid sequence attached to the site. The ribosome

enables the binding of mRNA, tRNAs, as well as various translation factors. The complete 70S ribosome (50S and 30S) contains three tRNA binding sites, termed the A-site, P-site, and E-site. During the elongation process, the tRNAs successively pass through the A-, P-, and then E-sites before dissociating from the ribosome.[48] Simultaneously, the ribosome undergoes massive rearrangements in which the subunits rotate relative to each other to accommodate for these translocations. During this rotation, the A and P site tRNAs move from 'classic' to 'hybrid' states: the anticodon ends stay in their original A and P sites and the acceptor ends move to the P and E sites.[48,49] GTP hydrolysis of EF-G then promotes translocation of the mRNA along the ribosome, coupled to a large intra-subunit rotation of the 30S head, after which the ribosome subunits rotate back for the next cycle.[50] Structural and biochemical studies have revealed many of the atomic-level changes that occur during this complex, multifaceted process and new details reshape models, raise questions, and provide potential new local sites for antibiotics.

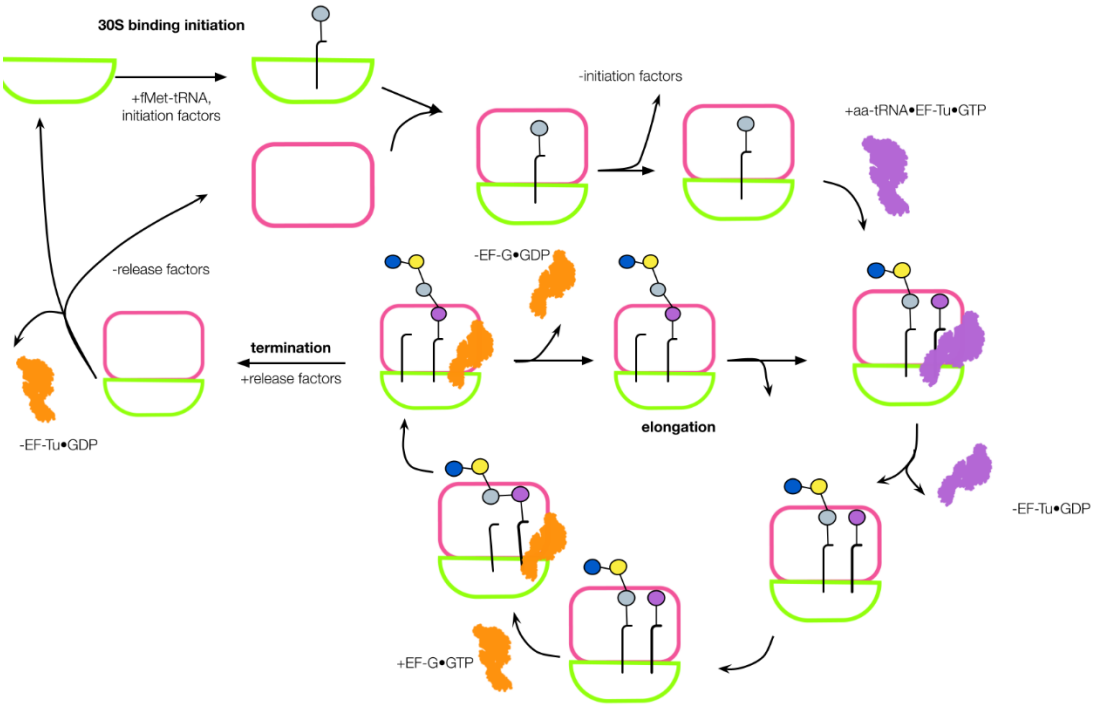


Figure 1.3 The translating bacterial ribosome interactions at a glance.

The bacterial ribosome undergoes various conformational changes and interactions for the translation of an mRNA sequence to a polypeptide chain. During 30S binding initiation (orange), initiation factors, and initiator-tRNA, and mRNA bind to the small ribosomal subunit. Then, the large ribosomal subunit joins. During elongation, elongation factor EF-Tu delivers an aminoacyl-tRNA to the ribosomal A-site. Given proper mRNA-tRNA codon-anticodon matching, the ribosome undergoes a series of conformational changes to accommodate the tRNA and facilitate peptide bond formation. Elongation factor EF-G promotes the translocation of the bound tRNAs and mRNA, while dissociating from EF-G. This cycle repeats until a stop-codon signals the ribosome to begin termination and recycling.[51]

While crystal structures obtained in the beginning of the 21st century demonstrated the importance of RNA to ribosome function, ribosomal proteins remain a critical infrastructure for ribosome structure and function. Ribosomal proteins usually contain one or several globular domains with elongated loops or N and C-terminal tails, which have considerable intramolecular mobility. While loops comprise only 18% of the structure of ribosomal proteins, they occupy 44% of the total surface of protein-RNA interactions.[48] Data analysis of bacterial ribosome structures have calculated that the loops and tails content of lysine and arginine are two-three times higher than in the globular parts of the protein.[52,53] The abundance of these charged residues is critical for charged interactions with the negatively charged RNA backbone. This abundance is also well situated as a target for lysine-reactive NHS/NHP-based cross-linker reactivity. The large number of unique inter-protein interactions have made the ribosome a frequent target of structural XL-MS in various studies.[38,54,55] The highly mobile nature of some ribosomal proteins has highlighted some of the hurdles in traditional cryo-EM and X-ray crystallography, while illustrating important structural validation applications of XL-MS. XL-MS is emerging as a complementary and beneficial method for understanding dynamic protein topologies and interactions to advance protein structural biology.

1.5 AIMS OF THIS WORK

A wealth of protein-denaturation MS-based methods have provided various avenues on the unbiased study of drug-target engagement assays. Although these assays capture a protein or protein residue's stability response to a drug or ligand, they often rely on a single-protein (TPP) or single-peptide readout (SPROX). In the case of SPROX, the reliance on methionine peptides limits the coverage of information. To better catalogue our understanding in protein-ligand protein structural stability, this dissertation focuses on the development and validation of interfacing protein-denaturation experimental design with a XL-MS workflow and analysis. The use of quantitative XL-MS should expand the coverage of peptides available to assess protein-ligand engagement with increased residue reactivity as well as multiple binary cross-linked peptide-pairs as opposed to single peptides. To accomplish this feat, denaturing cross-linking mass spectrometry for protein unfolding combines protein-denaturation with iqPIR XL-MS technologies. First, protein-denaturation and cross-linking mass spectrometry for protein unfolding is applied to a standard protein and known ligand to assess the capacity of this technique in a protein and ligand-protein complex. Then, the application of this method was evaluated in a protein lysate. We assess our understanding of the gentamicin-Pseudomonas ribosome drug-lysate system with qXL-MS and protein-denaturation and qXL-MS for protein unfolding. These results demonstrate how protein-denaturation and qXL-MS enables detection of structural stability changes in translation binding factors and nitrogen-based metabolism proteins and represents a valuable new tool in the protein-denaturation MS-based toolkit.

Chapter 2. A PROOF-OF-CONCEPT OF PROTEIN-DENATURATION AND QXL-MS IN BSA

2.1 ABSTRACT

Recently, several mass spectrometry methods have utilized protein structural stability for the quantitative study of protein-ligand engagement. These protein-denaturation approaches evaluate ligand-induced denaturation susceptibility changes with a MS-based readout. These techniques, which include thermal proteome profiling (TPP) and stability of proteins from rates of oxidation (SPROX), each have advantages and disadvantages in their application. Here, we report the combination of protein-denaturation principles with quantitative cross-linking mass spectrometry using isobaric quantitative protein interaction reporter technologies. This method enables the evaluation of ligand-induced protein engagement through analysis of cross-link relative ratios across chemical denaturation. As a proof-of-concept, we found ligand-stabilized cross-linked lysine pairs in well-studied bovine serum albumin and ligand bilirubin. These links map to known binding sites Sudlow Site I and subdomain IB. We propose that protein-denaturation and qXL-MS can be combined with similar peptide-level quantification approaches, like SPROX, to increase the coverage information profiled for facilitating protein-ligand engagement efforts.

2.2 INTRODUCTION

For many years, protein target drug discovery efforts have relied on mass spectrometry (MS)-based affinity capture techniques. While affinity purification strategies have been useful for identifying protein targets, covalent modification of the drug is required and potentially changes binding properties.[13] More recently, a series of approaches exploiting the link between protein folding, thermodynamic stability, and ligand-induced changes have been developed. These methods, such as thermal proteome profiling (TPP), stability of proteins from rates of oxidation

(SPROX), and pulse proteolysis (PP), rely on ligand-induced changes on the biophysical properties of proteins when exposed to increasing denaturation.[19-21] To date, TPP has been the most popular approach for detecting protein-ligand interactions. Building from cellular thermal shift assay (CETSA), the TPP workflow exposes the lysates from cells or ligand-treated cells to a thermal gradient.[18] Proteins will unfold, aggregate, and eventually precipitate during this heat treatment based on their inherent stability and interaction with other partners. By subjecting the soluble fraction to isobaric labeling through tandem mass tags, the MS-based output allows for an unbiased search for ligand targets by observed shifts in calculated melting curves. Despite the body of target discoveries made by TPP, its protein-level readout cannot provide information on ligand binding at the regional or domain level. The peptide-level readout of SPROX and PP therefore compliment the characterization of protein-ligand interactions. However, SPROX and PP are limited in detecting methionine-containing and semi-tryptic peptides, respectively, and expanding the readout of peptide-level bottom-up approaches would further our ability to catalogue protein-target interactions.

Chemical cross-linking mass spectrometry (XL-MS) complements current peptide-level readout methods by capturing residue-level interactions on a large-scale. The reaction of a small molecule with two reactive groups, separated by a spacer with a defined length, forms covalent bonds between residues that are in close spatial proximity. The identified cross-links serve as "molecular rulers" that impose distance restraints and serve as a basis for subsequent computational modeling of protein three-dimensional structure. XL-MS has been applied to protein complexes and living cells to probe protein topologies and protein-ligand conformational changes.[40,42] However, the use of XL-MS to quantify ligand-induced structural stability shifts, to our knowledge, has not been documented.

Serum albumin is a major protein component of blood plasma involved in the transportation of long chain fatty acids, steroids, ions, and exogenous compounds and drugs. Serum albumin also functions in regulating osmotic pressure and protection against oxidative stress. It is a highly soluble and stable protein with a molecular weight of 66.5 kilodaltons. X-ray crystallographic structures of human albumin has revealed it consists primarily of α-helices and no β-sheet with 17 pairs of disulfide bridges.[56] BSA shares structural homology with human serum albumin in its "heart shaped" characterization with three homologous yet structurally distinct domains (I, II, III) connected by flexible loops. Each domain includes ten helices organized into two subdomains (A, B) that are built from six and four α-helices.[57-59] The "heart shape" metaphor, generated from the positioning of domains I-II and II-III, allow for several binding sites.

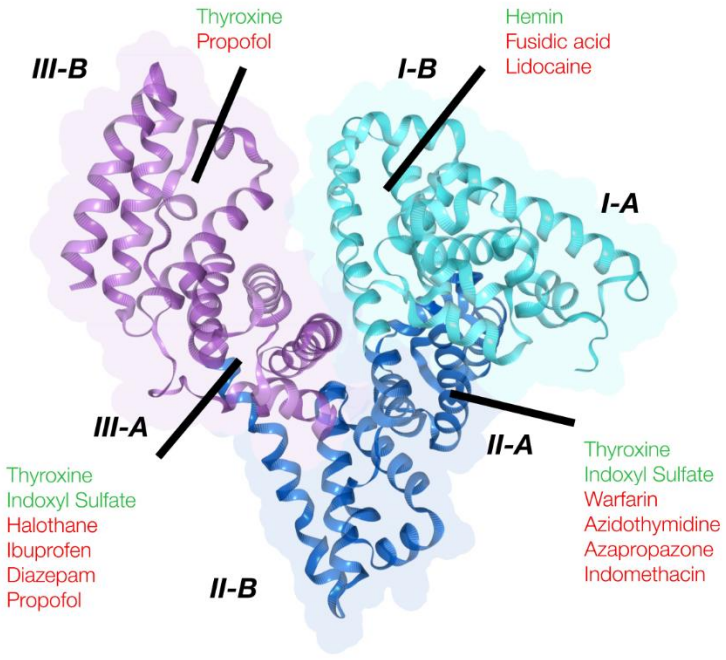


Figure 2.1 Structural annotations of the drug-binding specificity of serum albumin. The crystal structure of human albumin (PDB 4S5S) is represented to highlight the three domains (I, cyan; II, blue; III, purple). Identified binding sites of endogenous (green) and exogenous (red)

drugs for which crystal structures have been generated are annotated in each of the subdomains and connective loops. Annotations from Sand et al.[56]

Structural information on serum albumin and drug interactions has emerged in piecemeal fashion by the strong labors of classical biochemical and analytical chemistry methods (**Figure 2.1**). Many drugs with acidic or electronegative features usually bind at one of two primary sites, located in subdomains IIA and IIIA, respectively.

In this work we describe protein-denaturation with quantitative cross-linking mass spectrometry, a general strategy that combines protein-denaturation with quantitative XL-MS (qXL-MS) workflows for the quantitation of cross-link ratios generated during protein unfolding. Quantitative XL-MS with isobaric quantitative Protein Interaction Reporter (iqPIR) has previously been performed to demonstrate conformational changes upon protein-drug binding.[60] Wang et al.'s XL-MS in the presence of chaotropic agents characterized protein unfolding by MS1-based measurements.[61] Here, we expand application of MS2-based qXL-MS with protein-denaturation to detect ligand-induced structural stability shifts. When protein-denaturation and qXL-MS is tested on bovine serum (BSA) and known ligand bilirubin, we accurately identify ligand-induced stabilization on the domain-level.

2.1 MATERIALS AND METHODS

All reagents were obtained through Fisher (Waltham, MA) unless otherwise noted. Bovine serum albumin (heat shock fraction, pH 7, $\geq$98%) was purchased from Sigma-Aldrich (Saint Louis, MO) and used without further purification. Isobaric cross-linker was synthesized in house as described by Chavez et al.[39] Sequencing grade-trypsin was purchased from Promega (Madison, WI). C18 Sep-Pak cartridges for peptide desalting were purchased from Waters (Milford, MA).

BSA denaturation

The experimental workflow used in this study is shown in **Figure 2.2**. BSA (15 μM) was prepared in 170 mM disodium phosphate (Na_2HPO_4) buffer pH 8 and incubated with increasing concentrations of urea at room temperature for 10 min. For ligand-incubated samples, bilirubin stocks were prepared at 1 mM in dimethyl sulfoxide. The BSA-bilirubin solutions were mixed at a 1:1 molar ratio (30 μM). Mixtures were then incubated for 30 min at room temperature before the addition of urea and concentration adjustment in sodium phosphate buffer. The final BSA concentration in all samples were 15 μM for all experiments.

BSA intrinsic fluorescence

Fluorescence measurements were measured using a Cytation5 microplate reader (Biotek). All spectra were measured at 25 $\pm$ 0.1 °C. For measurements involving bilirubin, intrinsic fluorescence was measured by exciting the protein solution at 488 nm and the emission spectra were recorded in 530 nm. For all other measurements, excitation and emission were performed at 285 and 340 nm, respectively.[62] The fluorescence intensity data are averages of four independent measurements.

Cross-linking and sample preparation

BSA solutions were subjected to chemical cross-linking with 1 mM of iqPIR by adding either iqPIR reagent (reporter 808 or 812 m/z) from a concentrated stock in dimethyl sulfoxide. The crosslinking reaction was performed at room temperature for 45 min with constant shaking, as previously described.[39] The reaction was quenched upon addition of ammonium bicarbonate for 10 min. Cross-linked BSA solutions were then mixed 1:1 (iqPIR-808:iqPIR-812) (mol:mol) to the native, non-denaturing cross-linked condition to generate 100 μg total mixtures.

For reduction and alkylation, mixtures were incubated with 5 mM Tris-(2-carboxyethyl)phosphine (TCEP) from a 500 mM stock for 30 min followed by 10 mM iodoacetamide from a 500 mM stock for 45 min. Protein mixtures were digested overnight with a 1:200 ratio of trypsin to protein at 37°C. The resulting peptides were desalted by solid phase extraction using C18 SepPak cartridges. Dried cross-linked peptide material were reconstituted in 300 µl of 0.1% formic acid, 4% acetonitrile in water.

LC-MS analysis

Cross-linked peptide samples were analyzed by LC-MS/MS in technical triplicate using a nanoAcquity UPLC system (Waters) coupled to an QE Exactive Plus mass spectrometer (Thermo Scientific). Aliquots of 3 µl were loaded onto a 3 cm x 100 µM inner diameter fused silica trap column packed with a stationary phase consisting of 5 µM Reprosil C8 particles with 120 Å pores (Dr Maish GmbH) at a flow rate of 2 µL/min for 10 min. Peptides were then separated by reverse-phase chromatography analytical column (60 cm x 75 mM) by applying a linear gradient from 82% solvent A (water containing 0.1% formic acid) and 18% solvent B (acetonitrile containing 0.1% formic acid) to 60% solvent A and 40% solvent B over 1 hour at a flow rate of 300 nL/min. The column temperature was set at 55°C.

The QE Exactive was operating using a top five data-dependent acquisition method of ions with a charge state from +4 to +8 with a resolving power setting of 70,000 for MS1 and MS2 scans. Additional settings include a 1e6 AGC value and 100 ms maximum ion time for MS1 scans and 5e4 AGC value and 300 ms maximum ion time for the MS2 scans. Ions selected for MS2 were dynamically excluded from further selection for 30 s.

For identification, raw files were converted to mzXML format and searched for PIR mass relationships with mango.[63] The resulting MS2 files were searched using comet against a sequence

database containing forward and reverse sequences.[64] The database consisted of BSA and natural variant BSA-A214T along with a background of 4389 known false positive proteins from *Escherichia coli*.[65] The same databased was used to search the mzXML for looplink peptide pairs using comet. Resulting pepXML files were analyzed with XLinkProphet and filtered to an estimated false discovery rate of 1% at a nonredundant peptide pair level.[66]

For quantification, analysis was performed as described by Chavez et al.[39] Cross-links with a log2ratio having 95% confidence less or equal to 1 were considered quantified with high confidence. All analyses of cross-link log2ratios were calculated using R 4.0.2. K-means clustering of log2ratios was accomplished using the R-package kml.[67] Protein bar diagrams were generated in xiNet.[68]

Cross-linking identification and quantification are available on XLinkDB as dataset ureaXL_BSA_wBR_tech3_noNormalize3_Bruce.[69] Looplinks are available as dataset ureaXL_BSA_wBR_loop_tech3_noNormalize4_Bruce.

2.2 Results and Discussion

An experimental approach was developed with qXL-MS to measure protein unfolding with increasing chaotrope concentration in the presence and absence of ligands to reveal how protein ligand interactions affect protein stability. To investigate this, we developed and applied an experimental workflow to BSA, a well-studied protein system with documented chemical denaturant induced unfolding data (**Figure 2.2**).[62,71,72]

Aliquots of protein or protein-ligand stock solutions were added to a series of cross-linking buffer samples containing increasing amounts of urea. Then, samples were cross-linked using a constant amount of iqPIR reagent. After the cross-linking reaction, no protein aggregates were

observed after high-speed centrifugation. Each cross-linked sample is mixed with a 0M urea, non-denatured sample prior at a 1:1 ratio (iqPIR-808:812, mol:mol) prior to tryptic digestion. As such, experimental replicates were performed with reciprocal iqPIR labeling schemes. Using this approach, a total of 24 cross-linked protein and protein-ligand samples were generated.

2.2.1 *Isobaric PIR enables quantitative, reproducible measurements of protein denaturation*

Analysis of triplicate LC-MS/MS injections of label-swapped samples identified 531 non-redundant cross-link peptide pairs at an estimated FDR of 1%, corresponding to 368 lysine site pairs. From the total network, 31% of crosslinks mapped to inter-domain links. Over 90% were quantified in at least one condition. To validate the qXL-MS for protein unfolding strategy, we compared the quantitation of cross-links quantified in both label-swapped experimental replicates. Analysis of cross-link quantitation was observed to be highly reproducible with R^2 values of 0.47 to 0.95 (**Figure 2.3**). Representative MS2 results for cross-link quantifications are shown in **Figure 2.6**. Quantitative ion information from all experimental and technical replicates were consolidated to obtain a single log2ratio for each crosslink at a single urea concentration.

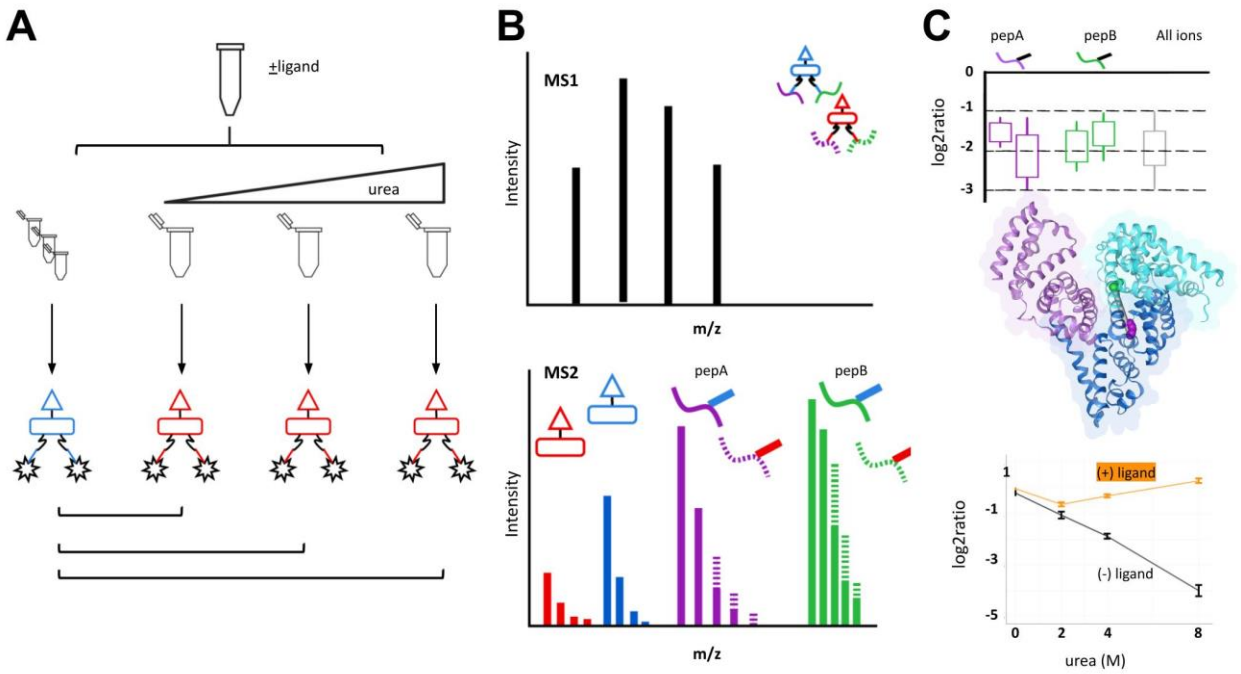


Figure 2.2 Denaturing quantitative XL-MS workflow.

A) Chemical denaturation of BSA. A BSA stock solution is divided and diluted with increasing urea. (B) Cross-linking of urea exposed protein. Solutions are crosslinked with iqPIR reagent (reporter 808 or 812). The same crosslinking design was done for BSA prior incubated with a known ligand. After crosslinking, samples are mixed with a non-denaturing control at a 1:1 ratio. Mixtures are reduced, alkylated, and digested prior to data acquisition. (C) LC-MS. Cross-linked peptide mixtures were prepared for LC-MS/MS using mango. (C) Data analysis. The data was searched using comet with crosslink validation using XLinkProphet. Crosslinking data was uploaded to XLinkDB for crosslink peptide pair quantitation viewing and structure mapping. Example peptide K211-K245 (EK211VLTSSAR-LSQK245FPK) was mapped to BSA structure (PDB: 4F5S).

To validate conformational changes induced by urea exposure, BSA solutions were examined by fluorescence spectrometry (**Table 2.1**). BSA has two tryptophan residues that possess intrinsic fluorescence and are sensitive to their local environment. Trp-134 is located on the hydrophilic surface of subdomain IB. Trp-212 is located within the hydrophobic binding

pocket of IIA, which possesses principle ligand-binding site Sudlow's Site I.[57] The intensity of their fluorescent emission band decreases markedly with increasing urea. The frequently proposed mechanism in this observance has been that unfolding induces conformational changes on the two tryptophan's local environments for fluorescent quenching (**Figure 2.3**).[62,73,74]

This distribution of cross-linked log2ratios was observed to increase with urea exposure. A negative control of non-denaturing 0M urea BSA mixed with itself did not alter the distribution of cross-ink log2ratios. The absolute fold change increases from two-fold to sixty-four-fold towards 8M urea exposure (**Figure 2.3**). Cross-link and looplink ratios compared to the negative control showed significant changes during denaturing conditions (*p-value* = 0.05) (**Figure 2.3**). To exclude the possibility that this was an artifact of increased variance, we asked whether the 95% confidence interval increased relative to denaturant. Quantitation confidence intervals were significant only at 8M urea exposure (**Figure 2.7**, mean ΔCI = +0.12, p-value = 1e-4, paired Wilcoxon test). The increased error represented 1% of the full range of log2ratios calculated at 8M urea exposure (**Figure 2.3**), supporting crosslink quantitation changes are the result of large-scale conformational changes during denaturation. These observed changes occurred when crosslinking before reduction and alkylation, with intact repeating structural disulfide bonds. Links between helices connected by disulfide bonds that were quantified by protein-denaturation qXL-MS demonstrated no log2ratio changes greater than two-fold at the majority of assayed points of protein denaturation (**Figure 2.8**).

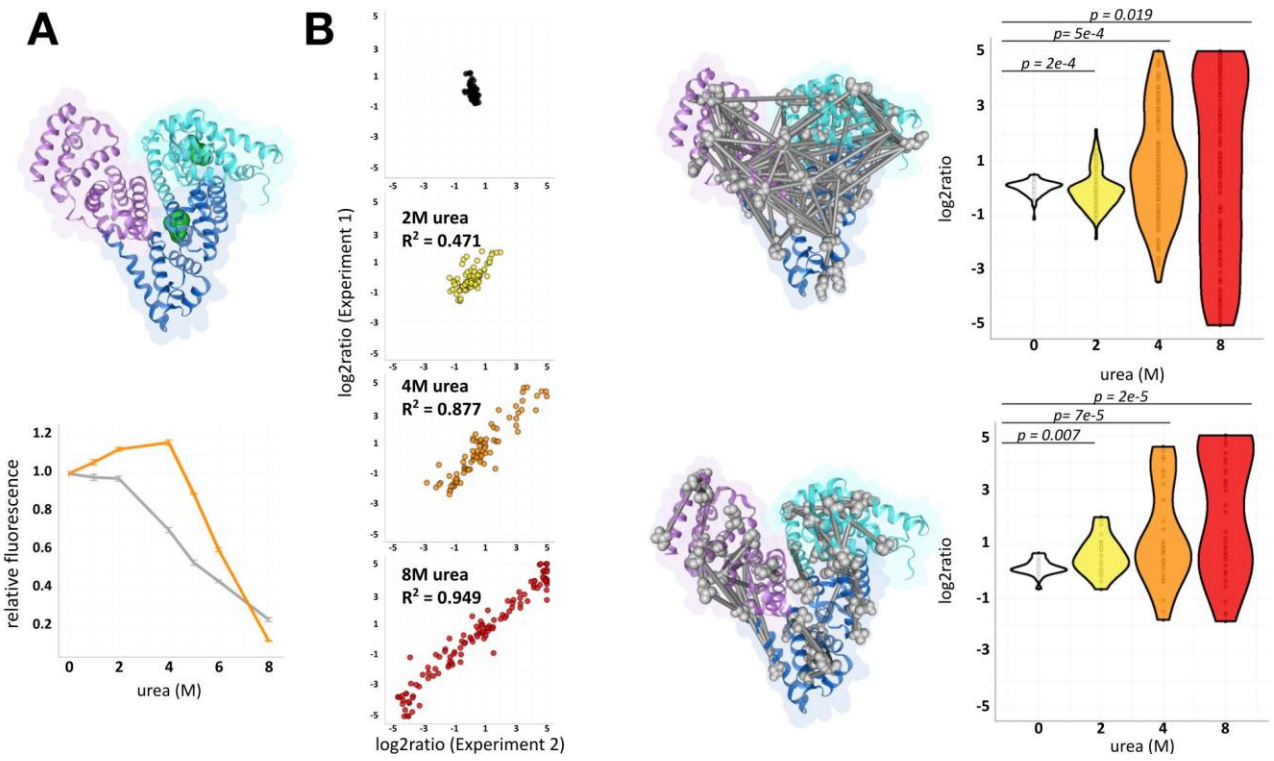


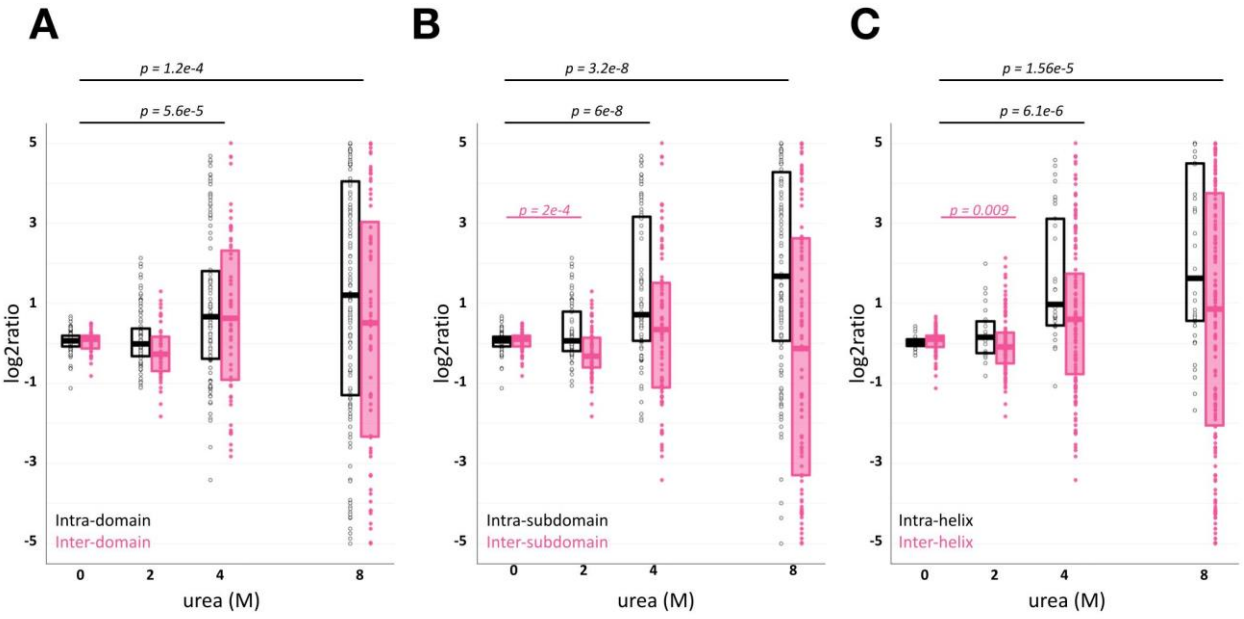
Figure 2.3 Evaluation of denaturation quantitative cross-linking mass spectrometry. (A) Structure of BSA (PDB: 4F5S). Domains 1, 2, and 3 are indicated with the following color code: cyan, blue, and purple. Intrinsic fluorescent residues Trp-134 and Trp-212 are represented as green spheres. Relative intrinsic fluorescence of BSA with and without bilirubin (BSA, black; BSA-bilirubin, orange). Upon urea exposure, loss of signal reflects changing Trp environments.[62,74] (B) Correlation of label-swap experimental qXL-MS for protein unfolding replicates. Identified mapped cross-linked peptide pairs (top), looplinks (bottom) displayed as grey bars. Distribution of quantified cross-linked ratios in BSA.

2.2.2 *Protein denaturation with qXL-MS demonstrates local domain and subdomain specific unfolding information*

The cross-link peptide relative ratios generated by qXL-MS for protein unfolding report on the biophysical properties to which the constraints map. Previous XL-MS studies on denatured BSA demonstrated loss of inter-domain cross-links and increase of near-sequence cross-links by k-means clustering of MS1-based spectral counts.[61] Similar analysis to our MS2-based qXL-MS

for protein unfolding data also shows sharp losses of inter-domain cross-links (**Figure 2.4)**. However, significant ratio differences were only observed in intra-domain increases across all quantified cross-links (**Figure 2.4**, mean Δlog2ratio = +1.09, p-value = 1e-4).

We next analyzed ratios on the subdomain and helix level in order to understand deeper structural stability features with urea exposure. We expected increasing denaturant to decrease inter-subdomain cross-link ratios. Additionally, intact disulfide bridges during chemical cross-linking should introduce constraints for intra-subdomain cross-links. We found significant increases of intra-domain links (**Figure 2.4**, mean Δlog2ratio = +1.788, p-value = 3e-8) as a consequence of urea exposure. However, inter-domain cross-link log2ratios decreases were



interpreted as not statistically significant (**Figure 2.4**, mean ΔCI = -0.0698). Similar results were seen for links mapped on the helix-level (**Figure 2.4**). Taken together, qXL-MS for protein

Figure 2.4 Characterizing BSA denaturation by qXL-MS for protein unfolding

Characterizing BSA denaturation by qXL-MS for protein unfolding. Each dot represents an individual crosslink or looplink relative ratio at a given urea exposure. Relative ratios are separated on the intra- (black) and inter- (pink) structure mapping on the domain (A), subdomain (B), or helix (C) level.

unfolding data captures the BSA denaturation landscape that subsists of increased intra-structure contacts and decrease of their inter-structure counterparts.

2.2.3 *Protein denaturation with qXL-MS allows identification of bilirubin-stabilized cross-links in the known ligand binding subdomains*

Lastly, the protein-denaturation and qXL-MS was used to investigate BSA conformational properties induced by the binding of known ligand bilirubin. Serum albumin binding to bilirubin, a catalytic degradation product of hemoglobin, has been extensively studied.[62,73,75-77] Previous studies on the binding of BSA to bilirubin have suggested a 1:1 ligand:protein ratio (mol:mol) generally occupies only the primary binding site.[73,75,77] Reciprocal-labeled experimental replicates of qXL-MS BSA-bilirubin solutions showed similar quantitative reproducibility to BSA (**Figure 2.5**). Intrinsic fluorescent readings after 15 min of bilirubin incubation showed a marked increase at 4-5M urea exposure as previously observed (**Figure 2.3**).[62,72] This observed spectroscopic shift has suggested structural rearrangements by Trp-134 and Trp-212 without integrity losses.[62]

To identify a set of cross-links that exhibited ratio differences given ligand incubation, we focused on 4M urea exposure data given the fluorescent signal differences. **Figure 2.5** shows the calculated 2-way ANOVA for each cross-link given bilirubin incubation. Applying a 2-fold cutoff filter and Bonferonni corrected p-value < 0.05, we measured significant stability changes for eleven cross-links. Lysine residues of IB and IIA represented the entirety of cross-links whose ratios increased in the presence of bilirubin (**Figure 2.5**). The observed ligand sensitivity at IIA are consistent with known binding site Sudlow Site I. While fluorescent and circular dichroism studies have placed the primary bilirubin binding site at IIA, X-ray crystallography of human serum albumin with bilirubin places the high affinity site at IB. While the primary binding site has not yet been fully clarified, both sites are supported by the qXL-MS for protein unfolding data.

Nonetheless, observed ligand induced structural changes can occur beyond the specific binding site.

We next analyzed whether bilirubin incubation affected log2ratios for links involving known ligand binding site Sudlow Site II (IIIA) in addition to Sudlow Site I and subdomain IB. We quantified changes in intra-subdomain links mapped to these regions that reached statistical significance (**Figure 2.5**, paired Wilcoxon signed rank test). On average, at 4M urea exposure, prior bilirubin incubation had stabilizing effects on IB and IIA cross-link relative abundance (mean Δlog2ratio = +0.42 and +0.45, p-value = 0.034 and 0.027, respectively). Additionally, we observed a destabilizing effect in IIIA (mean Δlog2ratio = -0.87, p-value = 1.5e-05). The direction of these bilirubin incubation effects persisted at 8M urea exposure (**Figure 2.5**). Taken together, these results show how patterns of MS/MS-quantified cross-links reveal regions relevant to structural stability changes upon protein-complex ligand binding.

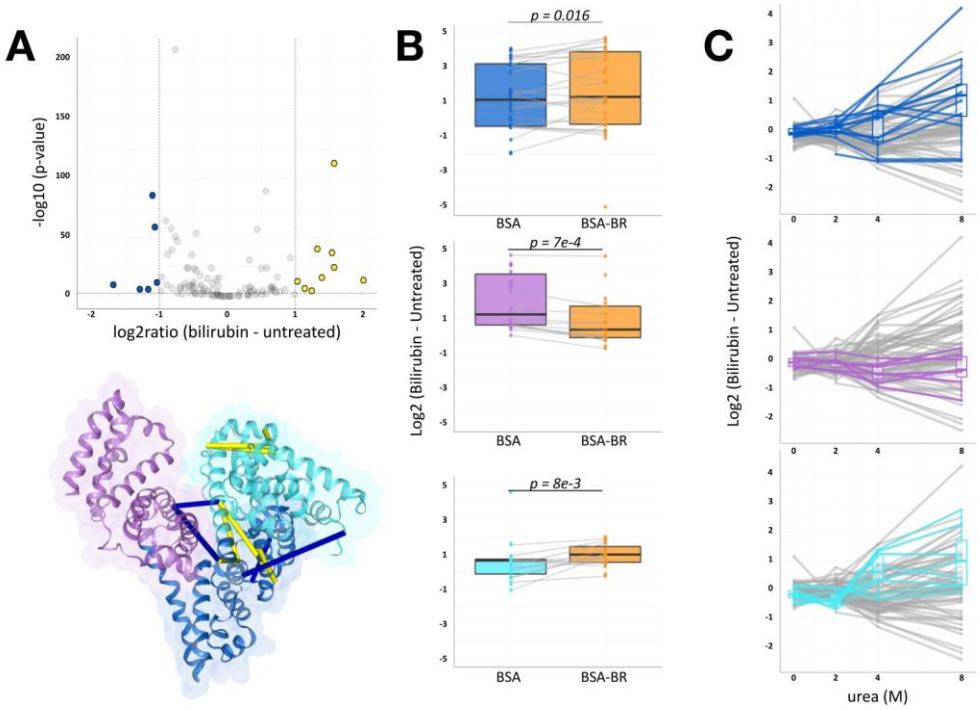


Figure 2.5 BSA-bilirubin experiments of qXL-MS for protein unfolding highlight IB, IIA denaturation stabilization

(A) Volcano plot of 111 quantified BSA cross-link ratios at 4M urea with and without bilirubin (log2 [bilirubin - apo]) versus statistical significance. Significance threshold set to 1e-2 and converted to -log10(p-value). The vertical and horizontal dotted lines show cut-off for fold-change and p-value, respectively. Blue and yellow points indicate cross-linked peptide ratios with decreasing and increasing quantitation, respectively. These cross-linked lysine pairs are mapped onto the BSA structure (PDB: 4F5S). (B) Boxplot showing log2ratio changes of subdomain intra-links at 4M urea with and without bilirubin (orange) pretreatment. Median log2ratios across quantified links are shown. The total number of quantified paired sites per group was: IB n=14, IIA, n=24, IIIA n=12 and p-values for group comparisons were computed using paired Wilcoxon signed rank test. (C) Log2 difference denaturation curve obtained by qXL-MS for protein unfolding. Each quantified cross-link represents 1 line. Highlighted crosslinks map to intra-links within subdomain IB (cyan), IIA (blue), or IIIA (purple).

2.3 Concluding Remarks

A new approach, protein-denaturation and qXL-MS, combines protein-denaturation with quantitative XL-MS sample workflows for the quantitative assessment of protein-ligand engagement studies. First, we demonstrate protein-denaturation and qXL-MS enables reproducible quantification of cross-link changes that occur during urea induced protein unfolding. Analyses that explore BSA denaturation patterns by quantified cross-link peptide pair data correlate with previously published structural data and the loss of inter-domain cross-links.[61] Second, the application of protein-denaturation and qXL-MS to the BSA-bilirubin complex identifies ligand-stabilization of Sudlow Site IB and IIA during chemical denaturation. Although these characterizations have been identified in previous studies, our findings present, to our knowledge, the first use of MS2-based cross-linked peptide-pair quantitative readout for denaturant-based, protein-ligand conformational assessments. The strategy presented has the potential to be used with other protein-denaturation MS-based approaches in studying drug-target interactions on the cellular scale.

2.4 Acknowledgements

The authors thank all members of the Bruce Lab for helpful insights and suggestions in the course of this work. This project was supported by the following grants from the National Institutes of Health 5R01GM086688 and 1R35GM136255.

2.5 SUPPLEMENTAL MATERIALS

Table 2.1 Intrinsic fluorescence raw values

		BSA				BSA + BR			
wavelength (nm)	urea (M)	read1	read2	read3	read4	read1	read2	read3	read4
285,340	blank	905542	929100	941877	954277	964577	974777	978515	982012
	0M	1872350	1913862	1917623	1943602	935371	953342	959569	957943
	1M	1848377	1901654	1911590	1912605	960342	970801	957962	968613
	2M	1853717	1887760	1900256	1909050	980559	968915	986866	986548
	4M	1623144	1660588	1636913	1609713	1045383	1046835	1035523	1036251
	5M	1443980	1469245	1486383	1484604	1086800	1088126	1076233	1081879
	6M	1367241	1382496	1384268	1383285	1125213	1124910	1113191	1103956
	8M	1171097	1189576	1194937	1198420	1056541	1064910	1057553	1049156
488,530	blank	2414	2480	2606	2547	2578	2661	2674	2628
	0M	1826	1895	1903	1877	29072	29807	30167	30195
	1M	1838	1953	1966	1934	30792	31623	31678	31636
	2M	1879	2013	1965	1941	32797	33626	33395	33256
	4M	2032	2098	2064	2032	34756	34375	34129	33618
	5M	1992	2050	2068	2093	26903	27110	26886	27019
	6M	2034	2078	2129	2132	19219	19153	18954	18819
	8M	1950	1986	2077	2029	6123	6106	6098	6104

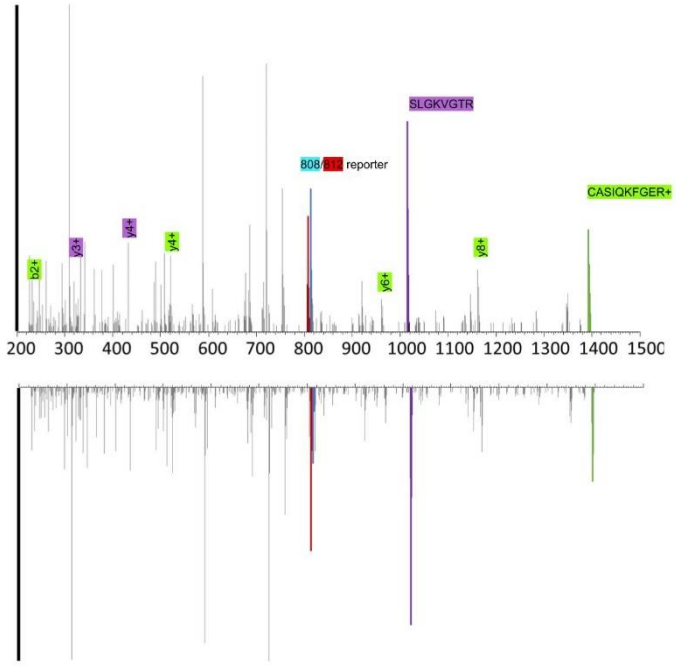


Figure 2.6 BSA crosslink spectra between reciprocally labeled experimental replicates. (A) Spectra of crosslink K228-K455 (CASIQK^{228}FGER- SLGK^{455}VGTR) in reciprocally labeled experiment 1 (top) and experiment 2 (bottom) displayed as mirror images. Peaks for reporter ions (iqPIR-808, red; iqPIR-812, blue), peptide CASIQK^{228}FGER (green), and peptide SLGK455VGTR (purple), are highlighted

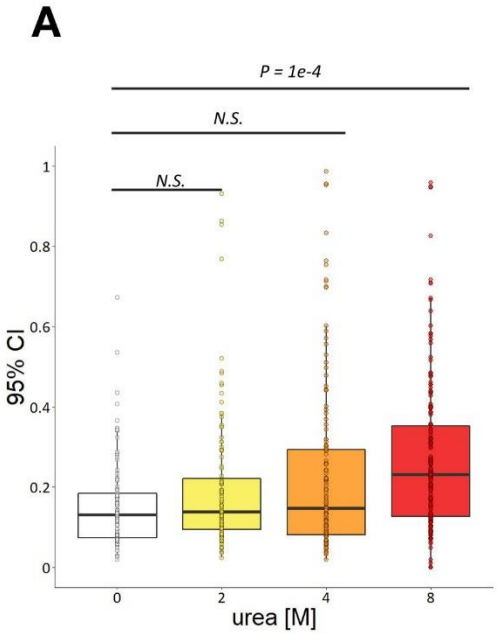


Figure 2.7 Quantitation confidence does not decrease with urea exposure.

The distribution of associated 95% confidence intervals for quantified cross-links (Fig 2D) were plotted with urea exposure. Errors were compared by paired Student's t-test with Bonferroni multiple test correction (N.S. = not significant, ***P<0.01). The median confidence interval between 0M and 8M urea exposure was 0.12 (B)

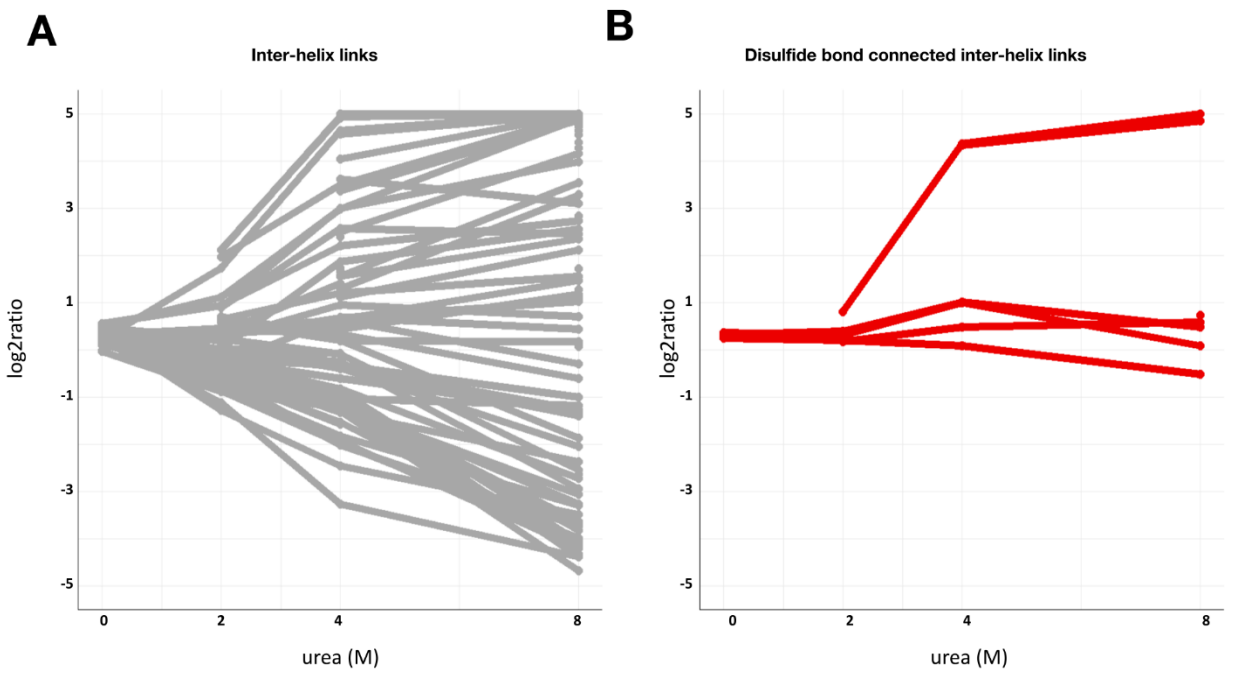


Figure 2.8 . qXL-MS data obtained on cross-link peptide pairs of lysines on disulfide linked helices.

(A) Urea concentration-dependent trends in cross-linked peptide pairs are identified by statistical filtering (3 of 4 concentrations quantified within a 95% confidence interval less than 1). (B) The average log2 values of cross-linked species mapped to helices connected by disulfide bonds are highlighted in red.

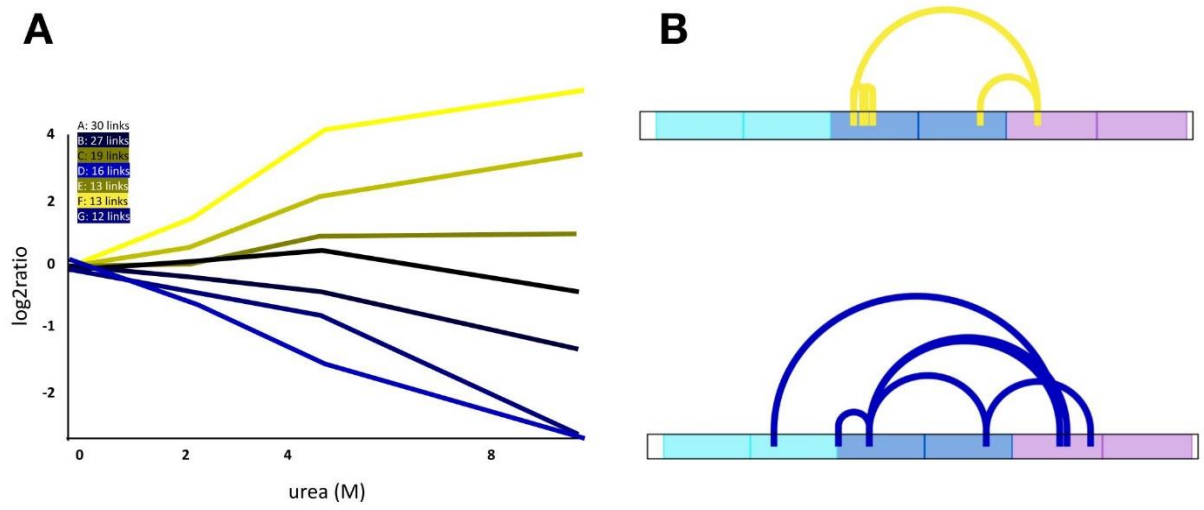


Figure 2.9 . K-means analysis of BSA urea structural stability by MS2-based quantitative cross-linking.

Urea concentration-dependent trends in cross-linked peptide pairs are identified by statistical filtering (3 of 4 concentrations quantified within a 95% confidence interval less than 1) and longitudinal k-means clustering. The average log2 values and number of cross-linked species present in each cluster are given. (B) Domain level cross-link maps of clusters that showed the largest chaneges, generated with XiNet, illustrate inter- and intra- domain differences between the clusters.

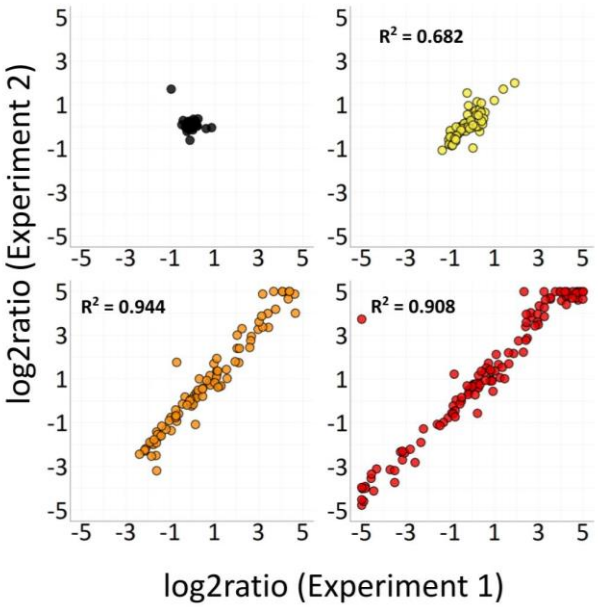


Figure 2.10. Reproducibility of reciprocal labeling in BSA-bilirubin experiments. Correlation of crosslink quantitation between two experimental replicates are shown at 0M (black), 2M (yellow), 4M (orange), and 8M (red) urea exposure. Pearson's R2 are given for each condition, indicating the reproducible observations of urea-induced unfolding and qXL-MS measurements.

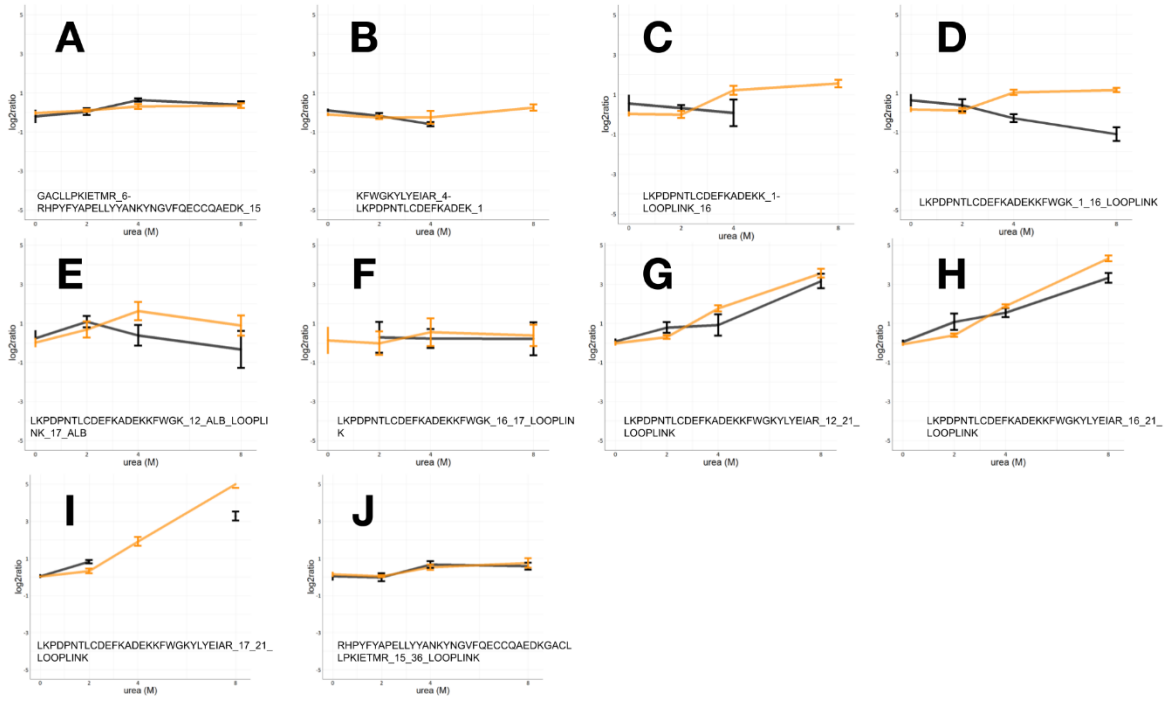


Figure 2.11. qXL-MS data obtained on cross-link peptide pairs of lysine residues in domain IB.

(A-I) Shown are the relative log2ratios obtained for each peptide pairs at increasing urea concentrations for BSA (black) or BSA-bilirubin (orange). Error bars represent 95% quantitation confidence given all relevant calculated MS2-based ion pairs across experimental duplicate and technical triplicate. While many cross-linked peptide level changes are insensitive or unaltered by bilirubin interaction, cross-linked peptide K140-K155 (D) how increased urea-induced level changes in the presence of bilirubin. Thus, bilirubin binding with BSA appears to affect protein stability within the IB domain as indicated by qXL-MS.

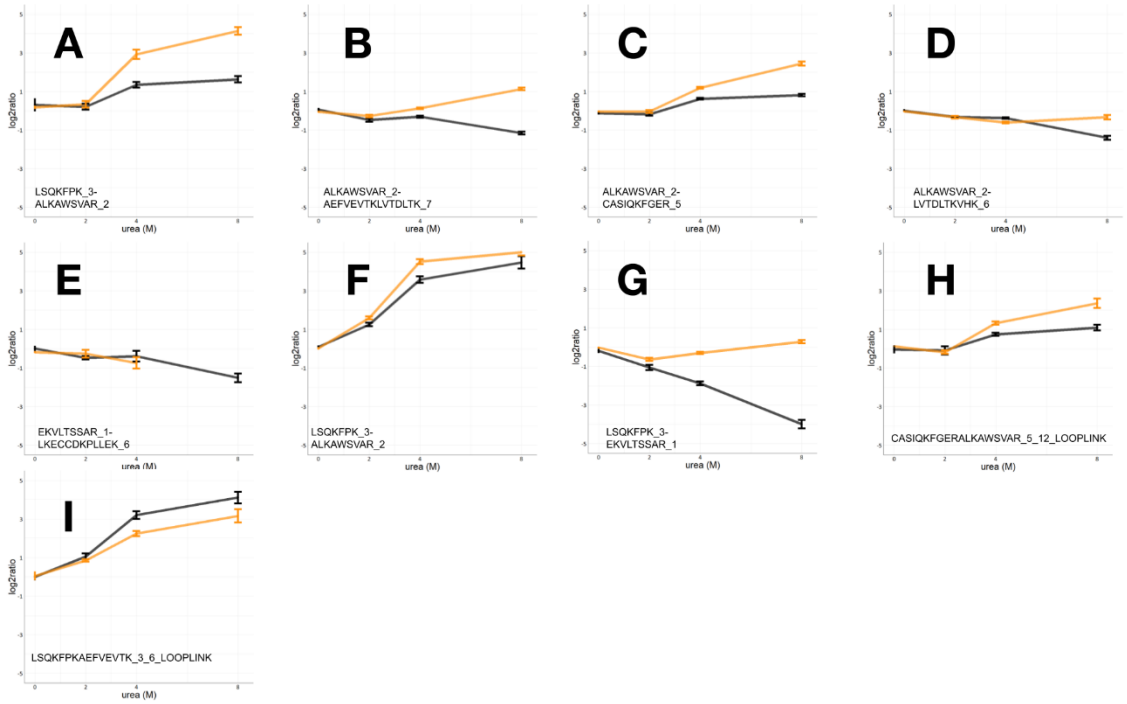


Figure 2.12. qXLMS data obtained on cross-link peptide pairs of lysines residues in domain IIA.

(A-J) Shown are the relative log2ratios obtained for each peptide pairs at increasing urea concentrations for BSA (black) or BSA-bilirubin (orange). Error bars represent 95% quantitation confidence given all relevant calculated MS2-based ion pairs across experimental duplicate and technical triplicate.

Chapter 3. APPLICATION OF PROTEIN-DENATURATION QXL-MS IN *P. AERUGINOSA* WITH GENTAMICIN

3.1 ABSTRACT

The discovery and validation of new antibiotics have slowed against rising antibiotic resistance on a global level. During the past decade, multiple mass spectrometry-based approaches have been developed for the large-scale evaluation of drug-target and off-target engagement. Quantitative cross-linking mass spectrometry for protein unfolding is one of several methods recently established to evaluate protein-ligand engagement based on the increased protein structural stability upon ligand binding. Here, we extend this technique to identify Pseudomonas protein stability changes involving the aminoglycoside gentamicin. The known binding interaction of gentamicin with the ribosome in a cell lysate is successfully detected and quantified. Changes in known ribosome binder Elongation factor G and nucleobase-containing compound metabolic proteins were also detected. These results demonstrate that quantitative cross-linking mass spectrometry for protein unfolding is a suitable tool for high-throughput ligand binding studies in a complex biological sample.

3.2 INTRODUCTION

Emerging antibiotic resistance represents a growing healthcare crisis as the rate of developing new antibiotics has not caught up with the pace of spreading antibiotic resistance.[46] Discovery of new antibiotics is urgently needed, but developing new drugs is a lengthy, costly, and often ineffective process. Even for historically established antibiotics, there is debate over how target engagement signals cell death or arrest.[78] The most successful single-agent antibiotics classes engage multiple protein targets and reflect multifactorial antimicrobial strategies. Consequently, clinical strategies against multidrug resistance have relied on drug combinations inaccessible by single-agent regiments.[79]

The ribosome is a common target of antibiotics in its essential task in orchestrating the translation of the genetic code into proteins.[47] The coordinated process requires the binding, interaction, and readout between the small and large ribosomal subunits, messenger RNA, transfer

RNA, and host of translation factors. These interactions are accompanied by a series of conformational movements of the ribosome and translation factors for the fidelity and speed of the translation process.[49]

Aminoglycosides, frequently used for the treatment of Gram-negative infections, ultimately target the translation machinery.[80] The class of antibiotics was first introduced into clinical use after being isolated from *Streptomyces griseus*, with several other members introduced over the following years.[81] The aminoglycosides can be divided into three classes depending on its ring structure (streptamine or 2-deoxystreptamine) and glycosidic linkages (4,5- or 4,6-disubstituted 2-deoxystreptamine).[82] Although resistance-conferring rRNA mutation experiments suggest the classes of aminoglycosies have different preferential specificities to the ribosome, all have been observed to bind to the 16S ribosomal rRNA of the 30S subunit. Specifically, structures have shown binding to the A-site, thus altering its conformation.[83] In turn, the aminoglycoside suite of antibiotics have been observed to reduce translation fidelity to induce mistranslation, inhibit transfer RNA's translocation, and inhibit ribosome recycling.[84-87]

Pseudomonas aeruginosa is an opportunistic pathogenic Gram-negative bacterium commonly associated with nosocomial infections, especially in immunosuppressed and cystic fibrosis hosts.[79] *P. aeruginosa* is prevalent because of its association with hospital acquired infections as well as long survival on abiotic surfaces.[88] This microorganism is widely investigated for its clinical relevance, ability to form strong biofilms, and antibiotic resistance. The growing number of multi-drug and pan-drug resistant strains of *P. aeruginosa* is increasing as resistance has been reported to commonly used drugs such as aminoglysocides, fluoroquinolones, and carbapenems. For example, *P. aeruginosa* strains have been observed to inactivate aminoglycosides through modifications such as acetylation and phosphorylation. And, growing individual aminoglycoside-resistant isolates have been observed to carry multiple modification enzymes.[89] Alternatively, reduced uptake and increased efflux of aminoglycosides have been implicated particularly in cystic fibrosis isolates.[90] The growing genotypic events to support molecular mechanisms of antibiotic resistance generate increasingly limited treatment options and exacerbate the need for alternative pathways for drug-target engagement.

Here, we adapt the protein-denaturation with quantitative XL-MS (qXL-MS) approach to *P. aeruginosa* and the aminoglycoside gentamicin. The XL-MS approach is a rapidly developing structural approach for the low resolution, yet large-scale information on protein and protein

complex in the form of distance restraints between cross-linked amino acid residues.[91] The protein-denaturation and qXL-MS workflow is an MS-based approach for measuring structural stability of proteins and protein-ligand complexes with an increasing concentration of chemical denaturant, and has been previously validated using a model protein and a known ligand. We reason protein-denaturation with qXLMS can provide orthogonal phenotyping of a clinically relevant pathogen and further our understanding of an established antibiotic. We report protein-denaturation with qXLMS detects gentamicin-treatment changes with the ribosome, translation interactor Elongation factor G, as well as proteins in the nucleobase-containing compound metabolism. The result of this work also help establishes the generality of the protein-denaturation and qXL-MS in a complex sample.

3.1 METHODS AND MATERIALS

Materials and bacterial strains

All reagents for sample preparation were obtained through Fisher (Waltham, MA) unless otherwise noted. The *P. aeruginosa* PAO1 strain was used.

Cross-linking and sample preparation

The qXL-MS for protein unfolding protocol was modified to be compatible with *P. aeruginosa*. Bacterial cells were grown overnight at 37°C in LB media and diluted 100-fold into 1 L of fresh Luria broth. Cultures were grown aerobically at 37°C with shaking until a desired optical density (OD_{600}) of ~0.7. Cells were pelleted at 4000 x g for 10 min, washed three times with 20 ml phosphate buffered saline, and resuspended in 170 mM disodium phosphate (Na_2HPO_4) buffer (pH 8) supplemented with protease inhibitor (Roche). Cells were then lysed by cryo-milling (5 × 30 Hz, 1 min).

Cell lysate aliquots of 0.5 mL were diluted at a 1:1 ratio to the desired concentration of urea (0, 2, 4, 8 M). After 10 minutes of incubation, solutions were cross-linked with iqPIR (reporter 808 or reporter 812) from a concentrated stock in dimethyl sulfoxide to a final

concentration of 10 mM. The crosslinking reaction was performed at room temperature for 45 minutes with constant shaking, as previously described. Cross-linked solutions were then mixed 1:1 (iqPIR-808 : iqPIR-812) (mg:mg) to the native, non-denaturing cross-linked condition to generate 3 mg total mixtures. Cross-linked protein mixtures were reduced, alkylated, and digested overnight at 37°C with trypsin (Promega) (1:200) (protein:trypsin). Resulting peptides were desalted by solid phase extraction using C18 SepPak cartridges. Dried cross-linked peptide material were first enriched by peptide size exclusion chromatography and fractionated by strong cation exchange before capture with monomeric avidin immobilized on UltraLink resin. The eluate of enriched cross-linked peptides was concentrated in a vacuum centrifuge and reconstituted in 0.1% formic acid, 4% acetonitrile in water.

LC-MS analysis

Cross-linked peptide samples were analyzed by LC-MS/MS in technical duplicate using a nanoAcquity UPLC system (Waters) coupled to an QE Exactive Plus mass spectrometer (Thermo Scientific). Aliquots of 4 µl were loaded onto a 3 cm x 100 µM inner diameter fused silica trap column packed with a stationary phase consisting of 5 µM Reprosil C8 particles with 120 Å pores (Dr Maish GmbH) with a flow rate of 2 µL/min for 10 min. Peptides were then separated by reverse-phase chromatography analytical column (60 cm x 75 mM) by applying a linear gradient from 82% solvent A (water containing 0.1% formic acid) and 18% solvent B (acetonitrile containing 0.1% formic acid) to 60% solvent A and 40% solvent B over 2 hours at a flow rate of 300 nL/min. The column temperature was set at 55°C.

The QE Exactive was operating using a top five data-dependent acquisition method of ions with a charge state from +4 to +8 with a resolving power setting of 70,000 for MS1 and MS2 scans. Additional settings include a 1e6 AGC value and 100 ms maximum ion time for MS1 scans

and 5e4 AGC value and 300 ms maximum ion time for the MS2 scans. Ions selected for MS2 were dynamically excluded from further selection for 30 s.

LC-MS/MS data files in RAW format were converted to .mzXML and searched for PIR mass relationships within a 10 ppm tolerance using mango.[63] The resulting MS2 files were searched using comet against the Uniprot with *P. aeruginosa* PAO1 database on March 7, 2021 containing both forward and reverse protein sequences.[64] Resulting pepXML files were validated with XLinkProphet and filtered to an estimated false discovery rate of 1% at a nonredundant peptide pair level.[66] Quantification of cross-linked peptide pairs was performed using iqPIR algorithms as described by Chavez et al.[92] The results were uploaded to XLinkDB available as dataset ureaXL_PAO1_wgen_tech2_noNormalize_Bruce. Cross-linked pair log2 ratios and associated p-values based on the Student's t-test on each quantified ion for every cross-link were downloaded from XLinkDB and used in calculations using R 4.0.2. K-means clustering of log2ratios was accomplished using the R-package kml.[67] Cross-links were mapped on available structures with SASD distances calculated by Jwalk.[93]

3.2 RESULTS

The general strategy used in this work is outlined in **Figure 3.1**. In this strategy, a complex biological mixture of proteins is subjected to qXL-MS for protein unfolding analyses in the presence or absence of drug. Each sample is added to a series of cross-linking buffer with an increasing amount of urea. The samples were subjected to chemical cross-linking with isobaric quantitative PIR (reporter-808 or reporter-812) reagent for the same amount of time. After the cross-linking reaction is quenched, the samples were then paired with a vehicle-treated, 0M urea, non-denatured sample prior at a 1:1 ratio prior to digestion. For the application of qXL-MS for

protein unfolding on a complex mixture, resulting cross-linked peptide pairs are further enriched before data acquisition. To achieve this, we implemented enrichments based on the three features of iqPIR molecule. First, we performed peptide size exclusion chromatography to enrich for higher molecular mass cross-linked peptide pairs. Then, we utilized peptide strong cation exchange chromatography to further enrich for the higher charge state ($\geq 4^+$) of inter-peptide cross-linked pairs. And third, fractions were enriched for the biotin tag by incubation with monomeric avidin beads.

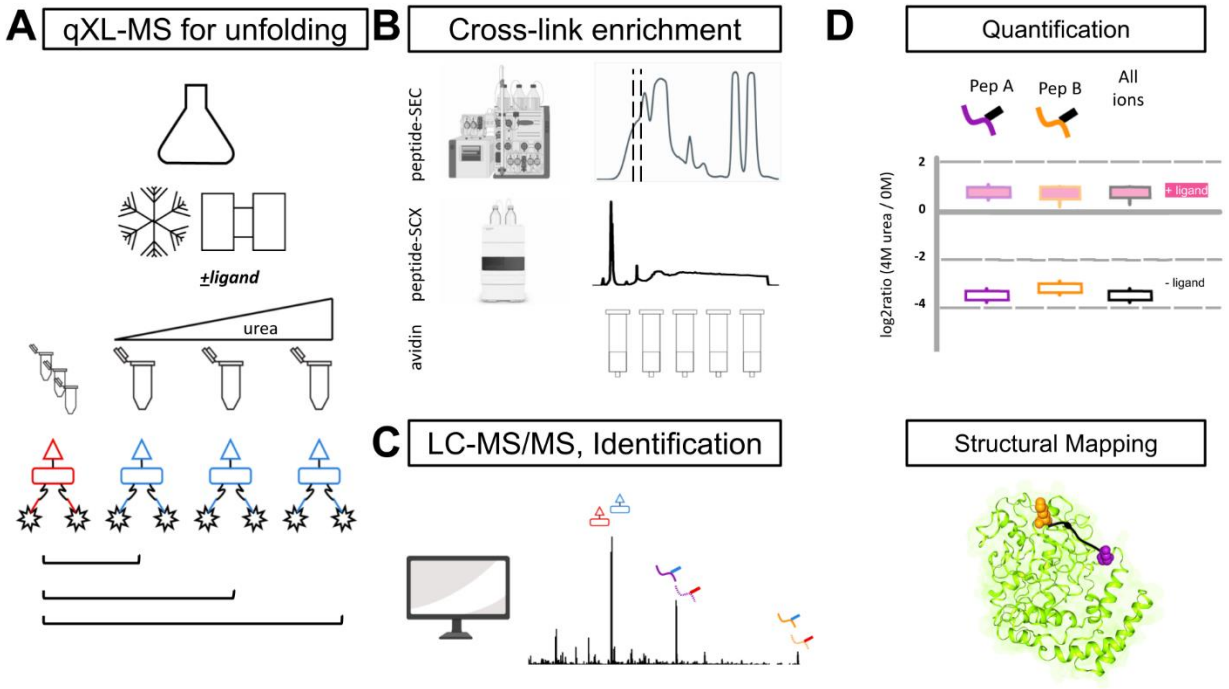


Figure 3.1. Schematic representation of protein-denaturation with quantitative XL-MS used in this study.

(A) Protein-denaturation qXL-MS chemical treatment. A stock lysate was subjected to two treatments, (± drug), and each is further divided and mixed with increasing urea. Samples were cross-linked with iqPIR-808 or iqPIR-812. After crosslinking, samples were mixed with a non-denaturing control at a 1:1 ratio. Mixtures are reduced, alkylated, and digested prior trypsin digestion and sample processing. (B) Cross-link enrichment. Cross-linked peptides were purified by a combination of peptide size exclusion chromatography (SEC), strong cation

exchange (SCX), and avidin affinity. (C) LC-MS. Cross-linked peptide mixtures were prepared for LC-MS/MS using Mango.[63] The data was searched using Comet with crosslink validation using XLinkProphet.[64,66] (D) Identified cross-linked peptides were quantified using in-house processing scripts.[92] Crosslinking data were uploaded to XLinkDB for quantitation viewing and structure mapping.[69] Example peptide K126-K293 (GEAK^{126}IGTTGR-VGSGPFPTELFDDVGAYLAK^{293}R) was mapped to the AlphaFold structure of Pseudomonas protein purA.[94]

3.2.1 *Protein denaturation and qXL-MS enable reproducible quantification of Pseudomonas protein structural changes during urea denaturation*

Mango analysis resulted in 2,224 non-redundant cross-links identified at an estimated false discovery rate (FDR) of 1% or less at the non-redundant cross-linked peptide pair level using probabilistic modeling with XLinkProphet. The links represent 358 protein pairs, including 80% intra-protein linkages. To validate the qXL-MS for protein unfolding strategy, we compared the quantitation of cross-links quantified in both label-swapped experimental replicates. Analysis of urea exposed cross-link quantitation was observed to be highly reproducible with R^2 values increasing from 0.47 to 0.95 (**Figure 3.2**). Quantitative ion information from all experimental and technical replicates were consolidated to obtain a single log2ratio for each crosslink at a single condition. To identify statistically meaningful cross-link changes responsive to urea exposure and gentamicin treatment, we applied multiple levels of statistical filtering. Quantified cross-links were filtered for at least four contributing ion pair ratios and a 95% confidence intervals of less than one. All cross-link data have been uploaded into XLinkDB and are accessible as the following dataset: ureaXL_PAO1_wgen_tech2_noNormalize_Bruce.

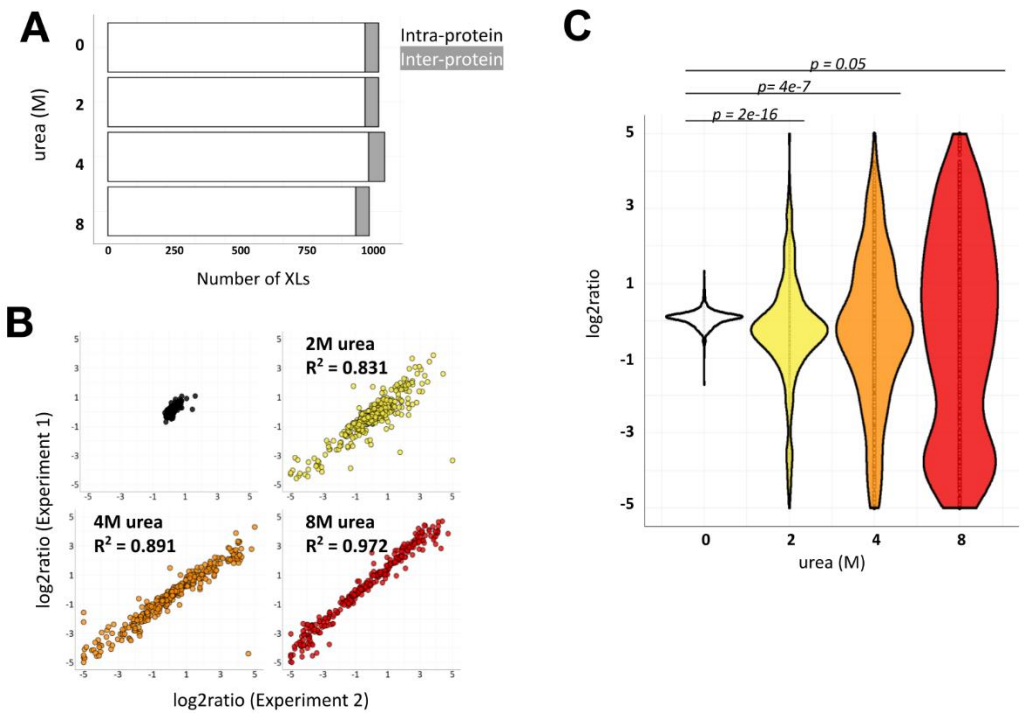


Figure 3.2 Evaluation of quantitative cross-linking mass spectrometry for protein unfolding in a complex bacterial lysate.

(A) Bar charts comparing peptide identification results obtained across urea conditions. White and grey bars denote intra-protein and inter-protein links, respectively. (B) Cross-link peptide pair quantitation correlates between label-swap experimental replicates. (C) Distribution of quantified cross-link ratios from a bacterial lysate for different urea exposures. Each point represents a unique cross-link quantitation.

No significant differences in identified inter-protein links were observed across urea samples (**Figure 3.2**). The mixing of isobaric cross-linked samples to a non-denaturated sample did not decrease the amount of inter-protein cross-links identified. Quantification of these inter-protein links shows a relative decrease of their abundance across increased urea exposure (**Figure 3.7**). Interestingly, the observed bimodal distribution of quantified inter-protein cross-links at 2M urea exposure separates the inter-protein links of the RNA polymerase (mean Δlog2ratio = -3.82, n = 9 cross-links) and succinyl-CoA synthesase (mean Δlog2ratio = -0.61, n = 6 cross-links) complexes. Other protein-denaturation MS-based methods observed similar denaturation curves

for co-complex proteins in the same compartment exhibiting.[23] It is worth noting that biochemical studies of bacterial succinyl-CoA still observed enzyme activity at 2M urea.[95]

The distribution of cross-linked log2ratios was observed to increase with urea exposure (**Figure 3.3**). We asked whether the proportion of cross-links that were significantly non-zero increased with the amount of urea. Cross-links with at least a two-fold change and Bonferroni adjusted p-value < 0.05 were considered at each urea concentration. From the lowest and highest concentrations of urea tested, the significantly changed quantified cross-link population changed from 25% to 62% (**Table 3.1**). While the 95% confidence interval was observed to increase relative to urea exposure (**Figure 3.6**, mean $\Delta CI = +0.13$, p-value = 2e-16, paired Student's t test), the increased error represented less than 2% of the full range of log2ratios calculated at 8M urea exposure.

3.2.2 *Denaturation of the bacterial ribosome*

The 70S ribosome is a highly conserved structure that contains over fifty distinct ribosomal proteins that interact with the ribosomal RNAs and each other.[96] Despite the catalytic activity of the rRNA in protein synthesis, the analysis of ribosomal proteins and their variation or deficiency has contributed to their importance for ribosome binding, translation fidelity, and overall function. Quantified ribosomal intra-protein cross-links filtered for those quantified in at least three of the four urea concentrations were subject to longitudinal k-means clustering on the log2ratios against urea. Previous XL-MS work with BSA in chemical denaturant shows that primary sequence distance correlated to decreased abundances of two-linked residues with more denaturant exposure.[61] With this in mind, the primary sequence distance between clusters was statistically significant (**Figure 3.3**, p = 7e-5, two-way ANOVA).

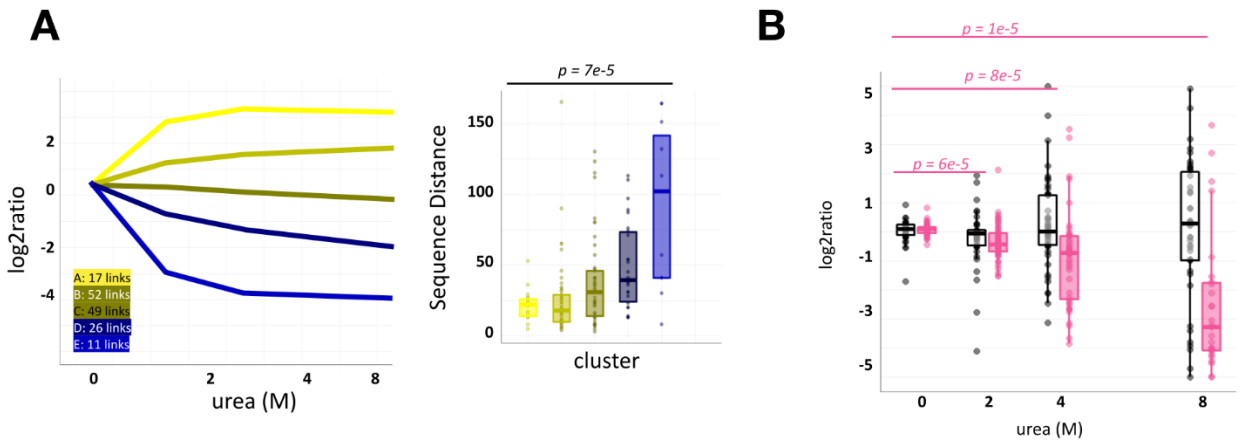


Figure 3.3 Characterizing Pseudomonas ribosome denaturation by qXL-MS for protein unfolding.

(A) Quantified intra-protein cross-links of the 70S ribosome were classified to five clusters by K-means clustering. The primary sequence distance for each cross-linked residue pair was calculated and plotted for each cluster. The distribution of the cluster's population is plotted as a box-and-whisker and their median center represented. Calculated primary sequence distances were compared by ANOVA. (B) Quantified cross-links are separated on the intra-domain (black) and inter-domain (pink) level. Each dot represented a quantified individual cross-link with the population plotted as a box-and-whisker.

We wondered if this primary sequence distance observance would extend to structural domain-level trends. Cross-linked residues were mapped back to their Uniprot-annotated domain and links were binned as intra-domain or inter-domain. Cross-links with either one or both missing domain annotations were excluded from this analysis. We observed a significant decrease of inter-domain cross-links with increasing amount of urea (Δlog2ratio = -3.3 [8M urea], p-value = 1e-5) and no similar shift for intra-domain links (**Figure 3.3**).

We next asked whether secondary structure affected urea sensitivity. Previous biochemical studies found that helical structures are biased to unfold after β-sheets during urea exposure.[97,98] Although many Pseudomonas proteins lack high-resolution crystal structures, we leveraged the AlphaFold Protein Structure Database to better annotate cross-linked lysine residues with

secondary structural information.[94] Cross-linked peptide pairs were considered mapped to an α-helix if their primary sequence distance was no more than 19 residues, at least one residue was annotated with a helical structure, and the other as a bend or turn; β-sheets were analyzed in the same way. Links between a helix and sheet constituted 22% of the population and were excluded from analysis. We identified 15% of quantified cross-links had helical structure mappings. This proportion for either secondary structure did not change when filtering for the statistically significant non-zero cross-links at any urea concentration (**Figure 3.8**). Although these data do not coincide with previous findings that helical structures unfold after β-sheets during urea exposure, the large-scale nature of this dataset could be corroborated with other peptide-level protein-denaturation MS methods. [97,98]

3.2.3 *Effects of gentamicin on the ribosome structural interactome*

Without the addition of urea, we performed quantitative XL-MS to detail the effects of gentamicin-related Pseudomonas ribosome interactome changes. The binding of gentamicin to RNA helix 44 and helix 69 of the small and large ribosomal subunit, respectively, has been well characterized at the genetic, biochemical, and structural level.[83,87] Gentamicin- and vehicle-treated Pseudomonas lysates were cross-linked with iqPIR reagents, paired, processed, and analyzed. Between two labeled swapped biological replicates and 202 quantified cross-links, we observed correlation at R^2 = 0.625 (**Figure 3.4**). Quantitative ion information from both experiments and their technical replicates were consolidated to obtain a single log2 (gentamicin/vehicle) ratio for each cross-link for a total of 456 gentamicin-related quantified cross-links. Quantified cross-linked peptide pairs were filtered for at least a two-fold change difference and Bonferonni adjusted p-value of 0.05 to identify 87 cross-links showing statistically significant differences from gentamicin incubation alone (**Figure 3.4**). Of these statistically different cross-

links, 60% of these cross-links were mapped to the 70S ribosome. The negative control of vehicle-treated cross-linked lysate mixed with itself showed suggested no significant changes (**Figure 3.9**).

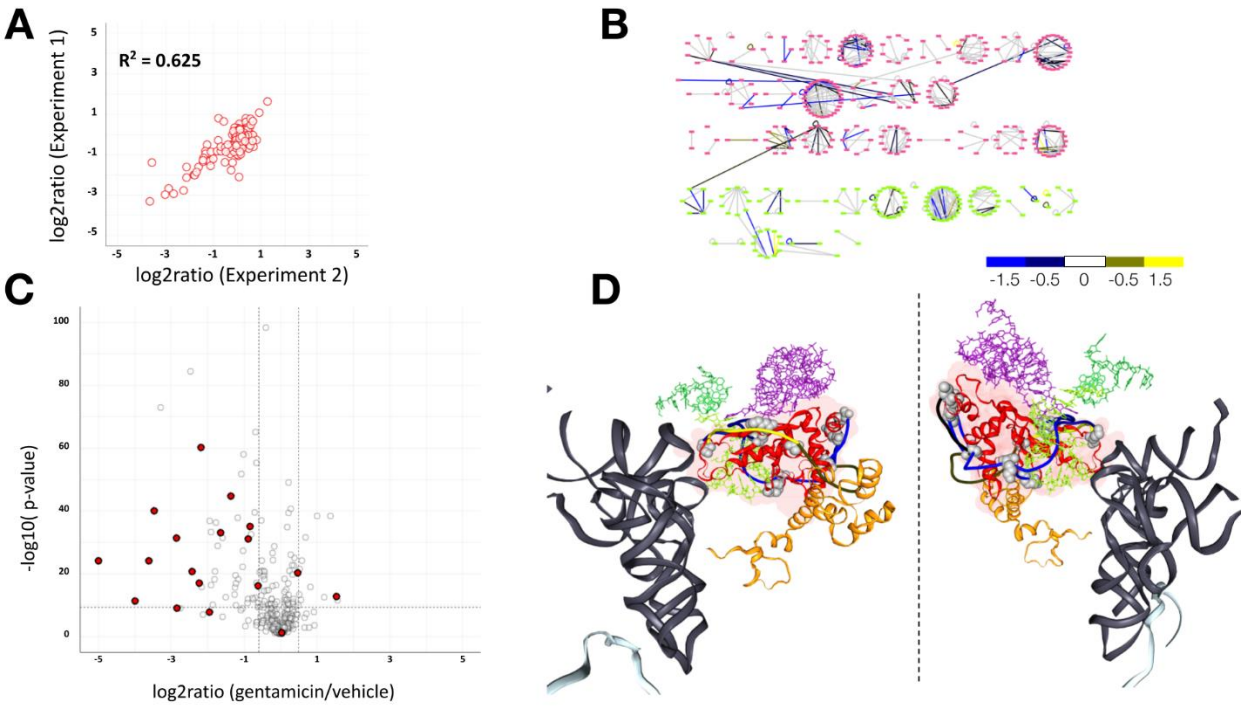


Figure 3.4 Gentamicin induced structural change on the ribosome by qXL-MS.
(A) Scatterplot of 202 quantified cross-link peptide pairs quantified in both label-swapped experiments. (B) Interaction network illustrating cross-linked ribosome lysine residues as nodes and cross-links as edges between the 30S (green) and 50S (pink) proteins. Edges are colored by their log2ratio(gentamicin/vehicle). (C) Volcano plot of quantified cross-linked peptide ratios log2(gentamicin/vehicle) versus statistical significance (p = 0.05). Student's t-test was applied to a cross-link's relevant quantified ion pairs and p-values were adjusted for multiple hypothesis testing by Bonferroni correction. Cross-links mapped to ribosomal proteins of extensive molecular rearrangements during translocation are highlighted in red.[49] (D) Cross-linked lysine pairs of L5 are mapped to the 70S ribosome (PDB: 7SSL). mRNA and tRNA are shown in grey and silver, respectively. 5S rRNA, 23S rRNA helix 84 and helix 85 are shown in purple, light green, and dark green, respectively. Ribosomal protein S13 is represented in gold and L5 in red with its solvent accessible surface highlighted. Quantified intra-protein cross-links are colored by their log2ratio (gentamicin/vehicle). The image is given with 180° y-axis rotation.

To validate our data, we first assessed an assortment of proteins known to be involved with the translating ribosome. Single molecule fluorescence resonance energy transfer and dye-labeled tRNA analysis have been used to observe gentamicin-induced delay of the translating ribosome elongation translocation cycle.[99,100] Crystallographic characterization of the ribosome during translation has captured snapshots of this dynamic process which the ribosome and all its functional ligands undergo conformational changes. These atomic models show that the reconfiguration of tRNAs and translation of the ribosome are influenced by ribosomal proteins L2, L5, L14, S13, and S19.[49] As anticipated, cross-links mapped to these proteins accounted for over 25% of the significantly decreased cross-link population (**Figure 3.4**).

Although over 60Å away from gentamicin's binding site on the 30S ribosome, a majority of L5 intra-protein cross-links were observed to decrease with gentamicin. L5 is conserved across biological kingdoms and has been experimentally identified as critical for intermolecular ribosomal RNA interaction.[101] The structure of L5 consists of five-stranded antiparallel β -sheet and four α-helices whose folding topology closely resembles the ribonuclear protein domain.[102] Early work in ribosomal macrostructure identified the existence of inter-subunit bridges and later experiments highlighted and contextualized their conformations with respect to the larger conformational changes.[103] The L5-S13 B1b bridge is the only identified protein-protein intersubunit bridge at this time.[104,105] While no L5-S13 inter-protein cross-links were detected in this study, cross-linking measurements quantified increase in intra-molecular cross-links in RL5 between K59-K79 with a log2(gentamicin/vehicle) = 3.03, indicating these sites were more frequently available for crosslinking in during gentamicin treatment (**Figure 3.4**). Mapping these sites onto the RL5 crystal structure correspond to domains α3 and β2-β3 turn, respectively. Alanine-scanning mutagenesis of L5 found mutations in both of these regions decreased 5S RNA

binding.[102] The decreased cross-links K88-K120 (Δlog2ratio = -1.37) and K120-K133 (Δlog2ratio = -2.84) engage the 23S h44. The single increased K59-79 link, as well as decreased K88-63 and K88-K79 are situated towards h45. The combination of these links support L5 conformational changes and the overall flexibility of the central perturbance.[105]

Surprisingly, 10% of all significantly decreased ribosomal cross-links mapped to protein L15 despite no linkage to its mobility in initial cryo-EM studies **(Figure 3.10**).[49] Investigations in identifying ribosomal proteins responsible for the peptidyl transferase center initially implicated L15, but its requirement is still up to contention. Nonetheless, L15 is considered a critical hub for assembly and formation of central perturbance, whose chemical footprinting has been used as a probe for ribosome structure.[106,107]

3.2.4 *Effects of gentamicin on Pseudomonas protein structural stability*

In the case of ligand-induced structural stability changes detected by protein-denaturation and qXL-MS, the observed log2ratio for a cross-link peptide-pair at a specific urea concentration with and without ligand can be compared by a ratio of ratios. To provide greater statistical comparison between these observed ratio differences, a two-way ANOVA was performed on all quantified ion pairs relevant to each cross-link at a given drug- or vehicle- treated urea concentration. The volcano plot generated from these calculations is shown in **Figure 3.5**.

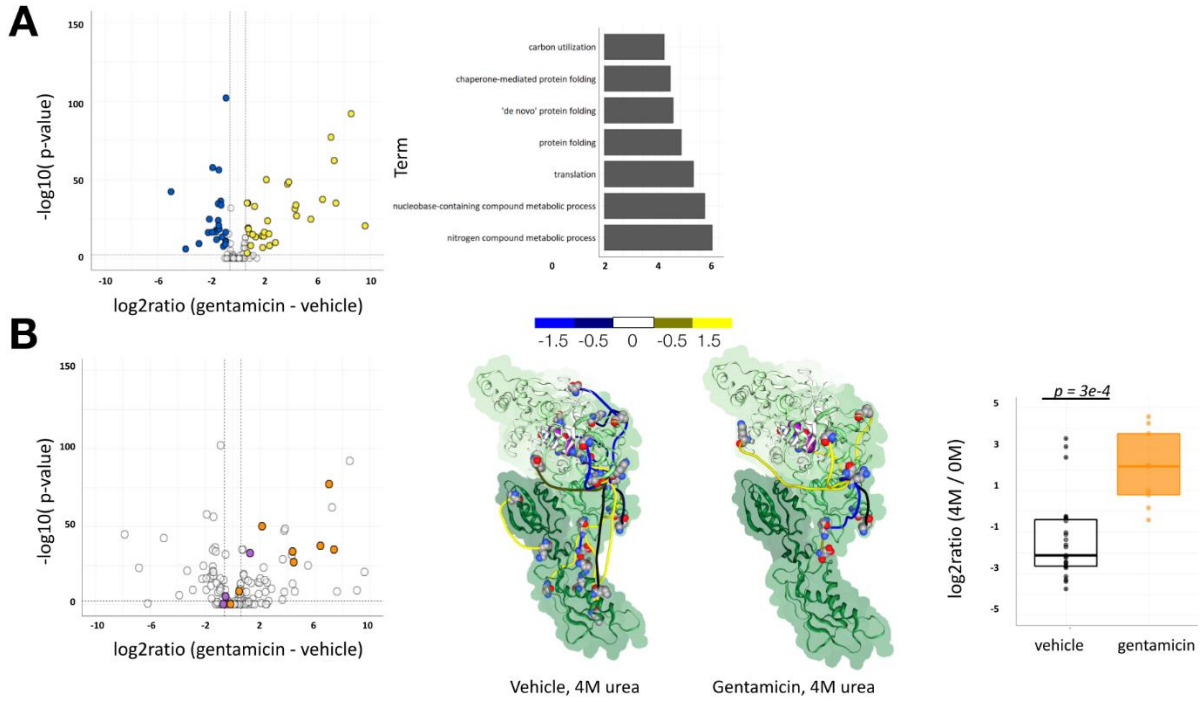


Figure 3.5 qXL-MS for protein unfolding of gentamicin-incubated Pseudomonas lysate identifies gentamicin-stabilized links in Elongation factor G.

(A) Volcano plot of qXL-MS data for protein unfolding scores at 4M urea exposure versus statistical significance (p = 0.05). Two-way ANOVA was applied to a cross-link relative ratio with or without drug and p-values were adjusted for multiple hypothesis testing by Bonferroni correction. Links filtered as gentamicin-destabilized (n = 27, blue) and stabilized (n = 34, yellow) are highlighted. Panther functional annotation Gene Ontology (GO) enrichment analysis of proteins with statistically significant cross-links. Bar plots represent the log-transformed p-value of GO terms. (B) Cross-linked peptides of EF-G (orange) and EF-Tu (purple) are highlighted on the qXL-MS for protein unfolding volcano plot. Quantified cross-links are shown and p-values for group comparisons were computed using Wilcoxon signed rank test. Cross-links are mapped to the EF-G structure (PDB 7SSL). EF-G domains are uniquely colored with cross-link SASD distances colored by log2ratio. Cross-link SASDs distance are mapped and colored by their specific log2ratios.

Gene ontology (GO) over-representation analysis of proteins with statistically significant urea-denaturation structural stability provided deeper insights to the biological processes where gentamicin proteins participate (**Figure 3.5**). Gentamicin-effector localization overlaps with established ribosomal translation machinery (**Figure 3.5**). Data analysis of qXL-MS for protein unfolding of the over-enriched translation machinery, excluding ribosomes, revealed seven changed links that mapped to Elongation factor G (**Figure 3.5**). EF-G is a universally conserved translational five-domain GTPase that promotes the translocation of tRNA through the ribosome. .[104,108,109] Single-molecule fluorescence resonance energy transfer studies that demonstrate stabilization of the ribosomal classical state correlate with their inhibition of EF-G catalyzed translocation.[99] qXL-MS data on EF-G detected losses only of K143-K437 (Δlog2ratio = -1.12; p-value = 7e-4, two-way ANOVA). The shortest SASD of this cross-link passes through the known GDP binding site in domain I (**Figure 3.11**). However, protein-denaturation and qXL-MS data revealed gentamicin-induced urea stability of domain I and other N-terminal domains (**Figure 3.5**). Moreover, the same K143-K437 link exhibited gentamicin-induced stabilization at 4M urea (Δlog2ratio = +7.04; p-value = 4e-17, two-way ANOVA). Overall, the altered log2(urea/native) values for EF-G cross-links saw a gentamicin-induced stabilization effect (mean Δlog2ratio = +3.44; p-value = 3e-4, two-way ANOVA). Point mutations across the entirety of *Pseudomonas* EF-G have observed various levels of gentamicin resistance.[110] The protein-denaturation and qXL-MS data here further supports the importance of EF-G in the gentamicin mode of action despite no studies documenting direct gentamicin-EF-G binding.

Our GO analysis results also showed that protein folding and nitrogen-based compound metabolism were overrepresented in the proteins whose cross-links exhibited gentamicin-induced changes of structural stability (**Figure 3.5**). The enrichment of protein folding is not surprising

given aminoglycosides have been observed to reduce translation fidelity and induce mistranslation.[84-87] Global mRNA analyses of gentamicin-treated *Bacillus subtilis* observed upregulation of heat shock genes.[111] In *Escherichia coli*, overexpression of GroEL/GroES provided increased survival following aminoglycoside exposure.[112] The observed enrichment of nitrogen-based compound metabolism reflects the growing importance of metabolic functions.[113] Metabolomic studies of *P. aeruginosa* treated with aminoglycoside amikacin and polymyxin demonstrated their synergistic effect interfered with the cell envelope biogenesis, tricarboxylic acid pathway, and arginine metabolism.[114] One of our strongest observed cases of gentamicin-induced protein stability was observed between K48-K147 on phosphoglucosamine mutase (glmM), an essential enzyme for cell wall precursor UDP-N-acetylglucosamine (mean Δlog2ratio = +8.56; p-value = 1e-40, two-way ANOVA).[115] A recent computational model of *E. coli* antimicrobial resistance and their genetic determinants suggested the importance of *glmM* for gentamicin survival.[116] Our results are in line with their *in silico* predictions and provide experimental support to the importance of glmM in aminoglycoside bacterial survival.

3.3 Concluding Remarks

This study demonstrates an extended use of quantitative XL-MS for protein unfolding to a Pseudomonas cellular lysate. The use of qXL-MS provides measurements to detect changes on the ribosomal structure with gentamicin exposure. The use of qXL-MS and urea on the gentamicin-Pseudomonas system enables large-scale quantitation of protein structural stability and further highlights the importance of Elongation factor G for the elongating ribosome. These data support the use of qXL-MS for protein unfolding as a suitable approach for quantitative drug-target structural stability analyses. The complex nature of MS-based approaches in assessing target engagement relies on a diverse repertoire of available methodologies when evaluating

therapeutic candidates. Utilizing this strategy in addition to current methods may provide more meaningful patterns to ultimately assay and validate a potential compound against an entire cellular proteome.

3.4 ACKNOWLEDGEMENTS

We thank all the members of the Bruce Lab, as well as the University of Washington Proteomics Resource, helpful discussions. Dr. Colin Manoil provided advice and *P. aeruginosa* strain PAO1. This project was supported by the following grants from the National Institutes of Health 5R01GM086688 and 1R35GM136255.

3.5 SUPPLEMENTAL MATERIALS

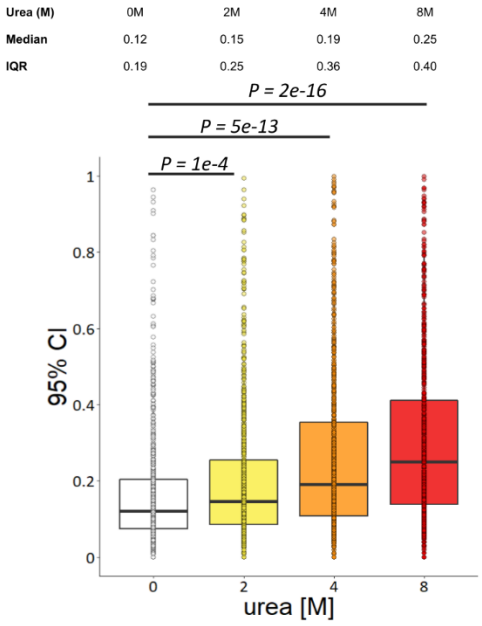


Figure 3.6 Distribution of 95% confidence interval for cross-link log2ratios with increasing denaturant concentration

All associated errors were compared by paired Student's t-test with Bonferroni multiple test correction.

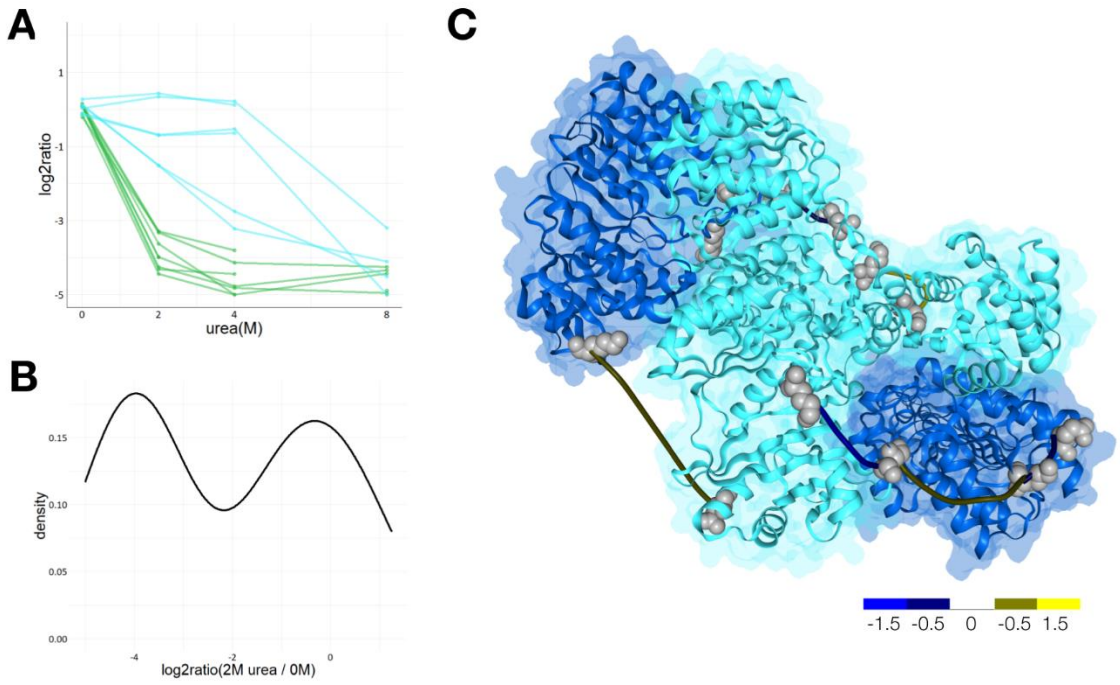


Figure 3.7 Co-complex inter-protein links follow similar trends at low urea.

(A) Protein-denaturation and qXL-MS data obtained on vehicle-treated inter-protein cross-link peptide pairs of RNA polymerase (green) and succinyl-CoA synthease (blue). Each dot is a quantified cross-link mapped to two separate proteins with lines connecting log2ratios of the same cross-link across different urea concentrations. (B) Density plot showing the distribution of inter-protein cross-links log2ratios quantified at 2M urea exposure. (C) Mapping of SucC-SucD cross-links to the succinate-CoA ligase crystal structure (PDB 1JKJ). The crystal structure is colored according to protein (SucC blue, SucD cyan) and cross-linked residues represented as grey spheres. Cross-linked residues are connected by SASD and colored accordingly to their log2ratio (2M urea / 0M urea).

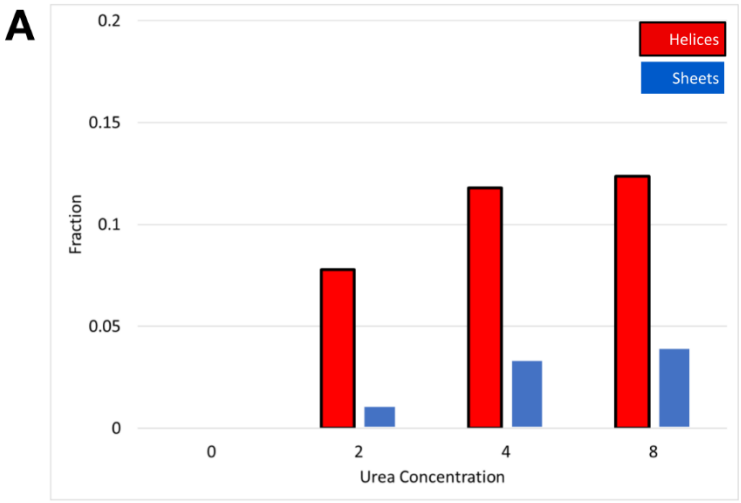


Figure 3.8 Differences of helical and sheet secondary structure as detected using protein-denaturation and qXL-MS.

Fraction of significant links quantified as non-zero (log2ratio > ±1 and adjusted p-value < 0.05) that map to either an alpha helix (red) or beta sheet (blue).

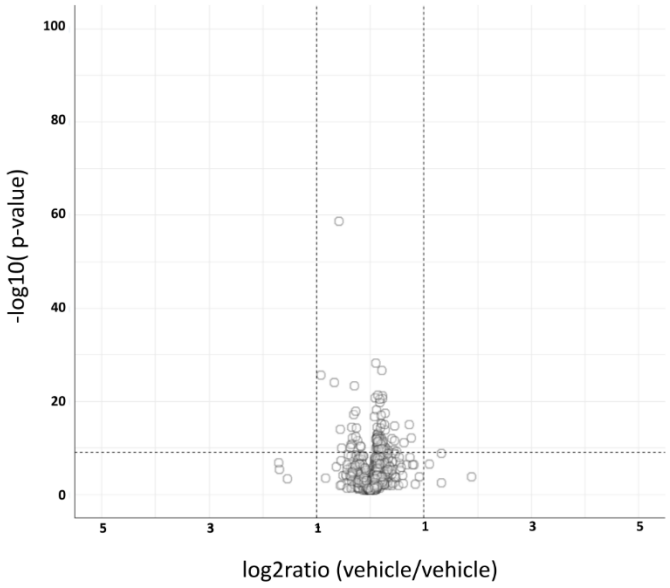


Figure 3.9 Vehicle-treated Pseudomonas control lysate shows no significant interactome changes by quantitative XL-MS.

Volcano plot of quantified cross-linked peptide ratios log2(vehicle/vehicle) versus statistical significance (p = 0.05). Student's t-test was applied to a cross-link's relevant quantified ion pairs and p-values were adjusted for multiple hypothesis testing by Bonferroni correction.

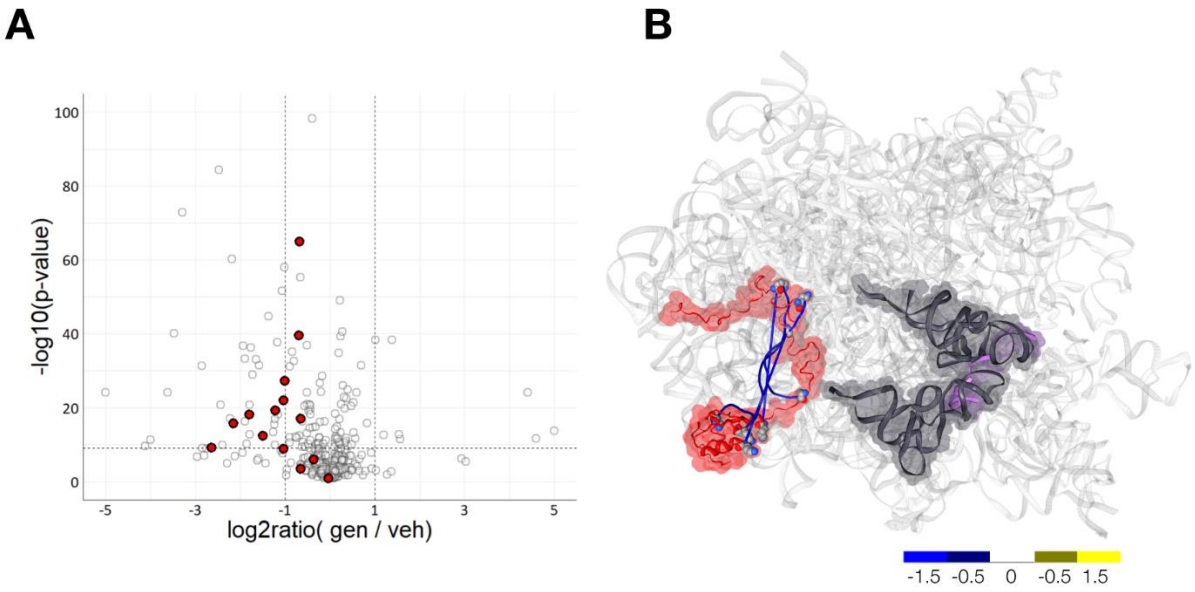


Figure 3.10 RL15 exhibits gentamicin-induced structural changes.

(A) Volcano plot of quantified cross-linked peptide ratios log2(gentamicin/vehicle) versus statistical significance (p = 0.05). Student's t-test was applied to a cross-link's relevant quantified ion pairs and p-values were adjusted for multiple hypothesis testing by Bonferroni correction. Cross-links mapped to ribosomal protein L15 are highlighted in red. (B) L15 structure (PDB 7SSL) displayed as a molecular cartoon (red) against the ribosome structure. Cross-linked lysine residues are shown as grey space-filled residues with cross-link SASDs displayed as colored bars by their log2(gentamicin/vehicle) ratio. The tRNAs and mRNA are indicated by black and purple highlights.

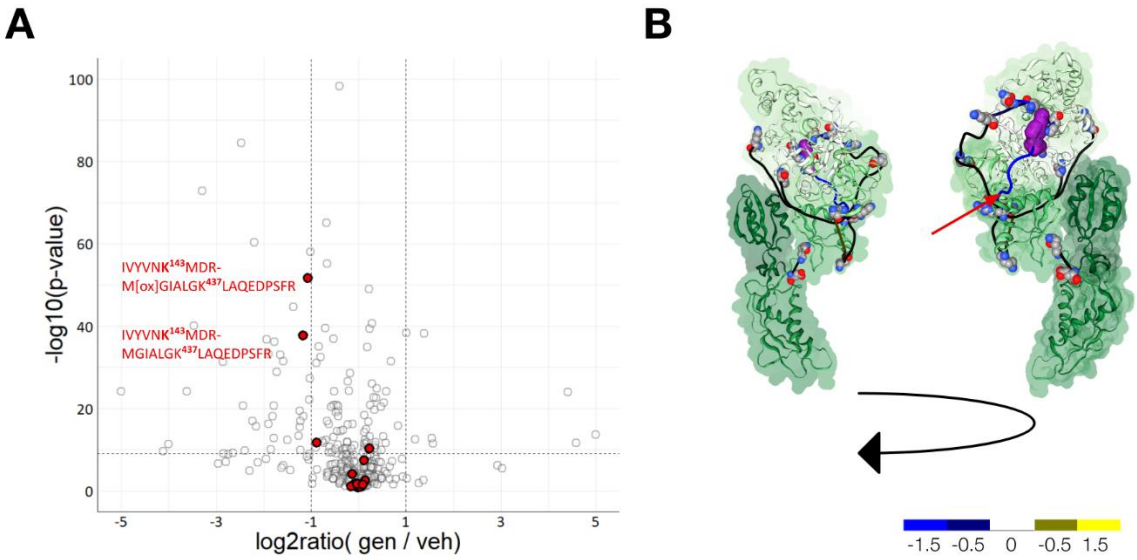


Figure 3.11 Gentamicin-induced quantitative XL-MS identifies decrease in K143-K437. (A) Volcano plot of quantified cross-linked peptide ratios log2(gentamicin/vehicle) versus statistical significance (p = 0.05). Student's t-test was applied to a cross-link's relevant quantified ion pairs and p-values were adjusted for multiple hypothesis testing by Bonferroni correction. Cross-links mapped to Elongation factor G are highlighted in red. (B) EF-G structure (PDB 7SSL) displayed as a molecular cartoon. Domains are colored different shades of green, cross-linked lysine residues are shown as grey space-filled residues with cross-link SASDs displayed as colored bars by their log2(gentamicin/vehicle) ratio. The GDP binding site is marked in purple with a red arrow indicating the K143-K437 link. The structure is rotated 180° to provide different perspectives.

Table 3.1 Summary table of cross-linked peptide pairs with significant non-zero values by qXL-MS across urea

	Urea concentration			
	0	**2**	**4**	**8**
total quantified XLs	682	716	676	650
statistically significant ≥ 2-fold changed XLs	0	182	301	389
percent	0	25	45	60

Table 3.2 Summary Table of qXL-MS ratios for cross-linked peptide pairs with and without gentamicin at 4M urea exposure

Cross-linked peptide pair					log2ratio (4M urea treatment / 0M urea)		
Protein	LysA	LysB	PeptideA	PeptideB	Vehicle	Gentamicin	p_value
LPXC	142	261	EVSVEEGDKR_8	GFKSGHALNNQLLR_2	-2.28	-3.89	1.91E-08
DHAS	257	285	FKNPIPVDGICVR_1	LNKDVPLTDIEGLISQHNPWVK_2	-0.42	-1.69	1.50E-16
ODP1	299	323	AKDGAYVR_1	FWDPLFAKDTAGLLQQR_7	0.75	-1.37	1.66E-11
IF2	759	800	FKDDVAEVR_1	SPKFGAIAGCMVTEGMVHR_2	-2.14	-3.77	5.24E-08
RL1	37	71	LLAELSTIKFK_8	GATVLPNGTGKSVR_10	1.7	-1.61	3.95E-10
RS10	30	59	FTVLISPHVNKDAR_10	LIDQSTQEIVETAKR_13	1.14	0.27	5.08E-45
RL23	44	8	KAVESLFSVK_0	VFKVLLGPHISEK_2	0.44	-0.63	0.000457
RL2	147	183	KVLAECR_0	NIPVGSTVHGIELKPGK_13	-1.78	-3.66	8.62E-26
RL2	125	95	TAHIALLKYADGER_7	GVAAGDQLISGIGAPIKAGNSM[147.04]PLR_16	0.88	-0.66	0.021728
RL22	41	49	VGEALNLLAFSSKK_12	KVLESAVANAEHNEGADVDDLK_0	0.1	-0.8	6.03E-08
RL24	26	42	VLKVLADDR_2	LVVGGVNLIKR_9	-0.18	-1.63	3.30E-11
RL5	120	174	AFKFPFR_2	GLNAKSFDGR_4	-1.77	-3.22	7.80E-16
RL5	120	174	AFKFPFRN_2	GLNAKSFDGR_4	-1.76	-3.2	6.07E-10
RS14	28	50	WNAQVALQKQPR_8	AELKAIIANPNSSAEER_3	0.54	-0.36	1.63E-05
RS8	108	76	YFEGKPVIEEVKR_11	GGLGVSIVSTNKGVM[147.04]TDR_11	0.22	-1.68	0.000101
RL15	29	70	FGFVSLKAMDR_6	GIGSGLGKTGGRGHK_7	0.57	-1.64	7.67E-08
RL15	109	129	GIAATKGAR_5	VKVMLSGEVGR_1	-2.3	-3.19	3.14E-05
RL15	63	96	LPKFGFVSLK_2	VEGDVVSLQTLKDANLINQHVQR_11	3.02	1.11	6.27E-08
RL15	96	96	VEGDVVSLQTLKDANLINQHVQR_11	VEGDVVSLQTLKDANLINQHVQR_11	1.56	-0.02	6.42E-06
RL15	109	96	VKVMLSGEVGR_1	VEGDVVSLQTLKDANLINQHVQR_11	2.81	1.65	1.15E-06
RL15	109	96	VKVM[147.04]LSGEVGR_1	VEGDVVSLQTLKDANLINQHVQR_11	2.75	1.84	0.000143
PUR4	59	89	FGTISPWSSKASDIAR_9	LLKYGPSVPVQEPSGR_2	4.18	-0.83	3.78E-19
RL19	64	88	KISNGVGVER_0	TFQTYSPIVDSLSVKR_14	0.08	-1.15	2.08E-15
KAD	23	79	ITEADCAKGFLFDGFPR_7	FITEKFGIPQISTGDMLR_4	-0.73	-2.12	9.92E-09
KAD	13	23	VILLGAPGAGKGTQAR_10	FITEKFGIPQISTGDM[147.04]LR_4	3.18	0.29	8.58E-05
KAD	13	79	VILLGAPGAGKGTQAR_10	ITEADCAKGFLFDGFPR_7	1.82	-2.08	0.002545
ENO	404	418	VSKYNQLLR_2	IEEQLGAKAPYR_7	3.29	-3.53	7.17E-11
RS1	117	158	VIKLDQK_2	VKGGFTVDVNGIR_1	0.46	-0.94	8.58E-10
TIG	146	183	KQNTR_0	IDGEAFAGGSAKGTLLVLGSGR_11	3.52	-4.35	3.25E-20
HTPG	202	304	GLKLYVQR_2	KYSDHIALPIELPK_0	-2.02	-3.42	4.27E-25
ARUC	117	164	VFLANSGAEANEAAFKLAR_15	YSDGFGPKFEGITHVPYNDLEALK_7	-0.9	0.56	0.026827
RPOA	25	1215	AKITLEPLER_1	AMLKLADLPESGQM[147.04]R_3	-3.81	3.82	0.000285

RPOA	25	1215	AKITLEPLER_1	AM[147.04]LKLADLPESGQMR_3	-4.07	5	0.00012
RPOA	25	1215	AKITLEPLER_1	AM[147.04]LKLADLPESGQM[147.04]R_3	-5	4.61	1.02E-09
FMT	291	46	GQKLM[147.04]PSAVK_2	LQLPGGKPLAFADLYNSR_6	-2.98	1.33	2.14E-14
ILVE	235	283	GPVTEKLQK_5	NTILTLAAEHGFKLVEK_12	-0.64	0.12	5.94E-16
EFTU	317	57	AFDQIDNAPEEKAR_11	FECEVYVLSKEEGGR_9	-0.7	0.58	5.17E-15
BRAC	139	157	GYKLIFR_2	TIGLDNMQGPVAGKFIAER_13	2.19	4.04	6.89E-07
NIRS	165	513	VVQPEYNKR_7	KQLNDLDLPNLFSVTLR_0	-1.51	0.45	8.63E-07
CH60	117	132	AVAAGMNPM[147.04]DLKR_11	ATVAIVAQLKELAKPCADTK_9	-0.85	-0.14	0.033616
DHAS	117	257	KVIDQALDAGTR_0	FKNPIPVDGICVR_1	-0.56	0.35	1.09E-07
SYD	514	575	TQSAGDVM[147.04]TQAPGSVDGKALR_17	VLGIDEAEQEEKFGFLLDALK_11	-3.74	0.79	0.000242
SUCD	272	282	FAALQDAGVKTVR_9	SLADIGKALAELTGWEVK_6	-0.74	0.06	3.65E-09
PPIA	149	162	DFGYAVFGKVVR_8	GMGVVDQIAKVPTTR_9	0.3	1.11	7.62E-09
ATPB	140	80	VIDLVCPFAKGGK_9	GLNVDSTGAAISVPVGKATLGR_16	-1.87	1.79	2.91E-13
PURA	126	293	ARGEAKIGTTGR_5	VGSGPFPTELFDDVGAYLAKR_19	-3.94	-1.1	4.73E-05
PURA	126	293	GEAKIGTTGR_3	VGSGPFPTELFDDVGAYLAKR_19	-3.63	0.75	1.87E-06
RL9	111	35	DIAEAVSAAGYPLEKAEVR_14	NFLLPQGKATVATAENVAAFEAR_7	-1.75	2.01	2.90E-21
RL9	111	22	VANLGNLGDKVNIKGGYAR_13	DIAEAVSAAGYPLEKAEVR_14	-2.94	2.57	1.42E-11
PUR9	281	428	SFLKPACVIVK_3	APTEQEIHDLIFAWKVAK_14	-3.89	-1.88	6.66E-08
GLMM	187	48	VLIGKDTR_4	VVLDCANGATYKIAPSVFR_11	-4.3	4.26	1.02E-40
RL1	14	37	AIAEKVVAGK_4	LLAELSTIKFK_8	-0.61	0.39	0.000332
RS7	25	36	KAVAER_0	YGSQILAKFMNHVMESGK_7	1.09	2.45	1.18E-06
EFG1	375	437	EEIKEVR_3	MGIALGKLAQEDPSFR_6	-2.22	2.16	2.10E-15
EFG1	375	437	EEIKEVR_3	M[147.04]GIALGKLAQEDPSFR_6	-2.22	2.23	1.71E-12
EFG1	143	437	IVYVNKMDR_5	MGIALGKLAQEDPSFR_6	-2.49	4.55	3.53E-34
EFG1	143	437	IVYVNKM[147.04]DR_5	MGIALGKLAQEDPSFR_6	-2.66	3.74	5.08E-17
EFG1	143	437	IVYVNKM[147.04]DR_5	M[147.04]GIALGKLAQEDPSFR_6	-3.15	4.24	5.75E-16
EFG1	485	703	YAEAPSNIVEALVKK_13	EFGVEANIGKPQVAYR_9	-1.37	0.79	1.80E-22
RL22	41	48	AAEIMKK_5	VGEALNLLAFSSKK_12	-0.29	0.85	2.92E-07
RL22	16	41	ISAQKAR_4	VGEALNLLAFSSKK_12	-0.83	1.57	0.000349
RL5	120	88	GLNAKSFDGR_4	EGWPIGVKVTLR_7	-0.03	0.68	5.04E-16
RL5	120	133	GLNAKSFDGR_4	GNYSM[147.04]GVKEQIIFPEIDYDKIDALR_7	-1.46	2.38	8.64E-22
RL15	109	129	GIAATKGAR_5	VKVM[147.04]LSGEVGR_1	-2.29	0.23	8.15E-07
TAL	76	76	FAVAVGKDILGVIPGR_6	FAVAVGKDILGVIPGR_6	-2.37	1.41	0.000332
ACON2	221	229	AKGFPVAYVGDVVGTGSSR_1	DGIEPVQPGSVGPLKQIEAVK_14	-2.25	-0.35	0.001189
HTPG	102	48	ELISNASDAADKLR_11	EDVVTHLGTIAKSGTADFLK_11	-3.42	3.85	9.55E-28
HTPG	277	360	SALTKR_4	VEGKLEYTSLLYVPGR_3	-0.81	1.45	4.70E-11
KDSA	257	257	LNKLEAFLSQLK_2	LNKLEAFLSQLK_2	-1.51	0.84	1.96E-07

Chapter 4. CONCLUDING REMARKS

Mass spectrometry-based methods have fundamentally altered the throughput in how we study and analyze protein structure. Structural MS has contributed insights with relatively smaller amounts of input material and without the need of a highly purified system. Recent developments of protein-denaturation MS-workflows have brought a new generation of techniques that have enabled deeper study of protein-ligand interactions on the proteome-scale. Symbolized by TPP, previous work that required some prior knowledge of interaction partners and high affinity antibodies can now be done at large-scale without prior knowledge. Although TPP quantifies protein-level solubility loses, there have been opportunities made capturing residue-level, denatured-state conformational topologies. Techniques like SPROX show success at capturing residue-level solvent accessibility across the denaturation curve, but remains limited to low-abundance residue methionine.

To provide alternative techniques, I interfaced the protein-denaturation experimental design with XL-MS workflows. The main development of this protein-denaturation with quantitative XL-MS workflow has three aspects: 1) the proof-of-concept of such a method in a model protein, 2) the application of this method to a complex cellular lysate, and 3) the demonstrated utilization of this workflow as another protein-denaturation MS-based technique to assay protein-ligand engagement. The utilization qXL-MS for protein unfolding is a new perspective in probing protein structural stability topologies that can be used with existing methods to provide a more complete understanding of protein target engagement.

Protein-denaturation with quantitative XL-MS adds to the growing suite of current generation protein-denaturation MS-based target-engagement methodologies by extending the application of isobaric quantitative PIR technologies. Yet, the chosen measurement of relative

cross-linked peptide pair ratios positions itself in a unique space between the two fields. On one hand, qXL-MS for protein unfolding provides more reactive sites and 'molecular rulers' to current solvent-denaturation and protein footprinting and residue-specific techniques. The combination of these techniques could potentially be more precise in identifying sub-global denaturation states. However, qXL-MS for protein unfolding curves cannot always be reliably fit nor estimated in a certain direction unlike the precipitation-based losses of protein or oxidation-based loses of native methionine. Moreover, stepping back from the *in vivo* advantages of XL-MS for *in vitro* lysate applications begets old questions on sample preparation optimization.

I confronted several technical hurdles while adapting the protein-denaturation and qXL-MS protocol for a complex lysate. First, I encountered lysate labeling requires additional, alternative clean-up steps to deplete hydrolyzed cross-linker prior to biotin-tag avidin enrichment. While increased input material followed with filter aided sample preparation seems the most direct, the binary isobaric nature of current cross-linker design makes such an experimental design unrealistic. We can now utilize a peptide SEC enrichment step to produce a concentrated peak of cross-linked peptide pairs as well as hydrolyzed cross-linker. The enriched fraction can be taken to the established methods of peptide SCX and biotin capture. However, alternative gradients posit the potential of optimized pepSEC methods to remove the need for SCX, lengthy dry down steps, and competition during avidin enrichment.

The second hurdle was the enrichment of cross-links peptide pairs from a complex lysate. At this point in time, late 2021, the current workflow reliably quantifies cross-links from a proteome of ~5000 proteins. When this method was applied to a HeLa cell lysate, cross-links were identified but far fewer (~20%) were quantified with the confidence filtering. The current limitation confines the effective extent of this project and its application to full cellular mammalian

systems. Can we adapt qXL-MS for protein unfolding to identify off-target effects of highly interactive therapeutics such as cyclosporin? The use of click-chemistry and lysine-reactive labels have seen success with two-fold or more unique identified lysines, but such evaluation has not been performed with iqPIR. Can the throughput of protein-denaturation and qXL-MS for protein unfolding be increased with recent iqPIR technologies? To increase the amount of channels, the next generation of isobaric PIR has since been extended to 6-plex iqPIR.[92] In thermal-denaturation methods, Proteome Integral Stability Alteration (PISA) has increased throughput in which the soluble proteins from different temperature points are combined to estimate the area under the protein melting curve.[117] Recently, solvent-PISA applies a similar workflow with a chemical precipitant approach.[118]

In addition to the work in our lab, there have been various methods to facilitate protein-target identification via structural stability studies. These methods have deepened our understanding of the modes of action of drugs that have not been completely elucidated. More importantly, the current generation of MS-based methods complement each other well to be used as a combinatorial target identification approach. Overall, advances in quantitative protein structural mass spectrometry techniques will together provide increased chances of discovering and validating therapeutics and combinatorial therapies.

4.1 REFERENCES

1. Anfinsen CB. Principles that govern the folding of protein chains. Science 1973;181(4096):223-30.
2. Ellis RJ, Hartl FU. Principles of protein folding in the cellular environment. Curr Opin Struct Biol 1999;9(1):102-10.
3. Englander SW, Mayne L. The nature of protein folding pathways. Proc Natl Acad Sci U S A 2014;111(45):15873-80.
4. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646-74.
5. Fitzgerald MC, West GM. Painting proteins with covalent labels: what's in the picture? J Am Soc Mass Spectrom 2009;20(6):1193-206.
6. Feng X, Liu X, Luo Q, Liu BF. Mass spectrometry in systems biology: an overview. Mass Spectrom Rev 2008;27(6):635-60.
7. Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 2009;11:49-79.
8. Boersema PJ, Kahraman A, Picotti P. Proteomics beyond large-scale protein expression analysis. Curr Opin Biotechnol 2015;34:162-70.
9. Moffat JG, Vincent F, Lee JA, Eder J, Prunotto M. Opportunities and challenges in phenotypic drug discovery: an industry perspective. Nat Rev Drug Discov 2017;16(8):531-543.
10. Wang L, Chance MR. Protein Footprinting Comes of Age: Mass Spectrometry for Biophysical Structure Assessment. Mol Cell Proteomics 2017;16(5):706-716.
11. Mandell JG, Baerga-Ortiz A, Falick AM, Komives EA. Measurement of solvent accessibility at protein-protein interfaces. Methods Mol Biol 2005;305:65-80.
12. Gingras AC, Gstaiger M, Raught B, Aebersold R. Analysis of protein complexes using mass spectrometry. Nat Rev Mol Cell Biol 2007;8(8):645-54.
13. Budayeva HG, Cristea IM. A mass spectrometry view of stable and transient protein interactions. Adv Exp Med Biol 2014;806:263-82.
14. Shrake A, Ross PD. Ligand-induced biphasic protein denaturation. J Biol Chem 1990;265(9):5055-9.
15. Schellman JA. The thermodynamic stability of proteins. Annu Rev Biophys Biophys Chem 1987;16:115-37.
16. Brandts JF, Lin LN. Study of strong to ultratight protein interactions using differential scanning calorimetry. Biochemistry 1990;29(29):6927-40.
17. Senisterra G, Chau I, Vedadi M. Thermal denaturation assays in chemical biology. Assay Drug Dev Technol 2012;10(2):128-36.
18. Martinez Molina D, Jafari R, Ignatushchenko M, Seki T, Larsson EA, Dan C, Sreekumar L, Cao Y, Nordlund P. Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. Science 2013;341(6141):84-7.

19. West GM, Tang L, Fitzgerald MC. Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass spectrometry-based strategy. Anal Chem 2008;80(11):4175-85.
20. Savitski MM, Reinhard FB, Franken H, Werner T, Savitski MF, Eberhard D, Martinez Molina D, Jafari R, Dovega RB, Klaeger S and others. Tracking cancer drugs in living cells by thermal profiling of the proteome. Science 2014;346(6205):1255784.
21. Park C, Marqusee S. Pulse proteolysis: a simple method for quantitative determination of protein stability and ligand binding. Nat Methods 2005;2(3):207-12.
22. Kaur U, Meng H, Lui F, Ma R, Ogburn RN, Johnson JHR, Fitzgerald MC, Jones LM. Proteome-Wide Structural Biology: An Emerging Field for the Structural Analysis of Proteins on the Proteomic Scale. J Proteome Res 2018;17(11):3614-3627.
23. Mateus A, Bobonis J, Kurzawa N, Stein F, Helm D, Hevler J, Typas A, Savitski MM. Thermal proteome profiling in bacteria: probing protein state. Mol Syst Biol 2018;14(7):e8242.
24. Mateus A, Kurzawa N, Becher I, Sridharan S, Helm D, Stein F, Typas A, Savitski MM. Thermal proteome profiling for interrogating protein interactions. Mol Syst Biol 2020;16(3):e9232.
25. Adhikari J, Fitzgerald MC. SILAC-pulse proteolysis: A mass spectrometry-based method for discovery and cross-validation in proteome-wide studies of ligand binding. J Am Soc Mass Spectrom 2014;25(12):2073-83.
26. Geer Wallace MA, Kwon DY, Weitzel DH, Lee CT, Stephenson TN, Chi JT, Mook RA, Dewhirst MW, Hong J, Fitzgerald MC. Discovery of Manassantin A Protein Targets Using Large-Scale Protein Folding and Stability Measurements. J Proteome Res 2016;15(8):2688-96.
27. Xu Y, Wallace MA, Fitzgerald MC. Thermodynamic Analysis of the Geldanamycin-Hsp90 Interaction in a Whole Cell Lysate Using a Mass Spectrometry-Based Proteomics Approach. J Am Soc Mass Spectrom 2016;27(10):1670-6.
28. Tran DT, Adhikari J, Fitzgerald MC. StableIsotope Labeling with Amino Acids in Cell Culture (SILAC)-based strategy for proteome-wide thermodynamic analysis of protein-ligand binding interactions. Mol Cell Proteomics 2014;13(7):1800-13.
29. Zeng L, Shin WH, Zhu X, Park SH, Park C, Tao WA, Kihara D. Discovery of Nicotinamide Adenine Dinucleotide Binding Proteins in the Escherichia coli Proteome Using a Combined Energetic- and Structural-Bioinformatics-Based Approach. J Proteome Res 2017;16(2):470-480.
30. DeArmond PD, West GM, Huang HT, Fitzgerald MC. Stable isotope labeling strategy for protein-ligand binding analysis in multi-component protein mixtures. J Am Soc Mass Spectrom 2011;22(3):418-30.
31. Chavez JD, Mohr JP, Mathay M, Zhong X, Keller A, Bruce JE. Systems structural biology measurements by in vivo cross-linking with mass spectrometry. Nat Protoc 2019;14(8):2318-2343.
32. Chavez JD, Lee CF, Caudal A, Keller A, Tian R, Bruce JE. Chemical Crosslinking Mass Spectrometry Analysis of Protein Conformations and Supercomplexes in Heart Tissue. Cell Syst 2018;6(1):136-141.e5.
33. Navare AT, Chavez JD, Zheng C, Weisbrod CR, Eng JK, Siehnel R, Singh PK, Manoil C, Bruce JE. Probing the protein interaction network of Pseudomonas aeruginosa cells by chemical cross-linking mass spectrometry. Structure 2015;23(4):762-73.

34. Tang X, Bruce JE. A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. Mol Biosyst 2010;6(6):939-47.
35. Bich C, Maedler S, Chiesa K, DeGiacomo F, Bogliotti N, Zenobi R. Reactivity and applications of new amine reactive cross-linkers for mass spectrometric detection of protein-protein complexes. Anal Chem 2010;82(1):172-9.
36. Zhang H, Tang X, Munske GR, Tolic N, Anderson GA, Bruce JE. Identification of protein-protein interactions and topologies in living cells with chemical cross-linking and mass spectrometry. Mol Cell Proteomics 2009;8(3):409-20.
37. Zheng C, Yang L, Hoopmann MR, Eng JK, Tang X, Weisbrod CR, Bruce JE. Cross-linking measurements of in vivo protein complex topologies. Mol Cell Proteomics 2011;10(10):M110.006841.
38. Zhong X, Navare AT, Chavez JD, Eng JK, Schweppe DK, Bruce JE. Large-Scale and Targeted Quantitative Cross-Linking MS Using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. J Proteome Res 2017;16(2):720-727.
39. Chavez JD, Keller A, Mohr JP, Bruce JE. Isobaric Quantitative Protein Interaction Reporter Technology for Comparative Interactome Studies. Anal Chem 2020;92(20):14094-14102.
40. Chavez JD, Liu NL, Bruce JE. Quantification of protein-protein interactions with chemical cross-linking and mass spectrometry. J Proteome Res 2011;10(4):1528-37.
41. Chavez JD, Schweppe DK, Eng JK, Zheng C, Taipale A, Zhang Y, Takara K, Bruce JE. Quantitative interactome analysis reveals a chemoresistant edgotype. Nat Commun 2015;6:7928.
42. Chavez JD, Schweppe DK, Eng JK, Bruce JE. In Vivo Conformational Dynamics of Hsp90 and Its Interactors. Cell Chem Biol 2016;23(6):716-26.
43. Chavez JD, Keller A, Zhou B, Tian R, Bruce JE. Cellular Interactome Dynamics during Paclitaxel Treatment. Cell Rep 2019;29(8):2371-2383.e5.
44. Wippel HH, Fioramonte M, Chavez JD, Bruce JE. Deciphering the architecture and interactome of hnRNP proteins and enigmRBPs. Mol Omics 2021;17(4):503-516.
45. Zhou B, Caudal A, Tang X, Chavez JD, McMillen TS, Keller A, Villet O, Zhao M, Liu Y, Ritterhoff J and others. Upregulation of mitochondrial ATPase inhibitory factor 1 (ATPIF1) mediates increased glycolysis in mouse hearts. J Clin Invest 2022;132(10).
46. Brown ED, Wright GD. Antibacterial drug discovery in the resistance era. Nature 2016;529(7586):336-43.
47. Wilson DN. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nat Rev Microbiol 2014;12(1):35-48.
48. Ling C, Ermolenko DN. Structural insights into ribosome translocation. Wiley Interdiscip Rev RNA 2016;7(5):620-36.
49. Gao H, Sengupta J, Valle M, Korostelev A, Eswar N, Stagg SM, Van Roey P, Agrawal RK, Harvey SC, Sali A and others. Study of the structural dynamics of the E coli 70S ribosome using real-space refinement. Cell 2003;113(6):789-801.
50. Gao YG, Selmer M, Dunham CM, Weixlbaumer A, Kelley AC, Ramakrishnan V. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. Science 2009;326(5953):694-9.
51. Achenbach J, Nierhaus KH. Translocation at work. Nat Struct Mol Biol 2013;20(9):1019-22.

52. Brodersen DE, Clemons WM, Carter AP, Wimberly BT, Ramakrishnan V. Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: structure of the proteins and their interactions with 16 S RNA. J Mol Biol 2002;316(3):725-68.
53. Klein DJ, Moore PB, Steitz TA. The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. J Mol Biol 2004;340(1):141-77.
54. Lauber MA, Reilly JP. Structural analysis of a prokaryotic ribosome using a novel amidinating cross-linker and mass spectrometry. J Proteome Res 2011;10(8):3604-16.
55. Tüting C, Iacobucci C, Ihling CH, Kastritis PL, Sinz A. Structural analysis of 70S ribosomes by cross-linking/mass spectrometry reveals conformational plasticity. Sci Rep 2020;10(1):12618.
56. Sand KM, Bern M, Nilsen J, Noordzij HT, Sandlie I, Andersen JT. Unraveling the Interaction between FcRn and Albumin: Opportunities for Design of Albumin-Based Therapeutics. Front Immunol 2014;5:682.
57. Sudlow G, Birkett DJ, Wade DN. The characterization of two specific drug binding sites on human serum albumin. Mol Pharmacol 1975;11(6):824-32.
58. Kragh-Hansen U. Evidence for a large and flexible region of human serum albumin possessing high affinity binding sites for salicylate, warfarin, and other ligands. Mol Pharmacol 1988;34(2):160-71.
59. He XM, Carter DC. Atomic structure and chemistry of human serum albumin. Nature 1992;358(6383):209-15.
60. Wippel HH, Chavez JD, Keller AD, Bruce JE. Multiplexed Isobaric Quantitative Cross-Linking Reveals Drug-Induced Interactome Changes in Breast Cancer Cells. Anal Chem 2022;94(6):2713-2722.
61. Wang JH, Tang YL, Gong Z, Jain R, Xiao F, Zhou Y, Tan D, Li Q, Huang N, Liu SQ and others. Characterization of protein unfolding by fast cross-linking mass spectrometry using di-ortho-phthalaldehyde cross-linkers. Nat Commun 2022;13(1):1468.
62. Tayyab S, Sharma N, Mushahid Khan M. Use of domain specific ligands to study urea-induced unfolding of bovine serum albumin. Biochem Biophys Res Commun 2000;277(1):83-8.
63. Mohr JP, Perumalla P, Chavez JD, Eng JK, Bruce JE. Mango: A General Tool for Collision Induced Dissociation-Cleavable Cross-Linked Peptide Identification. Anal Chem 2018.
64. Eng JK, Jahan TA, Hoopmann MR. Comet: an open-source MS/MS sequence database search tool. Proteomics 2013;13(1):22-4.
65. Leptos KC, Sarracino DA, Jaffe JD, Krastins B, Church GM. MapQuant: open-source software for large-scale protein quantification. Proteomics 2006;6(6):1770-82.
66. Keller A, Chavez JD, Bruce JE. Increased sensitivity with automated validation of XL-MS cleavable peptide crosslinks. Bioinformatics 2019;35(5):895-897.
67. Genolini C, Falissard B. KmL: a package to cluster longitudinal data. Comput Methods Programs Biomed 2011;104(3):e112-21.
68. Combe CW, Fischer L, Rappsilber J. xiNET: cross-link network maps with residue resolution. Mol Cell Proteomics 2015;14(4):1137-47.
69. Keller A, Chavez JD, Eng JK, Thornton Z, Bruce JE. Tools for 3D Interactome Visualization. J Proteome Res 2019;18(2):753-758.

70. Ryan MC, Stucky M, Wakefield C, Melott JM, Akbani R, Weinstein JN, Broom BM. Interactive Clustered Heat Map Builder: An easy web-based tool for creating sophisticated clustered heat maps. F1000Res 2019;8.
71. M a B, Tie Z, Z ou D, L i J, W ang W. Urea- and thermal-induced unfolding of bovine serum albumin. Modern Physics Letters B 2006;20(29):1909-1916.
72. Katz S, Denis J. Structural transformations in serum albumin as demonstrated by urea-perturbation technique. Biochem Biophys Res Commun 1967;28(5):711-7.
73. Ni Y, Su S, Kokot S. Spectrofluorimetric studies on the binding of salicylic acid to bovine serum albumin using warfarin and ibuprofen as site markers with the aid of parallel factor analysis. Anal Chim Acta 2006;580(2):206-15.
74. Khan MY, Agarwal SK, Hangloo S. Urea-induced structural transformations in bovine serum albumin. J Biochem 1987;102(2):313-7.
75. Goncharova I, Orlov S, Urbanová M. The location of the high- and low-affinity bilirubin-binding sites on serum albumin: ligand-competition analysis investigated by circular dichroism. Biophys Chem 2013;180-181:55-65.
76. Zunszain PA, Ghuman J, McDonagh AF, Curry S. Crystallographic analysis of human serum albumin complexed with 4Z,15E-bilirubin-IXalpha. J Mol Biol 2008;381(2):394-406.
77. Tatlidil D, Ucuncu M, Akdogan Y. Physiological concentrations of albumin favor drug binding. Phys Chem Chem Phys 2015;17(35):22678-85.
78. Dwyer DJ, Belenky PA, Yang JH, MacDonald IC, Martell JD, Takahashi N, Chan CT, Lobritz MA, Braff D, Schwarz EG and others. Antibiotics induce redox-related physiological alterations as part of their lethality. Proc Natl Acad Sci U S A 2014;111(20):E2100-9.
79. Kunz Coyne AJ, El Ghali A, Holger D, Rebold N, Rybak MJ. Therapeutic Strategies for Emerging Multidrug-Resistant Pseudomonas aeruginosa. Infect Dis Ther 2022;11(2):661-682.
80. Garneau-Tsodikova S, Labby KJ. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. Medchemcomm 2016;7(1):11-27.
81. Woodruff HB. Selman A. Waksman, winner of the 1952 Nobel Prize for physiology or medicine. Appl Environ Microbiol 2014;80(1):2-8.
82. Serio AW, Keepers T, Andrews L, Krause KM. Aminoglycoside Revival: Review of a Historically Important Class of Antimicrobials Undergoing Rejuvenation. EcoSal Plus 2018;8(1).
83. Borovinskaya MA, Pai RD, Zhang W, Schuwirth BS, Holton JM, Hirokawa G, Kaji H, Kaji A, Cate JH. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. Nat Struct Mol Biol 2007;14(8):727-32.
84. Wasserman MR, Pulk A, Zhou Z, Altman RB, Zinder JC, Green KD, Garneau-Tsodikova S, Cate JH, Blanchard SC. Chemically related 4,5-linked aminoglycoside antibiotics drive subunit rotation in opposite directions. Nat Commun 2015;6:7896.
85. Davies J, Gorini L, Davis BD. Misreading of RNA codewords induced by aminoglycoside antibiotics. Mol Pharmacol 1965;1(1):93-106.
86. Hirokawa G, Kiel MC, Muto A, Selmer M, Raj VS, Liljas A, Igarashi K, Kaji H, Kaji A. Post-termination complex disassembly by ribosome recycling factor, a functional tRNA mimic. EMBO J 2002;21(9):2272-81.

87. Misumi M, Nishimura T, Komai T, Tanaka N. Interaction of kanamycin and related antibiotics with the large subunit of ribosomes and the inhibition of translocation. Biochem Biophys Res Commun 1978;84(2):358-65.
88. Moradali MF, Ghods S, Rehm BH. Lifestyle: A Paradigm for Adaptation, Survival, and Persistence. Front Cell Infect Microbiol 2017;7:39.
89. Poole K. Aminoglycoside resistance in Pseudomonas aeruginosa. Antimicrob Agents Chemother 2005;49(2):479-87.
90. Poole K. Pseudomonas aeruginosa: resistance to the max. Front Microbiol 2011;2:65.
91. Chavez JD, Wippel HH, Tang X, Keller A, Bruce JE. In-Cell Labeling and Mass Spectrometry for Systems-Level Structural Biology. Chem Rev 2022;122(8):7647-7689.
92. Chavez JD, Keller A, Wippel HH, Mohr JP, Bruce JE. Multiplexed Cross-Linking with Isobaric Quantitative Protein Interaction Reporter Technology. Anal Chem 2021;93(50):16759-16768.
93. Bullock JMA, Thalassinos K, Topf M. Jwalk and MNXL web server: model validation using restraints from crosslinking mass spectrometry. Bioinformatics 2018;34(20):3584-3585.
94. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A and others. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res 2022;50(D1):D439-D444.
95. Joyce MA, Hayakawa K, Wolodko WT, Fraser ME. Biochemical and structural characterization of the GTP-preferring succinyl-CoA synthetase from Thermus aquaticus. Acta Crystallogr D Biol Crystallogr 2012;68(Pt 7):751-62.
96. WALLER JP. FRACTIONATION OF THE RIBOSOMAL PROTEIN FROM ESCHERICHIA COLI. J Mol Biol 1964;10:319-36.
97. Parui S, Jana B. Relative Solvent Exposure of the Alpha-Helix and Beta-Sheet in Water Determines the Initial Stages of Urea and Guanidinium Chloride-Induced Denaturation of Alpha/Beta Proteins. J Phys Chem B 2019;123(42):8889-8900.
98. Hua L, Zhou R, Thirumalai D, Berne BJ. Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. Proc Natl Acad Sci U S A 2008;105(44):16928-33.
99. Tsai A, Uemura S, Johansson M, Puglisi EV, Marshall RA, Aitken CE, Korlach J, Ehrenberg M, Puglisi JD. The impact of aminoglycosides on the dynamics of translation elongation. Cell Rep 2013;3(2):497-508.
100. Aguirre Rivera J, Larsson J, Volkov IL, Seefeldt AC, Sanyal S, Johansson M. Real-time measurements of aminoglycoside effects on protein synthesis in live cells. Proc Natl Acad Sci U S A 2021;118(9).
101. Perederina A, Nevskaya N, Nikonov O, Nikulin A, Dumas P, Yao M, Tanaka I, Garber M, Gongadze G, Nikonov S. Detailed analysis of RNA-protein interactions within the bacterial ribosomal protein L5/5S rRNA complex. RNA 2002;8(12):1548-57.
102. Nakashima T, Yao M, Kawamura S, Iwasaki K, Kimura M, Tanaka I. Ribosomal protein L5 has a highly twisted concave surface and flexible arms responsible for rRNA binding. RNA 2001;7(5):692-701.
103. Ning W, Fei J, Gonzalez RL. The ribosome uses cooperative conformational changes to maximize and regulate the efficiency of translation. Proc Natl Acad Sci U S A 2014;111(33):12073-8.

104. Frank J, Gao H, Sengupta J, Gao N, Taylor DJ. The process of mRNA-tRNA translocation. Proc Natl Acad Sci U S A 2007;104(50):19671-8.
105. Valle M, Zavialov A, Sengupta J, Rawat U, Ehrenberg M, Frank J. Locking and unlocking of ribosomal motions. Cell 2003;114(1):123-34.
106. Lieberman KR, Noller HF. Ribosomal protein L15 as a probe of 50 S ribosomal subunit structure. J Mol Biol 1998;284(5):1367-78.
107. Calvet LE, Matviienko S, Ducluzaux P. Network theory of the bacterial ribosome. PLoS One 2020;15(10):e0239700.
108. Czworkowski J, Wang J, Steitz TA, Moore PB. The crystal structure of elongation factor G complexed with GDP, at 2.7 A resolution. EMBO J 1994;13(16):3661-8.
109. Salsi E, Farah E, Dann J, Ermolenko DN. Following movement of domain IV of elongation factor G during ribosomal translocation. Proc Natl Acad Sci U S A 2014;111(42):15060-5.
110. Bolard A, Plésiat P, Jeannot K. Mutations in Gene. Antimicrob Agents Chemother 2018;62(2).
111. Lin JT, Connelly MB, Amolo C, Otani S, Yaver DS. Global transcriptional response of Bacillus subtilis to treatment with subinhibitory concentrations of antibiotics that inhibit protein synthesis. Antimicrob Agents Chemother 2005;49(5):1915-26.
112. Goltermann L, Good L, Bentin T. Chaperonins fight aminoglycoside-induced protein misfolding and promote short-term tolerance in Escherichia coli. J Biol Chem 2013;288(15):10483-9.
113. Zampieri M, Enke T, Chubukov V, Ricci V, Piddock L, Sauer U. Metabolic constraints on the evolution of antibiotic resistance. Mol Syst Biol 2017;13(3):917.
114. Hussein M, Han ML, Zhu Y, Zhou Q, Lin YW, Hancock REW, Hoyer D, Creek DJ, Li J, Velkov T. Metabolomics Study of the Synergistic Killing of Polymyxin B in Combination with Amikacin against Polymyxin-Susceptible and -Resistant Pseudomonas aeruginosa. Antimicrob Agents Chemother 2019;64(1).
115. Tavares IM, Jolly L, Pompeo F, Leitão JH, Fialho AM, Sá-Correia I, Mengin-Lecreulx D. Identification of the Pseudomonas aeruginosa glmM gene, encoding phosphoglucosamine mutase. J Bacteriol 2000;182(16):4453-7.
116. Pearcy N, Hu Y, Baker M, Maciel-Guerra A, Xue N, Wang W, Kaler J, Peng Z, Li F, Dottorini T. Genome-Scale Metabolic Models and Machine Learning Reveal Genetic Determinants of Antibiotic Resistance in Escherichia coli and Unravel the Underlying Metabolic Adaptation Mechanisms. mSystems 2021;6(4):e0091320.
117. Gaetani M, Sabatier P, Saei AA, Beusch CM, Yang Z, Lundström SL, Zubarev RA. Proteome Integral Solubility Alteration: A High-Throughput Proteomics Assay for Target Deconvolution. J Proteome Res 2019;18(11):4027-4037.
118. Van Vranken JG, Li J, Mitchell DC, Navarrete-Perea J, Gygi SP. Assessing target engagement using proteome-wide solvent shift assays. Elife 2021;10.

VITA

Martin T Mathay grew up in the _____. There he attended _____. Afterwards, he matriculated to Swarthmore College where he graduated with a ____. During that time, Martin performed research under the mentorship of Dr. Amy Vollmer and Dr. Alison Holiday. He investigated _____. Before starting graduate school, Martin discovered ___ for large-scale experimental data acquisition in the lab of ____ at _____. There he _____. In ____, after completing three lab rotations, Martin join the lab of ____ in the ____, where he pursued his PhD research on _____.