Top-Down Proteomics Strategies for the Characterization of Endogenous Cardiac Proteins and Protein Complexes

By
Emily A. Chapman

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Chemistry)

At the
UNIVERSITY OF WISCONSIN-MADISON
2025

Date of final oral examination: 01/21/2025

This dissertation is approved by the following members of the Final Oral Committee:

Ying Ge, Professor, Cell & Regenerative Biology and Chemistry
Song Jin, Professor, Chemistry
Amish Raval, Professor, Medicine
Lingjun Li, Professor, Pharmaceutical Sciences and Chemistry
Lloyd Smith, Professor, Chemistry

Top-Down Proteomics Strategies for the Characterization of Endogenous Proteins and Protein Complexes

by

Emily A. Chapman

To my grandfather, John,

Thank you for your unwavering love and support.

Go Badgers!

ACKNOWLEDGEMENTS

When I began searching for graduate schools, I was almost certain I didn't want to stay in the Midwest. Having grown up in the Chicago suburbs and then attending undergrad in Minneapolis, I was ready for a change. However, when I visited UW-Madison in early 2020, just before the pandemic, I knew immediately that Madison was where I belonged. During that graduate recruiting weekend, the kindness of graduate students, the welcoming atmosphere of the Chemistry department, and beauty of the campus left a lasting impression on me. I still remember driving back to Minneapolis after the recruiting weekend, thinking that no other school could compare and that Madison was where I wanted to call home for the next 5 years. A few months later, I found myself starting the Chemistry PhD program at UW-Madison, right in the middle of a pandemic. It was a strange time as there was no in-person welcome week, most rotations were online, and I even TA'd CHEM 103 online from my small studio apartment. Despite this challenging time, I was fortunate enough to join the Ge group, which had some pretty awesome people and exciting research happening.

Throughout my first year of graduate school, many older graduate students would often tell me how quickly the years would pass. Being only in my first year, I didn't quite believe them at the time. My TBE, RP, and thesis defense all seemed so far away and in the distant future. I never imagined that graduate school would go as fast as it did, but here I am, writing the acknowledgements section of my dissertation. A lot has happened over the past five years, and I am incredibly grateful to the people who have supported me throughout this journey. I would like to take some time to acknowledge those who have offered their constant love and support along the way.

To my mom and dad, thank you for your support and love over the years. You have taught me the value of hard work and the importance of perseverance. Whether it was driving 6+ hours to watch me play soccer or now visiting me in Wisconsin to catch a Badger football game, your efforts to be there for me have meant so much. Even though we're farther apart than we've ever been, I always look forward to my visits and the time we spend together, from record shopping and antiquing to visiting Arrington. I wouldn't be where I am today without your encouragement and guidance, and I'm truly grateful for everything you do.

To my grandfather John (Pops), I cannot thank you enough for your unconditional love and support throughout my life. Losing Pops at the end of my third year of graduate school was incredibly difficult, and I wish he could be here today to see me graduate. He was a true Wisconsinite, born and raised in Kenosha, and even spent some time at UW-Madison before transferring to Carthage College to finish his degree. Pops always encouraged me to work hard, strive for success, and never settle for less. He supported my passion for horseback riding and made it a priority to attend as many of my soccer games as he could, from high school through college. Pops was always my number one fan, and I will be forever grateful for his kindness, support, and love.

To my grandmother Carol (Nana), I want to also express my deepest gratitude for your unconditional love and support. You helped raise me, and I will forever cherish the memories we share, from going to horseback riding lessons together, eating PB&Js (with strawberry jelly), to watching our daily soap operas. I can always count on you to be there for me and to be ready at a moment's notice for a quick FaceTime call to catch up on life. You have always done your best to help and support me, and I am so grateful for that. You have been a constant source of positivity and love in my life, and I truly appreciate everything you have done for me.

To my Papa Bud and Nana Marla, I am so grateful for your unwavering love and support throughout the years. I truly appreciate the money you have sent me every month since I started college; it has been an incredible help, especially during graduate school. I also loved visiting you in Sacramento recently, your kindness and support made the trip feel truly special. I miss you both dearly and hope to come visit again sometime soon.

To the rest of my family, especially my brother Tyler, Uncle Mike, Aunt Beth, and Aunt Jana, thank you so much for everything over the years and for all the support you have given me. We always manage to have such a fun time together and create so many wonderful memories. I am looking forward to the next time I get to visit you all again soon. The moments we share are always so special, and I'm grateful for each one.

I want to give a special thank you to my fiancé, Summer. You have been my rock and my home throughout the last three years of graduate school, and I cannot thank you enough for your love and support, it has meant the world to me. It is funny how our paths have unknowingly crossed before, but it wasn't until the Chemistry soccer team at Keva that we were truly brought together. Over the years, we have created so many incredible memories together, from skiing and completing a century bike ride to running our first full marathon together, as well as traveling to amazing places like Italy. I am also so proud of you for taking the leap and going back to school for nursing. Your drive and passion in life inspire me, and I know you will be a fabulous nurse one day. I am excited for our future together and am so incredibly grateful to always have you by my side.

To my fellow Auggies, Nora, Juice, and Claire, I am so grateful that we have stayed really close friends after college even though now we are all scattered across different places. I'll always remember the late nights we spent watching the Bachelor and Farmer Wants a Wife or going out

to Applebee's and Target after a long day of studying. I will also never forget the memories we made together on the soccer field, from going to Nationals and running the beep test during preseason, to celebrating our wins in Dinkytown. Our recent trips to Fort Myers during the winter provided a much-needed escape from the cold and a nice break from graduate school. I am looking forward to creating more memories with you all and I can't wait to celebrate together soon.

To my new friends in Madison, it has been a blast getting to know all of you. I've met some incredibly cool people through rec soccer, playing on teams like The Bananas and Scott's Tott's. We have shared many nights at The Mason, talking about life, enjoying too many drinks, and eating way too much pizza. I have had so many unforgettable experiences with you all, such as hiking in the UP, curling, biking to New Glarus, and night skiing at Cascade, that have made my graduate school experience so much better. These were some of the best times of my life and I am so lucky to have shared them with you all.

To my friends in the Ge group, thank you for making graduate school not just bearable, but truly fun and entertaining. I have loved all of our scientific (and many non-scientific) discussions at lunch over the years, as they have been a great source of learning and laughter. I honestly couldn't have made it through graduate school without all of your support. To Tim T and David R., thank you for your invaluable guidance, mentorship, and friendship. You both were so incredibly generous with your time, and I am truly grateful for the wealth of knowledge you have shared with me. Your support and expertise have been essential in shaping me into the mass spectrometrist that I am today. To Timothy, my ultimate desk buddy and good friend, thank you for all the laughter and fun times we have shared. I will always remember our summer walks, biking, and cooking sessions together, all filled with great conversations and plenty of laughter. I can't express enough how much your friendship and support have meant to me. To Jake, my first

friend at UW-Madison and host during recruitment weekend, thank you for your guidance and for the many laughs throughout the years. You always knew how to bring a smile to my face and made time spent in (and out) lab more fun. To Eli, I can't thank you enough for introducing me to skate skiing and showing me how to wax my skis, as well as teaching me how to use MASH. To Reasoner, thank you for the nights out, the fun times in lab, and for being there when I needed someone to talk to. To Holden, a good southern gent, thank you for the great times both in and out of the lab. We can always count on you to keep things in check and provide tech support. To Gao, thank you for the many columns you have packed for me over the past several years, I would not have been able to finish my last project without your help. To all of my past and present desk buddies, Eli, Tim, Timothy, Holden, Ruby, Liam, and Gao thank you for letting me distract you when I needed a break and for always being there to bounce ideas and figures off of. To Brad, Matt, Jãán, Yareslie, Kwame, and Haotian, thank you for allowing me to work with you and it has been a wonderful experience being your mentor. Overall, it has been a great privilege to work with so many incredible past and present members of the Ge group and I also want to thank those that I did not already mention including Kyle, Melissa, Boris, Yanlong, Kalina, Scott, Mallory, Anna Grace, Andrew, Mandi, Kevin, David M., Ben, Elizabeth, Dr. Tim, Will, Kent, Morgan M., Sean, and Morgan G.

Graduate school would not have been possible without the support of my undergraduate advisors, Professor Michael Wentzel and Professor Vivian Feng. Working in Wentzel's research group ignited my passion for chemistry and research, while Vivian helped guide me through undergrad and introduced me to mass spectrometry. I am so incredibly grateful for their mentorship and would not be the scientist I am today without them.

Lastly, I am extremely thankful for the guidance and support provided by my graduate mentor, Professor Ying Ge. Her mentorship has contributed greatly to my successes, and I am proud of the work that we have published together. Ying has played a significant role in my growth as a scientist, and I will forever be grateful for the opportunities she has given me. I would also like to express my gratitude to my committee members, Professor Song Jin, Dr. Amish Raval, Professor Lingjun Li, and Professor Lloyd Smith, for their encouragement and support.

TABLE OF CONTENTS

ABSTRACT

Proteins are a highly diverse group of macromolecules and serve as the primary effectors of nearly all cellular functions, driving biological processes. The proteome is extremely multifaceted, owing to proteoforms—the diverse protein products arising from post-translational modifications (PTMs), alternative splicing, and genetic mutations. The complexity of the proteome is further exacerbated by the assembly of proteins into functional complexes which are highly dynamic entities that display substantial diversity in their composition, PTMs, and non-covalent interactions. Top-down mass spectrometry (MS)-based proteomics has emerged as a powerful technology for characterizing endogenous proteins, their proteoforms, and non-covalent complexes. However, the high dynamic range of the proteome, immense diversity of proteoforms, and low abundance of protein complexes in their native environments present challenges to studying endogenous proteins and protein complexes. Therefore, my doctoral work has sought to address these challenges by developing top-down MS-based proteomics strategies for the characterization of endogenous proteins and protein complexes directly from cardiac tissues. My methods aim to investigate biological heterogeneity at the proteoform level, elucidate protein structure-function relationships, and provide a comprehensive view of the endogenous heart proteoform landscape. **Chapter 1** introduces the role of proteoforms and protein diversity in biological systems and provides a comprehensive overview of denatured and native top-down MS-based proteomics, with a focus on applications to cardiovascular diseases. **Chapter 2** describes the implementation of denatured top-down MS (dTDMS) for the simultaneous quantification of sarcomeric proteoforms and protein expression in the myocardia of non-failing donors without heart diseases, compared to end-stage ischemic cardiomyopathy (ICM) patients. This approach uncovered significant PTM and expression-level changes in sarcomeric proteins that correlated

with left ventricular dysfunction in ICM patients, providing valuable insights into the molecular alterations underlying human ICM. **Chapter 3** describes a native nanoproteomics strategy for the enrichment and subsequent native top-down MS (nTDMS) analysis of the endogenous cardiac troponin complex, the master regulator and calcium (Ca^{2+}) sensitive switch, directly from human cardiac tissue. **Chapter 4** describes a direct infusion nTDMS platform which extends the application of Fourier-transform ion cyclotron resonance MS to directly characterize the structure, PTMs, and Ca^{2+} binding of endogenous proteins from human cardiac tissue lysate without prior separation methods. Finally, **Chapter 5** details a dTDMS approach to identify, quantify, and localize endogenous S-glutathionylation in ventricular myosin light chain 1 from human, swine, and mouse cardiac tissues. Overall, the studies detailed in this dissertation have comprehensively characterized endogenous proteins and protein complexes directly from cardiac tissues, providing valuable insights into the physiological and pathophysiological processes of the heart, while also establishing a framework for their structural characterization to reveal native assemblies and non-covalent interactions.

LIST OF FIGURES

LIST OF SUPPLEMENTAL FIGURES

LIST OF SUPPLEMENTAL TABLES

CHAPTER 1:

Introduction

1.1 Protein Diversity and the Role of Proteoforms

Proteins are a highly diverse group of macromolecules and serve as the primary effectors of cellular function, including catalysis, signaling, and structural support.[1, 2] Together, proteins constitute the proteome—an organized network of proteins that perform their roles at precise times and locations within the cell, often interacting with other proteins or biomolecules to carry out their functions.[3] Unlike the relatively static genome,[4, 5] the proteome is highly dynamic and complex due to proteoforms—diverse protein products resulting from protein post-translational modifications (PTMs), alternative splicing events, and mutations.[6-8] The complexity of the proteome is further increased by the assembly of proteins into functional complexes.[9, 10] Therefore, in the post-genomic era, proteomics—the study of proteins and their associated proteoforms—has become essential for the large-scale identification and quantification of proteins in biological systems, as well as for providing crucial insights into the underlying mechanisms and functions of proteins that are differentially expressed in both healthy and diseased states.[11, 12] However, the vast dynamic range of the proteome and immense diversity of proteoforms, which may exhibit functional differences, can make proteomic analyses challenging.[11]

1.1.1. Protein Structure

Proteins are assembled from a set of twenty different amino acids. Each of the twenty standard amino acid residues contain an α-carboxyl group, an α-amino group, and a unique side chain on the α-carbon that varies in size, charge, and polarity. Protein structure can be characterized at multiple levels of complexity, typically categorized into four distinct structural levels. The simplest level of protein structure, known as the primary structure, is the sequence of amino acids covalently linked together by peptide bonds to form a polypeptide chain. While some regions of the primary sequence may vary without compromising biological function, most proteins contain

essential regions with conserved sequences that are critical for determining their function and higher-order structures. The secondary structure of a protein refers to the local spatial arrangement of adjacent amino acid residues within a polypeptide chain that is stabilized by hydrogen bonds between the backbone atoms. Repetitive folding patterns that make up the secondary structure of a protein, such as α-helices and β-sheets, form through interactions between the amino and carbonyl groups of the backbone. The tertiary structure of a protein refers to the overall three-dimensional (3D) shape of the polypeptide chain, resulting from additional folding beyond the secondary structure. Protein 3D structure is stabilized by interactions between the side chains of amino acids, including hydrogen bonds, disulfide bridges, ionic bonds, Van der Waals interactions, and nonpolar/polar interactions. A protein typically has one or more stable 3D structures, or conformations that are integral to its biological function, where even minor changes in its structure can drastically alter the protein's function. Finally, proteins composed of more than one polypeptide chain or subunit have a quaternary structure, varying from simple dimers to large complexes.

1.1.2. Protein Post-Translational Modifications

PTMs are reversible or irreversible covalent modifications to the amino acid sequence of proteins that greatly increase the functional diversity of the proteome, with more than 500 unique PTMs having been identified.[13] Major classes of PTMs include phosphorylation, methylation, acetylation, and oxidative cysteine modifications. These protein modifications are involved in nearly all biological processes and are essential for regulating protein activity, localization, and stability by altering protein structure, function, and interactions with other biomolecules.[13, 14] However, aberrations in PTMs can lead to detrimental changes in protein structure and function, contributing to the onset and progression of diseases such as cardiovascular disease, cancer, and

neurological disorders.[11, 15] Therefore, characterizing protein PTMs in biological systems is crucial for understanding their roles in both physiological and pathophysiological processes.

Protein phosphorylation is a critical and reversible PTM that commonly occurs on serine, threonine, and tyrosine amino acid residues. Phosphorylation happens when protein kinases catalyze the transfer of a phosphate group from adenosine triphosphate (ATP) to the polar side chain of an amino acid residue in a target protein, while protein phosphatases mediate the removal of the phosphate group. Due to the reversible nature of protein phosphorylation, this modification is involved in the regulation of numerous cellular processes including cell signal transduction, cell growth, and apoptosis.[16, 17] Moreover, abnormal protein phosphorylation has been associated with several diseases.[18] For example, hyperphosphorylation of the neuronal microtubule-associated protein tau has been found to be a pathological PTM in Alzheimer's disease that can lead to degeneration of neurons in the brain.[19] Thus, altered protein phosphorylation states may serve as promising diagnostic, prognostic, and predictive disease biomarkers.

Methylation and acetylation are two other commonly detected PTMs. Methylation involves the addition of a methyl group from S-adenosylmethionine, primarily targeting the side chains of lysine and arginine residues, though alanine, cysteine, histidine, and asparagine residues can also be modified. Essential functional processes such as protein-protein interactions, transcriptional regulation, and protein stability can be regulated by protein methylation.[15] In contrast, acetylation involves the transfer of an acetyl group from acetyl-coA to lysine, serine, or threonine residues. Notably, acetylation occurs at lysine residues on the N-terminus of approximately 90% of human proteins.[20] Acetylation modulates many diverse biological processes, including signal transduction, cell metabolism, and DNA damage repair.[15, 21] For example, histone acetylation is an

important epigenetic modification that can alter chromatin structure to regulate gene expression.[22, 23]

Redox reactions involving reactive protein cysteine thiols can lead to various reversible or irreversible PTMs, induced by reactive oxygen species (ROS) or reactive nitrogen species (RNS).[24] Irreversible oxidative cysteine PTMs, such as the formation of sulfinic and sulfonic acids, are often indicative of oxidative stress.[25] Conversely, reversible oxidative cysteine PTMs, including nitrosylation, glutathionylation, and sulfenylation, play key roles in dynamic processes including redox signaling, cellular homeostasis, and enzymatic catalysis.[26] Aberrant reversible oxidative cysteine PTMs can also serve as indicators of oxidative stress and disease.[25, 27] Therefore, there is growing interest in characterizing oxidative cysteine modifications to better understand their functions within endogenous biological systems. However, the low abundance, labile nature, and diversity of these PTMs make their analysis challenging.

1.2 Top-Down MS-Based Proteomics Approaches for Characterization of Intact Proteins and Protein Complexes

Mass spectrometry (MS)-based proteomics has emerged as a valuable technology for the characterization of proteins, their proteoforms, and protein complexes in biological systems. Traditional MS-based proteomics has been carried out using the bottom-up proteomics (BUP) approach, where proteins are enzymatically digested into peptides, separated by liquid chromatography (LC), and analyzed by MS and tandem MS (MS/MS), providing protein identification, expression-level quantification, and PTM information.[28] While BUP is commonly used due to peptides being more soluble and easier to separate, this approach faces many challenges including limited sequence coverage from incomplete peptide recovery after enzymatic

digestion and loss of proteoform information due to the "peptide-to-protein" inference problem.[4, 8, 29]

Top-down proteomics (TDP) is a powerful approach for the identification, quantification, and characterization of proteoforms.[11, 30-34] In contrast to BUP, TDP analyzes intact proteins without digestion, preserving PTMs and protein structure. Top-down MS (TDMS) involves accurate intact molecular mass measurements and the fragmentation of whole proteoforms in the gas-phase using MS/MS to facilitate sequence analysis and localization of PTMs (**Figure 1.1**).[35-37] Notably, TDMS can be performed in either denaturing or native conditions. In denatured TDMS (dTDMS), proteins are prepared and electrosprayed under denaturing solution conditions typically using aqueous-organic solvent compositions with added acidic modifiers, which induce protein unfolding.[38] As a result, dTDMS is primarily performed to analyze protein primary sequence, quantify proteoforms, and site-localize PTMs (**Figure 1.1a.**).[39-42] While dTDMS can provide unique insights into the precise primary structure of a protein, it is also critical to characterize proteins and their non-covalently bound functional complexes with other proteins and/or ligands.[43] In native TDMS (nTDMS), proteins and their non-covalent complexes are electrosprayed under non-denaturing solution conditions, typically using the volatile buffer ammonium acetate.[38] The application of nTDMS offers valuable insights into protein sequence, PTMs and ligand binding sites (**Figure 1.1b**).[44-48] Additionally, nTDMS can elucidate the higher-order structure of proteins and protein complexes, providing information on solvent accessibility, subunit stoichiometry, and conformational dynamics (**Figure 1.1b**).[31, 49-52] Thus, nTDMS has emerged as an invaluable structural biology tool for the characterization of proteins and their non-covalent complexes under near-physiological conditions.

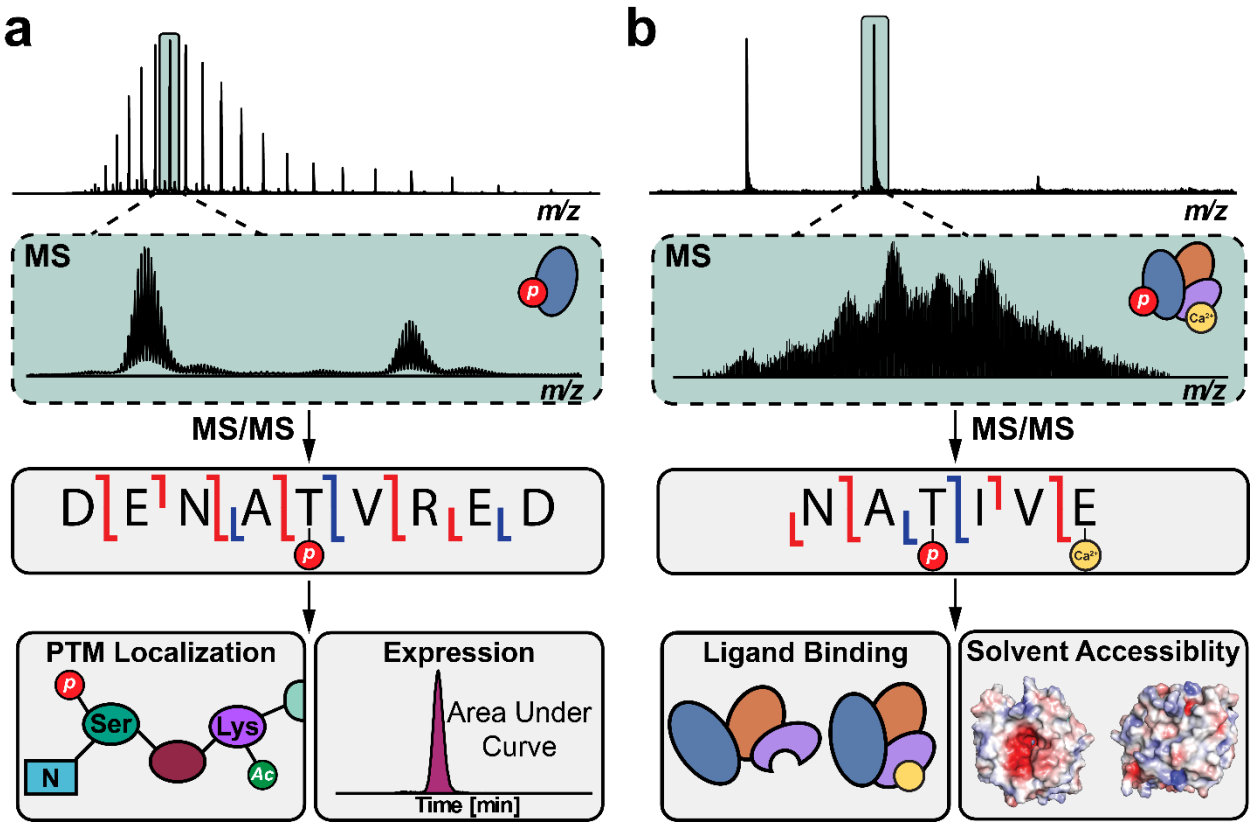


Figure 1.1. Schematic illustration comparing denatured and native top-down MS-based proteomics. **(a)** Denatured top-down MS-based proteomics analyzes intact proteins and their proteoforms to characterize protein primary sequence, localize PTMs, and quantify protein expression. **(b)** Native top-down MS-based proteomics analyzes intact proteins, their proteoforms, and non-covalent complexes in non-denaturing solution conditions to characterize PTMs, ligand binding sites and higher-order structure.

There are still many inherent challenges that TDP faces, including protein solubility, dynamic range, proteome complexity, analytical throughput, and intact protein data analysis.[30, 53, 54] For example, the proteome's high dynamic range (10^8-10^{12}) limits the detection of low-abundance proteoforms in complex biological matrices.[55] Moreover, the complexity of the proteome makes the analysis of large proteoforms (>30 kDa) or heavily modified proteins extremely difficult, often requiring separation of intact proteins and protein complexes prior to MS analysis.[30, 56-58] Nonetheless, TDP has become indispensable for various applications including global proteoform discovery in complex biological samples, biomedical research to uncover novel

disease biomarkers, and biopharmaceutical drug development for the characterization of protein-based biotherapeutics such as monoclonal antibodies.[34] Undoubtedly, continuous improvements in sensitivity, sample preparation, separation, and bioinformatics for TDP will deepen our understanding of the molecular mechanisms underlying diseases and advance the translation of TDP into the clinic.

1.2.1. Sample Preparation for Top-Down Proteomics

Sample preparation is an essential step for achieving reliable TDP data. Traditionally, sample preparation methods (e.g. homogenization and sonication) often use high concentrations of salts (>100 mM) and surfactants such as Triton X-100 and sodium dodecyl sulfate (SDS) to effectively extract and solubilize proteins.[33, 54] However, most salts and common surfactants cause signal suppression and compete with proteins for ionization, leading to the formation of adducted species during electrospray ionization (ESI)-MS analysis, and are thus MS-incompatible.[38] For example, SDS can cause complete signal suppression at levels as low as 0.01%.[59] Therefore, removal of MS-incompatible salts and surfactants is necessary prior to MS analysis.

To prepare protein samples for TDP analysis, incompatible salts and other impurities are commonly removed or exchanged into MS-compatible conditions using molecular weight cut-off (MWCO) ultrafiltration spin filters, size exclusion chromatography (SEC), or reversed-phase liquid chromatography (RPLC).[33, 34] Additionally, to remove surfactants used for solubilization of hydrophobic membrane proteins, protein precipitation is usually performed with chloroform/methanol or cold acetone.[60-62] However, protein precipitation may result in protein loss, incomplete removal of impurities, and sample variability.[63] Therefore, advances in the development of MS-compatible surfactants have been crucial for TDP. Specifically, a MS-compatible photocleavable anionic surfactant, 4-hexylphenylazosulfonate (Azo), has been

developed for TDP applications.[64] Azo can be easily synthesized and used as a replacement for SDS, enabling the solubilization of membrane proteins for the comprehensive characterization of PTMs by TDMS. Similarly, a nonionic, cleavable MS-compatible surfactant, n-decyl-disulfide-β-D-maltoside (DSSM) was developed to mimic n-dodecyl-β-D-maltoside (DDM), for TDP characterization of membrane proteins from cell lysates.[65] Furthermore, to prepare protein samples for TDP analysis, the introduction of artifactual proteoforms must be minimized. This can be achieved by adding protease and phosphatase inhibitors to buffers to prevent protein dephosphorylation and degradation, while also handling samples at 4°C to reduce artifactual oxidation of proteins.[66]

Particular attention must be given during native TDP sample preparation to preserve the higher-order structure of proteins and their non-covalent complexes. In nTDMS, the volatile salt ammonium acetate is commonly used to replace non-volatile salts, where ammonium can transfer a proton to either a protein or to acetate, resulting in the formation of gaseous ammonium and unadducted protein ions.[67] Samples can be further desalted using offline SEC spin filters or online SEC to separate proteins and protein complexes from MS-incompatible salts and small molecules. Meanwhile, membrane protein complexes must be extracted from their native environment using surfactants, such as DDM or oligoglycerol detergents (OGDs), which dissolve biological membranes and form soluble micelles by shielding the hydrophobic surfaces of proteins from water.[68, 69] The entire protein-micelle complex is then ionized, and the protein is released from the surfactant micelle in the mass spectrometer via in-source collisional activation.[46]

The high dynamic range inherent within the proteome complicates the detection of low-abundance proteins in complex biological matrices, often requiring front-end fractionation or enrichment strategies prior to TDMS analysis. Organelle fractionation (i.e. nuclear, cytosolic,

membrane, and mitochondrial) is an important strategy for enriching and analyzing subproteomes, enabling more targeted proteomics analysis.[70] There are three techniques by which subcellular fractionation can be performed including sequential extraction, differential centrifugation, and affinity capture. The separation of a complex biological sample into subproteomes offers two key advantages: more in-depth and comprehensive characterization of the subproteomes and the ability to gain additional insights into protein expression and PTMs within individual subcellular compartments.[71] For example, subcellular fractionation of whole cell lysate from the human cell line H1299 was used in a large-scale TDP study, which resulted in the identification of 347 mitochondrial proteins, many of which exhibited PTMs or variations in primary structure, thereby providing substantial coverage of the low molecular weight (MW) human proteome.[72]

Affinity purification (AP) is a widely used approach for the targeted isolation and purification of intact proteins and protein complexes, with antibodies being the most commonly used affinity reagents in biological research.[35, 73, 74] For example, a recent study using a stationary-phase-dissolvable native AP (SNAP) was developed to purify protein complexes for subsequent nTDMS analysis.[75] However, antibody-based methods are limited by batch-to-batch variability, high cost, poor chemical stability, and inability to target specific proteoforms.[76, 77] To address this problem, surface-functionalized superparamagnetic nanoparticles with a peptide affinity ligand (NP-Pep) were developed for denatured TDP analysis of cardiac troponin I (cTnI), a gold-standard cardiac biomarker, directly from serum.[78] The NP-Pep enrichment platform enabled highly sensitive and comprehensive proteoform-resolved cTnI analysis, while depleting highly abundant proteins such as human serum albumin. More recently, this NP-Pep method has been applied to structurally characterize the endogenous cardiac troponin (cTn) complex directly from human cardiac tissues using nTDMS, providing insights into cTn-calcium (Ca^{2+}) binding dynamics,

complex stoichiometry, and high-resolution mapping of the proteoform landscape.[79] Conceivably, this NP platform could be readily adapted to enrich other low-abundance protein targets when functionalized with the appropriate affinity ligand.

1.2.2. Intact Protein Separation

Major challenges for TDP include achieving in-depth proteome coverage and analyzing large proteoforms (>30 kDa), primarily due to the wide dynamic range of the proteome and the exponential decay in the signal-to-noise (S/N) ratio as protein MW increases.[56, 80] Thus, TDP workflows often necessitate high-resolution separation of intact proteins prior to TDMS analysis. LC-based methods, such as RPLC, SEC, ion-exchange chromatography (IEC), and hydrophobic interaction chromatography (HIC), are advantageous for TDP due to their direct compatibility with MS.

RPLC is the primary method employed in denatured TDP for online desalting, offline fractionation, and high-resolution separation prior to dTDMS analysis. In RPLC, denatured proteins are separated based on hydrophobic interactions using a nonpolar stationary phase and polar mobile phase, then eluted into the mass spectrometer using increasing concentrations of organic solvent.[34] As a result, the most polar proteins elute first, while the most hydrophobic proteins elute last. Although RPLC is commonly used for MS analysis of intact proteins, generally only the most abundant proteins and their associated proteoforms are separated, often failing to separate proteins larger than 50 kDa in complex biological samples.[56] Therefore, efforts to enhance RPLC separation of complex biological samples have focused on reducing column particle size and increasing column pressure and length.[81, 82]

SEC separates proteins based on size, with smaller proteins able to penetrate the pores of the stationary phase, while larger proteins do not enter the pores and elute first. There are several advantages with using SEC, including its high tolerance with a wide range of mobile phases in both denaturing and native conditions, preservation of biological activity of proteins, and reduced sample loss due to minimal interaction between the proteins and the stationary phase.[83] SEC often requires large sample amounts and is considered a low-resolution separation technique; however, SEC has still proven to be highly useful in TDP applicaitons.[56, 83-85] For example, SEC has been successfully used for rapid online buffer exchange (OBE), effectively separating protein complexes from non-volatile salts within a short amount of time, followed by subsequent native MS analysis.[86]

HIC and IEC are non-denaturing separation methods that can be directly coupled to MS, with both methods requiring high concentrations of salt in the mobile phase. In HIC, proteins are separated based on hydrophobicity using a decreasing salt concentration gradient with a stationary phase that has hydrophobic ligands (pentyl, hexyl, heptyl groups) attached to a hydrophilic underlayer.[87, 88] Bound proteins are typically eluted from the stationary phase from low to high hydrophobicity. By using volatile ammonium acetate as the salt in the mobile phase, HIC can be directly coupled online to MS, preserving native protein structure and enabling the analysis of monoclonal antibodies.[89] In contrast, IEC separates proteins based on surface charge characteristics using an increasing salt concentration gradient and charged stationary phase.[90, 91] A negatively charged stationary phase (cation exchange) will separate positively charged proteins, whereas a positively charged stationary phase (anion exchange), will separate negatively charged proteins. Recently, the use of online mixed-bed (containing both anion- and cation-exchange

material) IEC for nTDMS analysis of complex endogenous samples was reported, identifying endogenous proteins and protein complexes from human cardiac tissue up to 146 kDa.[92]

Multi-dimensional liquid chromatography (MDLC) is a powerful approach for overcoming the complexity of the diverse proteome, as it combines multiple orthogonal separation techniques to resolve proteins with diverse physicochemical properties to ultimately provide deeper proteome coverage.[30, 93] A strategy for TDP is the use of two-dimensional (2D) LC coupled to MS in an offline approach, where fractions are first collected from the first dimension and then analyzed in the second dimension that is interfaced directly with MS.[56, 85, 94, 95] For example, a novel 2D LC strategy combining HIC (first dimension) and RPLC (second dimension) was used to achieve high-resolution intact protein separation of a complex cell lysate.[95] Moreover, offline serial SEC (sSEC) which combines columns with varying pore sizes connected in series, has been coupled with RPLC-MS/MS to enable high-resolution size-based fractionation and detection of proteins up to 223 kDa.[56] Similarly, small-scale sSEC (s^3SEC) coupled to RPLC-MS/MS was developed to facilitate size-based separation of proteins extracted from minimal amounts of sample (1 mg tissue).[85] While 2D LC strategies can provide deeper coverage into the proteome, the use of additional dimensions of separation may hold promise for reducing the complexity of the proteome even further. Expanding on this concept, a 3D LC strategy combining HIC, IEC, and RPLC-MS was demonstrated, achieving a 14-fold increase in protein identifications compared to 2D approaches.[96] Although multi-dimensional separation techniques have advanced TDP, offline protein fractionation still remains time consuming, low-throughput, and requires relatively large sample amounts. Therefore, to achieve comprehensive proteome coverage, there remains a clear need for the further development of online MDLC strategies. Techniques such as multiple heart-cutting, active solvent modulation, and stationary-phase assisted modulation will likely be

essential for attaining deeper proteome coverage, particularly for non-denatured proteins and protein complexes.[97-99]

1.2.3. Instrumentation for Top-Down Proteomics

The main components of a mass spectrometer are the inlet system, ionization source, mass analyzer, and detector. Soft ionization techniques such as ESI and matrix-assisted laser desorption/ionization (MALDI) enables the transfer of biological analytes from condensed phase to gas phase, followed by their ionization.[100-102] ESI is advantageous over MALDI for intact protein analysis due to the production of multiply-charged ions, which lowers the mass-to-charge ratio (*m/z*) of the analyte to fall within the mass range of common mass analyzers.[103] Furthermore, these multiply-charged analytes can be efficiently fragmented in the gas phase using both energy- and electron-based fragmentation techniques. The development of nano-ESI, which reduces the emitter orifice to generate finer droplets for introduction into to the MS, has further expanded the application of ESI in TDP studies.[104] Nano-ESI offers several advantages, including minimal sample consumption and improved analytical sensitivity.[105, 106]

There are several instrumentation platforms with high-resolution mass analyzers available for TDP analysis. One of the most widely used mass analyzers in TDP is the quadrupole time-of-flight (QTOF) mass spectrometer, which is ideal for analyzing large proteins and protein complexes due to the nearly unlimited *m/z* range of the TOF and the quadrupole's ability to function as a mass filter or collision cell.[107, 108] For example, a QTOF mass spectrometer with a modified ESI source was used for native MS analysis of membrane protein complexes ranging from 127 to 232 kDa, as well as for the analysis of the 801 kDa chaperone GroEL protein.[109] More recently, the introduction of a hybrid trapped ion mobility QTOF (timsTOF) mass spectrometer, which positions ions in an electric field against a moving buffer gas, allows for the determination

of ion mobility values for structural analysis of native proteins and protein complexes.[41, 79, 110, 111] From these measured mobility values, the collision cross section (CCS) can be calculated using the Mason-Schamp equation, providing insights into protein shape and conformation.

The Fourier-transform ion cyclotron resonance (FTICR) and Orbitrap mass analyzers are used for their high-resolving power, which provides accurate mass measurements of intact proteins and protein complexes. FTICR MS has significantly advanced denatured and native TDP for characterizing proteoforms and their complexes in biological systems due to its ultrahigh resolving power, mass accuracy, and capability to perform various electron- and energy-based fragmentation techniques.[112, 113] The slower scan speed of FTICR MS does limit its efficiency for high-throughput LC-MS/MS analysis.[54] However, direct infusion FTICR MS workflows present a promising strategy for the rapid, in-depth characterization of proteoforms directly from complex biological samples.[114] Similarly, Orbitrap mass spectrometers with ultrahigh mass range (UHMR) have enabled TDP analysis of high MW proteins and protein complexes approaching one megadalton.[115] Additionally, the Orbitrap mass spectrometer can be coupled with LC for global proteoform analysis, as well as for measuring the binding of ligands to protein complexes.[116, 117]

1.2.4. Top-Down Tandem MS

TDP provides a comprehensive approach to characterize protein primary sequence, modifications, as well as the higher-order structure of protein complexes. In a MS/MS experiment, specific proteoforms of interest are first isolated and then fragmented using various fragmentation techniques to achieve backbone cleavages and generate sequence-informative fragment ions. Energy-based fragmentation techniques such as collisionally activated dissociation (CAD) and higher-energy collisional dissociation (HCD)[118] are the most widely used and robust techniques for TDMS analysis.[33, 54] In these techniques, selected precursor protein ions are accelerated to

collide with an inert gas, such as nitrogen, generating *b*- and *y*-type fragment ions from the cleavage of C-N amide bonds along the protein backbone. CAD and HCD are commonly employed in dTDMS studies to localize PTMs and fully characterize the sequence of a protein.[40, 58, 72, 119] However, because these methods tend to cleave the most labile bonds, they may inadvertently cleave labile PTMs. For nTDMS analysis of protein complexes, CAD was initially thought to induce unfolding and subunit dissociation, making this fragmentation technique unsuitable for structural characterization.[120] However, recent native TDP studies have demonstrated that CAD can provide valuable sequence information directly from native protein complexes, offering insights into non-covalent metal-binding and higher-order structural details.[79, 121]

Electron-based fragmentation techniques such as electron capture dissociation (ECD)[122] and electron transfer dissociation (ETD),[123] have become important alternatives to collision-based techniques, particularly for their ability to preserve labile PTMs. ECD and ETD induce non-specific cleavage of N-C_α bonds in the protein backbone, generating *c*- and $z^\bullet$-type fragment ions. ECD has been extensively used in dTDMS studies for mapping labile PTMs such as phosphorylation.[35, 74, 124-126] For example, ECD was used to localize phosphorylation sites on cTnI to Ser22/23 and determined the order of phosphorylation/desphosphorylation between these sites in normal and diseased heart tissues.[35] In addition to site-localizing PTMs, ECD has proven highly effective for the structural elucidation of proteins and protein complexes in nTDMS studies, characterizing subunit binding interactions, solvent accessibility, and non-covalent metal binding.[31, 127-130] While ECD and ETD fragmentation techniques offer numerous advantages for proteoform characterization, their dependence on charge state can limit fragmentation efficiency, particularly for lower charge states of native proteins and protein complexes.[131]

The advent of ultraviolet photodissociation (UVPD) using 193 nm photons has extended the utility of the TDP approach for characterization of intact proteins and protein complexes even further.[132, 133] Although the mechanism of UVPD is still not yet fully understood, extensive backbone cleavages occur when chromophores in the peptide backbone absorb high-energy photons and become electronically excited, leading to the generation of *b*-, *y*-, *c*-, and $z^{\bullet}$-type fragment ions, along with *a*-/*x*-type fragment ions from the cleavage of C_{α}-C bonds. Furthermore, unlike ECD and ETD, UVPD's fragmentation efficiency is independent of the charge state of the precursor ion. As a result, UVPD enables comprehensive characterization of protein sequences and the identification and localization PTMs in TDP analyses.[134]

1.2.5. Top-Down Proteoform Quantification

Quantitative analysis of proteoforms using TDP can reveal significant changes in protein expression and PTMs to uncover disease-associated proteoform alterations and identify potential biomarkers.[135] Three main quantitative approaches have been developed for TDP analysis: metabolic labeling, chemical labeling, and label-free quantification.[34, 136] Metabolic labeling involves differential isotopic labeling of proteins in vitro for the relative quantification of proteoforms expressed by cells cultured under various conditions.[136] Methods such as stable isotope labeling by amino acids (SILAC) and neutron encoding (NeuCode) SILAC, have shown great potential for TDP.[137-139] In contrast, chemical labeling uses isotopically labeled chemical tags, such as tandem mass tags, which covalently bind to specific amino acid residues, allowing for the quantification of intact proteoforms, typically at the MS/MS level.[136] Label-free quantification, which does not rely on isotopic labels or chemical tags, is the most widely used method in TDP for the relative quantification of proteoform intensity between LC-MS runs in complex biological samples.[35, 39, 40, 136, 140-142] Label-free quantification of intact proteoforms typically occurs at the

MS level by averaging across specific chromatographic peaks to identify proteoforms based on their unique charge state distributions and *m/z* ion intensities in the mass spectrum. The intensity of the proteoform of interest is then calculated by combining the ion intensities of multiple charge states using deconvolution algorithms, followed by statistical analysis to compare different groups. Recently, a label-free TDP platform was developed to simultaneously quantify protein expression and PTMs in complex biological samples.[143] In this platform, protein expression levels were determined by integrating the area under the curve (AUC) of extracted ion chromatograms (EICs), while the relative quantification of PTMs was based on the deconvoluted averaged mass spectra. Furthermore, this method was applied to simultaneously quantify sarcomeric proteoforms with high reproducibility, revealing molecular changes in PTMs and protein expression-levels that correlated with left ventricular (LV) dysfunction in ischemic cardiomyopathy (ICM) patients.[144]

1.2.6. Data Analysis

A fundamental aspect of TDP is the identification of MS/MS fragment ions produced from intact proteoforms, which often results in highly complex spectra with overlapping charge states.[145, 146] TDP data analysis typically involves spectral deconvolution, database searching, validation, and visualization.[33] Given the inherent complexity of high-resolution MS/MS spectra, advanced software tools are critical for deconvoluting spectral data and ensuring accurate database searching for proteoform identification. Deconvolution is crucial in TDP data analysis for simplifying complex spectra by transforming an isotope and charge state distribution into a single monoisotopic mass.[34] Several deconvolution strategies have been developed for TDP data analysis including eTHRASH,[147] TopFD,[148] pParseTD,[149] MSDeconv,[145] and SNAP. However, determining charge states for large MW proteins and protein complexes, particularly those lacking isotopic resolution in complex native mass spectra with overlapping charge states, poses a significant

challenge. To address this, the UniDec algorithm, which utilizes Bayesian deconvolution, effectively assigns charge states to generate a zero-charge mass distribution.[150] Additionally, UniDec offers tools for identifying potential PTMs and ligand-binding sites, making this deconvolution algorithm highly valuable for interpreting native mass spectra.[151] Database search algorithms for proteoform identification include MS-Align,[152] TopPIC,[148] Proteoform Suite,[153] ProSight PC,[154] and Mascot Top Down.[155] However, many of these spectral deconvolution and database search algorithms lack a user-friendly interface that is easily accessible. Therefore, MASH Native, a universal software that integrates multiple spectral deconvolution and database search algorithms into a single, user-friendly interface, was developed to process denatured and native TDP data.[156]

1.3 Applications of Top-Down Proteomics to Cardiovascular Disease

Cardiovascular disease (CVD) remains the leading cause of morbidity and mortality worldwide, imposing significant economic burdens.[157] CVD encompasses a broad range of diseases that contribute to or are associated with heart failure (HF), including cardiomyopathies, acute myocardial infarction (AMI), and atrial fibrillation. The etiology of CVD is highly heterogenous, with underlying causes varying significantly from patient to patient.[158, 159] Thus, understanding the molecular mechanisms underlying CVD is crucial for identifying novel disease biomarkers and developing therapeutic interventions.

The sarcomere is the basic contractile unit of the heart, composed of structural building blocks of the myocardium, including actin-based thin and myosin-based thick filaments, which are laterally bordered by the Z-disk (**Figure 1.2**).[160-162] Fluctuations in intracellular Ca^{2+} concentration induce conformational changes within the cTn complex and tropomyosin, supporting actin-myosin cross-bridge formation that regulates cardiac contraction and relaxation.[163, 164] The Z-disk protein

network cross-links myofilaments into highly organized 3D structures, stabilizing the sarcomere.[165, 166] Alterations in expression levels, PTMs, and Ca^{2+} handling of sarcomeric proteins play consequential roles in cardiac physiology and pathophysiology.[71, 167, 168] Therefore, to fully understand how these alterations influence CVD mechanisms, a holistic view of the endogenous heart proteoform landscape needs to be obtained.

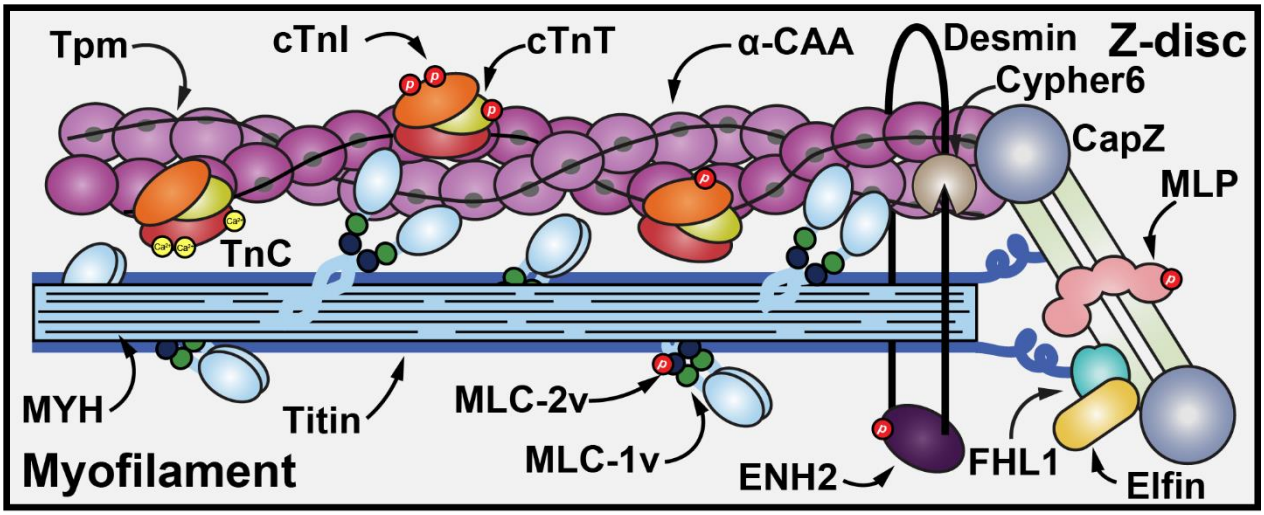


Figure 1.2. Schematic illustration of representative myofilament and Z-disc proteins in the cardiac sarcomere. Myofilament proteins are tropomyosin (Tpm), cardiac troponin I (cTnI), cardiac troponin T (cTnT), troponin C (TnC), alpha-cardiac actin (α-CAA), myosin heavy chain (MYH), titin, ventricular myosin light chain 2 (MLC-2v), and ventricular myosin light chain 1 (MLC-1v). Z-disc proteins are the enigma homolog 2 (ENH2), four-and-a-half LIM domains 1 (FHL1), elfin, muscle LIM protein (MLP), capZ, cypher, and desmin. Red circles represent phosphorylation, while yellow circles represent calcium (Ca^{2+}) binding.

1.3.1. Identification and Characterization of Cardiac Sarcomeric Proteoforms

TDP is uniquely suited to provide a 'bird's eye view' of the cardiac sarcomeric proteoform landscape, capturing biological heterogeneity at the intact proteoform level.[11, 135] By delivering whole-sequence information for intact proteins and their associated proteoforms, TDP enables the comprehensive identification and characterization of sarcomeric protein isoforms, polymorphisms,

and combinatorial PTMs.[35, 126, 169-172] Most notably, dTDMS has been used to comprehensively map all PTMs of cTnI across species, including mouse, rat, pig, and human hearts.[35, 124, 171, 173-175] In human cTnI, high-resolution top-down analysis localized phosphorylation sites to Ser22/Ser23 in heart tissue from non-failing and chronic HF patients.[35] Interestingly, this study found differences in the phosphorylation/dephosphorylation order between the two sites in cTnI in non-failing and diseased human tissues, with Ser22 being phosphorylated prior to Ser23 in non-failing hearts, and Ser23 being dephosphorylated prior to Ser22 in diseased hearts.[35]

While phosphorylation remains the most extensively characterized PTM in sarcomeric proteins, many other important PTMs need to be identified and fully characterized. For instance, Bayne et al. employed dTDMS to identify and localize deamidation of human ventricular myosin light chain 2 (MLC-2v) at Asn13, a modification found to be atrial-specific.[119] In addition, S-glutathionylation (SSG) of ventricular myosin light chain 1 (MLC-1v) was recently identified in human, swine, and mouse cardiac tissues for the first time.[176] Using dTDMS, SSG was localized to Cys66 in human, Cys68 in swine, and Cys84 in mouse cardiac tissues.[176] Together, these studies demonstrate how TDP preserves the combinatorial nature of endogenous PTMs and remains the most effective method for simultaneously characterizing multiple co-occurring PTMs, including those that were previously unknown.

1.3.2. Quantification of Cardiac Sarcomeric Proteoforms

Quantitative analysis of proteoforms using TDP enables the detection of disease-associated alterations in protein expression and PTMs, providing unique insights into the molecular mechanisms underlying physiological and pathophysiological processes in the heart.[135] Phosphorylation of cTnI has been identified as a potential next-generation cardiac biomarker, as altered phosphorylation profiles of cTnI are associated with dysregulated cellular signaling during

the onset and progression of HF.[35] Many previous studies have quantified cTnI phosphorylation in non-failing donor hearts compared to failing hearts using denatured TDP.[35, 39, 78, 144] Across these studies, cTnI has consistently been found to be predominately bis-phosphorylated (*pp*cTnI) in non-failing donor hearts and unphosphorylated in in diseased hearts.

The ability to simultaneously quantify the proteoforms of multiple endogenous sarcomeric proteins opens new possibilities for investigating how the coordinated signaling between myofilament and Z-disc proteins may influence the onset and progression of CVD. Peng et al. provided the first report of concerted changes in the phosphorylation of myofilament and z-disk proteins using dTDMS.[172] A significant, concerted decrease in the phosphorylation of myofilament proteins cTnI, MLC-2, and the z-disc protein enigma homolog 2 (ENH2) was observed in acutely infarcted swine myocardium, suggesting that coordinated signaling between myofilament and z-disc proteins may contribute to contractile dysfunction in AMI.[172] Similarly, a coordinated decrease in the phosphorylation of cTnI and ENH2 was observed in both human HCM[39] and ICM[144], suggesting potential dysregulation of the cAMP-dependent protein kinase A (PKA)-mediated pathway, as both proteins can be phosphorylated by PKA. In addition to quantifying disease-associated proteoforms, TDP can also enable the simultaneous quantification of protein expression.[143] For example, sarcomeric protein expression levels were found to be significantly altered in human ICM tissues compared to non-failing donor tissues, with notable decreases observed in several sarcomeric proteins, including cTnI and ENH2 expression.[144]

The heart is highly heterogenous, exhibiting chamber-specific variations that are essential for regulating physiological and pathophysiological functions. Thus, quantifying chamber-specific variations in protein isoform expression and PTMs is essential for understanding cardiac function, elucidating disease mechanisms, and advancing biomarker discovery. Gregorich et al. investigated

chamber-specific and transmural variations in myofilament protein PTMs from swine cardiac tissues, focusing on the phosphorylation of cTnI, cardiac troponin T (cTnT), alpha-tropomyosin (α-Tpm) and MLC2 using quantitative dTDMS.[177] Notably, minimal transmural differences in the phosphorylation of these myofilament proteins were observed.[177] However, significant chamber-specific variations in the phosphorylation of cTnI and α-Tpm were observed, including lower basal phosphorylation of cTnI and higher phosphorylation of α-Tpm in the atria compared to the ventricles.[177] Chamber-specific variations have also been discovered for the MLC isoforms in adult non-failing human hearts using dTDMS which revealed that MLC-1v, rather than MLC-2v, exhibited ventricle-specific expression, whereas MLC-2a exhibited atria-restricted expression.[119] Overall, TDP offers an unbiased approach to characterizing chamber-specific variations in protein isoform expression and PTMs, emphasizing the importance of precise tissue sampling to address heterogeneity in the heart when investigating disease mechanisms and identifying proteoform-specific biomarkers.

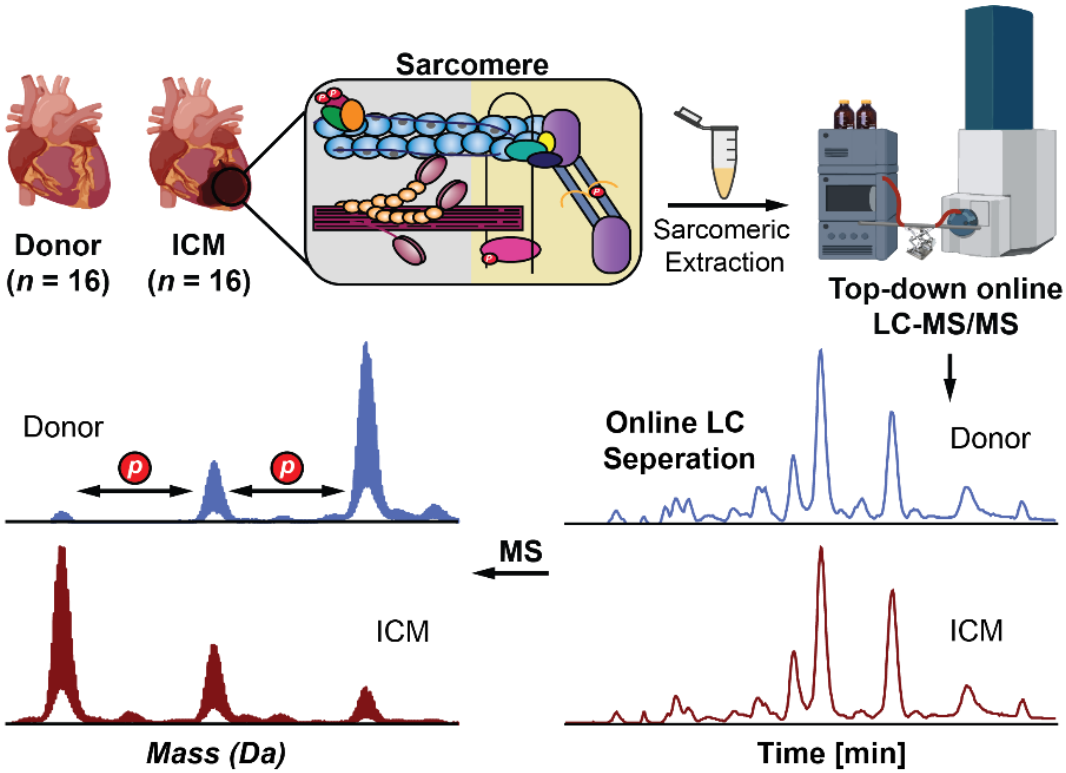
1.3.3. Structural Characterization of Endogenous Cardiac Proteins and Protein Complexes

Comprehensive characterization of the structure, dynamics, and Ca^{2+} binding properties of endogenous sarcomeric proteins and their non-covalent complexes is essential for understanding cardiac function and disease mechanisms. nTDMS has emerged as a powerful structural biology tool, that can localize PTMs, identify Ca^{2+}-binding sites, and elucidate the 3D structure of endogenous sarcomeric proteins and their non-covalent complexes in the gas-phase.[79, 92, 114] A native nanoproteomics platform was recently developed for the enrichment and subsequent nTDMS analysis of the endogenous cTn complex, the master regulator and Ca^{2+} sensitive switch of cardiac muscle.[79] Seventeen endogenous cTn proteoforms were identified directly from human cardiac tissue, while preserving PTMs including N-terminal acetylation and phosphorylation,

along with Ca^{2+} binding.[79] CAD fragmentation of the cTn complex enabled localization of Ca^{2+} binding regions within TnC, even revealing sequential binding of Ca^{2+} ions to the subunit.[79] Native TIMS-MS analysis also provided insights into the conformational changes that occur to the cTn complex upon Ca^{2+} binding during cardiac contraction.[79] nTDMS was further applied to the direct infusion of human cardiac tissue lysate for FTICR MS analysis, enabling the characterization of endogenous sarcomeric proteins, including MLC-2v, MLC-1v, and α-Tpm.[114] This analysis localized the Ca^{2+} binding region in MLC-2v and established a direct link between the structural features of these sarcomeric proteins and their functional roles within the sarcomere.[114] Nonetheless, the intricacies in native sample preparation and the heterogeneous nature of biological samples present significant challenges to studying endogenous proteins and their non-covalent complexes directly from cardiac tissues.[178] To overcome these challenges, further development of non-denaturing chromatographic strategies is needed to advance nTDMS analysis of endogenous proteins and provide more comprehensive insights into the human heart proteome. Fisher et al. recently demonstrated the use of online mixed-bed IEC for nTDMS analysis of endogenous proteins from human heart tissue lysate, detecting metal-binding and oligomeric proteins up to 146 kDa.[92] While this approach provides improved coverage compared to direct infusion nTDMS, MDLC will likely be necessary for achieving even deeper coverage into the human heart proteome.

CHAPTER 2:

Defining the Sarcomeric Proteoform Landscape in Ischemic Cardiomyopathy by Top-down Proteomics



Portions of this chapter are adapted from the following published manuscript:

Chapman, E.A.; Aballo, T.J.; Melby, J.A.; Zhou, T.; Price, S.J.; Rossler, K.J.; Lei, I.; Tang, P.C.; Ge, Y. Defining the Sarcomeric Proteoform Landscape in Ischemic Cardiomyopathy by Top-Down Proteomics. *Journal of Proteome Research* **2023**, *22* (3), 931-941. DOI: https://doi.org/10.1021/acs.jproteome.2c00729. Copyright © 2023, reused with permission from American Chemical Society.

2.1 Abstract

Ischemic cardiomyopathy (ICM) is a prominent form of heart failure but the molecular mechanisms underlying ICM remain relatively understudied due to marked phenotypic heterogeneity. Alterations in post-translational modifications (PTMs) and isoform switches in sarcomeric proteins play important roles in cardiac pathophysiology. Thus, it is essential to define sarcomeric proteoform landscape to better understand ICM. Herein, we have implemented a top-down liquid chromatography (LC)-mass spectrometry (MS)-based proteomics method for the identification and quantification of sarcomeric proteoforms in the myocardia of donors without heart diseases (n = 16) compared to end-stage ICM patients (n = 16). Importantly, quantification of post-translational modifications (PTMs) and expression reveal significant changes in various sarcomeric proteins extracted from ICM tissues. Changes include altered phosphorylation and expression of cardiac troponin I (cTnI) and enigma homolog 2 (ENH2) as well as an increase in muscle LIM protein (MLP) and calsarcin-1 (Cal-1) phosphorylation in ICM hearts. Our results imply that the contractile apparatus of the sarcomere is severely dysregulated during ICM. Thus, this is the first study to uncover significant molecular changes to multiple sarcomeric proteins in the LV myocardia of the end-stage ICM patients using LC-MS-based top-down proteomics.

2.2 Introduction

Heart failure (HF) is the leading cause of death worldwide and produces immense economic costs.[179-182] Ischemic cardiomyopathy (ICM) is a prominent form of HF wherein left ventricular (LV) systolic dysfunction reduces blood flow to the heart, leading to oxygen deprivation and myocardial hypoxia.[183-185] However, the molecular mechanisms underlying ICM remain relatively understudied due to marked phenotypic heterogeneity.

The sarcomere, the basic contractile unit of the heart, contains the structural building blocks for myocardium including actin-based thin and myosin-based thick filaments laterally bordered by the Z-disk.[160-162] In the presence of intracellular calcium, conformational changes within the cardiac troponin complex and tropomyosin induce actin-myosin cross-bridge formation to regulate cardiac contraction.[163, 164] The Z-disk protein network helps cross-link myofilaments into highly organized three-dimensional structures to stabilize the sarcomere.[165, 166] Alterations in post-translational modifications (PTMs) in sarcomeric proteins such as phosphorylation, oxidation, degradation, and isoform switches play important roles in cardiac pathophysiology.[167, 168, 186-188] Thus, it is essential to quantitatively determine the changes in the PTMs and isoforms of sarcomeric proteins to better understand ICM.

Top-down mass spectrometry (MS)-based proteomics is a powerful tool for the analysis and characterization of diverse proteoforms, the myriad number of expressed protein products that arise from polymorphisms, alternative splice variants, and post-translational modifications (PTMs).[6-8, 11, 33] Distinct from the bottom-up MS approach in which intact proteins are enzymatically digested into peptides,[28] top-down MS analyzes intact proteins and labile PTMs. Additionally, tandem top-down MS (MS/MS) can fragment intact proteins to comprehensively characterize proteoforms from complex biological samples.[2, 30, 39, 78, 189] Quantitative analysis of

proteoforms using the top-down approach can reveal significant changes in isoform expression and relative abundance of PTMs across various heart disease phenotypes.[4, 135] We previously detected a decrease in phosphorylation of cardiac troponin I (cTnI) in ischemic human hearts using quantitative top-down MS/MS and affinity purification.[35, 190] Recently, our lab has developed a highly robust top-down LC-MS/MS method that can simultaneously quantify protein expression and PTM and applied to non-human skeletal and cardiac muscle tissues.[143] However, a study characterizing changes in the sarcomeric proteoform landscape of human ICM remains lacking.

Herein, we aimed to uncover sarcomeric proteoform changes in failing heart tissues from patients in end-stage ischemic heart failure and non-failing heart tissues from donors without heart disease to better understand ICM. We have implemented a top-down MS proteomics method for the simultaneous quantification of sarcomeric proteoforms with high reproducibility to reveal the molecular changes protein PTMs and isoforms correlated with LV dysfunction in ICM patients. Our results elucidate significant PTM and expression-level changes in myofilament and Z-disk proteins during ischemic heart failure. Collectively, our results demonstrate the unique advantages of top-down proteomics for defining the sarcomeric proteoform landscape and quantifying protein expression and PTMs in parallel, providing valuable insights into the molecular alterations underlying human ICM.

2.3 Results

2.3.1. Analysis of Sarcomeric Proteins in Donor versus ICM Tissues by Online Top-Down LC-MS/MS

The goal of this project was to reveal the molecular mechanisms underlying ICM using top-down MS-based proteomics. To do so, we implemented an online top-down LC-MS/MS

method to simultaneously quantify sarcomeric protein expression and modification levels between LV tissue from non-failing donor hearts and left ventricular apex from ICM tissues collected during LVAD surgeries (**Figure 2.1**). We have demonstrated that LV and apex tissues in donor hearts (n = 4) have similar sarcomeric proteoform levels (**Figure S2.1**), which allows sarcomeric proteoform comparison between donor LV and ICM apex tissues in this study.

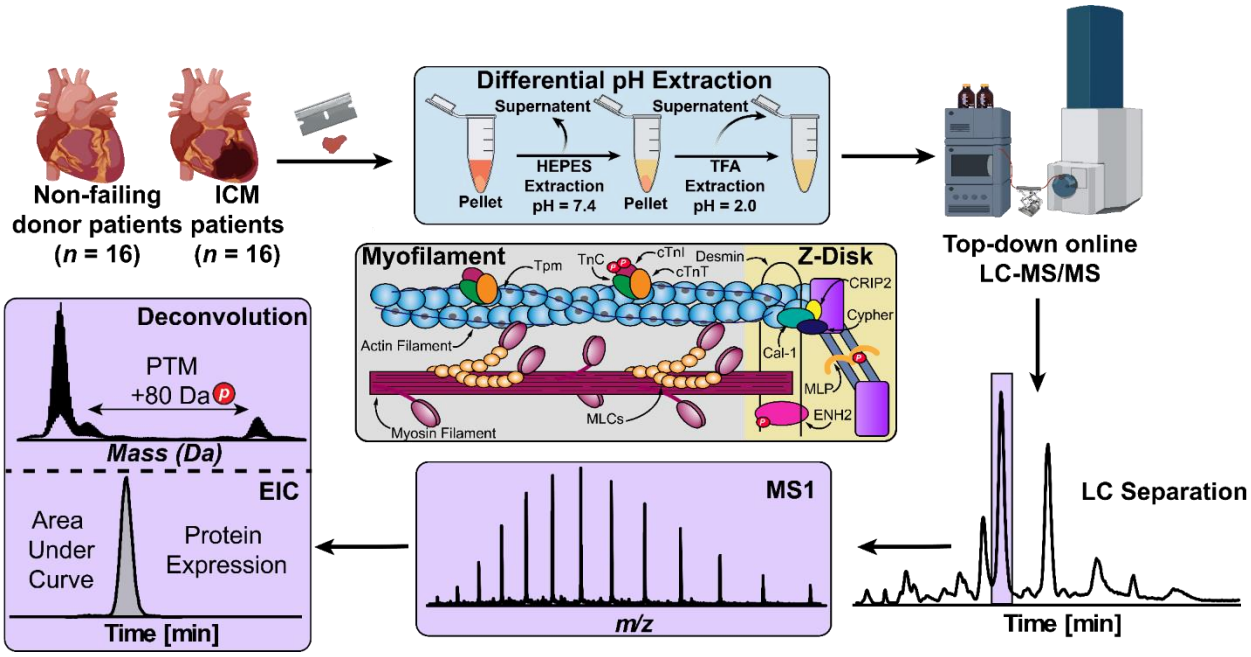


Figure 2.1. Label-free top-down proteomics workflow for the simultaneous quantification of sarcomeric protein expression and modification in ICM. Sarcomeric proteins were extracted from ~10 mg of non-failing donor (n = 16) and ICM (n = 16) human cardiac tissues using a differential pH-based extraction. Tissue was first homogenized in HEPES buffer (pH = 7.4) to remove cytosolic proteins. After centrifugation, the remaining pellets were homogenized in TFA buffer (pH = 2.0) to enrich for sarcomeric proteins. Intact proteins (400 ng) were then separated by online reversed-phase LC and MS data (MS^1) were acquired using a Bruker maXis II quadrupole-time-of-flight (QTOF) mass spectrometer. All data analysis was performed using Bruker DataAnalysis v. 4.3. Protein modifications were quantified by calculating the relative abundance of the most abundant proteoforms from the deconvoluted spectra. Protein expression was quantified by finding the area under the curve (AUC) of the EIC.

The whole procedure including tissue homogenization, sarcomeric protein extraction, and LC-MS/MS analysis can be performed in less than 3 h, demonstrating a fast and high-throughput

method to simultaneously characterize protein expression and PTMs. SDS-PAGE was used to visualize proteins in the HEPES versus TFA extractions (**Figure S2.2**). The results showed that our differential pH-based extraction method successfully depletes cytosolic proteins and enriches sarcomeric proteins in donor and ICM cardiac tissues. Our online top-down LC-MS/MS method revealed a diverse sarcomeric proteoform landscape in donor and ICM tissue samples (**Figure 2.2**).

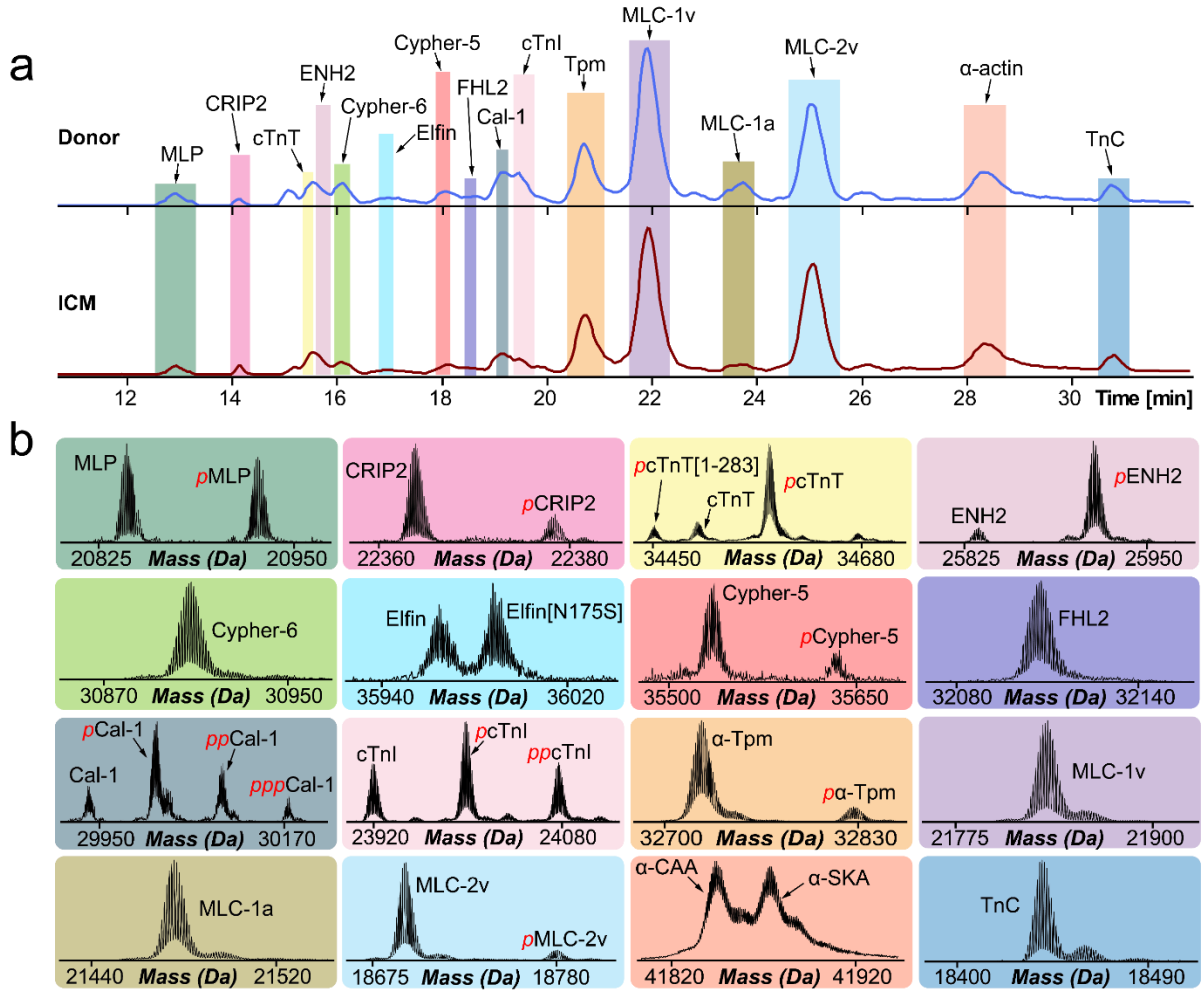


Figure 2.2. Diverse proteoform landscape detected in donor and ICM human cardiac tissues. (a) Representative base peak chromatograms (BPC) showing high-resolution separation of intact sarcomeric and Z-disk proteins muscle lim protein (MLP), cysteine-rich protein 2 (CRIP2), cardiac troponin T (cTnT), enigma homolog 2 (ENH2), cypher-6, elfin, cypher-5, four and a half LIM domain protein 2 (FHL2), calsarcin-1, cardiac troponin I (cTnI), tropomyosin (Tpm), ventricular myosin light chain-1 (MLC-1V),

atrial myosin light chain-1 (MLC-1a), ventricular myosin light chain-2 (MLC-2v), alpha-actin (α-actin), troponin C (TnC) in non-failing donor (blue trace) and ICM (red trace) tissues procured via left ventricular device surgery. **(b)** High-resolution deconvoluted mass spectra highlighting the variety of sarcomeric and Z-disk proteoforms observed. Red *p* and *pp* denote monophosphorylated and bisphosphorylated proteoforms, respectively.

Myofilament proteins identified include cTnI, cardiac troponin T (cTnT), troponin C (TnC), tropomyosin (Tpm) isoforms, α-actin isoforms, ventricular isoform of MLC-2 (MLC-2v), ventricular isoform of MLC-1 (MLC-1v), and atrial isoform of MLC-1 (MLC-1a). We also detected various Z-disk proteins such as enigma homolog 2 (ENH2), muscle LIM protein (MLP), cysteine-rich protein 2 (CRIP2), cypher-5, cypher-6, elfin, calsarcin-1 (Ca1-1), and four and a half LIM domains 2 (FHL2).[191, 192] Sarcomeric proteins were first identified based on their intact protein mass with a mass error tolerance within 4 ppm. A total of 35 human sarcomeric proteoforms were identified compared to 63 human sarcomeric proteins present in Uniprot. Conceivably, it is difficult to identify large proteins due to the limited fragments obtained from online LC-MS/MS data.[39, 56, 80] A full list of sarcomeric proteoforms identified in donor and ICM tissues can be found in **Table S2.1**.

For further protein characterization we employed online LC-MS/MS analysis with collisionally activated dissociation (CAD) fragmentation to generate *b*- and *y*-ions. LC-MS/MS was performed on all sarcomeric proteins detected. Fragment assignments were confirmed in MASH Native software to identify the ventricular isoforms of MLC-1 and MLC-2. The method applies for all sarcomeric proteins detected in the analysis. Uniprot databases were used to map the sequences of human MLC-1v and MLC-2v. Removal of N-terminal methionine and addition of N-terminal trimethylation was validated for human MLC-1v and MLC-2v which is in good agreement with previous literature.[193] The determination of the N-terminal modifications was

further validated by comparison of the theoretical fragment ion masses and the experimental fragment ion masses. Previous studies have reported the localization of mono-phosphorylated MLC-2v to Ser15 or Ser19.[194, 195] This data indicates that Ser19 is a site of MLC-2v phosphorylation due to sequence informative *b* and *y*-ions further demonstrating the power of top-down MS/MS. Total fragmentation of MLC-1v reveals 25 *b*-ions and 23 *y*-ions achieving 23% sequence coverage (**Figure S2.3**). Similarly, total fragmentation of MLC-2v reveals 14 *b*-ions and 22 *y*-ions achieving 21% sequence coverage (**Figure S2.4**). Overall, these data reveal how top-down MS/MS can be used to unambiguously sequence intact myofilament proteins for protein identification and characterization.

2.3.2. Simultaneous Quantification of Sarcomeric Protein Expression and Modification in ICM using Online Top-Down LC-MS/MS Method with High Reproducibility and Linearity

We next assessed the reproducibility and linearity of our top-down online LC-MS/MS method. To ensure our extraction protocol and LC-MS/MS method was reproducible we performed three extraction replicates using donor tissue sample (**Figure S2.5, Figure S2.6**). With the same amount of protein loaded for each biological replicate, overlaid base peak chromatograms (BPCs) and total ion chromatograms (TICs) show highly similar chromatographic profiles and MS signal intensities demonstrating the comparative reproducibility of our method.

To compare protein expression levels between ICM and donor samples using top-down quantitative analysis, we first evaluated the linearity of our instrument response with 250, 400, 500, 600, 750, 1000, and 1200 ng of total protein from the same donor tissue extract injected in triplicate.[143] EICs of representative sarcomeric proteins ENH2, cTnI, α-Tpm, MLC-1v, MLC-2v, and TnC were generated, and a mutual linear range of 250-1200 ng was established by measuring the summed abundance of the AUC of each EIC (**Figure 2.3a-b**). Injection replicates across each

amount of total protein loaded demonstrated excellent reproducibility, sensitivity, and linearity achieving R^2 values greater than 0.99. (**Figure 2.3c, Figure S2.7**). As the instrument responds linearly to differing amounts of total protein loaded, this analysis allows us to be confident in quantitating changes in expression level. Therefore, we can also be confident in quantifying changes in protein abundance between samples.

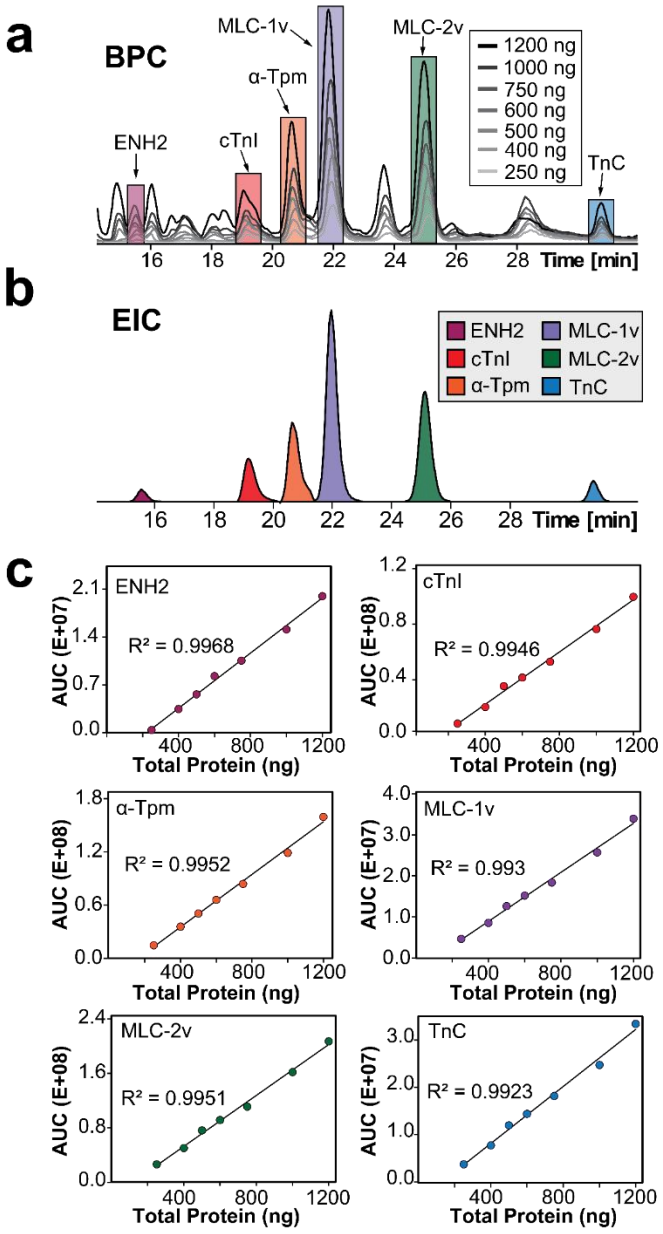


Figure 2.3. Mutual linear range response determination for key sarcomeric proteins. (a) Overlaid BPCs of 250, 400, 500, 600, 750, 1000, and 1200 ng of total protein loaded into the LC-MS demonstrating highly similar proteoform profiles. **(b)** EICs were generated for ENH2, cTnI, α-Tpm, MLC-1v, MLC-2v,

and TnC by combining the top 3-5 most abundant charge states of all proteoforms of the same protein. **(c)** Mutual linear range response determination for key sarcomeric proteins. AUCs of the individual sarcomeric proteins exhibit mutual linear correlation with 250-1200 ng of total protein injected with 3 injection replicates for each point.

2.3.3. Correlated Decrease in ENH2 and cTnI Expression and Modification in ICM Tissues

Significant alterations in PTM and expression level changes in ENH2 and cTnI were identified in ICM compared to non-failing donor samples. Three major cTnI proteoforms were observed in the deconvoluted mass spectra for donor and ICM samples including unphosphorylated cTnI, monophosphorylated cTnI (*p*cTnI), and bisphosphorylated cTnI (*pp*cTnI) (**Figure 2.4a**). Two major ENH2 proteoforms were observed in donor and ICM samples including unphosphorylated ENH2 and monophosphorylated ENH2 (*p*ENH2) (**Figure 2.4a**). We then quantified the expression of cTnI and ENH2 and observed decreased expression levels in ICM compared to donor tissues (**Figure 2.4b**). Total phosphorylation levels were quantified by calculating the relative abundance of proteoforms as their representative percentages among all detected proteoforms in the deconvoluted mass spectra (Ptotal). Total phosphorylation levels of ENH2 in donor tissues were between 0.29 to 1.00 mol Pi/mol protein and 0.10 to 0.80 mol Pi/mol protein in ICM tissues (**Figure 2.4c**). We observed a significant decrease in *p*ENH2, with total phosphorylation of ENH2 decreasing from 85% to 50% in ICM tissues compared to donor tissues. Total phosphorylation levels of cTnI in donor tissues were between 0.55 to 1.98 mol Pi/mol protein and 0.25 to 1.44 mol Pi/mol protein in ICM tissues (**Figure 2.4c**). Notably, we also observed a significant decrease in cTnI phosphorylation, with total phosphorylation of cTnI decreasing from 90% to 56% in ICM tissues compared to donor tissues. To determine if the decrease in phosphorylation of ENH2 and cTnI in ICM tissues were dependent on one another, we plotted a

linear correlation achieving a strong, positive Pearson correlation coefficient (r = 0.8926) and p < 0.00001 (**Figure 2.4d**).

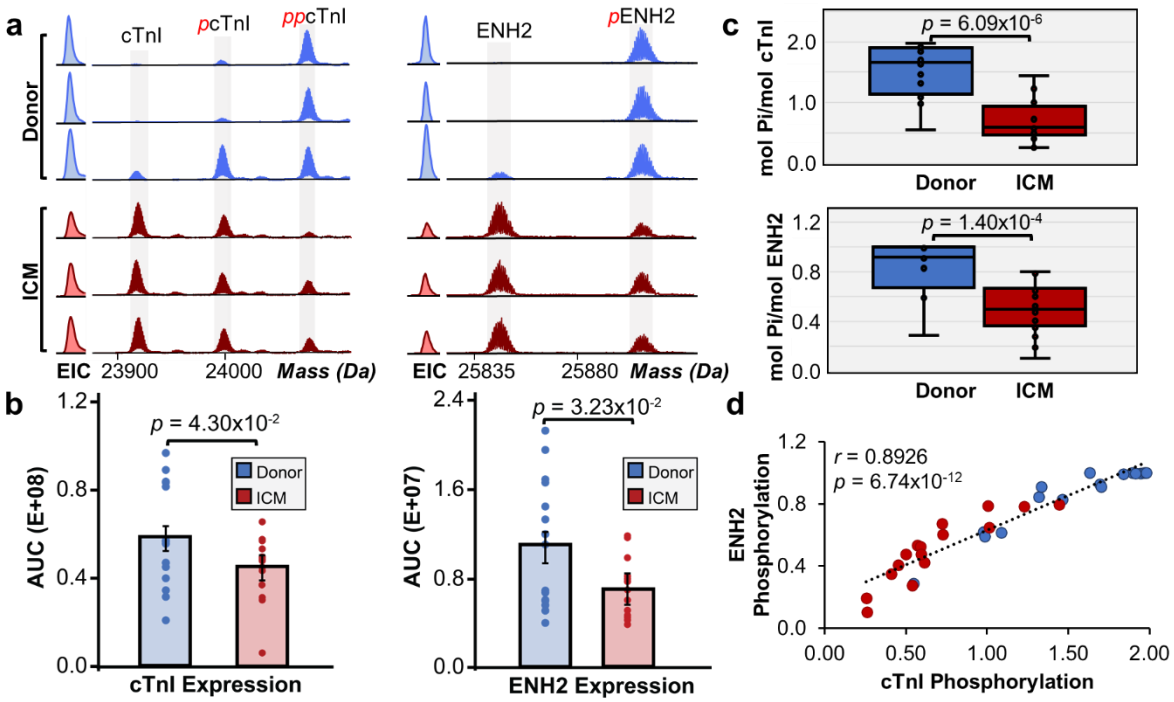


Figure 2.4. Coordinated decrease in cTnI and ENH2 phosphorylation in ICM. (a) Representative high-resolution deconvoluted spectra and EICs of cTnI and ENH2 in non-failing donor (blue) compared to ICM (red) tissues. Red *p* and *pp* denote monophosphorylated and bisphosphorylated proteoforms, respectively. **(b)** Quantification of cTnI and ENH2 expression. Groups were considered significantly different by paired-student *t* tests with $p < 0.05$. **(c)** Total protein phosphorylation calculated by mol pi/mol protein for cTnI and ENH2. Box, median and interquartile range (25%, 75%); whiskers, minimum and maximum values. Horizontal lines represent the median of the group. Groups were considered significantly different by paired-student *t* tests with $p < 0.001$. **(d)** Linear correlation between ENH2 and cTnI phosphorylation. The Pearson correlation coefficient (r = 0.8926) considers the groups strongly correlated and statistically significant with $p < 0.00001$.

These findings agree with a previous study in human hearts with hypertrophic cardiomyopathy (HCM) which implied that the coordinated decrease in ENH2 and cTnI phosphorylation is due to dysregulation of the cAMP-dependent protein kinase A (PKA)-mediated pathway.[39] PKA signaling is essential for cardiac function and plays important roles in cardiac contractility,

metabolism, and homeostasis. Inactivation of this pathway has previously been associated with impaired cardiac contraction and myocardial ischemia.[196, 197] β-blockers remain widely used as antihypertensives to manage cardiovascular symptoms such as high blood pressure and irregular arrhythmia. While β-blockers antagonize β-adrenergic receptors in which the PKA-mediated pathway is activated by, the effect of β-blockade on the PKA-mediated pathway remains incompletely understood.[198, 199] Nonetheless, our previous study investigated the effect of β-blockade on reduced cTnI phosphorylation and found no significant correlation between reduced phosphorylation in cTnI and β-blocker treatments.[35] Thus, the decreased PKA-mediated phosphorylation of cTnI and ENH2 in ICM hearts may be independent of β-blocker treatment.

2.3.4. Differential Isoform Expression of Tpm and α-Actin

Human Tpm is a thin-filament associated protein which is expressed as several isoforms that closely interact with cTnT and α-actin to regulate muscle contraction.[125, 169, 200] *Tpm1* encodes for both α-Tpm and κ-Tpm, whereas *Tpm2* and *Tpm3* encode for β-Tpm and γ-Tpm, respectively.[169, 201] Our previous study confirmed the presence of α-Tpm, κ-Tpm, and β-Tpm in the human heart with α-Tpm being the most abundantly expressed isoform.[169] The major isoforms of Tpm that we detected in this study were unphosphorylated α-Tpm, monophosphorylated α-Tpm (*p*α-Tpm), and monophosphorylated κ-Tpm (*p*κ-Tpm) (**Figure 2.5a**). The low abundance isoforms detected were unphosphorylated κ-Tpm, unphosphorylated γ-Tpm, and unphosphorylated skeletal β-Tpm (skβ-Tpm). We quantified the abundance of each isoform relative to the total abundance of all isoforms by finding the AUC for α-Tpm, κ-Tpm, and skβ-Tpm in donor versus ICM tissues (**Figure 2.5c**). γ-Tpm was too low in abundance and could not be accurately quantified. We observed a significant decrease in skβ-Tpm expression, with total skβ-Tpm expression decreasing from 2.9% to 1.5% of total Tpm abundance in ICM tissues

compared to donor tissues. The difference in isoform expression of α-Tpm and κ-Tpm in donor versus ICM tissues were not significant. The mechanism of how Tpm isoform ratios affect cardiac function remains unclear; however, we also detected a decrease in skβ-Tpm expression in HCM tissues using top-down MS/MS.[39] Thus, our data implies that changes in skβ-Tpm isoform expression in the heart may alter cardiac function and impair systolic function in ICM.

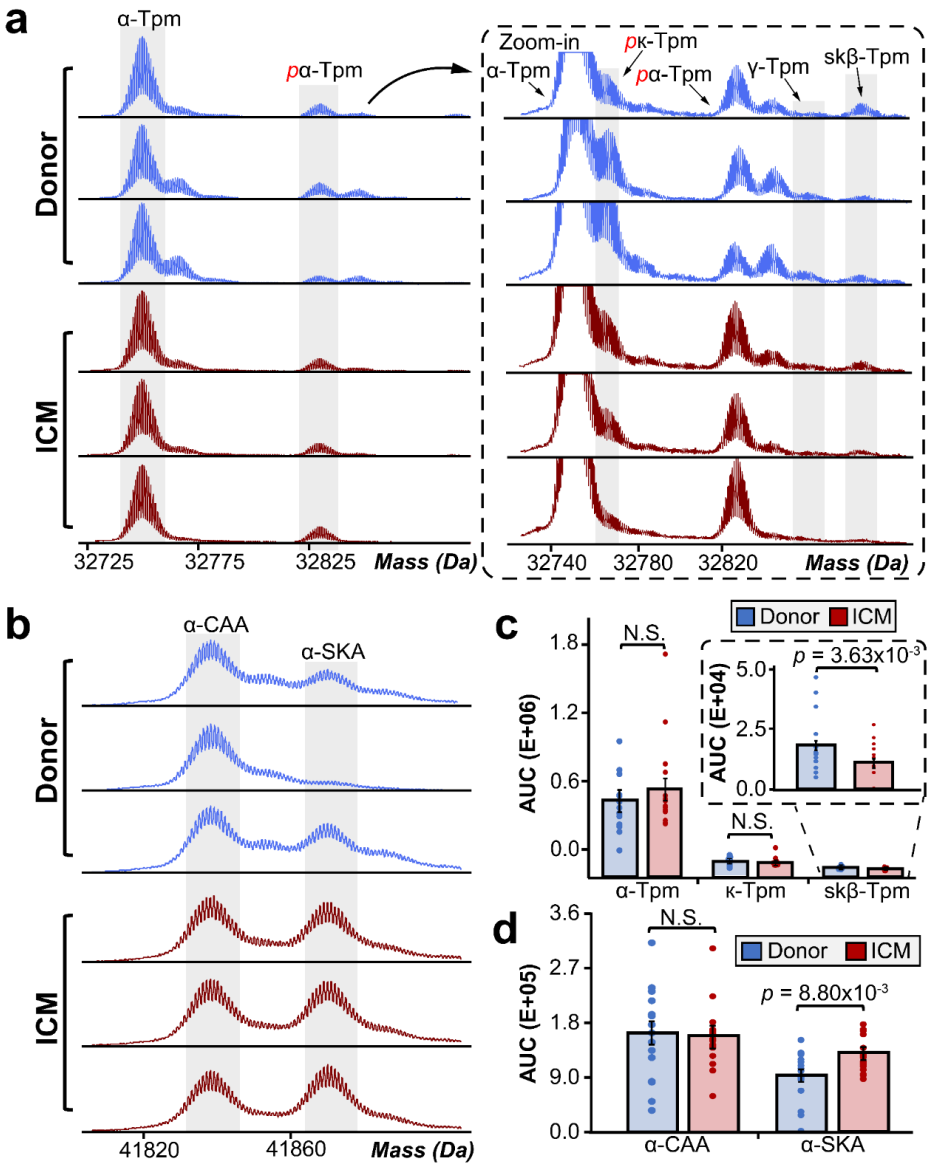


Figure 2.5. Differential isoform expression of tropomyosin and α-actin. (a) Representative high-resolution deconvoluted spectra of tropomyosin in non-failing donor (blue) compared to ICM (red) tissues. We identified four Tpm isoforms including α-Tpm, β-Tpm, κ-Tpm, and γ-Tpm. Red *p* denotes

monophosphorylated proteoforms. **(b)** Representative deconvoluted spectra of α-CAA and α-SKA in non-failing donor (blue) compared to ICM (red) tissues. (c) Quantification of Tpm and α-actin isoforms by finding the AUCs of each respective isoform. Groups were considered significantly different by paired-student *t* tests with $p < 0.005$.

In cardiac and skeletal muscle, *ACTC* and *ACTA* encode for skeletal α-actin (α-SKA) and cardiac α-actin (α-CAA) isoforms, respectively. These two isoforms are co-expressed in myocardium and play roles in sarcomere structure and integrity. Here, we observed both α-actin isoforms in donor and ICM tissues (**Figure 2.5b**). We observed a significant increase in α-SKA expression, with total α-SKA expression increasing from 36% to 45% of total α-actin abundance in ICM tissues compared to donor tissues (**Figure 2.5d**). Previously, our lab has found a significant increase in the relative abundance of α-SKA in non-failing donor hearts compared to failing dilated cardiomyopathy (DCM) hearts.[170] Therefore, these data agree with previous reports that the upregulation of α-SKA expression in failing hearts is a promising biomarker of heart disease.

2.3.5. Increased Phosphorylation of Full-Length cTnT in ICM Tissues

The deconvoluted mass spectra for cTnT showed multiple proteoforms including unphosphorylated cTnT, monophosphorylated cTnT (*p*cTnT), unphosphorylated truncated cTnT with a lysine cleaved from the C-terminus (cTnT [aa 1-286]), and monophosphorylated truncated cTnT (*p*cTnT [aa1-286]) (**Figure S2.8a**). We observed a significant increase in full-length cTnT phosphorylation, with total cTnT phosphorylation increasing from 79% to 86% ICM tissues compared to donor tissues (**Figure S2.8b**). There also appeared to be an increase in total phosphorylation of truncated cTnT [aa 1-286]; however, the change in total cTnT [aa 1-286] phosphorylation between donor and ICM tissues was not significant (**Figure S2.8c**). Similar to cTnI, cTnT can be phosphorylated by a variety of kinases and is a choice biomarker for the detection of cardiac injury due to its cardiac specificity.[202, 203] cTnT mutations are known to be a

common cause of many different cardiomyopathies and PTMs play a vital role in cardiac contractility.[203-205] Interestingly, we observed a homozygous and heterozygous cTnT polymorphism in four ICM and one donor samples (**Figure S2.9**). The polymorphism corresponded to a +28 Da mass shift, indicating a possible alanine (A) to valine (V) point mutation which has previously been reported as a cardiomyopathy variant at amino acid positions 38 or 114.[206, 207] However, being that this point mutation was also present in one donor tissue sample, it is unlikely this polymorphism to be a potential biomarker for ICM.

2.3.6. Increased Phosphorylation of Z-Disk Proteins in ICM Tissues

In addition to myofilament proteins, we also identified and quantified Z-disk proteins such as MLP and calsarcin-1. The deconvoluted mass spectra for MLP showed unphosphorylated MLP and monophosphorylated MLP (*p*MLP) (**Figure 2.6a**). The deconvoluted mass spectra for calsarcin-1 revealed multiple phosphorylated proteoforms including unphosphorylated calsarcin-1, monophosphorylated calsarcin-1 (*p*Cal-1), bisphosphorylated calsarcin-1 (*pp*Cal-1), and trisphosphorylated calsarcin-1 (*ppp*Cal-1) (**Figure 2.6c**). Total phosphorylation levels of MLP in donor tissues were between 0.0 to 0.44 mol Pi/mol protein and 0.02 to 0.66 mol Pi/mol protein in ICM tissues. We observed a significant increase in MLP phosphorylation, with total phosphorylation increasing from 19% to 46% of total MLP abundance in ICM tissues compared to donor tissues (**Figure 2.6b**). We also quantified phosphorylation of calsarcin-1 (**Figure S2.10**). The total phosphorylation levels of calsarcin-1 in donor tissues were between 0.73 to 1.72 mol Pi/mol protein and 0.96 to 1.75 mol Pi/mol protein in ICM tissues. Similar to MLP, we observed a significant increase in calsarcin-1 phosphorylation, with total phosphorylation of calsarcin-1 increasing from 86% to 90% of total calsarcin-1 abundance in ICM tissues compared to donor tissues (**Figure 2.6d**). MLP and calsarcin-1 are known to be involved in severe forms of

cardiomyopathy. Mutations in the *CSRP3* gene which encodes for MLP have been directly involved in the development of HCM and DCM.[208] Calsarcin-1 is a mediator of protein phosphatase activity and deficiency of this protein sensitizes the heart to calcineurin, a protein phosphatase that plays important roles in HCM when dysregulated.[209, 210] Nevertheless, the molecular mechanisms and effect of PTMs on MLP and calsarcin-1 remain relatively unknown. This is the first study to report increased phosphorylation of MLP and calsarcin-1 in ICM. The total phosphorylation of MLP and calsarcin-1 exhibited a negative linear correlation with a Pearson correlation coefficient r = -0.03, indicating that these two phosphorylation's are not correlated which are unlikely to be phosphorylated by the same kinase or closely interact with each other in the z-disk. Despite reports of MLP being significantly down regulated in DCM and ICM,[211] we did not observe a significant change in MLP expression.

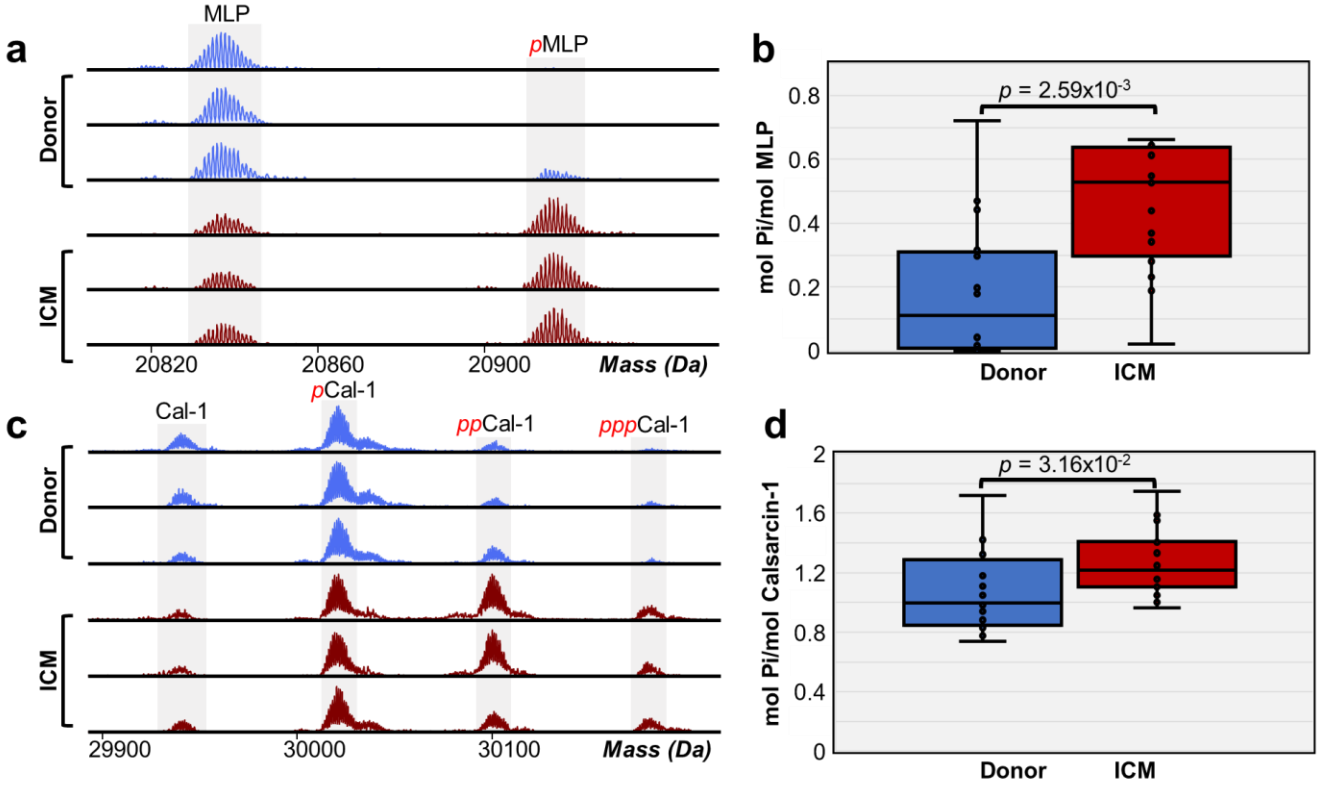


Figure 2.6. Increase in MLP and calsarcin-1 phosphorylation in ICM. (a) Representative high-resolution deconvoluted spectra of MLP in non-failing donor (blue) compared to ICM (red) tissues. Red *p*

denotes monophosphorylated MLP. **(b)** Total protein phosphorylation calculated by mol pi/mol protein for MLP. Box, median and interquartile range (25%, 75%); whiskers, minimum and maximum values. Horizontal lines represent the median of the group. Groups were considered significantly different by paired-student *t* tests with $p < 0.01$. **(c)** Representative high-resolution deconvoluted spectra of calsarcin-1 (Cal-1) in non-failing donor (blue) compared to ICM (red) tissues. Red *p, pp,* and *ppp* denotes mono-, bis-, and tris-phosphorylated proteoforms, respectively. **(d)** Total protein phosphorylation calculated by mol pi/mol protein for calsarcin-1. Box, median and interquartile range (25%, 75%); whiskers, minimum and maximum values. Horizontal lines represent the median of the group. Groups were considered significantly different by paired-student *t* tests with $p < 0.05$.

2.4 Discussion

Overall, in this study, we have focused on sarcomere subproteome and achieved simultaneous quantification sarcomeric protein expression and PTMs from non-failing donor and failing ICM human cardiac tissues. ICM is a highly heterogenous cardiovascular disease that is characterized by significant alterations to the left ventricle of the heart, impacting systolic function and inducing myocardial hypoxia. While ischemic heart disease is the leading cause of death worldwide, there are only very few proteomics studies that have been performed directly using human ICM samples due to the difficulty in obtaining human heart tissue samples.[212, 213] Roselló-Lletí et al. analyzed cardiac protein changes in ICM from human LV tissues using 2-dimensional gel electrophoresis followed by in-gel digestion and bottom-up MS, showing that proteins involved in cellular stress response, respiratory chain, and cardiac metabolism were altered.[212] More recently, Yi et al. quantitatively analyzed the LV proteome in human ICM using bottom-up LC-MS/MS and observed differentially expressed proteins related to metabolism, immune response, muscle contraction, and signal transduction.[213] Nonetheless, bottom-up proteomics suffers from many limitations including the protein interference problem and loss of vital information about PTMs and alternative splice variants due to proteolytic cleave of intact proteins into peptides.[8, 29]

On the other hand, top-down proteomics can directly analyze intact proteins allowing for identification and characterization of proteoforms with full sequence coverage.[6-8, 11, 33] Therefore, top-down proteomics is ideally suited for studying human ICM at the proteoform level to reveal molecular changes within the human sarcomeric proteome.

In this study, we have extracted sarcomeric proteins and utilized top-down online LC-MS/MS for simultaneous quantification of many sarcomeric proteoforms in failing human ICM tissues compared to non-failing donor tissues. Specifically, we discovered a significant decrease in cTnI and ENH2 phosphorylation with a significant increase in MLP, Calsarcin-1, and cTnT phosphorylation in failing ICM tissues compared to non-failing donor tissues (**Table S2.2**). We also found sarcomeric protein expression levels to be significantly altered in ICM tissues compared to non-failing donor tissues with a significant decrease in cTnI, ENH2, and β-Tpm expression and significant increase in α-SKA expression. Previously, our lab has employed a top-down MS combined with affinity purification of cTnI to characterize cTnI proteoforms in ICM and DCM tissues, which revealed that cTnI phosphorylation was greatly reduced in failing ICM and DCM hearts compared to brain-dead donor hearts, implying that phosphorylation of cTnI is a potential biomarker for chronic heart failure.[35] Moreover, in another previous study, Peng et al. generated swine models with acute myocardial infarction (AMI) and identified reduced phosphorylation in cTnI and MLC2 in the myofilaments and ENH2 in the Z-disk.[172] We do observe some similar proteoform changes in swine and human ischemia such as reduced phosphorylation of cTnI and ENH2. On the other hand, we have also observed unique differences only in human ischemia but not in swine ischemia such as significant increases in cTnT, calsarcin-1, and MLP phosphorylation, whereas reduced phosphorylation of MLC2 was observed in swine AMI. Additionally, our lab has previously published an extensive study of the sarcomeric proteoform

landscape in human HCM.[39] We found some similar proteoform changes in ICM and HCM such as decreased phosphorylation of cTnI and ENH2, whereas significant increases in MLP and calsarcin-1 phosphorylation were only observed only in ICM but not in HCM. Since human heart tissue samples in ICM and HCM were obtained from end-stage/late-stage heart failure patients, conceivably they share similarity likely due to the converging pathways and similar molecular processes that leads towards heart failure. Conversely, distinct proteoform changes observed are representative of their respective cardiomyopathies.

One major limitation of this study is the difficulty in detecting the large proteins (>50 kDa) using our online top-down LC-MS/MS approach mainly due to the co-elution of proteins during the 1DLC separation prior to MS/MS analysis and the exponential decrease in the signal-to-noise ratio as protein molecular weight increases.[30, 56] Moreover, it is difficult to identify large proteins due to the limited fragments from online LC-MS/MS data. Our lab has previously developed a top-down 2D-LC method using offline serial size exclusion chromatography (sSEC) to separate high MW proteins up to 223 kDa and coupling with top-down MS for detection of large proteoforms.[56] While the 2DLC sSEC method allows for high-resolution size-based fractionation of large intact proteins from complex protein mixtures, this offline 2DLC method was time consuming and required a large amount of samples. Here we only have a limited amount of ICM tissue samples which were collected during LVAD surgery. Despite these limitations, the major advantages of this online top-down LC-MS/MS method include high reproducibility, superior linearity, and unambiguous quantification of protein PTMs and expression in the sarcomeric subproteome from failing ICM and non-failing donor human cardiac tissues.

2.5 Conclusions

We employed a top-down MS proteomics method for the simultaneous quantification of sarcomeric protein expression and PTMs from ICM and donor human cardiac tissues. This study represents the first top-down LC-MS/MS-based proteomics study of human ICM, which uncovers a diverse sarcomeric proteoform landscape and molecular changes associated with end-stage ischemic heart failure. We identified five significant changes in phosphorylation and four significant changes in protein expression in ICM. Notably, we observed a coordinated decrease in the phosphorylation and expression of cTnI and ENH2, implying dysregulation in the PKA-mediated pathway during ischemic HF. Our results also revealed that Z-disk proteins are altered during ICM, as we observed a significant increase in phosphorylation of calsarcin-1 and MLP. Finally, our results showed differential isoform expression of α-SKA and skβ-Tpm in ICM. Overall, our results highlight the importance of characterizing ICM at the proteoform level to reveal molecular changes within the extremely complex cardiac proteome. In-depth understanding of the molecular consequences underlying ICM will provide valuable information on ICM disease progression and possible therapeutic interventions.

2.6 Methods

2.6.1. Chemicals and Reagents

All reagents were purchased from MilliporeSigma (Burlington, MA, USA) unless otherwise noted. Buffers were prepared with HPLC-grade water from Fisher Scientific (Fair Lawn, NJ, USA). Isopropanol and acetonitrile were purchased from Fisher Scientific (Fair Lawn, NJ, USA). Amicon, 0.5 mL cellulose ultra-centrifugal filters with a molecular weight cutoff (MWCO) of 10 kDa were purchased from MilliporeSigma.

2.6.2. Human Cardiac Tissue Collection

LV myocardium from nonfailing donor hearts with no history of heart diseases were used as control tissues (donor, n = 16). Apex myocardium from failing ICM hearts were collected during left ventricular assist device (LVAD) implantation in which left ventricular apex is removed (ICM, n = 16).[214] Donor heart tissue was stored in cardioplegic solution prior to dissection and snap-frozen immediately in liquid nitrogen and stored at -80 ℃. ICM heart tissue was stored on dry ice immediately after surgical removal and then stored at -80 ℃. Donor and ICM hearts were obtained from the University of Wisconsin (UW)-Madison Organ and Tissue Donation and University of Michigan, respectively. The procedures for the collection of human non-failing donor heart and ICM failing heart tissues were approved by the Institutional Review Board (IRB) of the UW-Madison and University of Michigan, respectively. Available clinical deidentified data including age, gender, cause of death, and medical history are listed for the heart tissues used in this study in **Table S2.3.**

2.6.3. Sample Preparation

Sarcomeric proteins were extracted from human heart tissues using a differential pH-based extraction as previously published.[66] First, ~10 mg of heart tissue was quickly homogenized at 4

℃ and washed twice in 100 μL of HEPES buffer (25 mM HEPES pH = 7.4, 60 mM NaF, 10 mM L-Methionine, 1 mM DTT, 1 mM PMSF, 1 mM Na_3VO_4, 1x Protease Inhibitor Cocktail, and 1x Phosphatase Inhibitor Cocktail) with a Teflon pellet pestle. The homogenate was centrifuged at 21,000 *g* for 30 minutes at 4 ℃. The pellets were re-dispersed in 50 μL of freshly prepared trifluoroacetic acid (TFA) buffer (1% TFA pH = 2.0, 1 mM TCEP, 10 mM L-Methionine). The homogenate was centrifuged at 21,000 *g* for 30 minutes at 4 ℃ and the supernatant was collected. All samples were normalized to 0.08 mg/mL protein in 0.2% formic acid with 2 mM L-methionine by the Pierce™ BCA assay prior to LC-MS/MS analysis.

2.6.4. Top-Down Data Acquisition

Top-down LC-MS/MS was carried out by using a NanoAcquity ultra-high pressure LC system (Waters) coupled to a high-resolution maXis II quadrupole time-of-flight mass spectrometer (Bruker Daltonics). 5 μL (400 ng) of total protein was injected onto a home-packed PLRP column (PLRP-S) (Agilent Technologies), 10-μm particle size, 500-μm inner diameter, 1,000 Å pore size) using an organic gradient of 5 to 95% mobile phase B (mobile phase A: 0.2% formic acid in water; mobile phase B: 0.2% formic acid in 50:50 acetonitrile:isoproponal) at a flow rate of 12 μL/min and temperature of 35 ℃. Column pressure was maintained between 700-1200 psi. Mass spectra were taken at a scan rate of 0.5 Hz over 530-2000 *m/z* range. A total of three replicate runs were collected for each concentration between 250-1200 ng to establish instrument sensitivity and reproducibility. Samples were randomized during processing and LC-MS/MS analysis to correct for batch effects.[215]

Data-dependent LC-MS/MS was performed on sarcomeric protein extracts. The three most intense ions in each mass spectrum were selected and fragmented by collision-activated dissociation (CAD) with a scan rate of 2 Hz from 200-3000 *m/z*. The isolation window for online

AutoMS/MS CAD was 10 *m/z*. The collision DC bias was set from 18 to 35 eV for CAD with nitrogen as the collision gas.

2.6.5. Data Analysis

LC-MS raw files were processed and analyzed with Compass DataAnalysis 4.3 (Bruker Daltonics) software. All chromatograms were smoothened by the Gauss algorithm with a smoothing width of 2.04 s. Sophisticated Numerical Annotation Procedure (SNAP) peak-picking algorithm (quality factor: 0.4; signal-to-noise ratio (S/N): 3.0) was used to determine the monoisotopic mass of all detected ions. Mass spectra were deconvoluted using the Maximum Entropy algorithm within the DataAnalysis 4.3 software. The resolving power for Maximum Entropy deconvolution was set to 60,000k for proteins that were isotopically resolved.

Protein modifications were quantified based on the ratio of the peak intensity of the proteoform to the summed peak intensities of all proteoforms of the same protein using the deconvoluted mass spectrum. In order to quantify protein expression, the top 3-5 most abundant charge state ions of all proteoforms of the same protein were combined to generate one extracted ion chromatogram (EIC) (**Table S2.4**). The area under the curve (AUC) of each EIC was then manually integrated using the DataAnalysis 4.3 software.

Top-down LC-MS/MS data was processed using MASH Native software with the topPIC (v. 1.5.4) algorithm embedded.[148, 156] The fragment ions in the MS/MS spectra were searched and assigned based on the canonical entries of the *Homo sapiens* UniProt (UniProtKB) protein database (UP000005640, 20,389 reviewed entries, version December 2022). Monoisotopic masses were used for all proteoform determinations, and all fragment ions were manually validated with a mass tolerance of 20 ppm.

2.6.6. Statistical Analysis

Statistical analysis for group comparison was completed using paired-student *t* tests to determine the level of statistical significance for the quantification of sarcomere protein modification and expression. All *p*-values at $p < 0.05$ were considered significant. All error bars indicated in the figures represent the mean ± standard error of the mean (S.E.M.). Statistical analysis for the coordinated decrease in enigma homolog 2 (ENH2) and cTnI phosphorylation was completed using a Pearson's correlation. The Pearson correlation coefficient (*r*) is considered strongly correlated if it is above 0.7.

2.7 Supplemental Information

2.7.1. Supplemental Tables

Table S2.1. Sarcomeric proteoforms identified. Gene, Uniprot ID, proteoform, modified forms, retention time (RT) in minutes, calculated most abundant mass, experimental most abundant mass, and mass error for the proteoforms we identified in this experiment. Online LC-MS/MS data allowed for confident assignment of high-abundance proteoform identifications. The low-abundance modified proteoforms were manually identified based on highly accurate intact mass measurements together with the prior knowledge on identified sarcomeric proteforms from previous publications[35, 39, 169, 193] or Uniprot. Abbreviations: muscle LIM protein (MLP); cysteine rich protein 2 (CRIP2); enigma homolog 2 (ENH2); cardiac troponin T (cTnT); four and a half LIM domains protein (FHL2); cardiac troponin I (cTnI); tropomyosin (Tpm); myosin light chain 1, ventricular isoform (MLC-1v); myosin light chain 1, atrial isoform (MLC-1a); myosin light chain 2, ventricular isoform (MLC-2v); cardiac α-actin (cα-actin); skeletal α-actin (sα-actin); troponin C (TnC); N-terminal acetylation (N-acetyl); phosphorylation (phospho); mono-phosphorylation (*p*); bis-phosphorylation (*pp*); tris-phosphorylation (*ppp*); N-terminal tri-methylation (N,N,N-trimethyl).

Gene	Uniprot ID	Proteoform	Modified Forms	RT (min)	Calc Most Abundant Mass (Da)	Expt Most Abundant Mass (Da)	Mass Error (ppm)
CSRP3	P50461-1	MLP	Met removal	12.6-13.3	20836.90	20836.88	1.0
CSRP3	P50461-1	*p*MLP	Met removal, phospho	12.6-13.3	20916.87	20916.88	0.5
CRIP2	P52943-1	CRIP2	Met removal, N-acetyl	12.6-13.3	22403.00	22403.00	0.0
CRIP2	P52943-1	*p*CRIP2	Met removal, N-acetyl	12.6-13.3	22482.97	22482.94	1.3
PDLIM5	Q96HC4-2	ENH2	Met removal, N-acetyl	15.3-15.9	25841.39	25841.31	3.1
PDLIM5	Q96HC4-2	*p*ENH2	Met removal, N-acetyl, phospho	15.3-15.9	25921.35	25921.31	1.5
TNNT2	P45379-6	cTnT [aa 1-286]	aa 1-286, Met removal, N-acetyl	15.3-15.9	34373.34	34373.31	0.9
TNNT2	P45379-6	*p*cTnT [aa 1-286]	aa 1-286, Met removal, N-acetyl, phospho	15.3-15.9	34453.31	34453.38	2.0
TNNT2	P45379-6	cTnT	Met removal, N-acetyl	15.3-15.9	34502.44	34502.38	1.7
TNNT2	P45379-6	*p*cTnT	Met removal, N-acetyl, phospho	15.3-15.9	34582.41	34582.38	0.9
LDB3	O75112-6	Cypher-6	Met removal, N-acetyl	16.9-17.4	30909.83	30909.81	0.6
PDLIM1	O00151-1	Elfin	Met removal, N-acetyl	16.9-17.4	35955.05	35955.00	1.4
PDLIM1	O00151-1	Elfin [N175S]	Met removal, N-acetyl, N175S	16.9-17.4	35982.06	35982.19	3.6
LDB3	O75112-5	Cypher-5	Met removal, N-acetyl	17.4-17.8	35548.24	35548.19	1.4
FHL2	Q14192-1	FHL2	Met removal, N-acetyl	18.4-18.9	32101.63	32101.56	2.2
MYOZ2	Q9NPC6-1	Calsarcin-1	N-acetyl	18.9-19.4	29940.11	29940.06	1.7
MYOZ2	Q9NPC6-1	*p*Calsarcin-1	N-acetyl, phospho	18.9-19.4	30020.08	30020.06	0.7
MYOZ2	Q9NPC6-1	*pp*Calsarcin-1	N-acetyl, 2x phospho	18.9-19.4	30100.04	30100.00	1.3
MYOZ2	Q9NPC6-1	*ppp* Calsarcin-1	N-acetyl, 3x phospho	18.9-19.4	30180.01	30180.13	4.0
TNNI3	P19429-1	cTnI	Met removal, N-acetyl	18.9-20.0	23918.83	23918.81	0.8
TNNI3	P19429-1	*p*cTnI	Met removal, N-acetyl, phospho	18.9-20.0	23998.79	23998.81	0.8
TNNI3	P19429-1	*pp*cTnI	Met removal, N-acetyl, 2x phospho	18.9-20.0	24078.76	24078.75	0.4
TPM1	P09493-1	α-Tpm	N-acetyl	20.2-21.2	32750.74	32750.69	1.5
TPM1	P09493-1	*p*α-Tpm	N-acetyl, phospho	20.2-21.2	32830.70	32830.63	2.1
TPM1	P09493-6	κ-Tpm	N-acetyl	20.2-21.2	32691.70	32691.69	0.3
TPM1	P09493-6	*p*κ-Tpm	N-acetyl, phospho	20.2-21.2	32771.67	32771.63	1.2
TPM3	P06753-1	γ-Tpm	Met removal, N-acetyl	20.2-21.2	32860.79	32860.75	1.2
TPM2	P07951-1	skβ-Tpm	N-acetyl	20.2-21.2	32892.62	32892.56	1.8
MYL3	P08590-1	MLC-1v	Met removal, N,N,N-trimethyl	21.4-22.6	21841.96	21841.94	0.9
MYL4	P12829-1	MLC-1a	Met removal, N,N,N-trimethyl	23.2-24.1	21474.90	21474.88	0.9
MYL2	P10916-1	MLC-2v	Met removal, N,N,N-trimethyl	24.6-25.6	18699.40	18699.44	2.1
MYL2	P10916-1	*p*MLC-2v	Met removal, N,N,N-trimethyl, phospho	24.6-25.6	18779.37	18779.38	0.5
ACTC1	P68032-1	α-CAA	Met removal, Cys removal, N-acetyl, methylation	27.8-29.0	41839.92	41839.94	0.5
ACTA1	P68133-1	α-SKA	Met removal, Cys removal, N-acetyl	27.8-29.0	41871.89	41871.88	0.2
TNNC1	P63316-1	TnC	N-acetyl	30.5-31.1	18443.58	18443.56	1.1

Table S2.2. Proteoform changes in ICM. Summary of significant proteoform changes in failing heart tissues from patients in end-stage ischemic heart failure (n = 16) compared to non-failing donor patients (n = 16). Groups were considered significantly different by paired-student t tests with $p < 0.05$.

Proteoform Change	ICM Compared to Donor	Fold Change	p-value
cTnI phosphorylation	Decreased	2.2	6.09E-06
ENH2 phosphorylation	Decreased	1.7	1.40E-04
MLP phosphorylation	Increased	2.5	2.59E-03
Calsarcin-1 phosphorylation	Increased	1.2	3.16E-02
cTnT phosphorylation	Increased	1.1	1.21E-02
Skeletal β-Tpm expression	Decreased	1.7	3.63E-03
Skeletal α-actin expression	Increased	1.4	8.80E-03
cTnI expression	Decreased	1.3	4.30E-02
ENH2 expression	Decreased	1.5	3.23E-02

Table S2.3. Available clinical data for donors and ICM patients. Donor heart and ICM tissue from left ventricular assist device (LVAD) surgeries have been coded to Donor# and ICM#, respectively. Samples were obtained from two sources: the University of Wisconsin (UW)-Madison Organ and Tissue Donation and the University of Michigan. The sample code (assigned for this publication) and Biopsy ID (assigned at the time of de-identification) are provided. Clinical data such as age, gender, cause of death, and medical history data are listed if available.

Sample Code	Biopsy ID	Age	Gender	Cause of Death	Medical History
Donor 1	D12	-	-	-	-
Donor 2	D03	-	-	-	-
Donor 3	D05	-	-	-	-
Donor 4	D10	-	-	-	-
Donor 5	YG07	44	M	Cerebrovascular/stroke	-
Donor 6	YG08	60	M	Head Trauma	No past history of heart disease; echo and coronary angiogram normal
Donor 7	YG06	-	-	Stroke	-
Donor 8	YG59	56	M	Braindead	Emphysema, high cholesterol, bipolar, depression, meth & marijuana use, PTSD, vitamin D deficiency, smoker
Donor 9	YG23	27	M	Head Trauma	No past history of coronary artery disease
Donor 10	YG25	61	M	Head Trauma	No past history of heart disease; echo and coronary angiogram normal
Donor 11	YG42	24	M	Burn Victim	No past history of heart disease; echo and coronary angiogram normal
Donor 12	YG43	57	M	Acute Respiratory Distress Syndrome (ARDS)/Pulmonary Emboli	-
Donor 13	YG49	47	M	Stroke	Type-1 diabetic, end-stage renal failure
Donor 14	YG53	65	F	Head Trauma	Trivial mitral valve regurgitation, mild hypertension, smoking, anxiety
Donor 15	YG54	62	M	Head Trauma	No past history of heart disease, had hypothyroidism and high cholesterol
Donor 16	YG57	64	F	Stroke	Hypertension, hyperlipidemia, smoking, and mild coronary artery and aortic valve calcification
ICM 1	LVAD 99	64	M	-	-
ICM 2	LVAD 20	60	M	Heart failure	Coronary artery disease
ICM 3	LVAD 37	51	M	-	Coronary artery disease

Table S2.3. continued.

ICM 4	LVAD 40	64	M	-	Coronary artery disease
ICM 5	LVAD 54	38	M	-	Coronary artery disease
ICM 6	LVAD 56	55	F	-	-
ICM 7	LVAD 61	29	M	-	Coronary artery disease
ICM 8	LVAD 68	64	M	-	-
ICM 9	LVAD 71	71	M	-	Triple vessel coronary artery disease
ICM 10	LVAD 78	66	M	-	-
ICM 11	LVAD 81	71	M	-	-
ICM 12	LVAD 87	58	F	-	-
ICM 13	LVAD 89	71	M	-	-
ICM 14	LVAD 90	68	F	-	-
ICM 15	YG22	36	M	-	-
ICM 16	YG09	-	F	-	-

Table S2.4. Generation of extracted ion chromatograms for sarcomeric proteins. Protein, modified forms, and extracted ions used for expression quantification when determining the mutual linear range of our method. Three extracted ions were used for low abundant proteoforms and five extracted ions were used for highly abundant proteoforms. Abbreviations: cardiac troponin I (cTnI); myosin light chain 1, ventricular isoform (MLC-1v); myosin light chain 2, ventricular isoform (MLC-2v); tropomyosin (Tpm); enigma homolog 2 (ENH2); troponin C (TnC); N-terminal acetylation (N-acetyl); phosphorylation (phospho); N-terminal tri-methylation (N,N,N-trimethyl).

Protein	Modified Forms	Extracted Ions for Expression Quantification (+/- 0.2 m/z)
cTnI	Met removal, N-acetyl	748.4378; 704.4726; 684.3767
	Met removal, N-acetyl, phospho	750.9405; 728.2137; 800.9330
	Met removal, N-acetyl, 2x phospho	777.7114; 753.9991; 688.9428; 730.6374; 803.5994
MLC-1v	Met removal, N,N,N-trimethyl	841.0841;809.9702;874.6852;911.0886;754.1784
MLC-2v	Met removal, N,N,N-trimethyl	850.9823;935.9796;813.5469;891.4557;985.1853
	Met removal, N,N,N-trimethyl, phospho	854.6166; 839.9766; 817.5043
α-Tpm	N-acetyl	840.7429; 780.7640; 762.6302; 799.7830; 819.7488
	N-acetyl, phospho	842.7948; 782.6649; 764.4892
ENH2	Met removal, N-acetyl	763.3985; 741.6150; 701.5574; 811.0497; 837.1796
	Met removal, N-acetyl, phospho	785.9555; 720.5417; 864.4494
TnC	N-acetyl	1025.6500; 1085.9245; 971.7224; 1153.7303; 1230.5803

2.7.2. Supplemental Figures

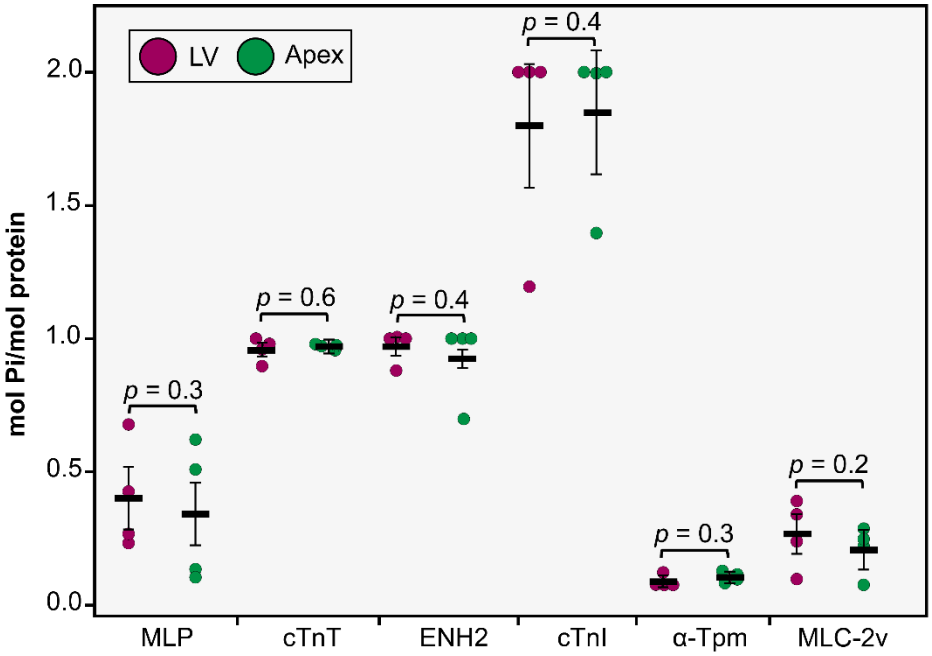


Figure S2.1. Sarcomeric proteoform levels between left ventricle and apex tissues in healthy donors are comparable. Total protein phosphorylation calculated by mol Pi/mol protein for MLP, cTnT, ENH2, cTnI, α-Tpm, and MLC-2v in LV (magenta) and apex (green) tissues from four healthy donor hearts (n = 4). Groups were considered significantly different by paired-student t tests with $p < 0.05$.

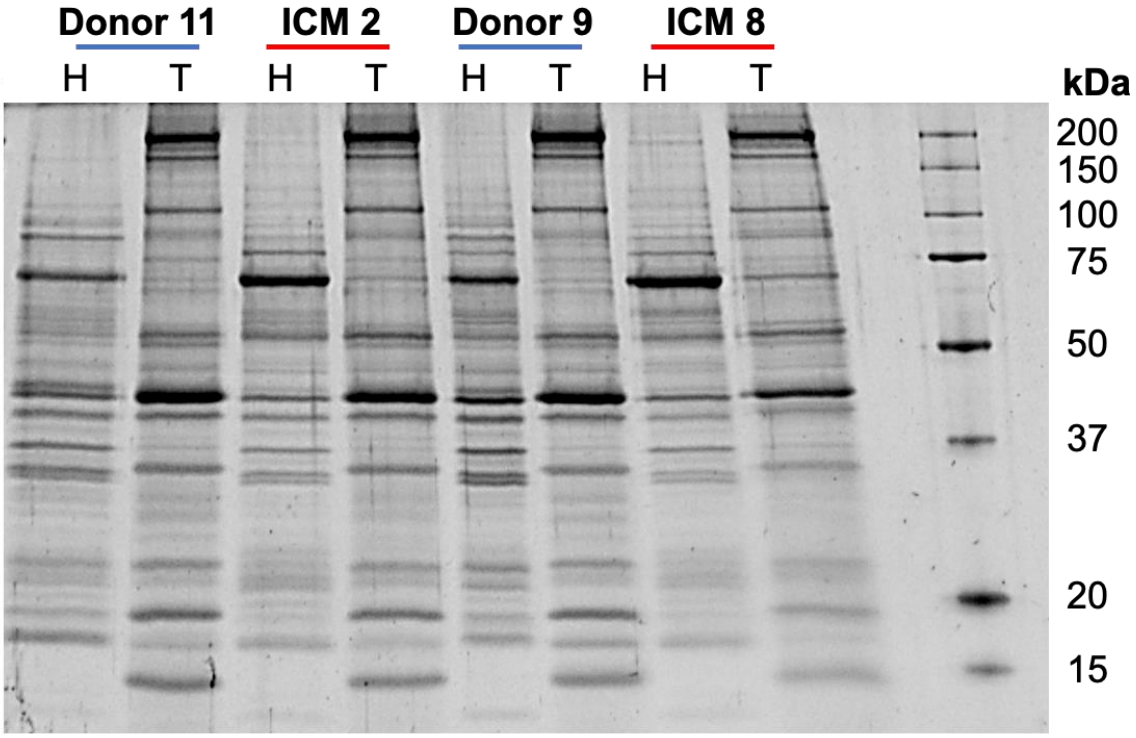


Figure S2.2. SDS-PAGE of HEPES (H) and TFA (T) extracts from donor and ICM human cardiac tissues. The results show depletion of cytosolic proteins during the HEPES (H) extraction and successful enrichment of sarcomeric proteins during the TFA (T) extraction from human cardiac tissues. 500 ng of total protein was loaded into each gel lane.

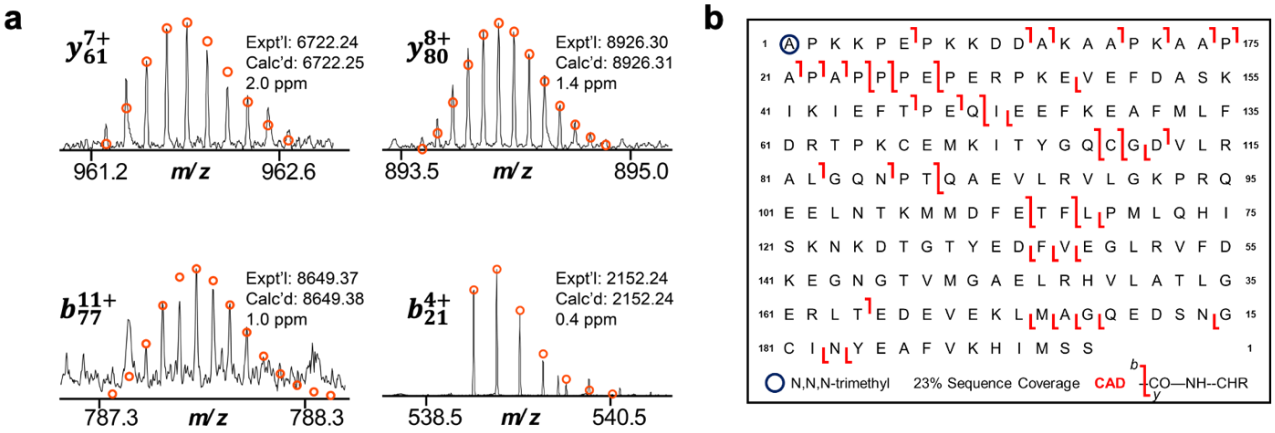


Figure S2.3. Online LC-MS/MS characterization of MLC-1v. (a) Representative CAD fragment ions. Circles represent the theoretical isotopic abundance distribution of the isotopomer peaks corresponding to the assigned monoisotopic mass. **(b)** Sequence map of MLC-1v. N-terminal trimethylation is circled in blue. For MLC-2v there were 25 *b* ions and 23 *y* ions achieving 23% sequence coverage.

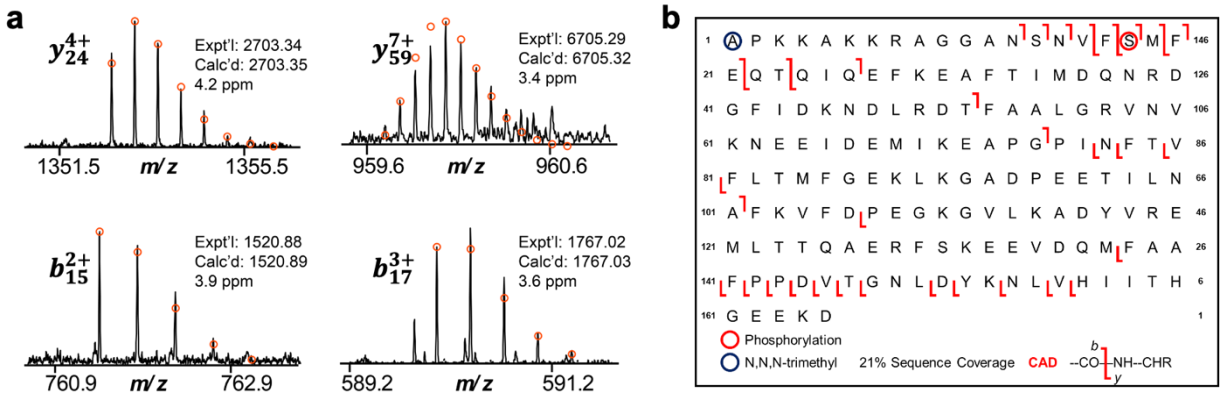


Figure S2.4. Online LC-MS/MS characterization of MLC-2v. (a) Representative CAD fragment ions. Circles represent the theoretical isotopic abundance distribution of the isotopomer peaks corresponding to the assigned monoisotopic mass. **(b)** Sequence map of MLC-2v. N-terminal trimethylation is circled in blue and phosphorylation is circled in red. For MLC-2v there were 14 *b* ions and 22 *y* ions achieving 21% sequence coverage.

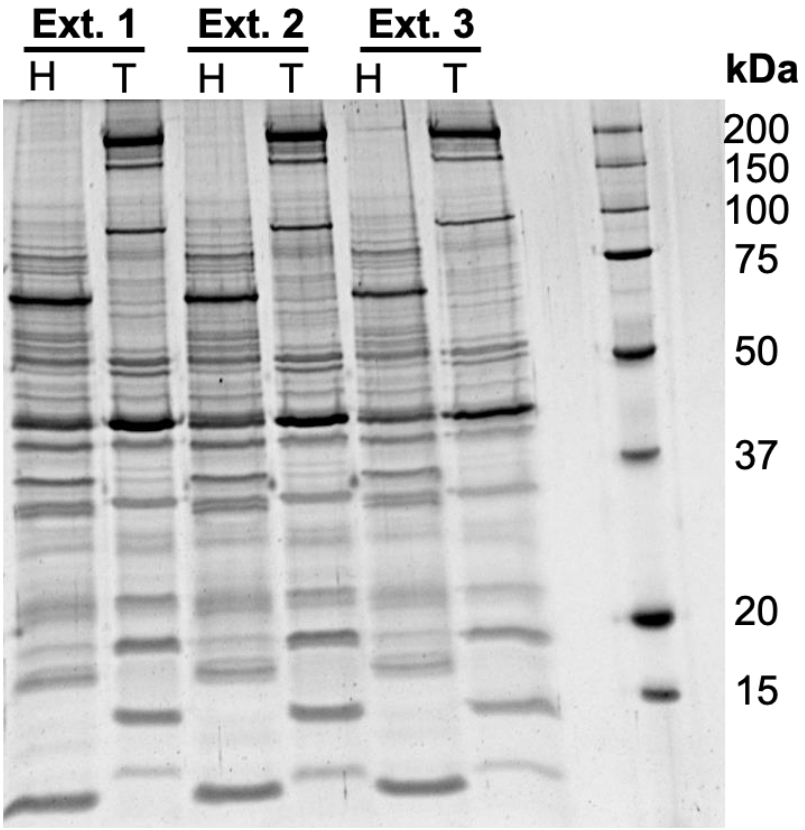


Figure S2.5. SDS-PAGE of HEPES (H) and TFA (T) from three extraction replicates from donor human cardiac tissue. The results show reproducible depletion of cytosolic proteins (H) and successful enrichment of sarcomeric proteins (T) from human cardiac tissues. 500 ng of total protein was loaded into each gel lane.

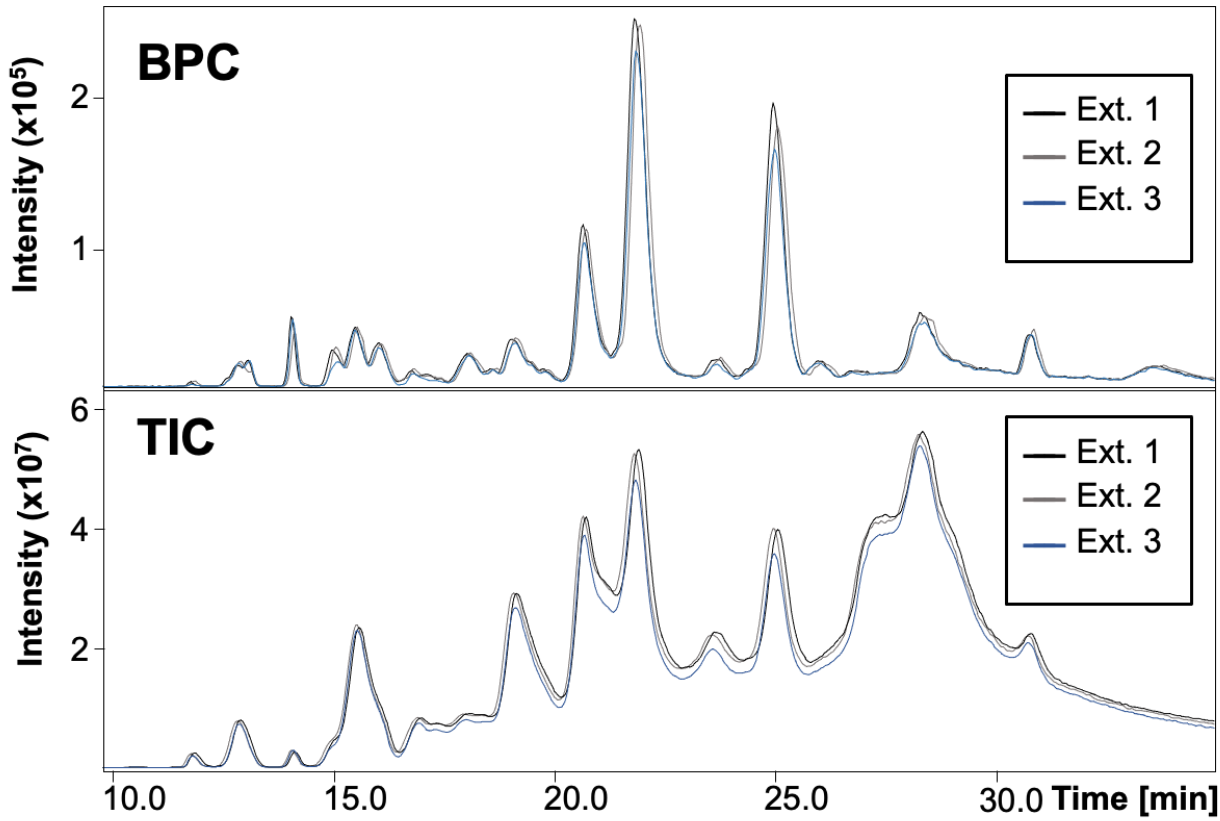


Figure S2.6. Extraction replicates showing highly similar and reproducible chromatographic profiles. Overlaid BPCs and TICs of three biological extraction replicates of donor cardiac tissue showing high reproducibility of the method.

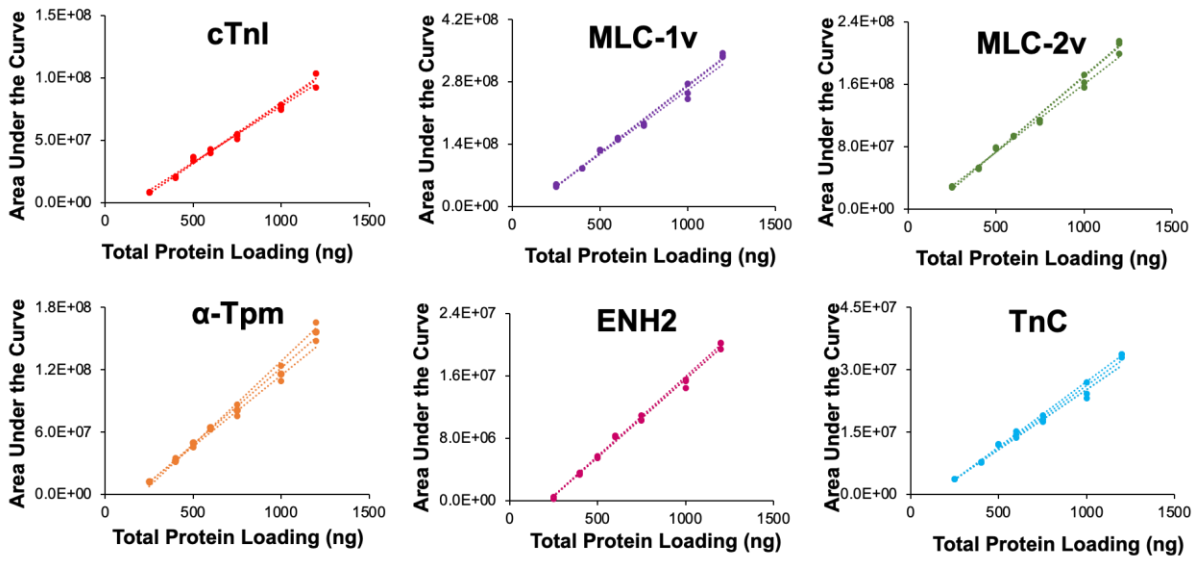


Figure S2.7. Mutual linear range response determination for major sarcomeric proteins. Individual linear fittings for cTnI, MLC-1v, MLC-2v, α-Tpm, ENH2, and TnC from 250 to 1200 ng total protein amount loaded to the LC-MS.

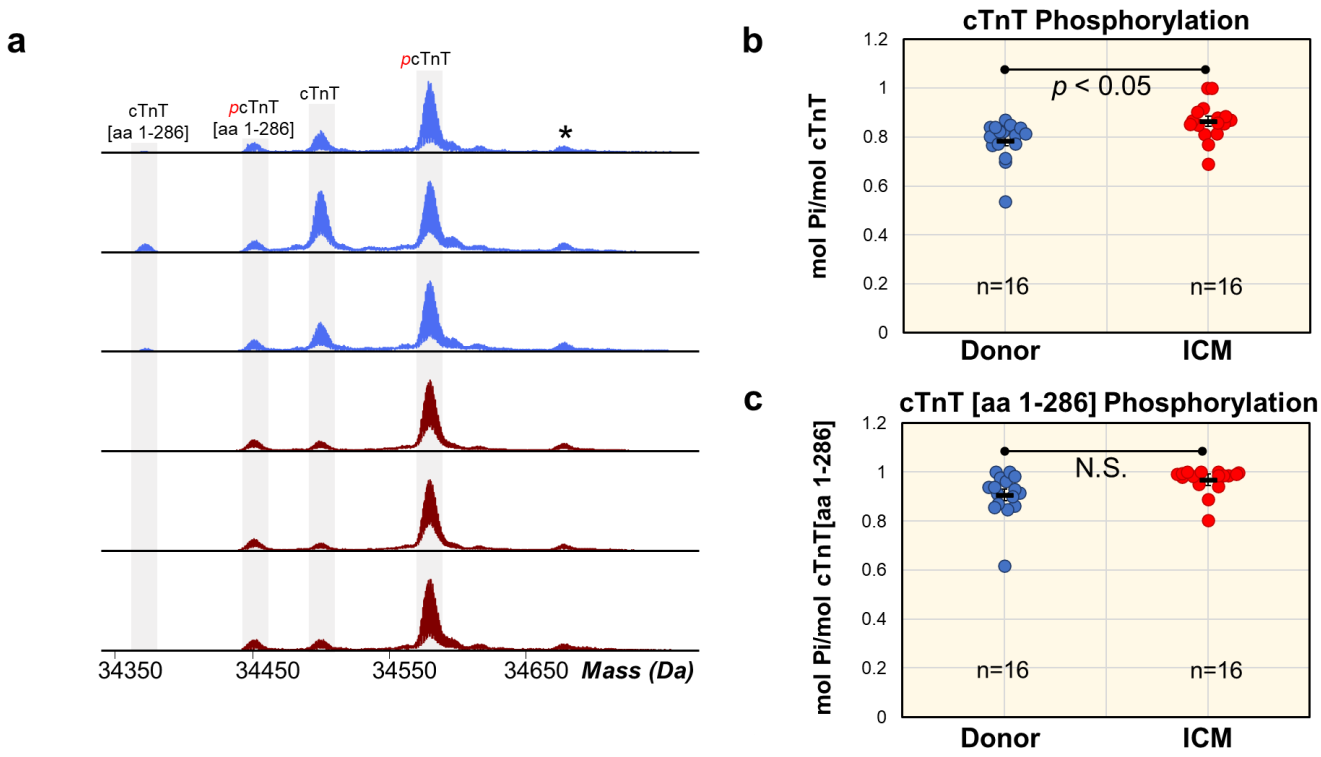


Figure S2.8. Quantification of full-length and truncated cTnT phosphorylation. (a) Representative deconvoluted spectra of cTnT in non-failing donor (blue) compared to ICM (red) tissues. We observed unphosphorylated cTnT, monophosphorylated cTnT (*p*cTnT), unphosphorylated truncated cTnT with a lysine cleaved from the C-terminus (cTnT [aa 1-286]) and monophosphorylated truncated cTnT (*p*cTnT [aa1-286]). Red *p* denotes monophosphorylated proteoforms. The asterisk represents a phosphate adduct. **(b)** Total protein phosphorylation calculated by mol Pi/mol protein for full-length cTnT in donor (n = 16) and ICM (n = 16) tissues. Horizontal bars represent the mean of the group and error bars represent the SEM in black. Groups were considered statistically different at $p < 0.05$. **(c)** Total protein phosphorylation calculated by mol Pi/mol protein for truncated cTnT[aa 1-286] in donor (n = 16) and ICM (n = 16) tissues. Horizontal bars represent the mean of the group and error bars represent the standard error of mean (SEM) in black. Groups were considered statistically different at $p < 0.05$.

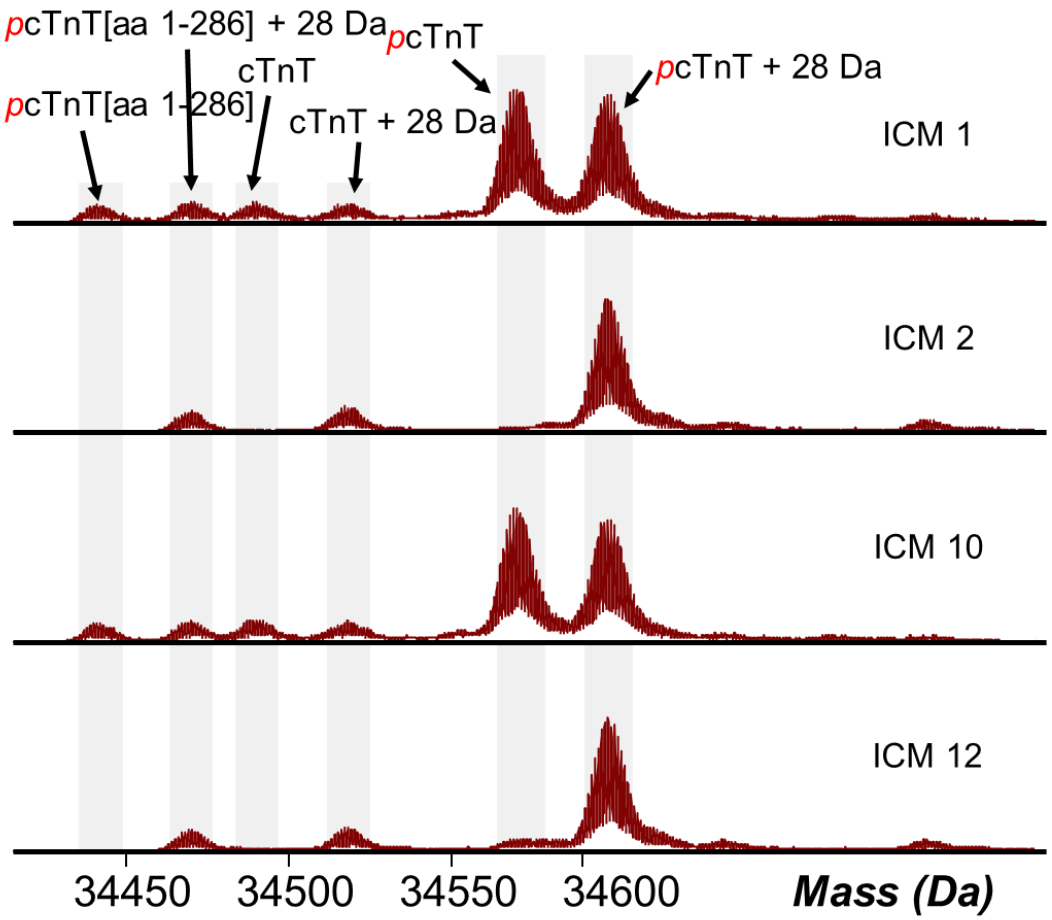


Figure S2.9. cTnT point mutation present in ICM samples. Representative deconvoluted spectra of cTnT homozygous and heterozygous polymorphism (+28 Da) observed in four ICM tissue samples.

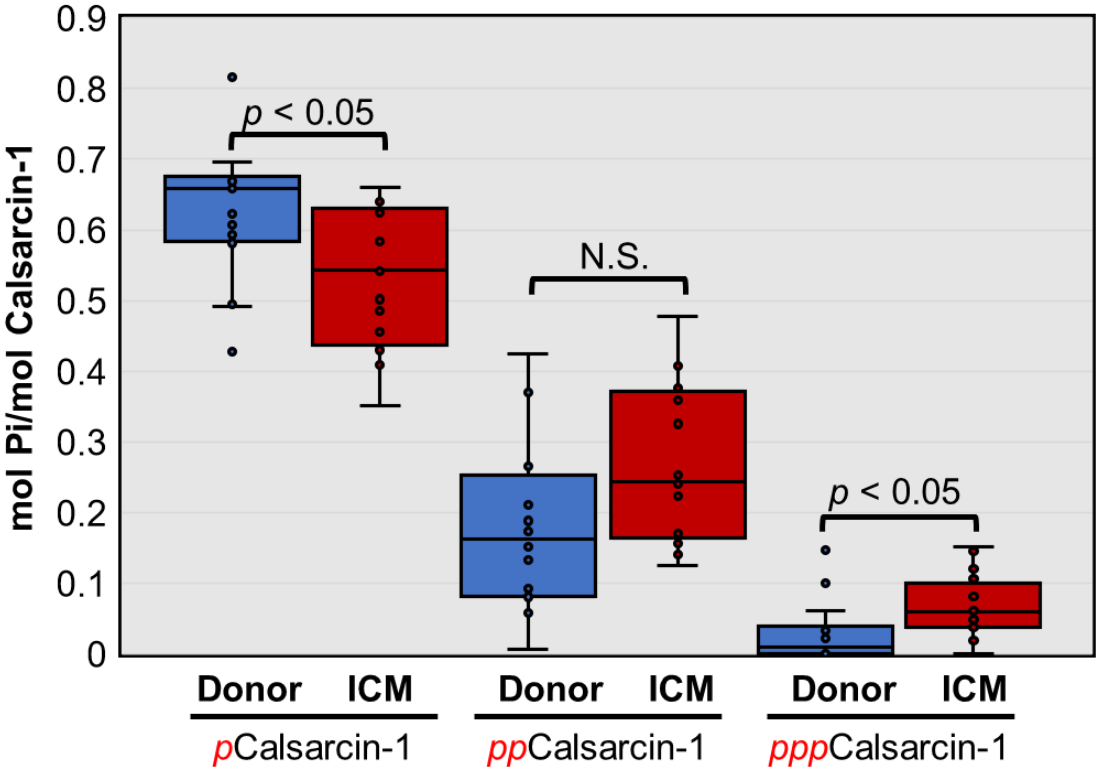
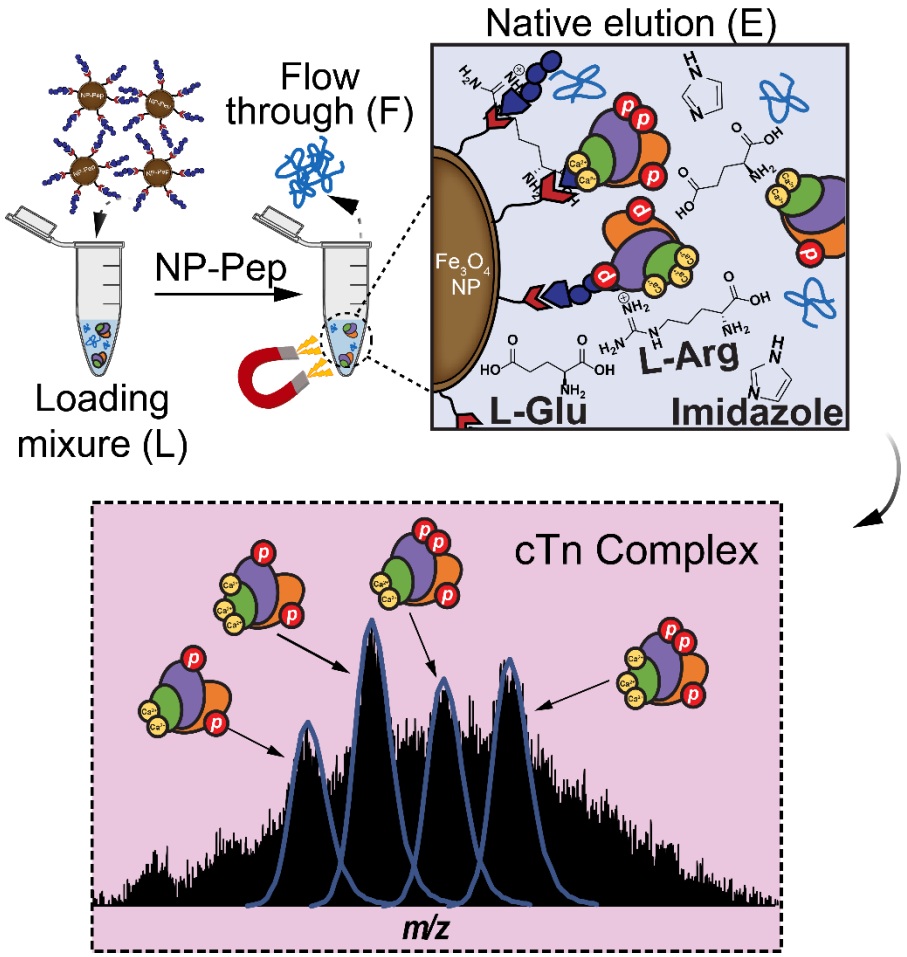


Figure S2.10. Quantification of calsarcin-1 proteoform phosphorylation. Total protein phosphorylation calculated by mol pi/mol protein calsarcin-1 for *p* calsarcin-1, *pp* calsarcin-1, and *ppp* calsarcin-1. Box, median and interquartile range (25%, 75%); whiskers, minimum and maximum values. Horizontal lines represent the median of the group. Groups were considered significantly different by paired-student *t* tests with $p < 0.05$.

CHAPTER 3:

Structure and Dynamics of Endogenous Cardiac Troponin Complex in Human Heart Tissue Captured by Native Nanoproteomics



Portions of this chapter are adapted from the following published manuscript:

Chapman, E.A.*; Roberts, D.S.*; Tiambeng, T.N.; Andrews, J.; Wang, M.; Reasoner, E.A.; Melby, J.A.; Li, B.H.; Kim, D.; Alpert, A.J.; Jin, S.; Ge, Y. Structure and Dynamics of Endogenous Cardiac Troponin Complexes in Human Heart Tissue Captured by Native Nanoproteomics. *Nature Communications* **2023**, *14* (1), 8400. DOI: https://doi.org/10.1038/s41467-023-43321-z. *denotes equal contribution.

3.1 Abstract

Protein complexes are highly dynamic entities that display substantial diversity in their assembly, post-translational modifications, and non-covalent interactions, allowing them to play critical roles in various biological processes. The heterogeneity, dynamic nature, and low abundance of protein complexes in their native states present challenges to study using conventional structural biology techniques. Here we develop a native nanoproteomics strategy for the enrichment and subsequent native top-down mass spectrometry (nTDMS) of endogenous cardiac troponin (cTn) complexes directly from human heart tissue. The cTn complex is enriched and purified using peptide-functionalized superparamagnetic nanoparticles under non-denaturing conditions to enable the isotopic resolution of cTn complexes, revealing their complex structure and assembly. Moreover, nTDMS elucidates the stoichiometry and composition of the cTn complex, localizes Ca^{2+} binding domains, defines cTn-Ca^{2+} binding dynamics, and provides high-resolution mapping of the proteoform landscape. This native nanoproteomics strategy opens a paradigm for structural characterization of endogenous native protein complexes.

3.2 Introduction

The vast majority of proteins within a cell assemble into protein complexes to perform specific functions and play crucial roles in regulating diverse biological processes.[10, 178] Thus, comprehensive characterization of the structure and dynamics of endogenous protein complexes is essential for understanding fundamental biological processes and disease mechanisms to develop new therapeutic interventions.[10, 178] Protein complexes are highly dynamic entities with substantial diversity in their assembly, post-translational modifications (PTMs), and non-covalent interactions. Moreover, endogenous protein complexes often exist in low abundance in their native states.[178] These present tremendous challenges to study their structure and dynamics using traditional structural biology techniques such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryo-electron microscopy (cryo-EM).[178] Native top-down mass spectrometry (nTDMS), a technique combining native MS[44, 47, 49, 216, 217] and top-down proteomics[6, 39-41, 58, 64], has emerged as a powerful structural biology tool for characterization of protein complexes.[31, 45, 48, 50, 51, 218] In nTDMS, intact proteins are introduced into the mass spectrometer under non-denaturing conditions, preserving the noncovalent interactions between protein subunits and their associated ligands as well as PTMs. The intact proteins are then fragmented in the gas-phase to map the PTMs and ligand binding sites.[31] nTDMS enables the structural characterization of macromolecular protein complexes, subunit stoichiometry, non-covalent interactions, as well as the analysis of their proteoforms – the diverse protein products arising from alternative splice isoforms, genetic variations, and PTMs.[6-8] However, so far only over-expressed recombinant or high-abundance proteins and protein complexes have been characterized by nTDMS. Significant challenges remain in the structural characterization of

endogenous protein complexes that are present in relatively lower-abundance due to the difficulty in isolating them and sensitivity required to resolve heterogenous complexes.[178]

Here, we have developed a native nanoproteomics platform integrating the native enrichment of protein complexes directly from tissues using surface functionalized superparamagnetic nanoparticles (NPs) with high-resolution nTDMS to characterize the structure and dynamics of endogenous protein complexes. Specifically, we applied this method to enrich and structurally elucidate the heterotrimeric cardiac troponin (cTn) complex (~77 kDa), a key protein complex essential to cardiac muscle contraction, directly from human heart tissues. The cTn complex is a master regulator of cardiac contraction and represents the Ca^{2+} sensitive switch[219, 220] of striated muscles assembled from three molecular subunits: troponin C (TnC), the Ca^{2+}-binding subunit; cardiac troponin I (cTnI), the actin-myosin inhibitory subunit; and cardiac troponin T (cTnT), the thin-filament anchoring subunit.[221, 222] Both cTnI and cTnT serve as gold standard biomarkers for diagnosing acute coronary syndrome due to their cardiac specificity and their release into the bloodstream following cardiac injury.[223] Moreover, the association of Ca^{2+} ions with the TnC subunit along with phosphorylation of the cTnI subunit together initiate a cascade of molecular events on the thin filament and induce actin-myosin cross bridge formation necessary for cardiac contraction.[224, 225] However, only partial structural information of the core domain of the human cTn complex in its Ca^{2+} saturated state has been obtained from conventional X-ray crystallography excluding the intrinsically disordered but functionally critical regions of the N- and C- terminal regions of both cTnI and cTnT.[226] Moreover, the cTn structure is highly dynamic due to Ca^{2+} binding[222, 227, 228] and PTMs[35, 39, 229] that regulates muscle contraction, yet traditional methods have not effectively captured these dynamic conformational changes, primarily because they are limited to visualizing the static state of the thin filament.[230]

Furthermore, recombinantly expressed cardiac troponin subunits are frequently utilized and incorporated into reconstituted cTns or cardiac thin filaments to investigate the structure-function relationships of the complex as well as Ca^{2+} sensitivity.[231-233] However, the lack of post-translational machinery in prokaryotic cell expression systems presents intrinsic limitations when recapitulating eukaryotic, post-translationally modified endogenous proteins such as phosphorylated cTn subunits, using recombinant protein systems. [234, 235].

In this work, we enrich and purify cTn complexes directly from human heart tissue and achieve isotopic resolution of endogenous cTn complexes to reveal cTn complex structure and assembly. Our results elucidate the stoichiometry and composition of the heterotrimeric cTn complex, define the conformational roles of cTn-Ca^{2+} binding dynamics, locate the Ca^{2+} binding domains (II-IV), and map the proteoform landscape with direct analysis of the stoichiometry of various proteoforms. Overall, this work demonstrates that nTDMS, enabled by native nanoproteomics, can comprehensively characterize the structure and dynamics of endogenous protein complexes.

3.3 Results

3.3.1. A Native Nanoproteomics Platform for the Enrichment and Comprehensive Characterization of Endogenous Protein Complexes

Our antibody-free native nanoproteomics platform integrates native enrichment and purification of endogenous protein complexes using peptide functionalized superparamagnetic iron oxide (magnetite, Fe_3O_4) NPs (NP-Pep) followed by comprehensive characterization using nTDMS (**Figure 3.1**). Building on our previous denatured NP enrichment study[78], we sought to establish a method that could enrich and purify native protein complexes directly from human

heart tissue. We first optimized the native protein extraction buffers to effectively extract intact protein complexes from human heart tissue using a high ionic strength lithium chloride (LiCl) buffer at physiological pH and a phosphate wash buffer for multiple depletions of highly abundant cytosolic proteins (**Figure 3.1a, Table S3.1**). Next, we hypothesize that the specifically designed peptide[236] on the NP surface that contains a combination of charged and aromatic residues that leads to the specific binding to the protein complex of interest, would be amenable to a competitive elution strategy using amino acids.[237-239] L-Arg and L-Glu are amino acids previously known to improve protein solubility and in-solution stability by forming protonated clusters,[238] and were included to enhance protein complex elution efficacy. Moreover, imidazole, a positively charged aromatic small molecule which functions as the side chain of histidine,[237] was added to disrupt the polar interactions of the peptide functionalized on the NP surface for the native competitive elution of NP-captured protein complexes, without disrupting the intermolecular interactions between the protein subunits. We found that the optimal buffer composition for effective native competitive elution of NP-captured protein complexes was a combination of 750 mM L-Arg, 750 mM imidazole, and 50 mM L-Glu (pH 7.5) as native elution efficiency of amino acid solutions increases with their concentration (**Figure 3.1a, Figure S3.1**).[240]

For a native proteomics workflow, the NP-Pep was incubated with sarcomeric protein mixtures, magnetically isolated to remove non-specifically bound proteins, and the bound protein complexes were eluted off the NP-Pep using the above optimized native elution buffer cocktail. Non-MS compatible buffers and salts were removed using either online or offline size-exclusion chromatography (SEC) to transfer the complexes into a MS-compatible ammonium acetate solution. Subsequently, the enriched protein complexes were subjected to various nTDMS techniques to characterize protein complex structure, assembly, and dynamics including: online

SEC for rapid screening of protein complexes (**Figure 3.1b**), ultrahigh-resolution Fourier transform ion cyclotron resonance (FTICR)-MS to characterize protein complex stoichiometry and proteoform landscape (**Figure 3.1c**), and trapped ion mobility (TIMS)-MS to resolve the structural dynamics of non-covalent interactions in regulating protein complex conformational change (**Figure 3.1d**). These techniques allowed for direct analysis of structural features revealing the molecular composition, stoichiometry, and non-covalent interaction of endogenous protein complexes (**Figure 3.1e-f**).

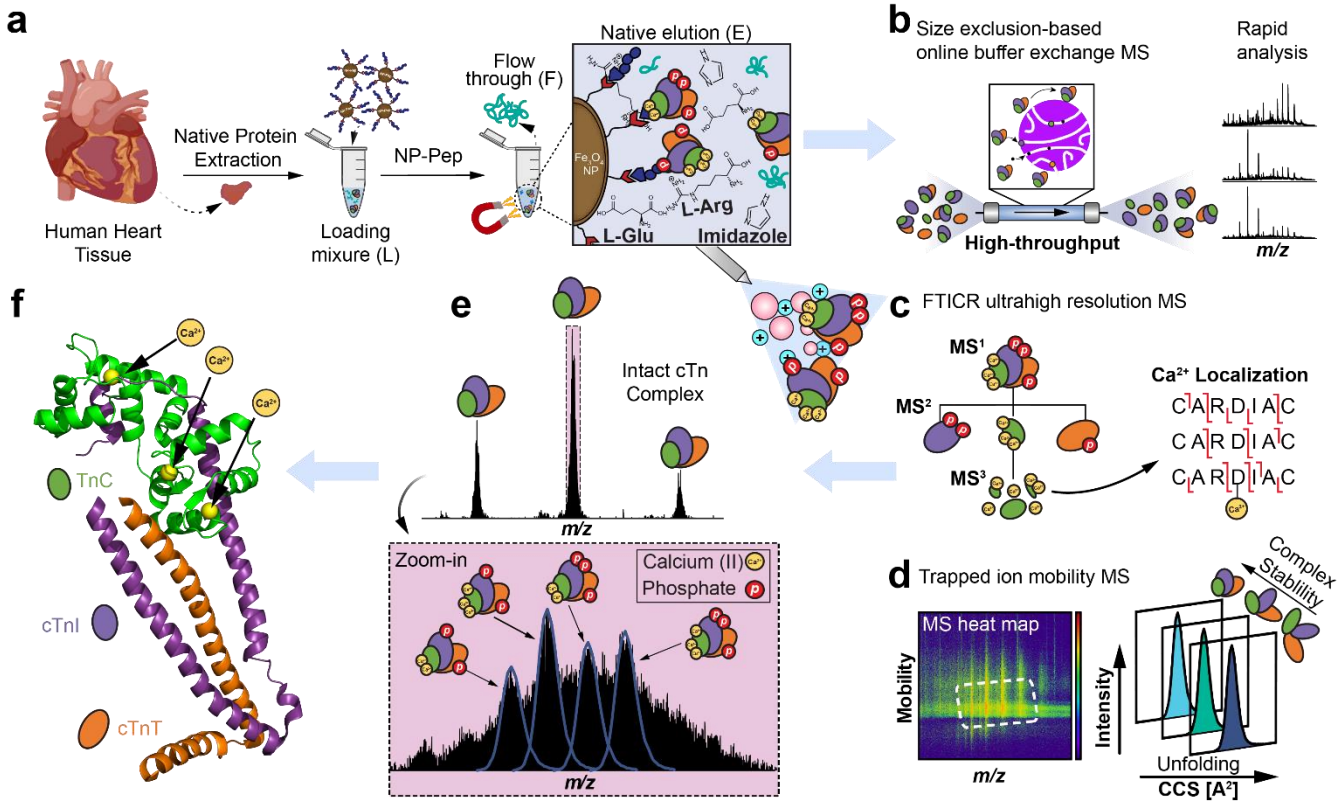


Figure 3.1. Native nanoproteomics platform for the characterization of endogenous cTn complexes from human heart tissue. (a-f). (a) Sarcomeric proteins were extracted using a high ionic strength lithium chloride (LiCl) buffer at physiological pH. The heart tissue extract (loading mixture, L) is then incubated with peptide-functionalized nanoparticles (NP-Pep). Following magnetic isolation, the nonspecifically bound proteins are removed as flow through (F). The bound protein complexes are then eluted (E) off the NPs using a native elution buffer containing L-glutamic acid, L-arginine, and imidazole. Native top-down MS (nTDMS) analysis of enriched protein complexes proceeds by either **(b)** native size exclusion chromatography (SEC) for online buffer exchange and rapid analysis, **(c)** ultrahigh-resolution Fourier

transform ion cyclotron (FTICR)-tandem MS (MS/MS) analysis for probing complex heterogeneity, stoichiometry, and localization of Ca^{2+} ions, or **(d)** trapped ion mobility spectrometry (TIMS) coupled with MS (timsTOF) for structural characterization of complex-Ca^{2+} binding dynamics. **(e)** High-resolution mass spectra of endogenous cardiac troponin (cTn) heterotrimeric complexes enriched by NP-Pep directly from human heart tissue and analyzed by nTDMS. **(f)** Structural representation showing the cTn heterotrimeric complex. Troponin C (TnC) is depicted in green, cardiac troponin I (cTnI) in purple, cardiac troponin T (cTnT) in orange, and Ca^{2+} ions in yellow. PDB: 1J1E.

3.3.2. Native Enrichment of Endogenous cTn Complex from Human Heart Tissue

To enrich cTn complexes from human heart tissues, the Fe_3O_4 NPs were functionalized with a 13-mer peptide (NP-Pep; HWQIAYNEHQWQC)[78, 236] with high binding affinity (K_d = 270 pM) towards cTnI under non-denaturing condition. Native cTn complex enrichment yielded TnC, cTnI, and cTnT in approximately a 1:1:1 ratio (**Figure S3.2, Figure S3.3a**), reflecting the composition of the heterotrimeric cTn complex in the sarcomere. Moreover, cTn could be enriched directly from human heart tissue across multiple NP-Pep synthetic batches and replicates (**Figure S3.2**). To demonstrate the enrichment and purification of all three cTn subunits while preserving their proteoforms, we further used an online reversed-phase liquid chromatography (RPLC) top-down tandem MS (MS/MS) method comparing the initial sarcomere protein loading mixture (L), the resulting flow through (F), and the final elution mixtures (E) (**Figure S3.3b**). Not only were all three subunits of the cTn complex, cTnI, cTnT and TnC, significantly enriched, but also the PTM profiles of endogenous cTnI, cTnT, and TnC were faithfully preserved, without introducing artifactual modifications (**Figure S3.4**). The top-down proteomics results provide a bird's eye view of the proteoform landscape of TnC, cTnT, and cTnI for direct analysis of the stoichiometry of their various proteoforms. Moreover, this native NP enrichment strategy was found to be highly reproducible across three independent donor heart tissues (**Figure S3.5**).

We next investigated whether the enriched endogenous cTn complex could be analyzed under native conditions by SEC-MS using an online buffer exchange (OBE) method[86, 241] for rapid analysis of native protein complexes after separation from MS-incompatible buffers (**Figure S3.6**). Upon analysis of the resulting SEC-MS chromatograms, protein complexes were effectively separated from non-volatile buffer components within 7 minutes (**Figure S3.6**). Additionally, heterotrimeric cTn complexes (z = 16+-21+) were reproducibly enriched across three independent elution mixtures from the same heart tissue sample highlighting the reproducibility of this SEC-OBE native MS method (**Figure S3.7**).

3.3.3. Structural Heterogeneity of Endogenous cTn Complex Revealed by nTDMS

To comprehensively characterize the endogenous heterotrimeric cTn complex, we employed a nTDMS approach using an ultrahigh-resolution FTICR mass spectrometer for unequivocal proteoform sequencing and protein complex characterization. Native mass spectra of the enriched cTn complex revealed a charge state distribution of 18+ to 21+ (3800 *m/z* to 4300 *m/z*) with the most abundant charge state (z = 19+) detected between 4050 *m/z* to 4080 *m/z* (**Figure 3.2, Figure S3.8**). In-depth examination of the endogenous cTn complex revealed four unique complex proteoforms comprised of both covalent and non-covalent modifications (**Figure 3.2b**). All heterotrimer cTn complex proteoforms were identified with high mass accuracy (< 2 ppm) (**Figure 3.2b, Table S3.2**). Significantly, the predominant forms of the heterotrimeric cTn complex were revealed to comprise of mono-phosphorylated cTnT, mono-phosphorylated and bis-phosphorylated cTnI, and TnC with three bound Ca^{2+} ions (most abundant cTn complex MW = 77136 Da). These results show that the human cTn complex exists in structurally diverse molecular compositions in the sarcomere with highly heterogenous covalent and non-covalent modifications that yield a suite of different intact assemblies.

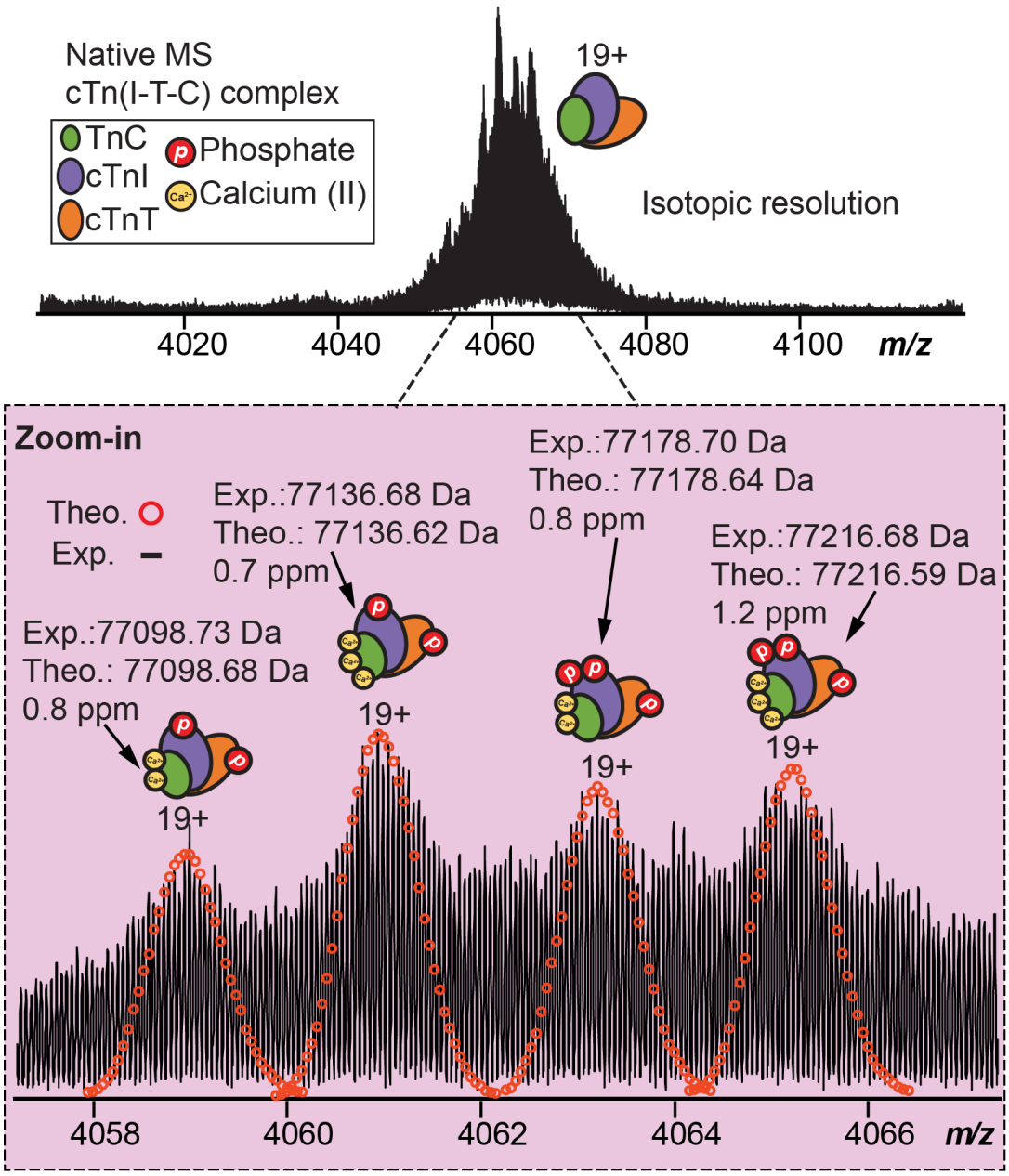


Figure 3.2. High-resolution native top-down MS analysis demonstrates heterogeneity of endogenous heterotrimeric cTn complexes. Native MS spectra of the endogenous cTn complex using ultrahigh-resolution FTICR-MS to achieve isotopic resolution. Most abundant charge state (z = 19+) is isolated. Zoomed-in mass spectra shows the isotopically resolved cTn complex proteoforms detected. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy. All individual ion assignments are within 2 ppm from the theoretical mass.

3.3.4. Sequence-Specific Structural Elucidation of cTn Complex Quaternary Structure through nTDMS

Next, we performed complex-up native MS analysis[242] with collisionally activated dissociation (CAD)[121] to elucidate the stoichiometry and composition of the heterotrimeric cTn complex (**Figure 3.3**). First, the intact heterotrimeric cTn complex (*z* = 18+-20+) was isolated with no dissociated cTn subunits present in the resulting mass spectra (**Figure 3.3a**). Subunit ejection of the cTn complex was then observed by CAD (**Figure 3.3b**). We detected the intact cTn(I-C) dimer at 2700 *m/z* to 3600 *m/z* (*z* = 12+-16+, MW = 42556 Da), cTnT monomer at 2500 *m/z* to 3200 *m/z* (*z* = 11+-14+, MW = 34580 Da), and TnC monomer at 2600 *m/z* to 3100 *m/z* (*z* = 6+-7+, MW = 18520 Da). Ejected cTnI monomer was not detected in appreciable abundance. We suspect that this may be due to maintaining native conditions during our MS analysis which seems to preserve the strong binding affinity of cTnI to TnC.[243, 244] Additionally, it is possible that the ejected cTnI monomer may appear at too low of a charge (*i.e.* too high of *m/z* range) for us to meaningfully detect during our complex-up native MS analysis.

To gain additional structural and sequence-specific information of the endogenous cTn complex, we performed complex-down analysis[242] of the dissociated cTn subunits (**Figure 3.3c-d**). The isolation of ejected cTnT monomers (*z* = 14+) revealed multiple proteoforms, including unphosphorylated cTnT, monophosphorylated cTnT (*p*cTnT), and phosphorylated cTnT with C-terminal Lys truncation (*p*cTnT [aa 1-286]) (**Figure 3.3c**). Further CAD fragmentation revealed fragments which suggest that the *C*-terminus of cTnT is more solvent exposed than the *N*-terminus forming the heterotrimer interface (**Figure 3.3c, Figure S3.9**). Next, the isolation of ejected cTn(I-C) dimer (*z* = 14+) revealed six unique proteoforms with different Ca^{2+} occupancy and phosphorylation states: unphosphorylated cTnI associated with TnC and 2 Ca^{2+}, unphosphorylated

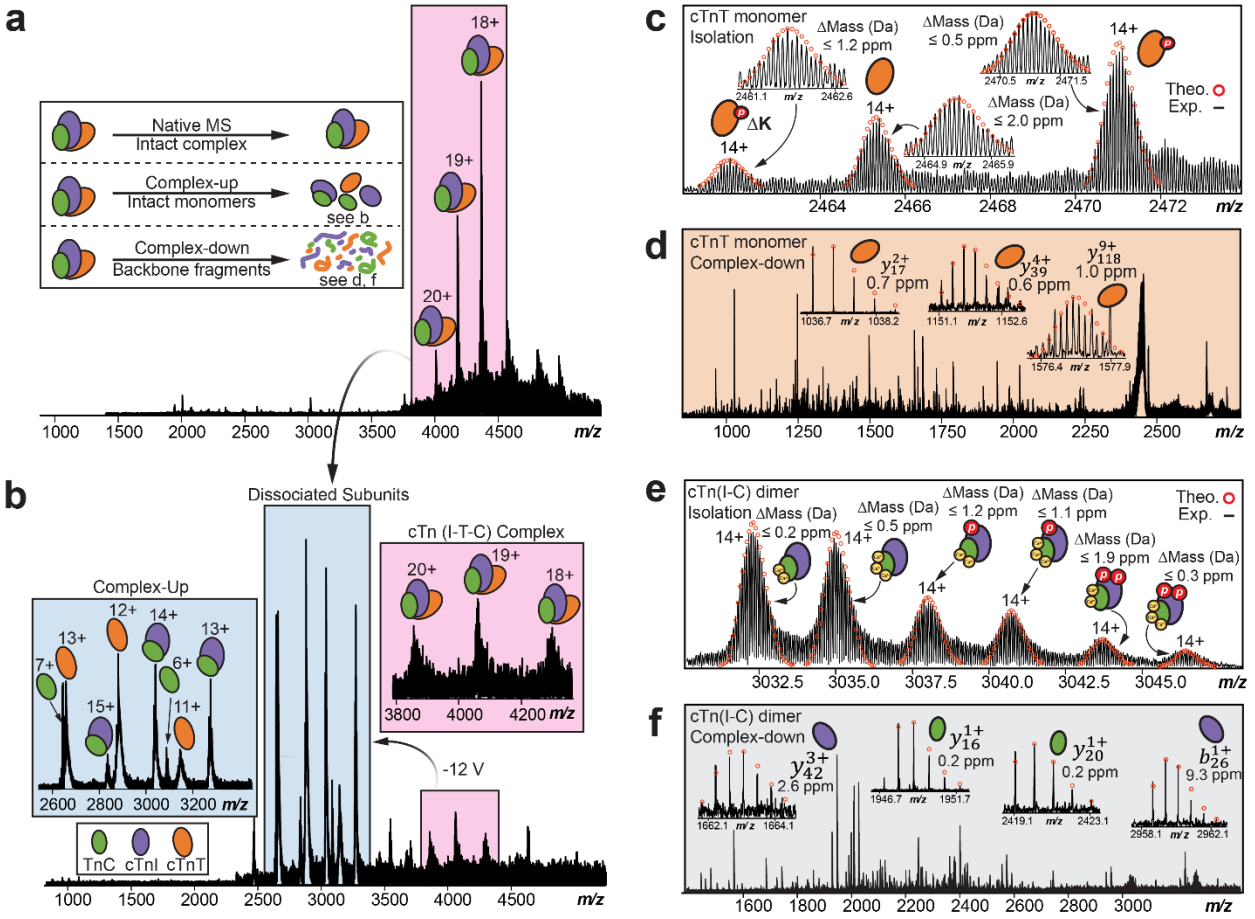


Figure 3.3. Native top-down MS (nTDMS) analysis for the characterization of cTn complexes directly from human heart tissue. (a-f). (a) Isolation of the heterotrimeric cTn complex achieved by ultrahigh-resolution FTICR-MS with charge states labeled. Inset shows three common approaches used for nTDMS analysis. **(b)** Complex-up analysis of the cTn complex ejects cTn monomers and dimer to reveal heterotrimeric complex stoichiometry. **(c)** Isolation of ejected cTnT native monomer (z = 14+) followed by complex-down analysis using collisionally activated dissociation (CAD) fragmentation. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy. All individual ion assignments are < 2 ppm from the theoretical mass. **(d)** Representative MS/MS spectra and CAD fragment ions ($y_{17}^{2+}, y_{39}^{4+}, y_{118}^{9+}$) obtained from the nTDMS analysis. **(e)** Isolation of ejected cTn native dimer (z = 14+) followed by complex-down analysis using CAD fragmentation. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy. All individual ion assignments are within 2 ppm from the theoretical mass. **(f)** Representative MS/MS spectra and CAD fragment ions (cTnI: y_{42}^{3+}, b_{26}^{1+}, TnC: y_{16}^{1+}, y_{20}^{1+}) obtained from the nTDMS analysis.

cTnI associated with TnC and 3 Ca^{2+}, monophosphorylated cTnI (*p*cTnI) associated with TnC and 2 Ca^{2+}, *p*cTnI associated with TnC and 3 Ca^{2+}, bisphosphorylated cTnI (*pp*cTnI) associated with TnC and 2 Ca^{2+}, and *pp*cTnI associated with TnC and 3 Ca^{2+} (**Figure 3.3d**). The cTn(I-C) dimer precursor ions (z = 14+) were further subjected to CAD fragmentation requiring high collision voltage (70 V) to obtain additional structural information (**Figure 3.3d, Figure S3.10**). We observed phosphorylation of dissociated cTnI monomer at Ser22 and Ser23, which are the targets of PKA-mediated phosphorylation.[35] Our nTDMS analysis also suggests that both the intrinsically disordered *C*- and *N*-termini of cTnI are more solvent exposed than the stable internal regions[231] that form the subunit-subunit interfaces of the cTn complex. While the relative stability of the different cTn molecular forms could be inferred based on how they responded to the nTDMS analysis, we could not infer exactly how Ca^{2+} binding and phosphorylation impacted the stability of the cTn complexes. Moreover, we did not observe lowly charged "stripped" cTn heterotrimer complexes that is expected at a much higher *m/z* likely due to the limitation of our FTICR mass spectrometer which is not optimized for detection of ions at very high *m/z*.[84]

3.3.5. Direct Localization of Ca^{2+} Binding Domains in Endogenous cTn Complex

We have characterized the three Ca^{2+} binding domains present in endogenous TnC using nTDMS (**Figure 3.4**). TnC is an EF-hand Ca^{2+}-binding protein that is an essential Ca^{2+} sensing molecular subunit in the heterotrimeric cTn complex.[227, 228] Human TnC consists of three functional metal-binding motifs (domains II-IV) that can be occupied by Ca^{2+} and are responsible for regulating cardiac muscle contraction.[235] Ca^{2+} binds to domain II with low-affinity (~10^{-5} M) and serves as the "regulatory" domain for initiating cardiac contraction.[245] In contrast, domains III and IV, referred to as "structural" sites, bind Ca^{2+} with high-affinity (~10^{7} M^{-1}).[246] Domains III and IV consistently remain saturated with Ca^{2+} during relaxation and contraction, whereas Domain

II is only occupied during contraction.[247] Notably, domains III and IV exhibit a slower Ca^{2+} exchange rate than domain II.[246, 248] The nTDMS isolation spectra of ejected TnC monomer (z = 8+) revealed multiple proteoforms of endogenous TnC in its Ca^{2+}-bound states (**Figure 3.4a**). Specifically, we observed TnC with 0, 1, 2, and 3 Ca^{2+} ions bound with baseline isotopic resolution and high mass accuracy. The relative proportion of singly, doubly, and triply bound Ca^{2+} states were found to be approximately 0.2, 0.5, and 0.1, respectively (**Figure S3.11**).

The essential amino acid regions necessary for an individual Ca^{2+} ion to associate with TnC domains were next determined through nTDMS analysis. Aspartic acid and glutamic acid have the highest binding affinities for Ca^{2+} at neutral pH.[235,249] To localize Ca^{2+} binding domains, TnC proteoforms were subjected to CAD to generate extensive backbone fragmentation through collisional activation ramping (**Figure 3.4b-c, Figure S3.12**). We found TnC domain III to be the least vulnerable Ca^{2+} binding to collisional activation, while domain II was found to be the most vulnerable Ca^{2+} binding region. To localize the primary region for Ca^{2+} binding that is least vulnerable to collisional activation, we isolated the TnC proteoform at 2312 *m/z* and performed CAD to yield product ions y_{52} + Ca^{2+}, y_{30}, b_{115} + Ca^{2+}, and b_{109} (**Figure S3.13**). Therefore, the primary Ca^{2+} binding domain was localized to ^{113}DLD^{115} in domain III. The next Ca^{2+} binding domain was localized to the structural region between ^{141}DKNND^{145} in domain IV by first isolating the TnC proteoform at 2316 *m/z* and then generating CAD product ions b_{140} + Ca^{2+}, b_{145} + 2 Ca^{2+}, y_{16}, and y_{22} + Ca^{2+} (**Figure S3.14**). Finally, the most vulnerable Ca^{2+} binding region was localized to regulatory domain II between ^{73}DFDE^{76} by isolating the TnC proteoform at 2321 *m/z* and generating CAD product ions b_{65}, b_{91} + Ca^{2+}, y_{85} + 2 Ca^{2+}, and y_{94} + 3 Ca^{2+} (**Figure S3.15**). This is the first study to localize endogenous TnC Ca^{2+} binding regions to domains II, III, and IV in the

cTn complex (**Figure 3.4d)**. Additionally, a molecular depiction of the divalent association of Ca^{2+} to amino acid residues ^{73}D and ^{76}E in TnC domain II is illustrated in **Figure 3.4e**.

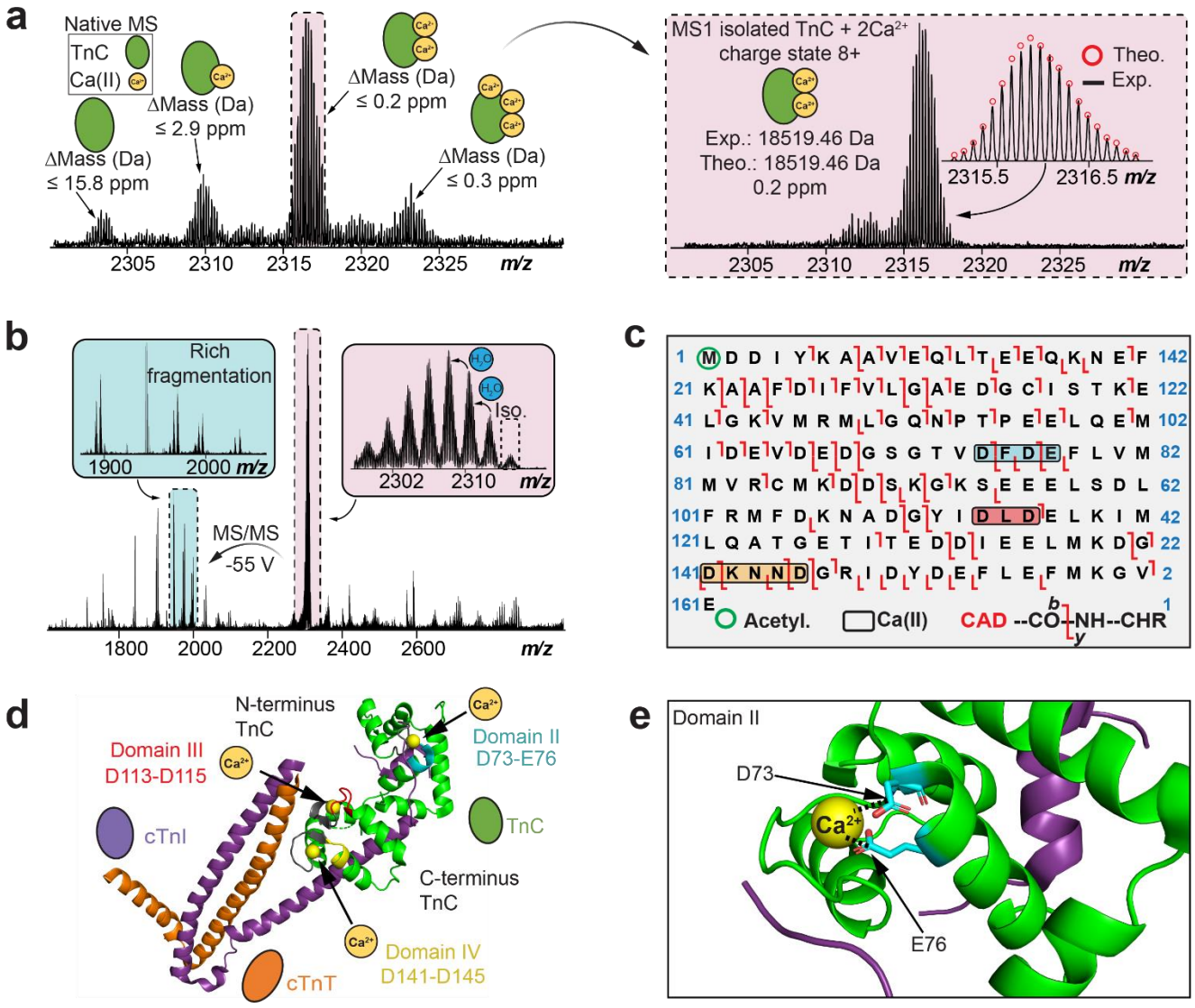


Figure 3.4. Native top-down MS (nTDMS) enables localization of endogenous Ca^{2+} binding domains within TnC. (a-e) (a) Isolation of native TnC monomer (z = 8+) showing proteoforms with differential Ca^{2+} binding using ultrahigh-resolution FTICR-MS. (Inset) Isolation of TnC + 2 Ca^{2+}proteoform (z = 8+). Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy proteoform characterization. All individual ion assignments are within 16 ppm from the theoretical mass. **(b)** MS/MS characterization of TnC + 2 Ca^{2+} proteoform isolated from the quadrupole window centered at 2316 *m/z* by collisionally activated dissociation (CAD) fragmentation. **(c)** nTDMS CAD fragmentation map showing the localization of three Ca^{2+} ions and N-terminal acetylation in

TnC monomer achieving ~ 63% total bond cleavage. We observed sequential binding of Ca^{2+} ions to TnC with Ca^{2+} first binding to domain III (red), then domain IV (yellow), and lastly domain II (blue). **(d)** Structural representation of the cTn complex with the experimentally defined Ca^{2+} binding domains (II-IV) highlighted (domain II, blue; domain III, red; domain IV, yellow, UniprotKB annotations, gray). TnC is depicted in green, cTnI in purple, cTnT in orange, and Ca^{2+} ions in yellow. PDB: 1J1E. **(e)** Magnified view of TnC Ca^{2+} binding in domain II showing possible coordination with amino acid residues D73 and E77. PDB: 1J1E.

3.3.6. Determination of cTn-Ca^{2+} Binding and Conformational Dynamics

Binding of Ca^{2+} to TnC has extensive effects to the heterotrimeric cTn complex function and structure in regulating cardiac contraction.[222, 226, 228, 250] The core of the heterotrimeric cTn complex maintains a stable conformation, while flexible regions undergo extensive conformational changes when Ca^{2+} binds to TnC.[230] This transition shifts the cTn complex from a 'closed' state where muscle contraction is prevented by cTnI binding to actin, to an active 'open state'. In this active state, the N-terminal domain of TnC opens, allowing for the binding of cTnI, and facilitating the interaction between actin and myosin, ultimately leading to cardiac muscle contraction. Due to the conformational heterogeneity and presence of intrinsically disordered regions along the heterotrimeric cTn complex, it is challenging to obtain crystal structures of cTn in its active and closed states upon Ca^{2+} binding using traditional structural biology techniques.[226] Therefore, to assess the intricate relationship between cTn-Ca^{2+} binding and conformational dynamics, we performed native TIMS-MS analysis. TIMS positions ions in an electrical field against a moving buffer gas to determine ion mobility ($1/K_0$) values for structural analysis of native proteins and protein complexes.[41, 110, 251] The measured mobilities are converted into collision cross-section (CCS) values to provide structural information on ion shape and size to evaluate conformational changes in the protein three-dimensional structure.[252, 253]

First, the native TIMS-MS parameters were optimized using bovine serum albumin (BSA, ~ 66 kDa) due to BSA having similar *m/z* and ion mobility regions as the cTn complex. Modifying the desolvation parameters proved critical for effective high-resolution ion mobility separation of BSA conformers (**Figure S3.16**). We next used TIMS-MS to separate and analyze the endogenous TnC monomer and cTn(I-C) dimer Ca^{2+}-bound proteoforms (**Figure 3.5**). The CCS values for TnC monomer with 0, 1, 2, and 3 Ca^{2+} ions bound were determined to be 1853, 1849, 1829, and 1844 $Å^2$, respectively (**Figure 3.5a-b**). The experimentally obtained CCS values revealed a more compact conformation for TnC than was predicted by IMPACT[254] calculation derived from available TnC crystal structures (**Table S3.3**). On the other hand, the CCS values for cTn(I-C) dimer with 2 (3623 $Å^2$) and 3 (3640 $Å^2$) Ca^{2+} ions were in excellent agreement with the calculated IMPACT CCS values (**Figure 3.5c-d, Table S3.3**). For TnC monomer, the structure remains in an open conformation in its apo and 1 Ca^{2+} bound states, whereas TnC shifts to a closed state upon 2 Ca^{2+} ions binding (**Figure 3.5e**). Interestingly, TnC shows a slightly more open conformation upon 3 Ca^{2+} ions binding. We reason that TnC with 2 Ca^{2+} ions bound forms a more compact conformation because Ca^{2+} structural binding domains III and IV are typically saturated at resting diastolic concentrations, but when the third Ca^{2+} ion binds to domain II, TnC adopts a more open conformation to prepare the *N*-terminal region of TnC for binding to the *C*-terminal region of cTnI to initiate cardiac muscle contraction.[228, 255] Likewise, binding of three Ca^{2+} ions to cTn(I-C) dimer shifts the dimer to a more open conformation when compared to the 2 Ca^{2+} bound state (**Figure 3.5f**). These results suggest that cTnI and TnC form a more open conformation when saturated with Ca^{2+} ions in preparation for cTnT engagement along the thin filament for subsequent cTn complex formation and cardiac muscle contraction.[256]

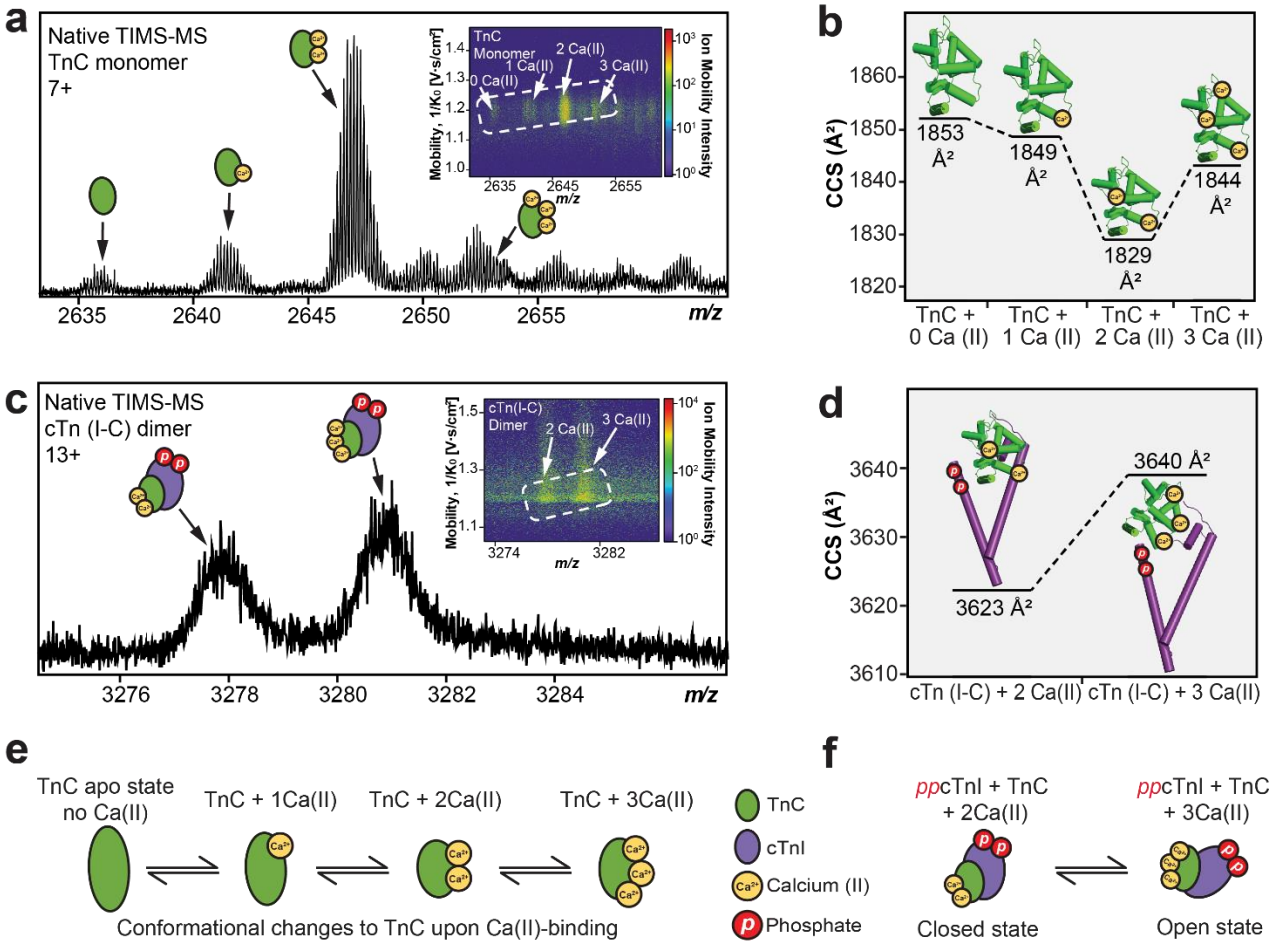


Figure 3.5. Ion mobility separation of TnC monomer and cTn dimer Ca^{2+} bound proteoforms by native TIMS-MS. (a-f) (a) TnC monomer proteoforms (z = 7+) from TIMS-MS analysis. (Inset) Associated ion mobility heat map showing differences in ion mobility and intensity for each different Ca^{2+} bound proteoform. **(b)** Collision cross section (CCS) of the TnC proteoforms obtained under nitrogen (N_2) drift gas. **(c)** cTn dimer proteoforms (z = 13+) from TIMS-MS. (Inset) Associated ion mobility heat map showing differences in ion mobility and intensity for each different Ca^{2+} bound proteoform. **(d)** CCS of the observed cTn dimer proteoforms obtained under N_2 drift gas. **(e)** Illustrations of TnC monomer and **(f)** cTn dimer conformational changes upon Ca^{2+} binding.

The structural roles of Ca^{2+} in maintaining the intact heterotrimer stability were further profiled by native TIMS-MS analysis (**Figure 3.6**). We added EGTA to sequester Ca^{2+} from the intact heterotrimer complex and probe the specific contribution of individual Ca^{2+} binding regions to cTn complex stability (**Figure 3.6a**). First, the heterotrimeric cTn complex (z = 18+ to 21+) was resolved by native TIMS-MS analysis (**Figure 3.6b-c**) without EGTA incubation. The CCS of the

native cTn complex with all Ca^{2+} occupied was determined to be 4880 $Å^2$ which was comparable to the CCS value calculated using the IMPACT[254] method (4192 $Å^2$) and is in good agreement with the previously reported partial crystal structure[226] that is truncated due to intrinsically disordered regions present within the cTn complex (**Table S3.3**). Under increasing concentrations of EGTA (25, 50, and 100 mM) to sequester cTn-bound Ca^{2+}, the intact cTn complex was gradually unfold, as evidenced by an increase in the average charge stage as well as a shift in the charge state envelope to lower *m/z* regions for the cTn complex. TIMS analysis of the cTn complex revealed protein conformer enhancement and loss of protein complex stability which correlated with increasing EGTA concentration (**Figure 3.6 d-f, Table S3.4**). Additionally, plotting the experimentally determined CCS values as a function of charge state for TnC monomer, cTn(I-C) dimer, and cTn(I-T-C) heterotrimer illustrates the linear relationship between charge state and native protein conformation for each cTn molecular form (**Figure S3.17**). While TIMS analysis offers comprehensive insights into protein complex size and stability, determining the precise extent of Ca^{2+} binding within the cTn complex remains challenging, as only an average mass for the ensemble can be attained. We suspect that, despite incubating the cTn complex with 100 mM of EGTA, Ca^{2+} may not be entirely stripped from the complex. The presence of cTnI significantly enhances the sensitivity of TnC to Ca^{2+} binding, potentially preventing complete Ca^{2+} removal from the cTn complex due to the equilibrium between EGTA-Ca^{2+} K_d and cTn-Ca^{2+} K_d being of a similar order. Overall, native TIMS-MS analysis of endogenous TnC monomer, cTn(I-C) dimer, and the cTn heterotrimer gas-phase structures provide insights into the conformational changes that occur upon cTn-Ca^{2+} binding during cardiac contraction.

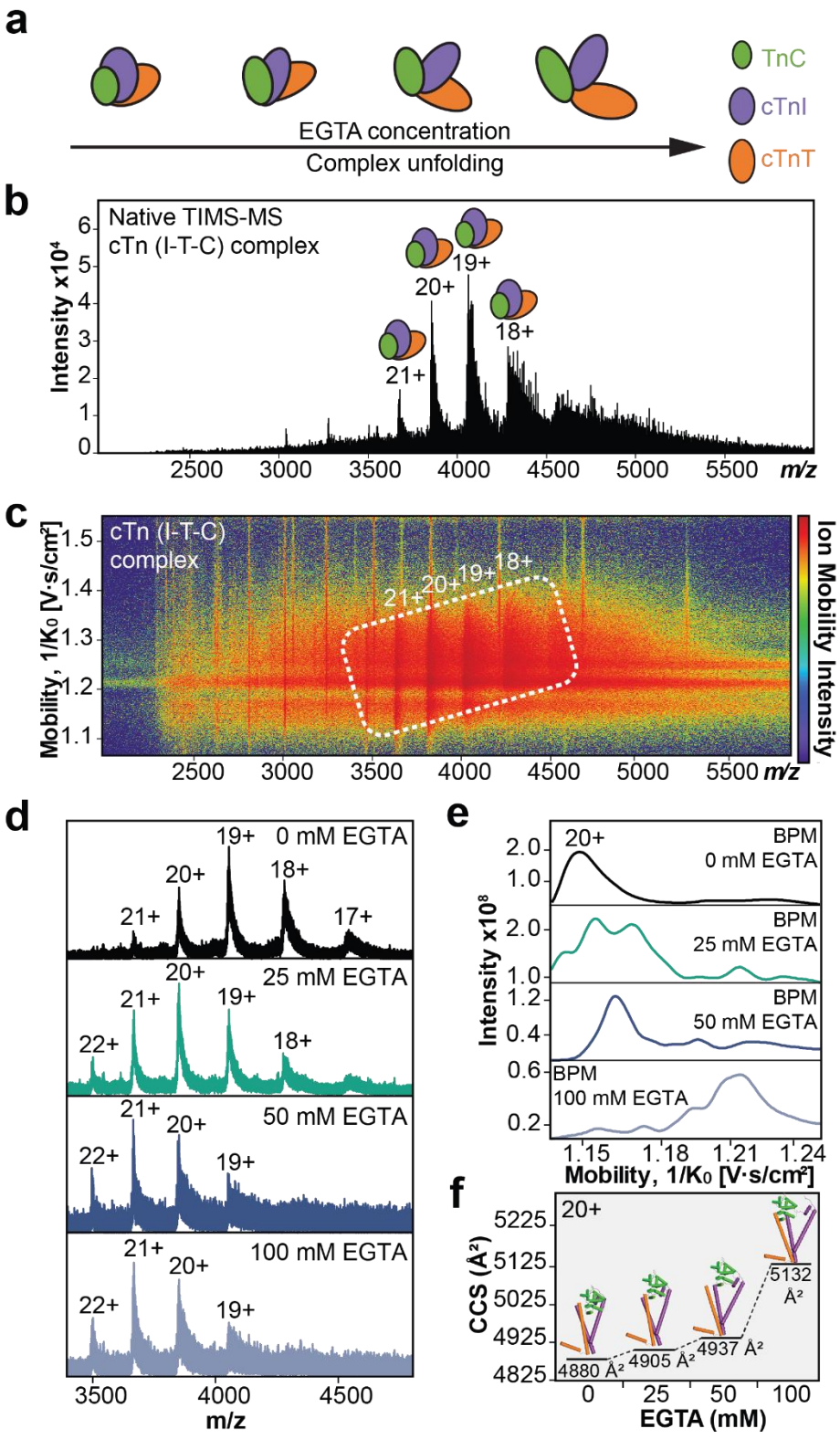


Figure 3.6. Addition of a Ca^{2+} chelator to the cTn complex provides insights into cTn-Ca^{2+} binding dynamics. (a-f) (a) Illustration of the cTn complex unfolding when incubated with increasing concentrations of EGTA; a Ca^{2+} ion chelator that sequesters cTn-bound Ca^{2+}. **(b)** Mass spectrum of the cTn

complex (z = 18-21+) by native TIMS-MS. **(c)** Associated ion mobility heat map corresponding to the data presented in b. **(d)** Enriched cTn complexes were incubated with 0, 25, 50, and 100 mM EGTA for 3h at 4°C and then analyzed by native TIMS-MS. **(e)** Base peak mobilogram spectra (BPM) of cTn complexes incubated with EGTA (z = 20+). **(f)** Collision cross section (CCS) analysis of cTn complexes incubated with EGTA (z = 20+) obtained under N_2 drift gas. The loss of protein complex stability directly correlates with increasing EGTA concentration (25 to 100 mM).

3.4 Discussion

Here we have developed a native nanoproteomics platform for enrichment/purification and structural characterization of endogenous protein complexes and their non-covalent interactions in their native state together with comprehensive proteoform mapping. This approach addresses challenges in the current nTDMS field in the isolation and analysis of endogenous protein complexes. By avoiding denaturation or digestion steps, nTDMS quickly emerges as a powerful structural biology tool providing insights into the endogenous protein complexes and their functional states.[178] However, nTDMS studies have primarily relied on purification strategies using overexpressed recombinant proteins and/or highly abundant proteins due to the difficulties in isolating endogenous protein complexes.[45, 48, 50, 129, 257, 258] To our knowledge there has been no previous nTDMS study to structurally characterize heterogenous endogenous protein complexes directly from heart tissue samples.

We chose to apply the native nanoproteomics method to characterize the structure and dynamics of the endogenous cTn complex given its high significance in cardiac function and clinical diagnosis. The structure of the cTn complex has been previously investigated by traditional structural biology techniques but were limited to resolving only the primary domains of the complex.[226, 230] Moreover, it is challenging to characterize the dynamic structural changes of endogenous cTn-Ca^{2+} binding events and PTMs directly from human samples by these

methods.[231, 232, 259, 260] While the influence of Ca^{2+} exchange on cTn stability and function has been suggested, previous studies have often employed recombinantly expressed cTn subunits to reconstruct cTn complexes or cardiac thin filaments. These constructs frequently lack PTMs and do not directly assess the entire cTn heterotrimeric complex. In contrast, our nTDMS results offer indispensable structural insights of the intact endogenous cTn complex. Notably, our native nanoproteomics study identified 17 endogenous cTn proteoforms directly from human heart tissue, while preserving PTMs including N-terminal acetylation and phosphorylation, along with non-covalent interactions such as Ca^{2+} binding with high resolving power and mass accuracy. We observed sequential binding of Ca^{2+} ions to TnC with the primary and secondary binding regions in domain III (^{113}DLD^{115}) and domain IV (^{141}DKNND^{145}) near the *C*-terminus, respectively. Whereas the tertiary Ca^{2+} binding region was localized to domain II (^{73}DFDE^{76}) at the *N*-terminus of TnC. Ca^{2+} binding regions along TnC have been previously reported on the UniprotKb database as ^{65}DEDGSGTVDFDE^{76} in domain II, ^{105}DKNADGYIDLDE^{116} in domain III, and ^{141}DKNNDGRIDY^{152} in domain IV. In contrast, we provide a substantial refinement to the precise Ca^{2+} binding region in TnC, to just within a few amino acid residues, which demonstrates the high-resolution capabilities of this native nanoproteomics approach for localizing non-covalent metal binding. Given that the binding of Ca^{2+} to domain IV occurs in a sequential manner subsequent to the initial Ca^{2+} binding event in domain III, our findings may hold considerable biological significance. Prior investigations have elucidated that mutations within the Ca^{2+} binding to domains III and IV substantially diminish the affinity of TnC for the regulatory region of cTnI.[261, 262] Consequently, disruption of the sequential Ca^{2+} binding event could initiate deleterious physiological consequences, including the potential development of cardiomyopathies. Finally,

our results resolve the structural roles of Ca^{2+} binding in regulating conformational changes in TnC monomer, cTn(I-C) dimer, and cTn(I-T-C) heterotrimer.

3.5 Conclusions

Overall, this native nanoproteomics approach opens new opportunities for the enrichment/purification and structural characterization of endogenous protein complexes to reveal their native assemblies, proteoform landscape, and dynamical non-covalent binding. This approach utilizes functionalized NPs designed to selectively bind to specific target proteins, and the enriched protein complexes can preserve their native state when enrichment and elution conditions are established. This native nanoproteomics platform can be readily adaptable to other protein targets when the appropriate affinity ligand is identified and utilized, which will require the development and optimization of new functionalized nanomaterials. We envision the integration of designer nanomaterials with nTDMS can serve as a powerful structural biology tool in the analysis of endogenous protein complexes in their native states that are complementary to traditional biophysical techniques.

3.6 Methods

3.6.1. Synthesis of N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide

3-Butynoic acid (2.10 g, 1 eq., 25 mmol), 2-chloro-1-methylpyridinium iodide (7.00 g, 1.1 eq., 27 mmol), and dichloromethane (250 mL) were added to a round bottom flask. Next, a solution of (3-aminopropyl)triethoxysilane (5.53 g, 1 eq., 25 mmol), N,N-diisopropylethylamine (6.46 g, 2 eq., 50 mmol), and dichloromethane (125 mL) was prepared separately and added to the stirring round bottom flask. The reaction mixture was allowed to reflux for 1 hour. The crude product was concentrated *in vacuo* and partially redispersed in ethyl acetate, allowing for the separation of BAPTES from the insoluble 2-chloro-1-methylpyridinium iodide by gravity filtration. The product was concentrated *in vacuo* and purified twice with column chromatography using a gradient of 50:50 to 100:00 ethyl acetate: hexane to yield BAPTES and its alkyne isomer as a clear, orange oil (4.80 g, 67% yield). To effectively isomerize the propargylic isomer to N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide, anhydrous acetonitrile (120 mL), BAPTES isomers (3.38 g, 1 eq., 11.78 mmol), and tribasic potassium phosphate (2.50 g, 1 eq. 11.78 mmol) were added to a round bottom flask. The reaction was allowed to stir at room temperature for 2 hours before gravity filtration and subsequent concentration *in vacuo*. The product was purified with flash column chromatography using a gradient of 70:30 to 100:00 ethyl acetate: hexane to yield a clear, yellow oil. The final product was concentrated *in vacuo* for further use; ^{1}H NMR (500 MHz, Chloroform-*d*) δ (ppm) 5.99 (s, 1H), 5.62 (t, *J* = 6.6 Hz, 1H), 5.20 (d, *J* = 6.7 Hz, 2H), 3.83 (q, *J* = 7.0 Hz, 6H), 3.31 (q, *J* = 6.7 Hz, 2H), 1.66 (m, 2H), 1.24 (t, *J* = 7.0 Hz, 9H), 0.66 (m, 2H); ^{13}C NMR (125 MHz, Chloroform-*d*) δ (ppm) 211.56, 164.33, 91.02, 80.33, 58.47, 42.02, 23.02, 18.30, 7.66; ESI-MS for $C_{13}H_{25}NO_4Si$ $[M+H]^+$ observed: 288.162 *m/z*, $[M+H]^+$ calculated: 288.162 *m/z*. (2.30 g, 68% yield).

3.6.2. Synthesis of Iron-Oleate Precursor

Iron oleate was synthesized using a previously established method.[78] In a typical synthesis, iron (III) chloride hexahydrate (10.8 g, 40 mmol) was first dissolved in a mixture of 80 mL ethanol and 60 mL nanopure water in a three-neck round bottom flask (500 mL) containing a Teflon-coated egg-shaped (1-1/4" x 5/8") magnetic stir bar. Sodium oleate (36.5 g, 120 mmol) was then quickly added to the iron chloride solution along with 140 mL n-hexane. The resulting solution was then allowed to stir until the sodium oleate was completely dissolved. Afterwards, the reaction solution was heated to 70 °C for a 4-h reflux under a N_2 blanket. Upon completion, the reaction solution was cooled to room temperature, and the upper organic layer containing the iron oleate was washed three times with 30 mL nanopure water in a 250 mL separatory funnel. After washing, the iron oleate was concentrated *in vacu*o. Finally, the resulting iron oleate was transferred into a 100 mL round bottom flask, connected to a Schlenk line, and placed under vacuum overnight. For storage, the iron oleate was well sealed in a glass vial and placed in a desiccator.

3.6.3. Synthesis of 8 nm Magnetite (Fe_3O_4) Nanocrystals

Iron oleate precursor and magnetite nanoparticles were synthesized following published protocol, with minor modifications.[78] Briefly, the NPs were synthesized using 10 mmol (9.0 g) iron oleate, 5.5 mmol (1.56 g) oleic acid, and a 4:1 ODE:TDE (40 g: 10 g) solvent mixture. The mixture was stirred and degassed on a Schlenk line at 110 °C for 3 h, heated to 300 °C at a heating rate of 3.3 °C/min under a nitrogen flow, and then allowed to boil for 30 min. After three centrifugation wash cycles (10,000 × *g*, 20 min) using EtOH, the resulting NPs were then dried under vacuum, weighed, and redispersed in n-hexane at a concentration of 20 mg/mL for further use.

3.6.4. Synthesis of Fe_3O_4-BAPTES NPs (NP-BAPTES)

In a typical optimized large-scale synthesis of silane functionalized NPs, Fe_3O_4 NPs (6 mL from a 20 mg/mL stock) were added to anhydrous n-hexane (300 mL) in a 500 mL round bottom flask equipped with a Teflon-coated egg-shaped magnetic stir bar (1–1/4" × 5") to achieve a total NP concentration of 0.4 mg/mL. After the reaction mixture was heated to 60 °C with stirring (900 rpm), BAPTES (1.65 mL) was added dropwise to the flask for a 0.55% (v/v) total concentration of trialkoxysilane reagent, followed by dropwise addition of a small amount of acetic acid (30 μL) for an acidic catalyst concentration of 0.01% (v/v). After reacting at 60 °C for 24 h, the precipitate was collected and washed one time with n-hexane, one time with n-hexane/acetonitrile (v/v, 4:1), and one more time with n-hexane via centrifugation (10,000 × *g*, 10 min) to remove excess silane molecules and surfactants. The NPs were then dried under vacuum for later use.

3.6.5. Synthesis of Fe_3O_4-BAPTES-Peptide NPs (NP-Pep)

10 mg of NP-BAPTES were added to a 4-dram vial and dispersed in 2 mL of acetonitrile. 10 mg of cTnI-binding peptide (HWQIAYNEHQWQC)[236] were added to a separate 4-dram vial and dissolved in 8 mL of nanopure water. The pH of the peptide solution was adjusted to pH 8.0 by the addition of 75 μL of 1.0 M ammonium carbonate buffer pH 9.0 during simultaneous water bath sonication. The pH-adjusted peptide solution was added into the 4-dram vial containing the NP dispersion under water bath sonication. The NP reaction mixture was allowed to react under sonication for 1 h and later collected into Eppendorf tubes for washing. The peptide-functionalized NPs were washed three times with nanopure water via centrifugation (10,000 × *g*, 5 min) and subsequently isolated magnetically with a DynaMag to remove unreacted peptide. The resulting peptide functionalized NPs were redispersed in nanopure water at a concentration of 5 mg/mL.

3.6.6. Human Cardiac Tissue Collection and Ethical Compliance

Left ventricular (LV) myocardium from non-failing donor hearts with no history of heart diseases were used and obtained from the University of Wisconsin Organ and Tissue Donation-Surgical Recovery and Preservation Services. Donor heart tissues were stored in cardioplegic solution prior to dissection and snap-frozen immediately in liquid nitrogen and stored at -80 ℃. We have complied with all ethical regulations related to this study. All experiments on human samples followed all relevant guidelines and regulations. The procedures for the collection of human donor heart tissues were reviewed and approved by the University of Wisconsin - Madison Institutional Review Board (Protocol number 2013-1264). Research with human tissue has been conducted according to the principles of the Declaration of Helsinki. Clinical characteristics of the donor heart tissues used can be found in **Table S3.1**.

3.6.7. Native Protein Extraction

Cardiac tissue was first homogenized on ice using a Polytron homogenizer in 10 volumes (mL/g tissues) of native wash buffer (5 mM NaH_2PO_4, 5 mM Na_2HPO_4 (pH 7.0), 100 mM NaCl, 125 mM L-Met (pH 7.5), 1 mM PMSF and 1X MS-Safe protease and phosphatase inhibitor cocktail) to deplete the highly abundant cytosolic proteins. The homogenate was centrifuged at 17,000 × g for 3 min at 4°C, and the supernatant was discarded. The washing and homogenization step was repeated once more, and then the supernatant was discarded. To extract proteins such as the cardiac troponin (I-T-C) complex from human heart tissues without potentially denaturing the proteins' tertiary or quaternary structure, a high ionic strength LiCl buffer at physiological pH was used (25 mM Tris (pH 7.5), 700 mM LiCl, 125 L-Met (pH 7.5), 1 mM PMSF and 1X MS-Safe protease and phosphatase inhibitor cocktail. The resulting pellet was washed, centrifuged, and homogenized in 5 vol (mL/g tissue) of LiCl native extraction buffer, then incubated at 4 °C for 10

min to extract the sarcomeric proteome. The homogenate was centrifuged at 17,000 × *g* for 3 min at 4°C and the supernatant containing the sarcomeric proteome was transferred to new Eppendorf Protein Lo-Bind tubes, and further centrifuged at 21,000 × *g* for 30 min at 4 °C to clarify the extracts. The supernatants were finally transferred to new Eppendorf Protein Lo-Bind tubes, snap-frozen in liquid nitrogen, and stored at -80°C.

3.6.8. Native Purification of Endogenous Protein Complexes using Peptide-Functionalized Nanoparticle (NP-Pep) Enrichment

NP-Pep (5 mg) originally dispersed in H_2O was redispersed in a 2 mL Eppendorf Protein Lo-Bind tube with 1 mL of equilibration buffer (25 mM Tris (pH 7.5), 700 mM LiCl, 15 mM L-Met). The NPs were then centrifuged at 21,000 × *g* for 2 min at 4 °C, isolated from the solution using the DynaMag, and the supernatant was removed. 1 mL of equilibration buffer was then added to the NPs, the mixture was sonicated and vortexed to prepare for protein loading. Protein loading mixture (L) from heart tissue extract was diluted to a final volume of 1 mL with a buffered solution (25 mM Tris (pH 7.5), 700 mM LiCl, 15 mM L-Met) to a total protein loading of 0.6 mg/mL and then the NP-Pep mixture was added to the protein loading mixture, at an NP concentration of 2.5 mg/ml. After this mixture was agitated on a nutating mixer at 4 °C for 1 h, the NPs were centrifuged at 21,000 × *g* for 2 min at 4 °C, and then isolated from the solution using the DynaMag. The supernatant was collected and saved as the flow-through (F) fraction. The isolated NPs were then washed three times with equilibration buffer (25 mM Tris (pH 7.5), 700 mM LiCl, 15 mM L-Met; 0.20 mL/mg NP-Pep) following the same centrifugation and magnetic isolation steps to remove unbound, nonspecific proteins. To elute the bound cTn(I-T-C) complex, 500 µL of 750 mM L-arginine, 750 mM imidazole, 50 mM L-glutamic acid (pH 7.5) was added, and allowed to incubate with NP-Pep with agitation on a nutating mixer at 4 °C for up to 10 min. After

centrifugation and magnetic isolation, the resulting supernatant was collected as the elution fraction (E). Prior to protein concentration, Amicon Ultra Centrifugal filters were equilibrated with 500 μL of 750 mM L-arginine, 750 mM imidazole, 50 mM L-glutamic acid (pH 7.5), and centrifuged at 14,000 × *g* for 5 min at 4 °C. The NP-Pep elution mixture was concentrated using the Amicon Ultra Centrifugal filter (30 kDa MWCO, 0.5 mL) at 14,000 × *g* for 10 min at 4 °C, and the concentrated protein mixture (approximately 40 μL) was either transferred to a chilled LC-MS vial for automated online buffer exchange into 200 mM ammonium acetate solution using size exclusion chromatography columns. For offline MS analysis, concentrated protein mixture was buffer exchanged into 200 mM ammonium acetate solution using Bio-Spin® columns with Bio-Gel® P-30.

3.6.9. Standard Protein Sample Preparation

Bovine serum albumin (BSA, 66 kDa) was purchased from Sigma-Aldrich (St. Louis, MO). Native protein samples were prepared by buffer exchanging into 200 mM ammonium acetate solution by washing the sample six times through an Amicon Ultra Centrifugal filter (10 kDa MWCO, 0.5 mL). The protein sample was then diluted to 20 μM in 200 mM ammonium acetate solution.

3.6.10. Calcium Binding Experiments using EGTA for cTn(I-T-C) Complex Stability

To probe the role and influence of Ca^{2+} co-factor association with respect to the stability of the cTn(I-T-C) complex, a metallic ion chelator with high affinity for Ca^{2+} (1.0 M EGTA, pH 7.5) was added to NP-Pep elution mixtures at various final concentrations (0, 25, 50, and 100 mM EGTA). After addition of EGTA, protein solutions were vortexed, pulse spun by tabletop centrifuge, and incubated on ice at 4 °C for 3 h. After incubation with EGTA, samples were concentrated, and buffer exchanged into 200 mM ammonium acetate solution for offline native MS analysis as described above.

3.6.11. Top-Down RPLC-MS/MS Analysis of cTnI, cTnT, and TnC Monomers

For top-down RPLC-MS/MS protein sample preparation of the cardiac protein L, F, and NP-Pep E, Amicon Ultra Centrifugal filters (10 kDa, 0.5 mL) were pre-equilibrated with a cartridge volume of 0.2% formic acid (FA) in water and centrifuged at 14,000 × *g* for 5 min at 4 °C. After the filters were equilibrated, protein sample was introduced, and the proteins were desalted by repeated dilution with 0.2% FA in water to the cartridge filter volume, and concentrated by centrifugation at 14,000 × *g* for 5 min at 4 °C. This dilution and concentration process was repeated for five additional times to effectively buffer exchange proteins into MS-compatible mobile phase and to remove non-volatile salts and buffers. Buffer exchanged samples were then transferred to chilled LC-MS vials for reverse phase (RP)LC-MS/MS analysis. Top-down RPLC-MS/MS was carried out by either using an Acquity ultra-high pressure LC M-class system (Waters) coupled to a high-resolution maXis II quadrupole time-of-flight (QTOF) mass spectrometer (Bruker Daltonics) or by using a nanoAcquity ultra-high pressure LC system (Waters) coupled to a high-resolution Impact II QTOF mass spectrometer (Bruker Daltonics). 600 ng of total protein was injected onto a home-packed PLRP column (PLRP-S) (Agilent Technologies), 10-µm particle size, 500-µm inner diameter, 1,000 Å pore size using an organic gradient of 20 to 65% mobile phase B (mobile phase A: 0.2% FA in H_2O; mobile phase B: 0.2% FA in 50:50 acetonitrile/isopropanol) at a constant flow rate of 12 µL/min. For the maXis II QTOF, mass spectra were taken at a scan rate of 0.5 Hz over 200-3000 *m/z* range. Ion source parameters, the end plate offset, and capillary voltage were 500 and 4500 V, respectively. The nebulizer gas pressure (N_2), dry gas flow rate, and dry gas temperature were set to 0.5 bar, 4.0 L/min, and 200 °C, respectively. For ion transfer parameters, the funnel RF, octupole RF, quadrupole energy, low mass cutoff, and in-source collision-induced dissociation (isCID) energy were optimized to 400 V, 400 V, 5 V, 500 *m/z,* and

10 V, respectively. Additionally, collision Vpp, pre-pulse storage time, and collision cell transfer time were set to 2000 Vpp, 15 µs, and 100 µs, respectively. For the Impact II QTOF, mass spectra were taken at a scan rate of 0.5 Hz over 500-2000 *m/z* range. Ion source parameters, the end plate offset, and capillary voltage were 500 V and 4500 V, respectively. The nebulizer gas pressure (N_2), dry gas flow rate, and dry gas temperature were set to 0.5 bar, 4.0 L/min, and 200 °C, respectively. For ion transfer parameters, the funnel RF, hexapole RF, quadrupole energy, low mass cutoff, and in-source collision-induced dissociation (isCID) energy were optimized to 400 V, 400 V, 5 V, 500 *m/z*, and 10 V, respectively. Additionally, collision Vpp, pre-pulse storage time, and collision cell transfer time were set to 2000 Vpp, 15 µs, and 100 µs, respectively.

3.6.12. Size Exclusion Chromatography (SEC) for Online Buffer Exchange (OBE) and Native MS Analysis of the cTn(I-T-C) Complex

SEC experiments were performed using a NanoAcquity ultra-high pressure LC system (Waters) coupled to a high-resolution maXis II quadrupole time-of-flight mass spectrometer (Bruker Daltonics). 1 µg of total protein was injected onto a PolyHYDROXYETHYL A column (PolyHEA) (PolyLC Inc), 2.1 mm internal diameter, 100 mm length, 5 µm particle size, and 200 Å pore size. Protein samples were separated isocratically with with 200 mM ammonium acetate solution at a flow rate of 28 µL/min for 10 min with a 'divert to waste' step programmed at 7 min. After 7 min, the SEC eluate consisted of non-volatile buffers and salts which were diverted from the electrospray ionization source to prevent fouling of the mass spectrometer. After every protein sample, the PolyHEA column was flushed with 1 column volume of HPLC H_2O, and re-equilibrated with 1 column volume of 200 mM ammonium acetate solution. Mass spectra were taken at a scan rate of 0.25 Hz over 500-8000 *m/z* range. For ion source parameters, the end plate offset, and capillary voltage were 500 V and 4500 V, respectively. The nebulizer gas pressure (N_2)

was set to 1.5 bar, the dry gas flow rate set to 5.0 L/min, and the dry gas temperature set to 200°C. For ion transfer parameters, the funnel RF, octopol RF, quadrupole energy, low mass cutoff, and isCID were set to 400 V, 800 V, 4V, 500 *m/z*, and 150-200 V, respectively. Additionally, collision Vpp, pre-pulse storage time, and transfer time was optimized to 4000 Vpp, 45 µs, and 145 µs, respectively.

3.6.13. nTDMS Offline FTICR Analysis of cTn Complexes

Samples were analyzed by nanoelectrospray ionization via direct infusion using a TriVersa Nanomate system (Advion BioSciences) coupled to a solariX XR 12-T Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS, Bruker Daltonics). For the nanoelectrospray ionization source using a TriVersa Nanomate, the desolvating gas pressure was set to 0.60 PSI and the voltage was set to 1.5-1.7 kV versus the inlet of the mass spectrometer. The source dry gas flow rate was set to 4 L/min at 180°C. For the source optics, the capillary exit, deflector plate, funnel 1, skimmer voltage, funnel RF amplitude, octupole frequency, octupole RF amplitude, collision cell RF frequency, and collision cell RF amplitude were optimized at 190 V, 200 V, 150 V, 120 V, 300 Vpp, 2 MHz, 600 Vpp, 1.4 MHz, and 2000 Vpp, respectively. Mass spectra were acquired with an acquisition size of 1to 4M-words of data in the mass range 200-8000 *m/z*. Ions were accumulated in the collision cell for 5 to 30 s, and a time-of-flight of 2 ms was used prior to their transfer to the ICR cell. For collisionally activated dissociation (CAD) tandem MS (MS/MS) experiments, the collision energy was varied from 55 to 70 V, ion accumulation was optimized to 5 to 50 s, and acquisition size varied from 1 to 4M-words of data.

3.6.14. Native Trapped Ion Mobility (TIMS) MS Analysis of cTn Complexes

Samples were directly infused by a syringe at a flow rate of ~3 µL/min into a timsTOF Pro mass spectrometer (Bruker Daltonics). For the MS inlet, the end plate offset and capillary voltage were

set to 500 V and 3800 V, respectively. The nebulizer gas pressure (N_2) was set to 1.5 bar with a dry gas flow rate of 6.0 L/min at 180°C. The tunnel out, tunnel in, and TOF vacuum pressures were set to 8.46E-01, 2.586, and 1.52E-07 mBar. TIMS funnel 1 RF was set to 350 Vpp, and TIMS collision cell energy was set to 125 V. An IMS imeX accumulation time of 5.0 ms and cycle ramp time of 350 ms were found to yield optimal resolving power. The TIMS accumulation time was locked to the mobility range (typically 1.05 - 1.55 $1/K_0$). In the MS transfer optics, the funnel 1 RF, funnel 2 RF, deflection delta, isCID energy, multipole RF, and quadrupole ion energy were optimized to 350 Vpp, 350 Vpp, 60 V, 80 eV, 550 Vpp, and 2 eV. Agilent tune mix was directly infused to calibrate MS and CCS values. For MS calibration, the MS resolution for the most abundant calibrant signal, 1821 *m/z*, was 58, 000. Calibrant points at 922, 1222, and 1522 *m/z* were used for TIMS calibration. The TIMS resolution for the most abundant calibrant signal, 1821 *m/z*, was 70.6 CCS/ΔCCS. For native MS spectral collection, the quadrupole low mass was set to 2000 *m/z* with a scan range of 2000-8000 *m/z*. The collision energy was set to 4 eV, with a 2000 Vpp collision cell RF, a 232 µs transfer time, and a prepulse storage time of 5.0 µs.

3.6.15. Data Analysis

Mass spectrometry data was collected using otofControl v. 4.3 and ftmsControl v. 2.1.0. All data were processed and analyzed using Compass DataAnalysis v. 4.3/5.3 software and MASH Native v. 1.1.[156] The sophisticated numerical annotation procedure (SNAP) peak-picking algorithm (quality factor 0.4; signal-to-noise ratio (S/N) 3.0; intensity threshold 500) was applied to determine monoisotopic mass of all detected ions. All chromatograms were smoothed by the Gauss algorithm with a smoothing width of 2.04s. Denatured mass spectra were deconvoluted using the Maximum Entropy algorithm within the DataAnalysis v. 4.3 software with the resolving power for deconvolution set to 40,000 or 60,000k. MS/MS data were output from the DataAnalysis

software and analyzed using Native MASH for proteoform identification, calcium localization, and sequence mapping. All fragment ions were manually validated with a mass tolerance of 20 ppm. For experimental ion mobility analysis by timsTOF Pro, the collisional cross section (CCS) in A^2 for a species of interest was determined via the Mason Schamps equation (Equation 1).

$$\text{CCS} = \frac{3}{16}\sqrt{\frac{2\pi}{\mu k_b T}}\frac{ze}{N_0 k_0} \quad \textbf{(1)}$$

where μ is the reduced mass of the ion−gas pair ($\mu = \frac{mM}{(m+M)}$, where *m* and *M* are the ion and gas particle masses), k_b is Boltzmann's constant, *T* is the drift region temperature, *z* is the ionic charge, *e* is the charge of an electron, N_0 is the buffer gas density, and k_0 is the reduced mobility. Theoretical CCS values were determined using the IMPACT method.[46]

3.6.16. Statistical Analysis

Statistical analysis for group comparison was completed using paired-student *t*-tests. All *p*-values at $p < 0.01$ were considered significant. All error bars indicated in figures represent the mean $\pm$ standard error of the mean (SEM).

3.7 Supplemental Information

3.7.1. Supplemental Tables

Table S3.1. Clinical characteristics of non-failing donor hearts used. Non-failing donor heart tissues were obtained from the University of Wisconsin (UW) Organ and Tissue Donation, Surgical Recovery and Preservation Services. Average and standard error of the mean (SEM) is reported for age. Number of donors (*n)* is indicated in parentheses after each reported percentage.

Clinical Characteristic	Precentage of Sample Cohort
Age	58 ± 8.1 years
	min: 43 years
	max: 65 years
Male	40% (*n* = 2)
Female	60% (*n* = 3)

Table S3.2. cTn proteoforms identified in native top-down MS experiments. Proteoform, modified forms, theoretical most abundant mass, experimental most abundant mass, and mass error for the proteoforms we identified. Modified proteoforms were manually identified based on highly accurate intact mass measurements and MS/MS data together with the prior knowledge on identified cTn proteforms from previous publications[35, 39, 78, 144] or Uniprot. Abbreviations: cardiac troponin T (cTnT, T); cardiac troponin I (cTnI, I); troponin C (TnC, C); N-terminal acetylation (N-acetyl); phosphorylation (phospho, *p*); bis-phosphorylation (2x phospho, *pp*); methionine (Met).

Proteoform	Modified Forms	Theo Mass (Da)	Expt Mass (Da)	Mass Error (ppm)
TnC	N-acetyl	18443.58	18443.87	15.8
TnC	N-acetyl, 1 Ca(II)	18481.47	18481.52	2.9
TnC	N-acetyl, 2 Ca(II)	18519.46	18519.46	0.2
TnC	N-acetyl, 3 Ca(II)	18557.41	18557.42	0.3
*p*cTnT [aa 1-286]	aa 1-286, Met removal, N-acetyl	34452.31	34452.35	1.2
cTnT	Met removal, N-acetyl	34500.44	34500.41	0.9
*p*cTnT	Met removal, N-acetyl, phospho	34580.41	34580.34	2.0
cTn(I-C)	TnC: N-actetyl, 2 Ca(II) cTnI: Met-removal, N-acetyl	42438.31	42438.32	0.2
cTn(I-C)	TnC: N-acetyl, 3 Ca(II) cTnI: Met-removal, N-acetyl	42476.25	42476.27	0.5
cTn(I-C)	TnC: N-actetyl, 2 Ca(II) cTnI: Met-removal, N-acetyl, phospho	42518.27	42518.32	1.2
cTn(I-C)	TnC: N-actetyl, 3 Ca(II) cTnI: Met-removal, N-acetyl, phospho	42556.21	42556.26	1.1
cTn(I-C)	TnC: N-actetyl, 2 Ca(II) cTnI: Met-removal, N-acetyl, 2x phospho	42598.23	42598.31	1.9
cTn(I-C)	TnC: N-actetyl, 3 Ca(II) cTnI: Met-removal, N-acetyl, 2x phospho	42636.18	42636.17	0.3
cTn(I-T-C)	TnC: N-actetyl, 2 Ca(II) cTnI: Met-removal, N-acetyl, phospho cTnT: Met-removal, N-acetyl, phospho	77098.68	77098.73	0.8
cTn(I-T-C)	TnC: N-actetyl, 3 Ca(II) cTnI: Met-removal, N-acetyl, phospho cTnT: Met-removal, N-acetyl, phospho	77136.62	77136.68	0.7
cTn(I-T-C)	TnC: N-actetyl, 2 Ca(II) cTnI: Met-removal, N-acetyl, 2x phospho cTnT: Met-removal, N-acetyl, phospho	77178.64	77178.70	0.8
cTn(I-T-C)	TnC: N-actetyl, 3 Ca(II) cTnI: Met-removal, N-acetyl, 2x phospho cTnT: Met-removal, N-acetyl, phospho	77216.59	77216.68	1.2

Table S3.3. Summary of native TIMS-MS CCS (Å^2) values for TnC monomer, cTn(I-C) dimer, and cTn(I-T-C) heterotrimer complex. Theoretical CCS values were determined using the IMPACT[254] method.

cTn	Theoretical CCS [Å^2]	Experimental CCS [Å^2]	Charge State [*z*+]
TnC	2149	1853	8+
TnC + 1Ca(II)	2141	1849	8+
TnC + 2Ca(II)	2141	1829	8+
TnC +3Ca(II)	2131	1844	8+
cTn (I-C) + 2Ca(II)	3630	3623	15+
cTn (I-C) + 3Ca(II)	3637	3640	15+
cTn(I-T-C) Complex	4192	4574	19+

Table S3.4. Summary of native TIMS-MS CCS (Å^2) values for the cTn(I-T-C) complex incubated with 0, 25, 50, and 100 mM EGTA.

EGTA Concentration (mM)	Charge state [*z*+]	Mobility [1/K0]	CCS [Å^2]
0	20	1.16	4880.3
25	20	1.166	4905.2
50	20	1.174	4936.7
100	20	1.22	5132.1

3.7.2. Supplemental Figures

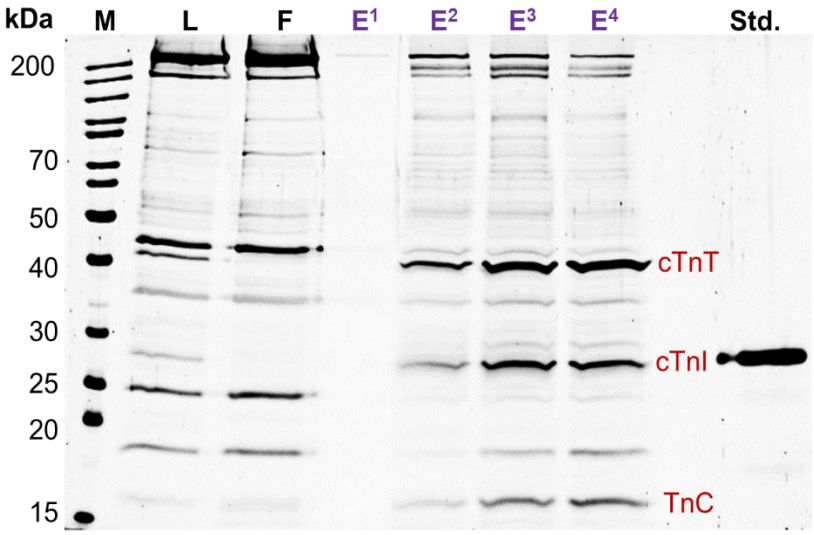


Figure S3.1. Optimization of native elution buffer solution for cTn enrichment with NP-Pep. SDS-PAGE visualizing sarcomeric proteins extracted from non-failing healthy donor heart tissue after the NP-Pep enrichment procedure. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. E^1: 100 mM L-arginine, 100 mM imidazole, 50 mM L-glutamic acid (pH = 7.5); E^2: 250 mM L-arginine, 250 mM imidazole, 50 mM L-glutamic acid (pH = 7.5); E^3: 500 mM L-arginine, 500 mM imidazole, 50 mM L-glutamic acid (pH = 7.5); E^4: 750 mM L-arginine, 750 mM imidazole, 50 mM L-glutamic acid (pH = 7.5). Higher concentrations are required for more effective elution from NP-Pep, suggesting a mechanism of competitive elution. M. protein marker, Std. endogenous cTnI protein standard.

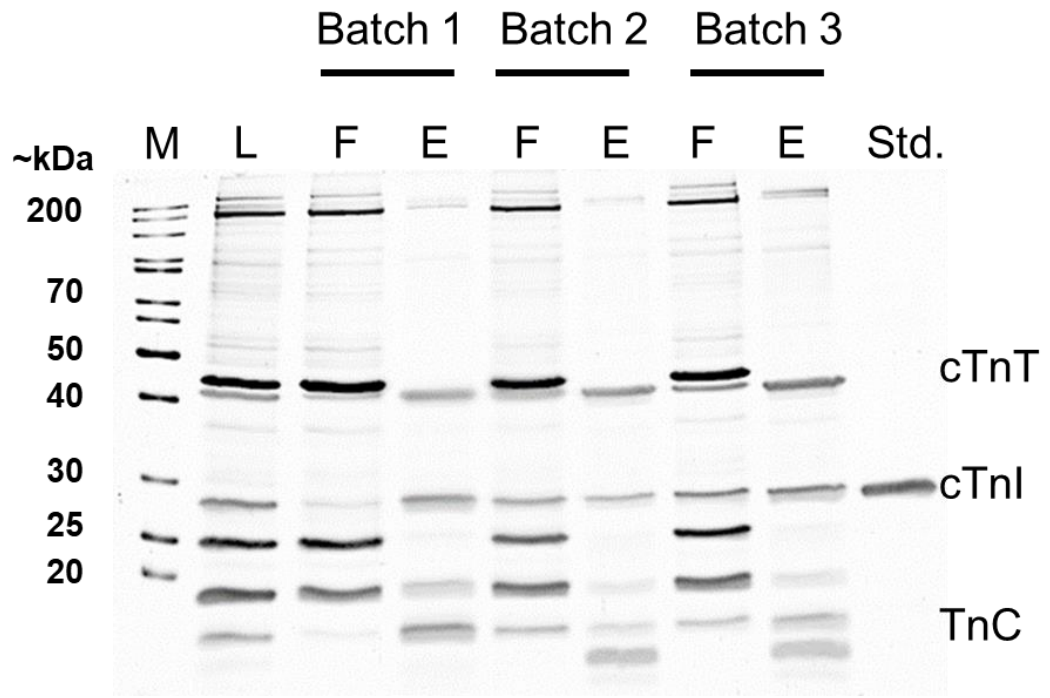


Figure S3.2. Inter-batch enrichment of cTn by NP-Pep. SDS-PAGE visualizing the cTn enrichment performance and demonstrating the high reproducibility obtained from three different synthesis batches of NP-Pep, using 750 mM L-Arg, 750 mM Imidazole, 50 mM L-Glu (pH = 7.5) as native elution buffer. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. M. protein marker, Std. endogenous cTnI protein standard.

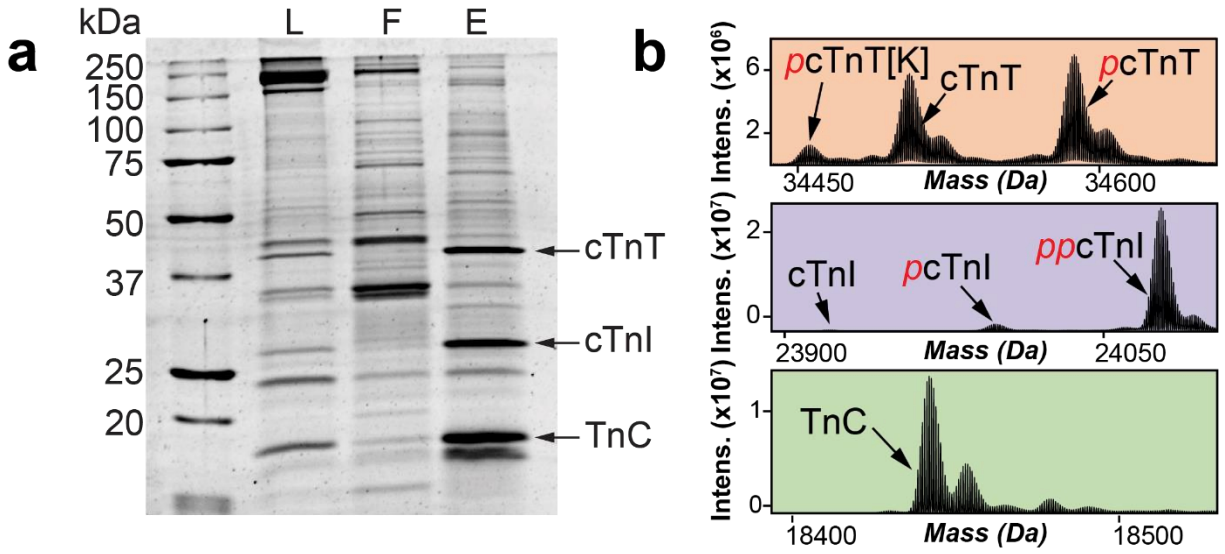


Figure S3.3. SDS-PAGE demonstrates successful enrichment of cTn complex subunits; cTnT (~35 kDa), cTnI (~24 kDa), and TnC (~18 kDa). (a) Loading mixture (L), flow through (F), and elution mixture after enrichment (E) were equally loaded on the gel (1 μg). The concentration of cTnI in SDS-PAGE gel bands was determined using software ImageJ. The recovery percentage was calculated as follows: % Recovery = [the amount of cTnI in the elution (after enrichment) / the original amount of cTnI in the loading mixture (before enrichment)] x 100%. The recovery of cTnI was 89%. **(b)** 600 ng of E was injected for reverse-phase liquid chromatography (RPLC)-MS analysis of cTn subunits. The deconvoluted mass spectra showing the birds eye view of the proteoform landscape for cTnT, cTnI, and TnC.

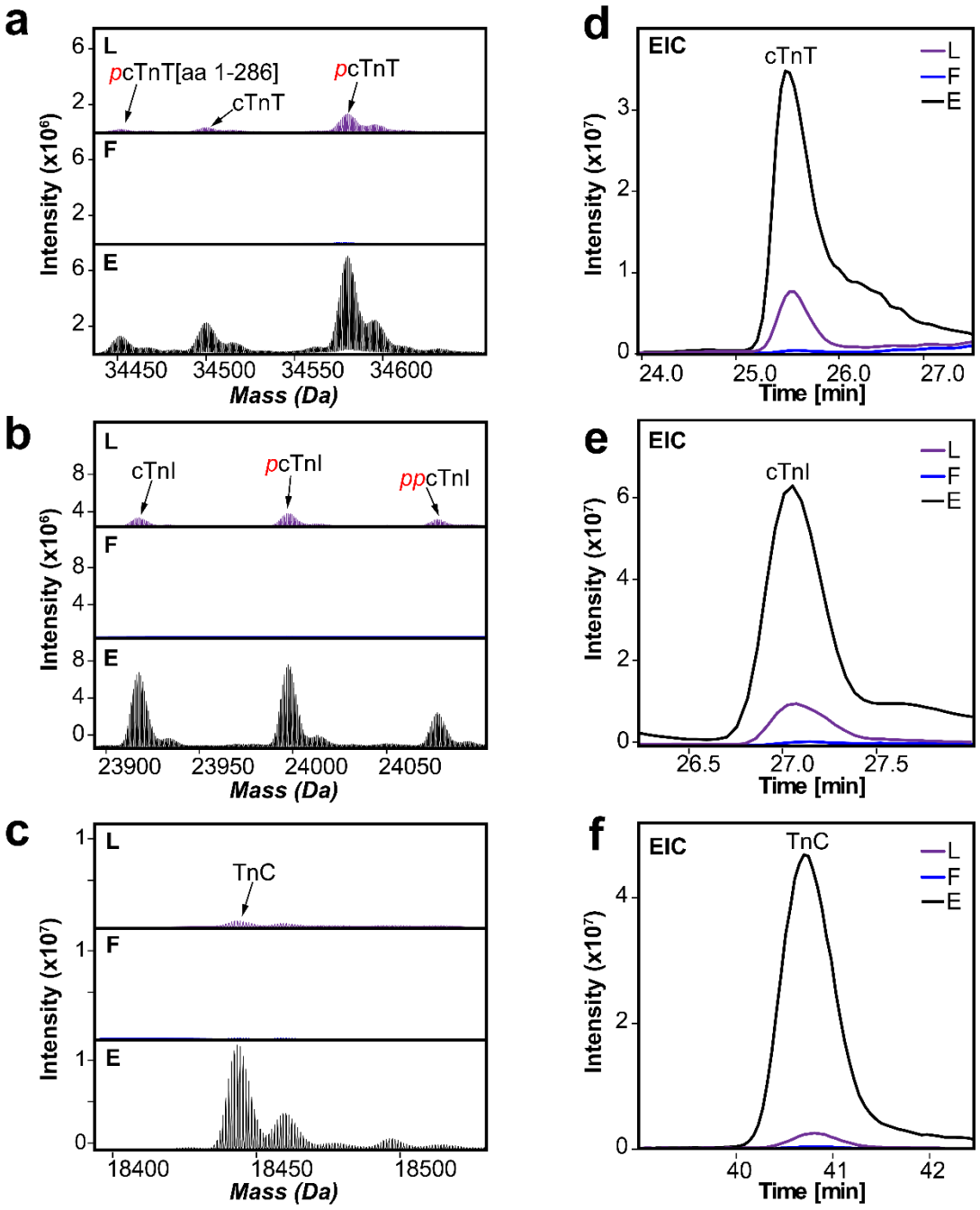


Figure S3.4. Evaluation of NP-Pep enrichment performance using RPLC-MS/MS. The deconvoluted top-down mass spectra of **(a)** cTnT, **(b)** cTnI, and **(c)** TnC proteoforms when loading mixture (L, purple), flow-through (F, blue), and elution mixture after native cTn enrichment (E, black) were equally loaded (600 ng) for online RPLC-MS/MS. Enrichment using NP-Pep demonstrates preservation of proteoforms while significantly enriching cTnT, cTnI, and TnC subunits compared to the initial L. Extracted ion chromatograms (EICs) of **(d)** cTnT, **(e)** cTnI, and **(f)** TnC for L, F, and E demonstrate successful enrichment of cTnT, cTnI, and TnC in E compared to the initial L.

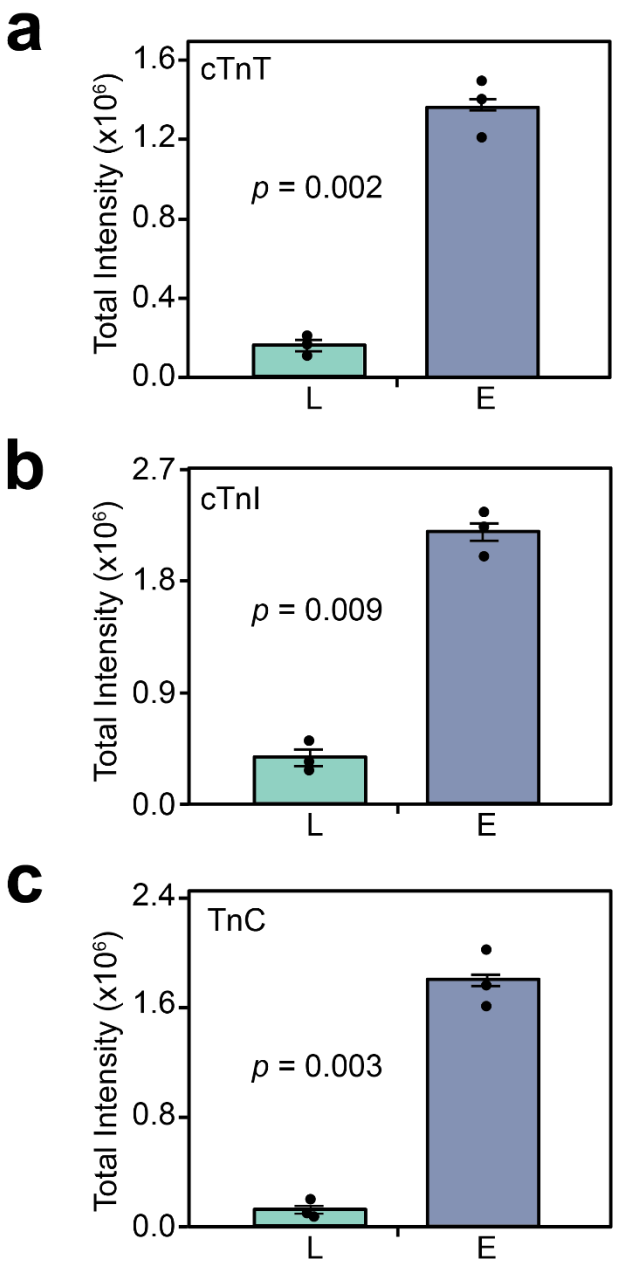


Figure S3.5. Reproducibility of the NP-Pep enrichment capturing cTn subunits in human heart tissues using RPLC-MS/MS. The total intensity of cTn subunits was calculated by summing individual ion intensities from the deconvoluted top-down mass spectra while normalizing for total protein amount (600 ng) injected between the loading mixture (L) and the elution mixture (E) corresponding to cTnT, cTnI, and TnC. cTnT **(a)**, cTnI **(b)**, and TnC **(c)** total intensity data in the loading mixture (L) and elution mixture (E) are representative of $n = 3$ independent donor heart enrichments with error bars indicating standard error of the mean. Groups were considered significantly different by paired-students *t*-tests with $p < 0.01$.

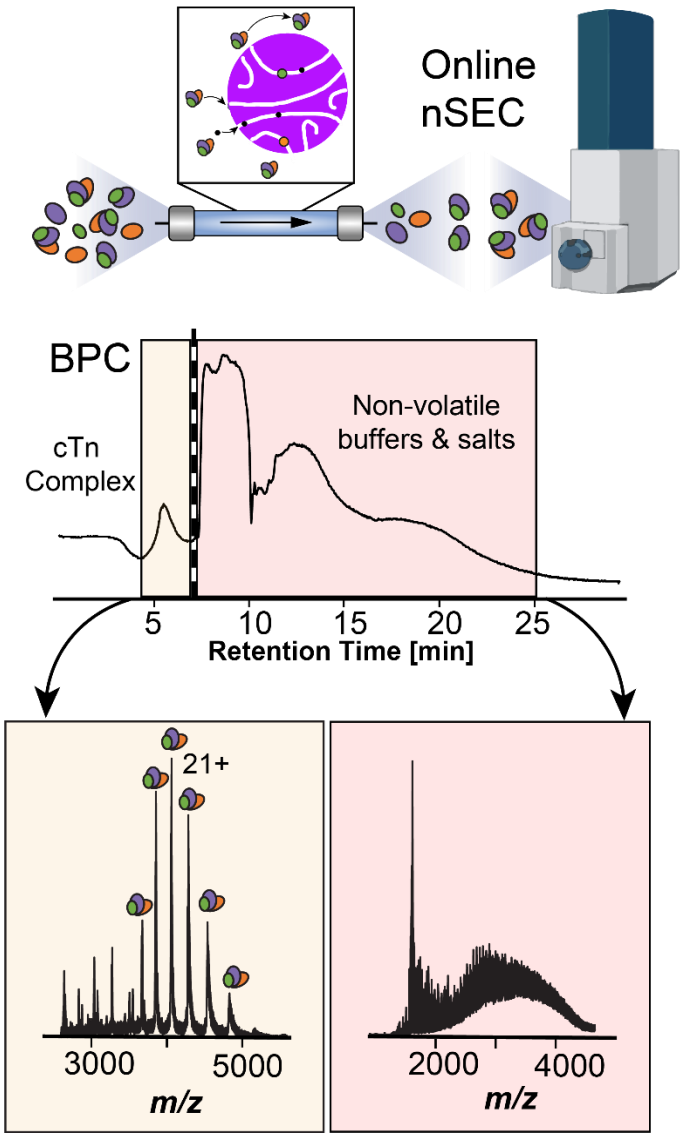


Figure S3.6. Native size exclusion chromatography (SEC) for online buffer exchange (OBE) enables MS analysis of protein complexes. Schematic representation of the native SEC-OBE workflow. Protein complexes are separated and buffer exchanged into 200 mM ammonium acetate using a PolyHYDROXYETHYL A column (PolyHEA) column coupled to a Bruker maXis II quadrupole time-of-flight mass spectrometer for native MS data collection. Base peak chromatogram (BPC) of enriched cTn mixture where the cTn complex (~77 kDa) elutes between 5-6 min. A 'divert to waste' step is programmed at 7 min to divert SEC eluate consisting of non-volatile buffer and salts from the electrospray ionization source.

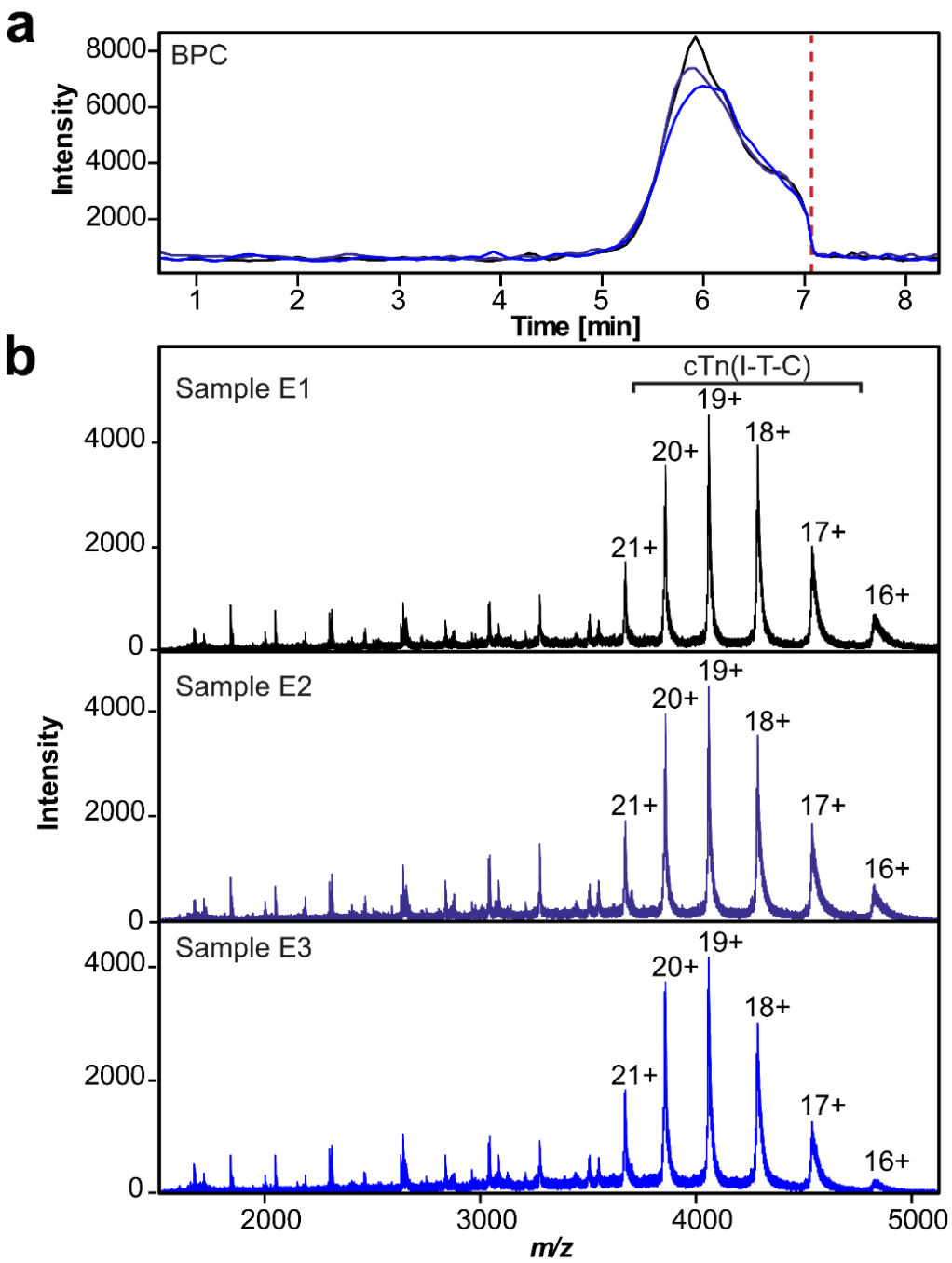


Figure S3.7. Reproducibility of native cTn enrichment and SEC-OBE native MS analysis. (a) Overlaid base peak chromatograms (BPCs) showing reproducibility of the intact cTn heterotrimer complex across three independent native enrichments using the same donor heart sample. Dashed red line is when the sample is diverted to waste to prevent non-volatile buffers and salts from fouling the electrospray ionization source of the mass spectrometer. **(b)** Native mass spectra of the three enrichment samples (E1, E2, and E3) demonstrating reproducible enrichment and SEC-OBE native MS analysis of the intact cTn heterotrimer complex (z = 16-21+).

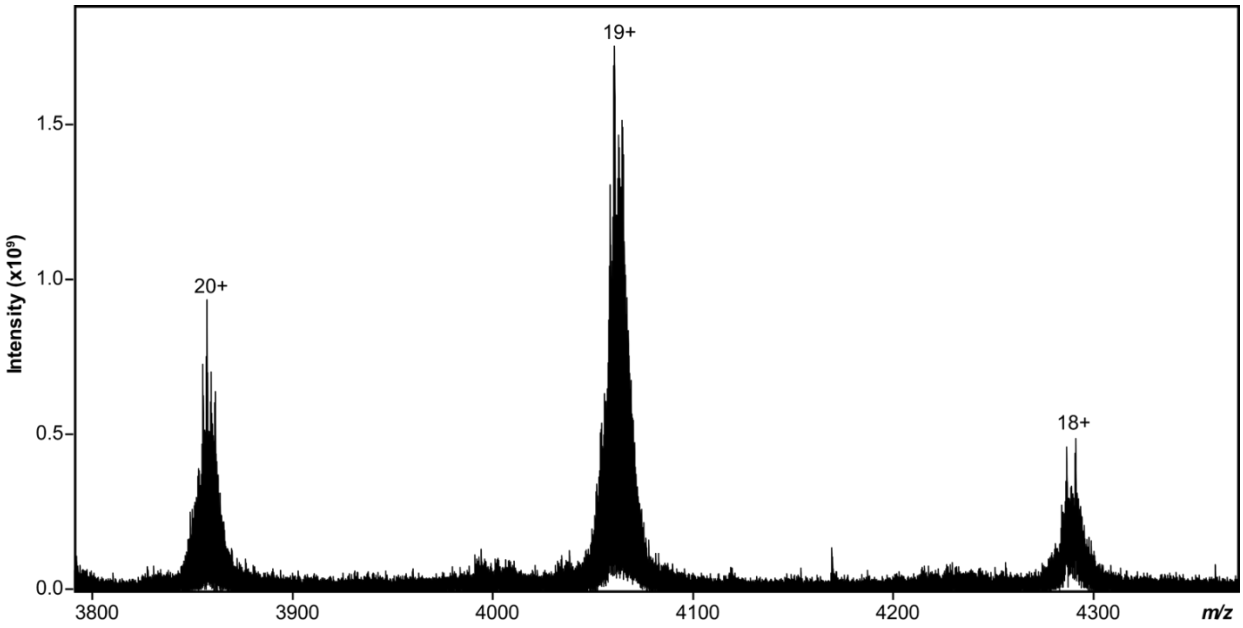


Figure S3.8. Isotopically resolved native MS spectra of endogenous cTn heterotrimer complex (~77 kDa). Native FTICR-MS mass spectrum of the heterotrimeric cTn complex directly enriched from human heart tissue. Charge states *z* = 18-20+ are shown.

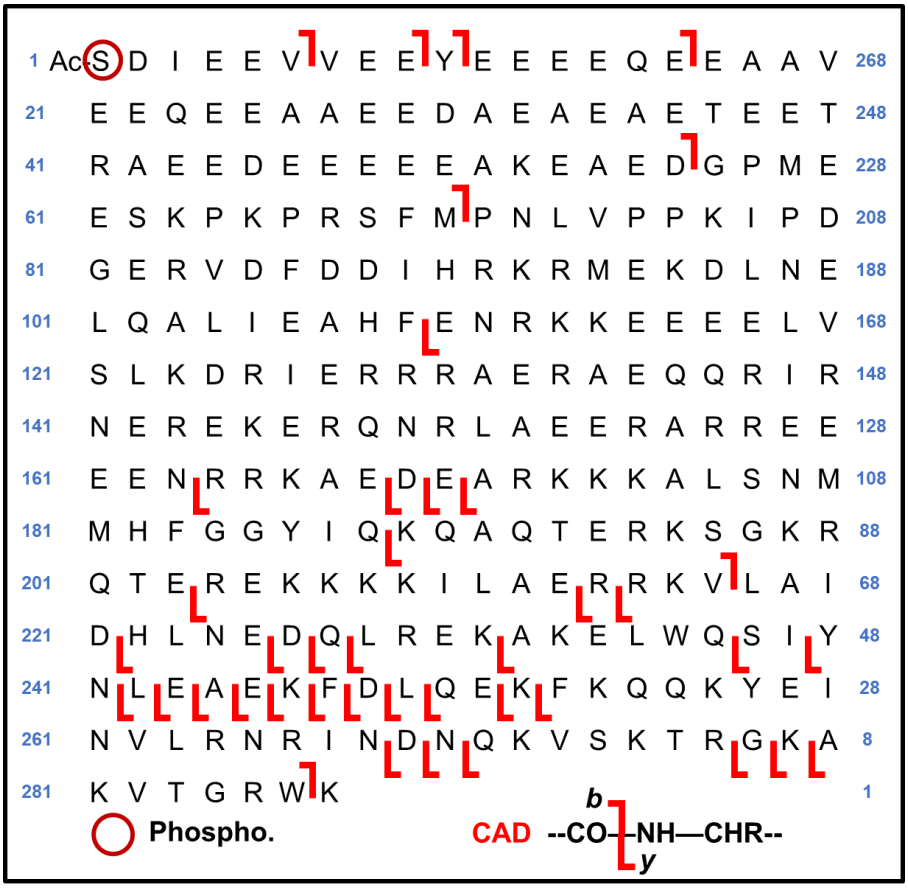


Figure S3.9. Native top-down MS (nTDMS) fragmentation map of cTnT monomer. Precursor ions were selected for collisionally activated dissociation (CAD) resulting in sequence informative b and y ions confirmed within a 20 ppm mass error tolerance. The sequence table with phosphorylation circled in red shows the characterization of cTnT monomer. The nTDMS data for cTnT generated 8 *b*-fragments and 33 *y*-fragments achieving 15% total bond cleavage.

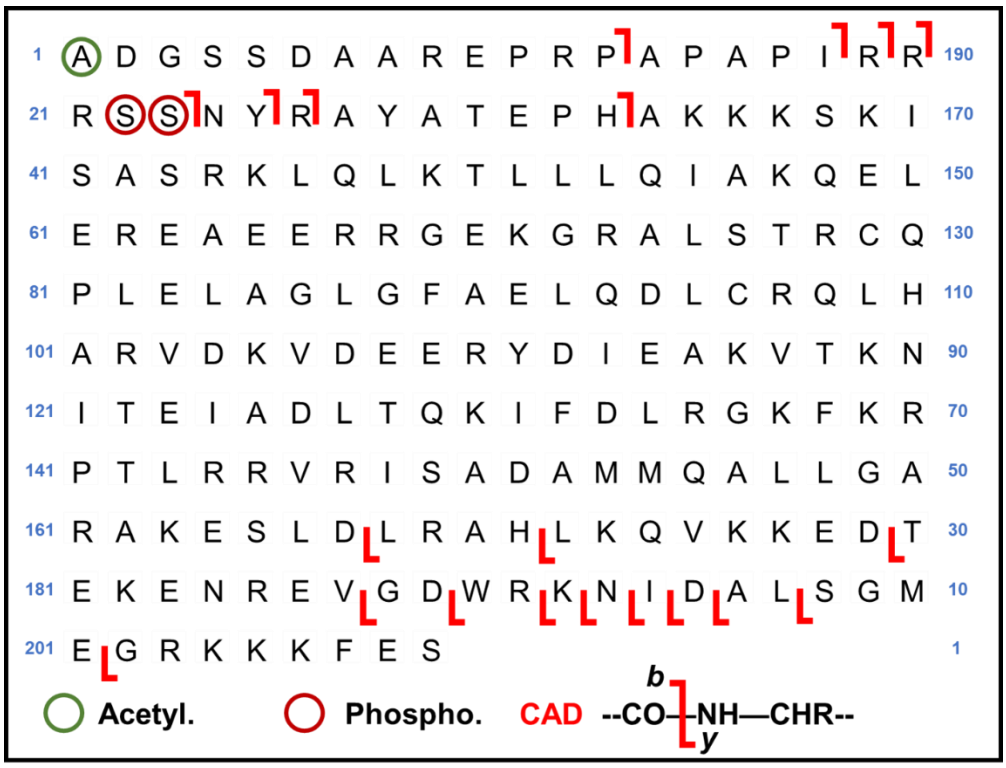


Figure S3.10. Native top-down MS (nTDMS) fragmentation map of cTnI monomer. Precursor ions were selected for collisionally activated dissociation (CAD) resulting in sequence informative b and y ions confirmed within a 20 ppm mass error tolerance. The sequence table with N-terminal acetylation circled in green and phosphorylation circled in red shows the characterization of cTnI monomer. The nTDMS data for cTnI generated 8 *b*-fragments, 12 *y*-fragments achieving 10% total bond cleavage.

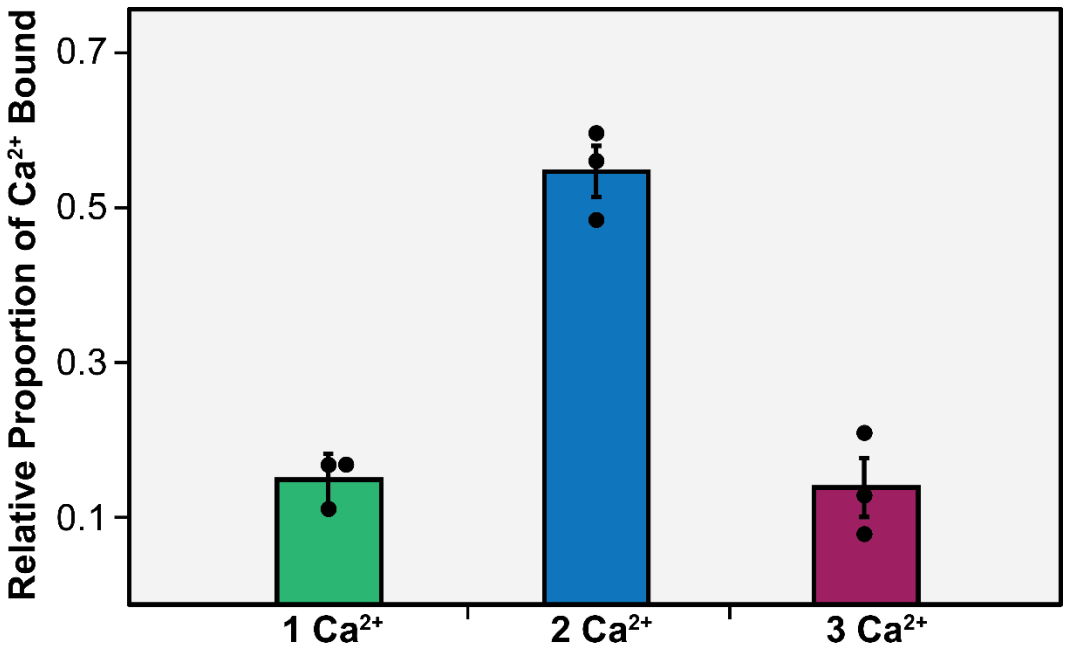


Figure S3.11. Relative proportion of Ca^{2+} bound in each TnC binding domain. Relative Ca^{2+} concentrations were quantified based on the ratio of the peak intensity of a Ca^{2+}-containing proteoform normalized to the summed peak intensities of all detected TnC proteoforms. Error bars indicate the mean ± standard error of the mean (S.E.M.). Data are representative of $n = 3$ independent experiments.

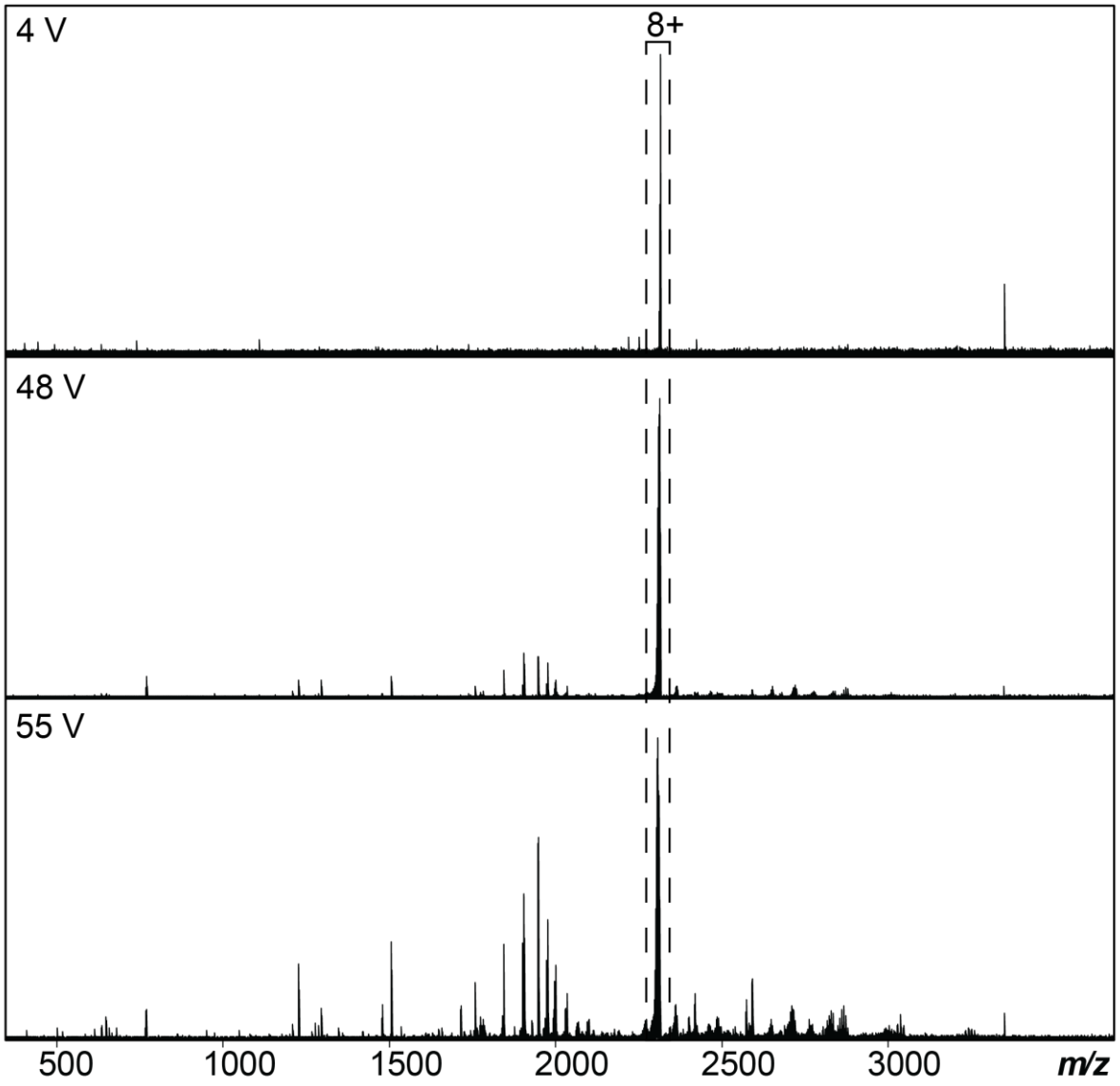


Figure S3.12. Collisional activation ramping of TnC monomer by native top-down MS. MS/MS characterization of the isolated TnC proteoform (z = 8+). A collisionally activated dissociation (CAD) energy of 4 V, 48 V, and 55 V were applied in the collision cell.

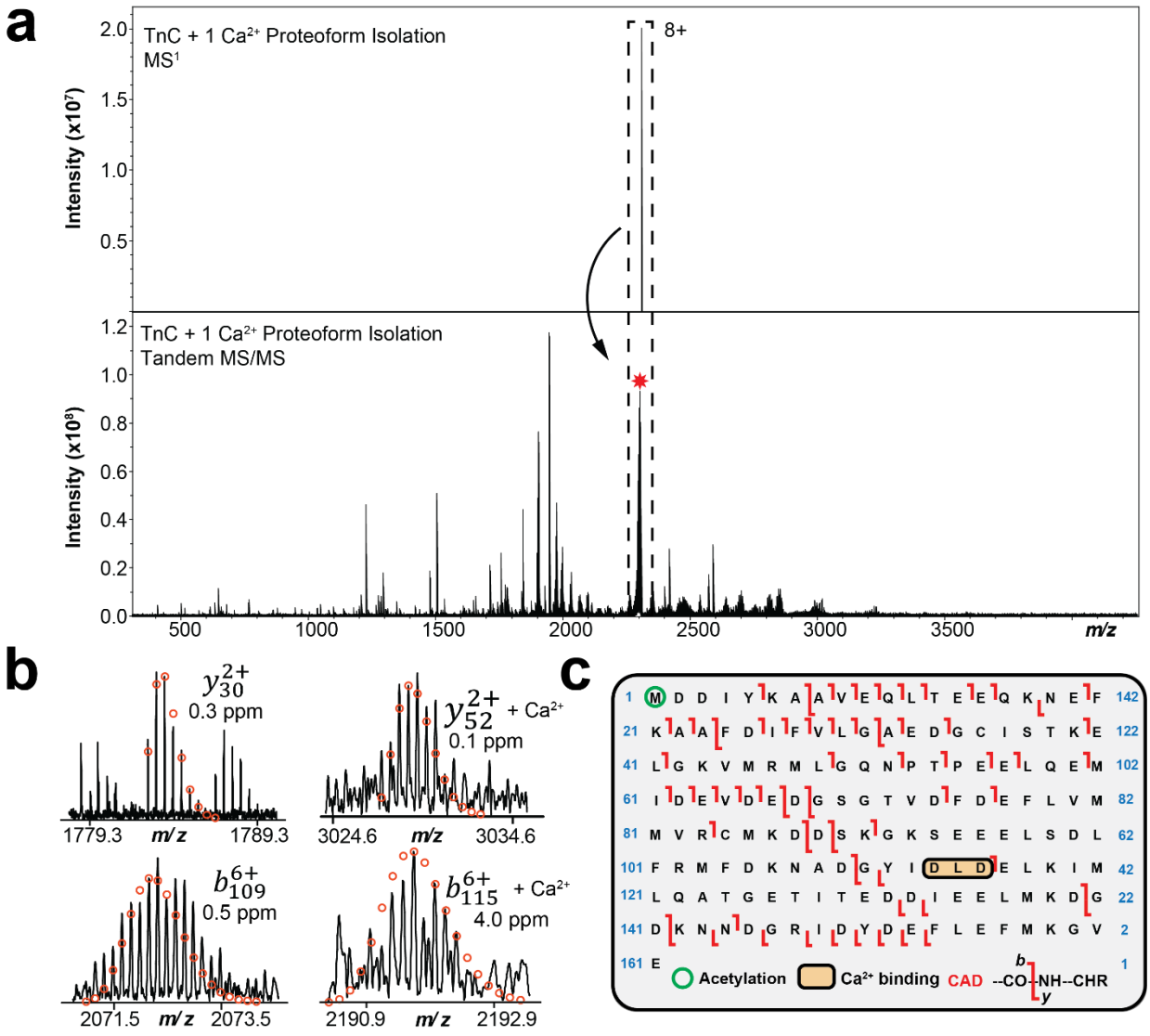


Figure S3.13. Localization of Ca^{2+} binding in TnC monomer domain III by native top-down MS (nTDMS). (a) Isolation of native TnC monomer + 1 Ca^{2+} (z = 8+, centered at 2312 m/z) from MS followed by nTDMS analysis of the isolated precursor. **(b)** Representative collisionally activated dissociation (CAD) fragment ions (y_{30}^{2+}, y_{52}^{2+}, b_{109}^{6+}, b_{115}^{6+}) obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion. **(c)** Ca^{2+} localized to D113-D115 in TnC domain III.

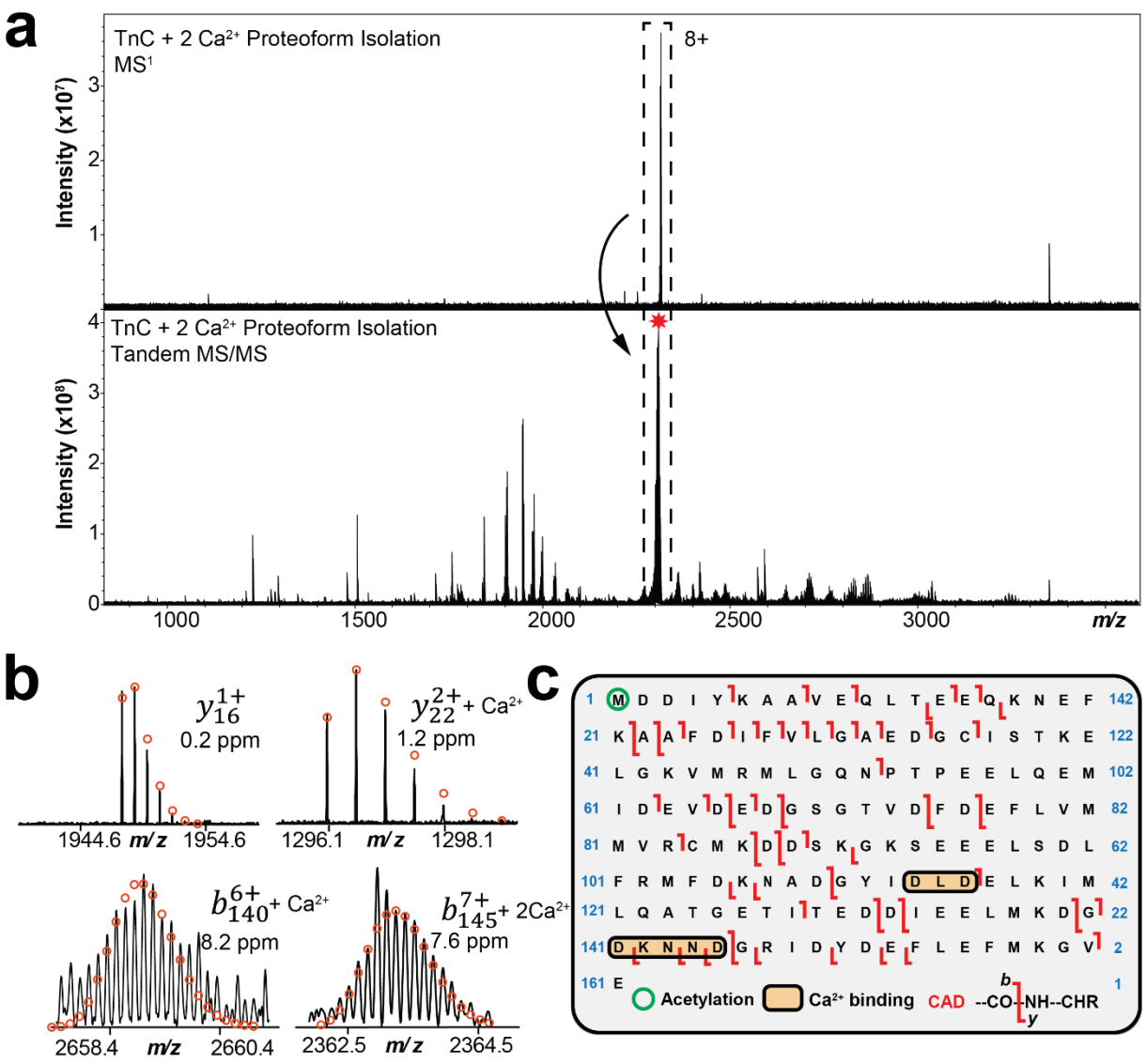


Figure S3.14. Localization of Ca^{2+} binding in TnC monomer domain IV by native top-down MS (nTDMS). (a) Isolation of native TnC monomer + 2 Ca^{2+} (z = 8+, centered at 2316 *m/z*) from MS followed by nTDMS analysis of the isolated precursor. **(b)** Representative collisionally activated dissociation (CAD) fragment ions (y_{16}^{1+}, y_{22}^{2+}, b_{140}^{6+}, b_{145}^{7+}) obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion. **(c)** Ca^{2+} localized to D141-D145 in TnC domain IV.

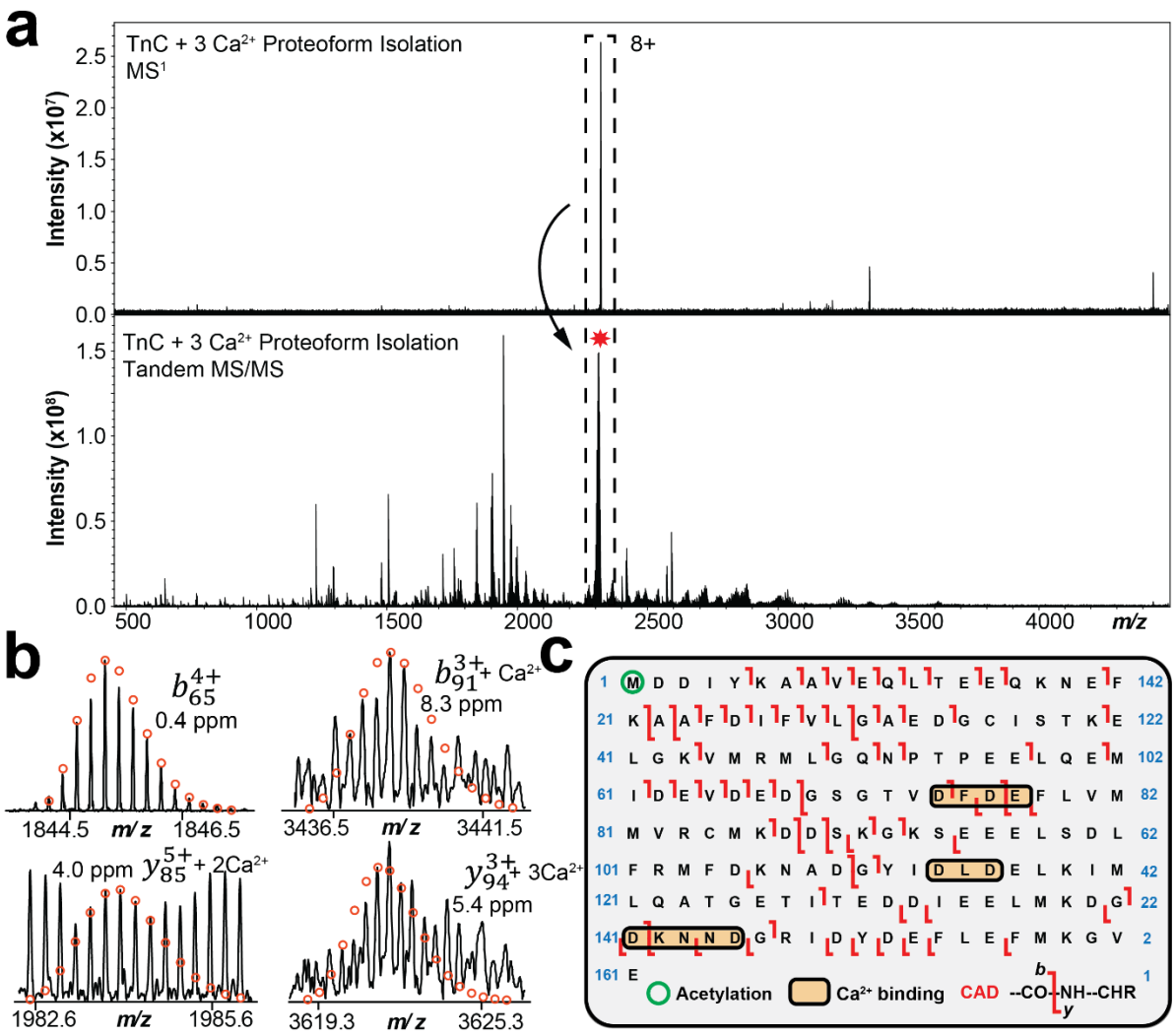


Figure S3.15. Localization of Ca^{2+} binding in TnC monomer domain II by native top-down MS (nTDMS). (a) Isolation of native TnC monomer + 3 Ca^{2+} (z = 8+, centered at 2321 *m/z*) from MS followed by nTDMS analysis of the isolated precursor. **(b)** Representative collisionally activated dissociation (CAD) fragment ions (b_{65}^{4+}, b_{91}^{3+}, y_{85}^{5+}, y_{94}^{3+}) obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion. **(c)** Ca^{2+} localized to D73-E76 in TnC domain II.

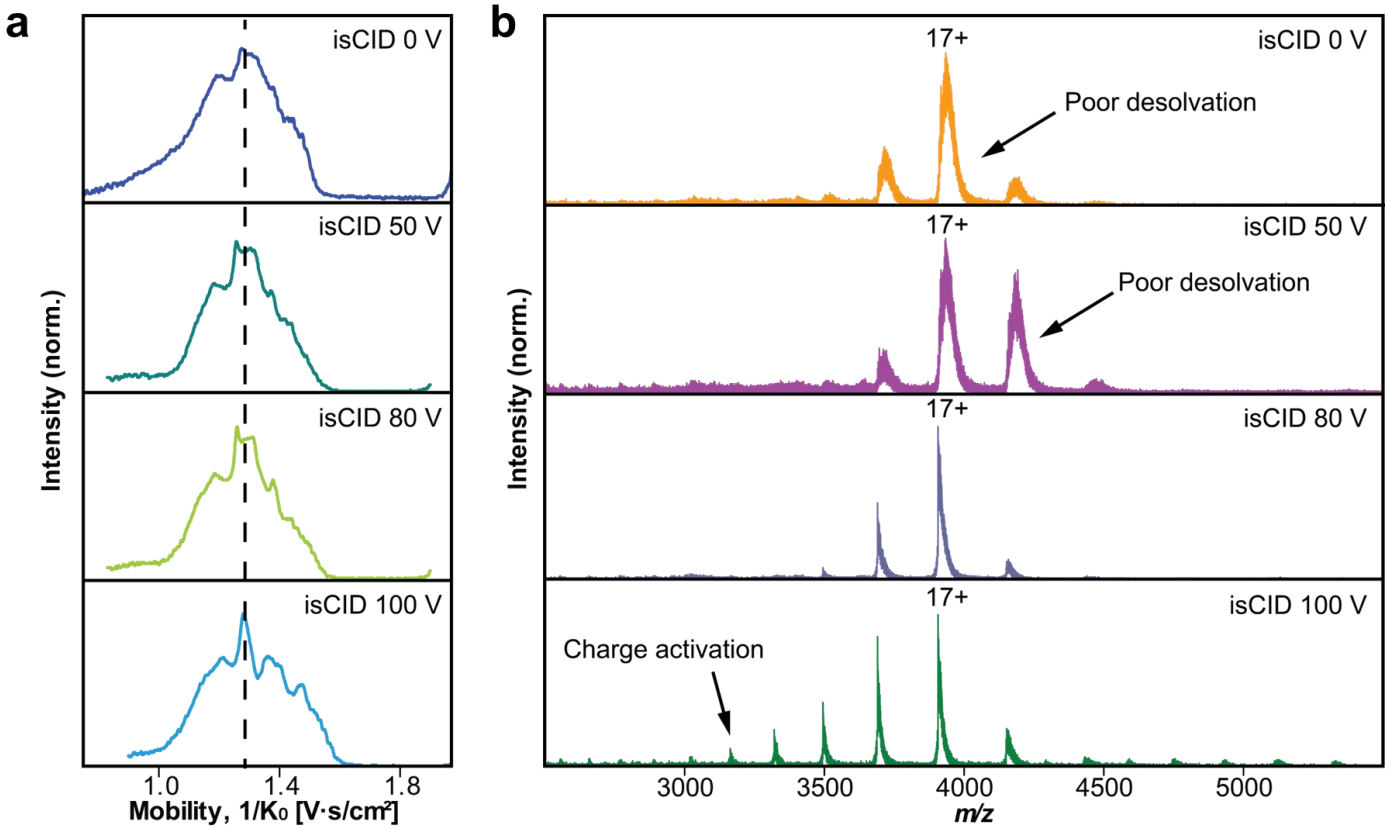


Figure S3.16. Optimization of TIMS parameters for high resolution native MS analysis of bovine serum albumin (BSA). (a) Normalized total ion mobility spectrum of native BSA showing distribution of $1/K_0$ (protein ion mobility) as a function of TIMS in-source collision-induced dissociation (isCID) (0-100V). **(b)** Normalized MS spectra of native BSA as a function of TIMS isCID. A TIMS isCID of 80 V yields the best desolvated gas-phase conformation of native BSA.

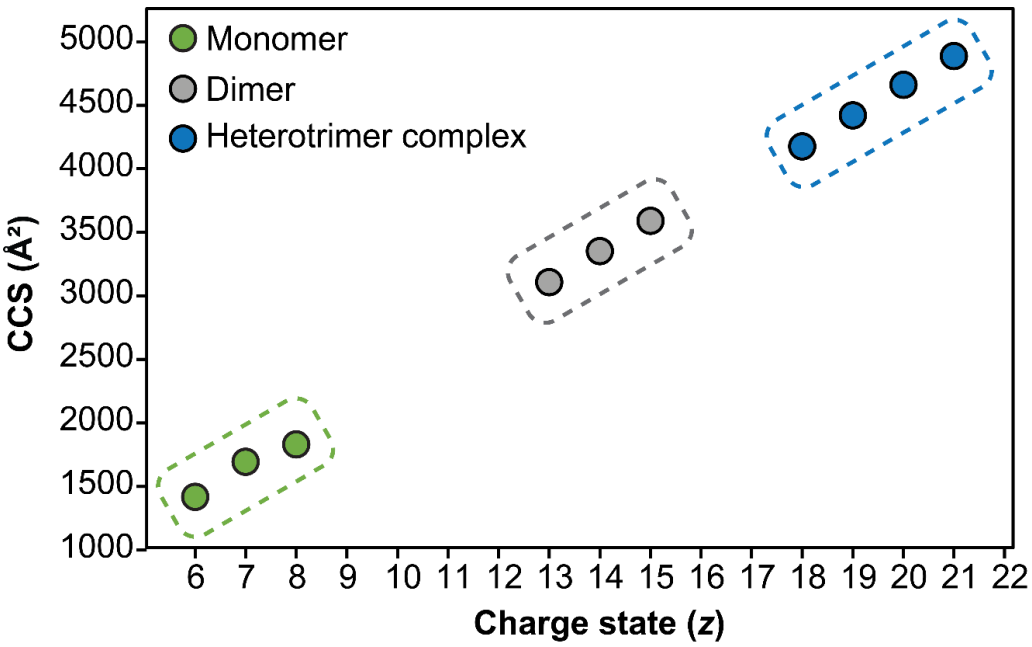
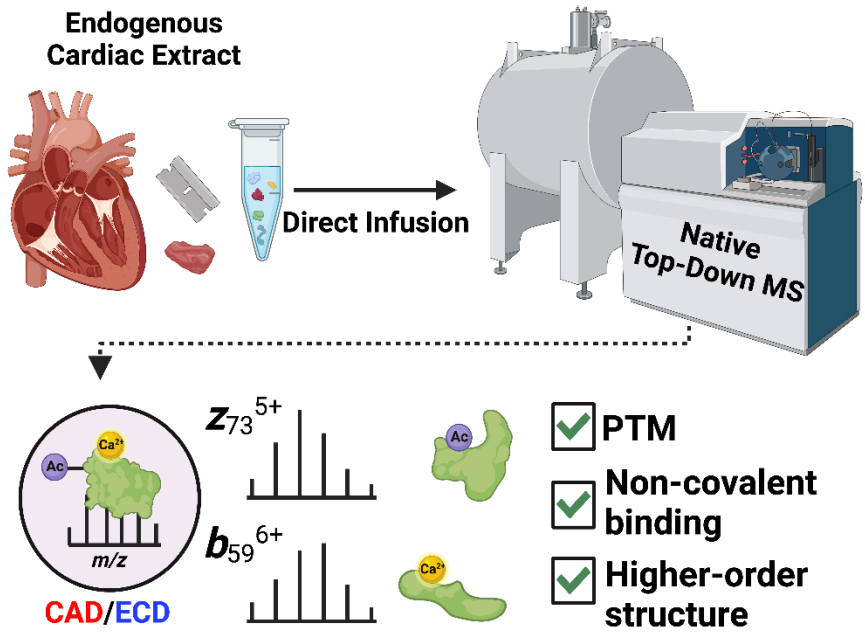


Figure S3.17. Plot of CCS value as a function of charge state by native TIMS-MS. Green circles indicate TnC monomer, grey circles indicate cTn(I-C) dimer charge states, and blue circles indicate cTn(I-T-C) heterotrimer complex charge states observed.

CHAPTER 4:

Native Top-Down Mass Spectrometry for Characterizing Sarcomeric Proteins Directly from Cardiac Tissue Lysate



Portions of this chapter are adapted from the following published manuscript:

Chapman, E.A.*; Li, B.H.*; Krichel, B.; Chan, H.; Buck, K.M.; Roberts, D.S.; Ge, Y. Native Top-Down Mass Spectrometry for Characterizing Sarcomeric Proteins Directly from Cardiac Tissue Lysate. *Journal of the American Society for Mass Spectrometry* **2024**, *35* (4), 738-745. DOI: https://doi.org/10.1021/jasms.3c00430. *denotes equal contribution.

4.1 Abstract

Native top-down mass spectrometry (nTDMS) is a powerful structural biology tool that can localize post-translational modifications (PTMs), explore ligand-binding interactions, and elucidate the three-dimensional structure of proteins and protein complexes in the gas-phase. Fourier-transform ion cyclotron resonance (FTICR)-MS offers distinct capabilities for nTDMS, owing to its ultra-high resolving power, mass accuracy, and robust fragmentation techniques. Previous nTDMS studies using FTICR have mainly been applied to over-expressed recombinant proteins and protein complexes. Here, we report the first nTDMS study that directly analyzes human heart tissue lysate by direct infusion FTICR-MS without prior chromatographic separation strategies. We have achieved comprehensive nTDMS characterization of cardiac contractile proteins that play critical roles in heart contraction and relaxation. Specifically, our results reveal structural insights into ventricular myosin light chain 2 (MLC-2v), ventricular myosin light chain 1 (MLC-1v), and alpha-tropomyosin (α-Tpm) in the sarcomere, the basic contractile unit of cardiac muscle. Furthermore, we verified the calcium (Ca^{2+}) binding domain in MLC-2v. In summary, our nTDMS platform extends the application of FTICR-MS to directly characterize the structure, PTMs, and metal-binding of endogenous proteins from heart tissue lysate without prior separation methods.

4.2 Introduction

The function of proteins is based upon their three-dimensional structure which can be modulated by metal-binding, post-translational modifications (PTMs), and genetic mutations.[178] Native top-down mass spectrometry (nTDMS) has emerged as a complementary structural biology tool alongside traditional techniques such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryogenic electron microscopy (cryoEM) for the structural characterization of intact proteins and their non-covalent interactions in the gas phase.[44, 49, 51, 52, 218, 263, 264] The application of nTDMS not only provides important insights into amino acid sequence, PTMs, and ligand-binding sites, but also elucidates the higher-order structure of proteins and protein complexes.[31, 45, 47, 258] Previous nTDMS studies have primarily been applied to purified proteins and protein complexes from over-expression systems, given the challenges associated with directly characterizing proteins from their endogenous environment.[48, 178, 265] This challenge is further exacerbated by the intricacies in native sample preparation and the inherent heterogeneous nature of biological sample matrices, both of which pose obstacles to obtaining clean, well-resolved native mass spectra.[38, 86]

Fourier-transform ion cyclotron resonance (FTICR)-MS becomes a powerful tool for nTDMS characterization of proteins and protein complexes due to its ultra-high resolving power, mass accuracy, and various fragmentation capabilities for comprehensive proteoform characterization.[41, 112, 113, 127] FTICR mass spectrometers are equipped with both electron- and energy-based tandem MS (MS/MS) fragmentation capabilities, which are used either individually or in combination to obtain sequence-specific and structural information during nTDMS analysis. Specifically, energy-[79, 121] and electron-based[127, 128, 266-268] fragmentation techniques can provide information on solvent-exposed regions and structural flexibility. This is based on the number of

N- and C-terminal fragment ions produced, as the solvent-exposed regions not confined by non-covalent interactions undergo preferred gas-phase fragmention.[269] Recently, our group has structurally characterized the endogenous cardiac troponin (cTn) heterotrimeric complex (I-T-C: ~77 kDa) directly from human heart tissue using a native nanoproteomics platform that involves directly infusing the enriched/purified cTn(I-T-C) complex onto a FTICR mass spectrometer for nTDMS characterization.[79] Specifically, we demonstrated that collisionally activated dissociation (CAD) generates *b*- and *y*-fragment ions during nTDMS analysis, providing information on solvent-exposed regions and subunit interfaces based on the number of N- and C- terminal fragment ions. Inspired by our previous study, we sought to investigate the possibility of direct characterization of native proteins from complex biological tissue samples using nTDMS without chromatographic separation strategies.

Herein, we utilize FTICR-MS as a platform for nTDMS analysis to characterize endogenous cardiac proteins directly from human heart tissue lysates via direct infusion without using chromatographic strategies (**Figure 4.1**). Human heart tissue lysate in the buffer surrogate ammonium acetate was directly infused using a nano-electrospray ionization source (nano-ESI) coupled to a FTICR mass spectrometer for nTDMS analysis (**Figure 4.1a**). Upon directly infusing human heart tissue lysate, the native mass spectra revealed different isotopic distributions that were representative of three cardiac proteins (**Figure 4.1b**). Cardiac protein ions were first isolated at the MS1 level to obtain highly accurate intact mass measurements and then fragmented by CAD and/or electron capture dissociation (ECD) for unambiguous proteoform characterization (**Figure 4.1c**). The resulting fragment ions map PTMs, Ca^{2+} binding, and higher order structures, offering structural insights into the function of these proteins present in endogenous heart tissues (**Figure 4.1d**). We have primarily detected and characterized contractile proteins in the sarcomere, the

basic contractile units in the cardiac muscle.[39, 144, 162, 168] Our results verify the calcium (Ca^{2+}) binding domain in the ventricular isoform of myosin light chain 2 (MLC-2v)[119, 270], while also revealing extensive backbone fragmentation in the unstructured N-terminal domain of the ventricular isoform of myosin light chain 1 (MLC-1v).[119] Moreover, we observed a marked flexibility within the N- and C-termini of alpha-tropomyosin (α-Tpm),[169, 200] a region where the cryo-EM structure lacks resolution. Overall, this work illustrates the power of nTDMS using FTICR-MS to characterize endogenous proteins, without conventional chromatography workflows, directly from complex biological mixtures.

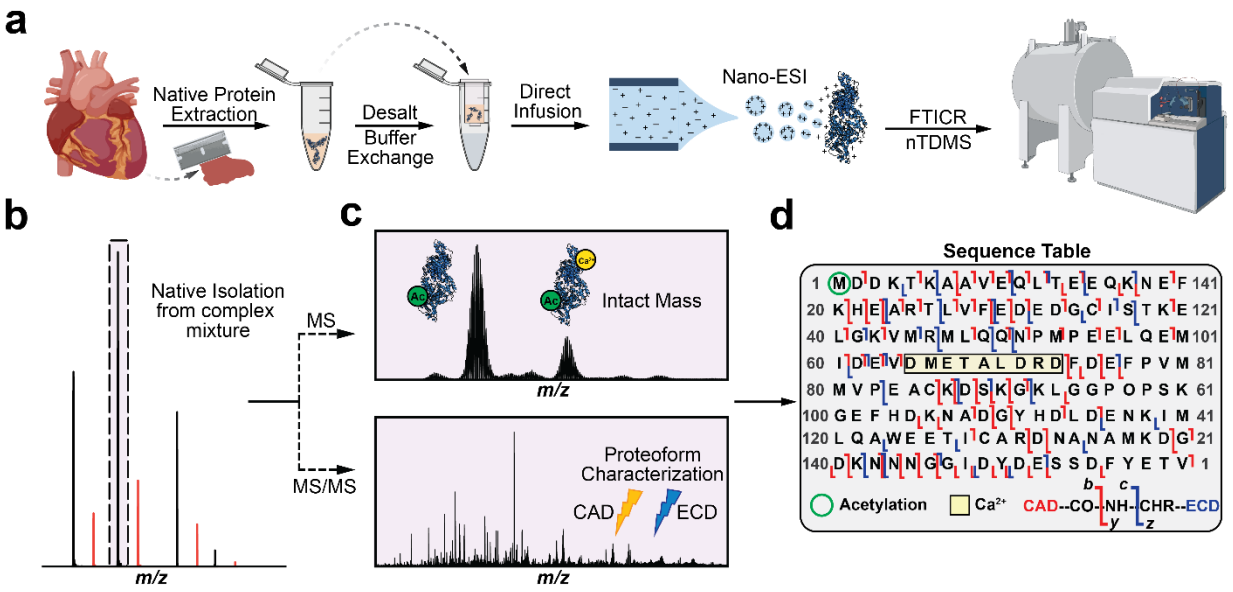


Figure 4.1. nTDMS workflow for the characterization of endogenous cardiac proteins. Illustration of the nTDMS workflow using a FTICR mass spectrometer. **(a)** Native proteins were first extracted from 500 mg of donor heart tissue using a LiCl buffer at physiological pH. The heart tissue extract was then buffer exchanged into 150 mM ammonium acetate using spin columns. Samples were analyzed by ultrahigh-resolution FTICR-MS via direct infusion using a nano-ESI source. **(b)** Native mass spectrum of endogenous proteins directly infused from human heart tissue extract. Individual charge states of proteins are isolated to acquire **(c)** MS and MS/MS data for intact mass analysis and proteoform characterization, respectively. **(d)** CAD and ECD fragment ions verify metal binding and PTMs to provide information on protein structure.

4.3 Results and Discussion

Inspired by our nTDMS analysis of enriched endogenous cTn complexes,[79] we investigated whether our nTDMS approach using an ultrahigh-resolution FTICR mass spectrometer could characterize endogenous proteins directly from cardiac tissue lysate without chromatographic separation strategies. Native proteins were extracted from 500 mg of human cardiac tissue using a native extraction protocol, as previously described.[79] Briefly, cytosolic proteins were first depleted using a phosphate wash buffer followed by protein extraction using a high ionic strength LiCl buffer at physiological pH. Cardiac tissue lysates were then prepared for nTDMS analysis by buffer-exchanging into the buffer surrogate ammonium acetate using spin columns. Using this extraction protocol and sample clean-up method, we were able to prevent unspecific adducts impairing signal of our native mass spectra. Moreover, when directly infusing cardiac tissue lysate to our FTICR mass spectrometer for nTDMS analysis, we observed various well resolved signals, exhibiting at least three isotopic resolved peak series that could be assigned to sarcomeric proteins: MLC-2v, MLC-1v, and α-Tpm between 1900-3000 *m/z* **(Figure 4.2)**.

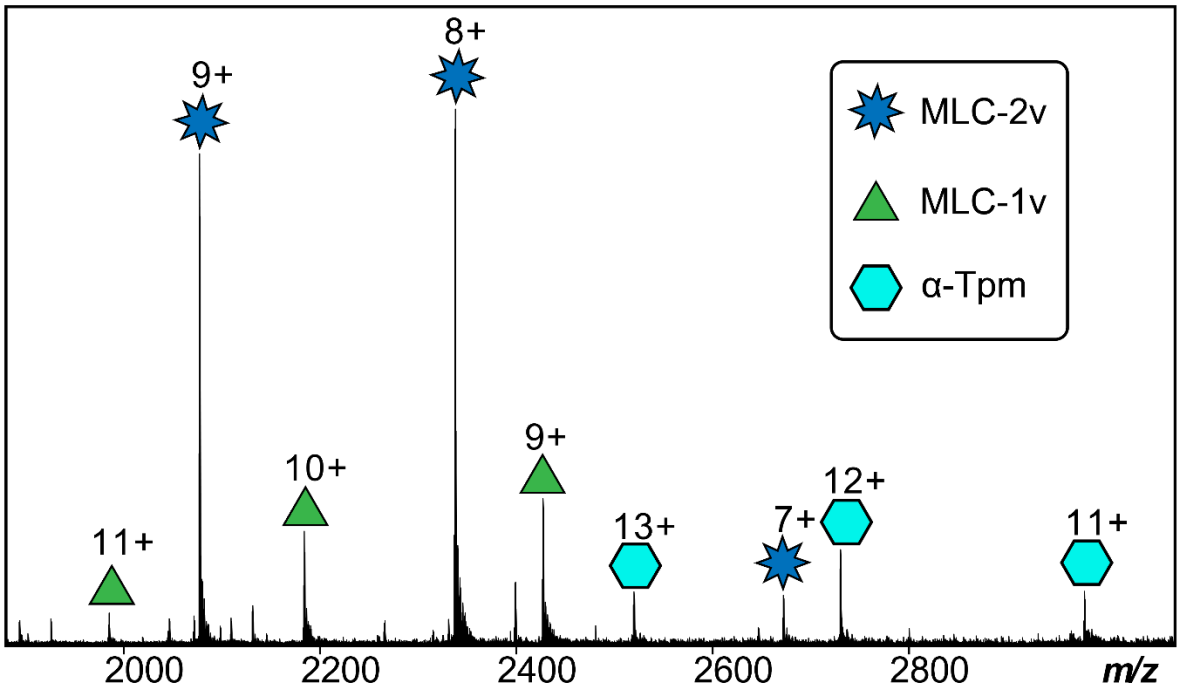


Figure 4.2. Direct infusion of cardiac tissue lysate for native FTICR-MS analysis. Native high-resolution FTICR-MS analysis revealed endogenous proteoforms of MLC-2v (*z* = 7-9+; blue star), MLC-1v (*z* = 9-11+; green triangle), and α-Tpm (*z* = 11+-13+; teal hexagon) directly from human heart tissue lysate between 1900-3100 *m/z*.

MLC-2v (~18.7 kDa) is an EF-hand Ca^{2+} binding protein that is known as the regulatory light chain and is a principal component of myosin, a hexamer that is the basic contractile unit of the sarcomere.[119, 270] Phosphorylation of MLC-2v by myosin light chain kinase occurs during Ca^{2+}-dependent cardiac contraction and helps to regulate myosin motility and function within the sarcomere.[271] MLC-2v consists of two predicted EF-hand domains between amino acid residues 28-73 and 102-160, respectively.[272] Based on previous literature[271] and the UniProtKB database, Ca^{2+} is predicted to bind to the first EF-hand domain in MLC-2v. However, the binding of Ca^{2+} to MLC-2v has not been previously subjected to a detailed biochemical analysis and is not present in the available cryo-EM structure. MLC-2v was the most abundant protein in the native mass spectra

upon directly infusing cardiac tissue lysate, with a charge state distribution from 7+ to 9+ between 2000 *m/z* and 2700 *m/z*. Upon native isolation of MLC-2v (z = 8+), three proteoforms were measured with high mass accuracy including the unphosphorylated (18688.37 Da), Ca^{2+} bound (18726.34 Da), and phosphorylated (18768.33 Da) proteoforms **(Figure 4.3a, Table S4.1)**. Next, CAD and ECD fragmentation was performed on the isolated charge state to gain MLC-2v sequence and structural information. Extensive backbone fragmentation generated 35 *b* and 23 *y* ions during the nTDMS analysis upon CAD fragmentation of MLC-2v **(Figure 4.3b-c, Figure S4.1)**. Notably, ECD fragments (17 *c* and 1 *z*● ions) were primarily reserved to the N-terminal domain of MLC-2v (**Figure 4.3b-c, Figure S4.2**). Collectively, the fragmentation pattern confirmed N,N,N-trimethylation to Ala1 and phosphorylation to Ser14, as previously reported.[119, 144] Analysis of MLC-2v by nTDMS also provided structural insights, offering clues to how its structure may affect its function. The abundance of N-terminal fragments, particularly the substantial amount of *c* fragment ions observed during ECD, strongly suggests that the N-terminal domain of endogenous MLC-2v is flexible and solvent-exposed. The near equal proportion of *b*- and *y*- fragment ions suggests that the C-terminus is also solvent exposed, which agrees with the CAD fragmentation data. The cryo-EM structure of MLC-2v (PDB: 5TBY, chain F) confirms that the N- and C-terminal domains are solvent exposed. Furthermore, the MLC-2v structure shows that the initial 50 amino acids in the N-terminus, where minimal secondary structure is defined, are more extended and unfolded compared to the last 50 amino acids in the C-terminus (**Figure S4.3**). The N-terminal extension contains the phosphorylation site of myosin light chain kinase, which initiates a conformational change in MLC-2v, subsequently promoting the interaction between myosin and actin.[273] Our observations reveal that the N-terminal extension of MLC-2v is

in a solvent-exposed state, thereby potentially optimizing its accessibility for PTMs and facilitating Ca^{2+} binding.

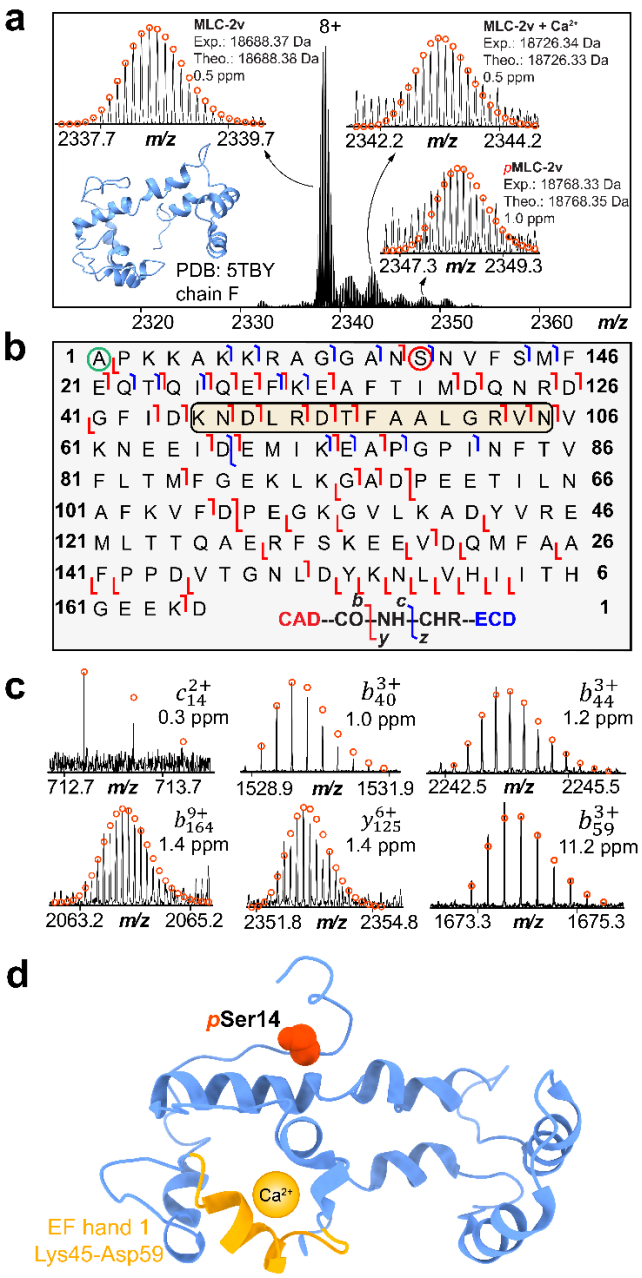


Figure 4.3. nTDMS characterization of endogenous MLC-2v. (a) MS1 native isolation spectra of endogenous MLC-2v (z = 8+) showing Ca^{2+}-bound and phosphorylated (*p*MLC-2v) proteoforms. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy proteoform characterization. **(b)** CAD and ECD combined

fragmentation map achieving 42% bond cleavage. The green and red circles are representative of N,N,N-trimethylation and phosphorylation, respectively. Ca^{2+} binds to MLC-2v in the yellow shaded region (Lys45-Asp59). **(c)** Representative fragment ions obtained from nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion. **(d)** Structure of MLC-2v with the Ca^{2+} binding region in EF-hand domain 1 (Lys45-Asp59) highlighted in yellow. The phosphorylation site on Ser14 is highlighted in red. PDB: 5TBY, chain F.

We further investigated Ca^{2+} binding to MLC-2v by isolating and subjecting the Ca^{2+} bound proteoform (z = 8+) to CAD fragmentation **(Figure S4.4)**. The monoisotopic mass of the MLC-2v proteoform with Ca^{2+} corresponds to the mass of MLC-2v with one bound Ca^{2+} (**Table S4.1**). Detailed nTDMS analysis yielded sequence informative product ions such as b_{44} with no Ca^{2+} bound and b_{59} with one Ca^{2+} bound **(Figure S4.4)**. Therefore, the Ca^{2+}-binding domain was localized to the N-terminal region between Lys45 and Asn59 where MLC-2v contains a predicted EF-hand, a common calmodulin derived Ca^{2+} binding motif **(Figure 4.3d)**. While the only cryo-EM structures of MLC-2v deposited were low resolution and absent of EF-hand domain typical amino acid coordination or electron density for Ca^{2+} binding. To our knowledge, this is the first study to experimentally verify the predicted Ca^{2+} binding site within MLC-2v derived directly from endogenous tissues.

The essential light chain of myosin, MLC-1v (~21.9 kDa), provides structural integrity to the sarcomere by facilitating actin-myosin cross bridge formation during cardiac contraction.[274] Unlike MLC-2v, MLC-1v does not bind Ca^{2+} at physiologically relevant concentrations.[271, 275] Charge states of MLC-1v from 9+ to 11+ were observed between 1900 *m/z* and 2500 *m/z* when directly infusing cardiac tissue lysate for nTDMS analysis. Native isolation of MLC-1v (z = 9+), revealed the most abundant proteoform to be N,N,N-trimethylated with a highly accurate monoisotopic mass of 21828.91 Da (**Figure 4.4a, Table S4.1**). CAD fragmentation was

performed, identifying 60 *b* and 16 *y* fragment ions to achieve 39% bond cleavage (**Figure S4.5**). The N,N,N,-trimethylation was suggested to be on Ala1, which has been previously reported by our group (**Figure 4.4b-c**).[119, 144] The fragmentation pattern revealed extensive backbone fragmentation within the first 50 amino acids of MLC-1v. Interestingly, the first 43 amino acids within the N-terminal extension are not resolved in the only protein data bank deposited cryo-EM structure of MLC-1v (PDB: 5TBY, chain D) which is typical for unstructured flexible regions. Our data indicates that, indeed, the N-terminal domain is highly solvent exposed compared to the C-terminal domain. This may be attributed to the first 43 amino acids in the N-terminal domain of MLC-1v being unstructured without a binding partner, but morphing into structure to form a complex with the C-terminal domain of actin to modulate actin-myosin cross-bridge formation during cardiac contraction.[276]

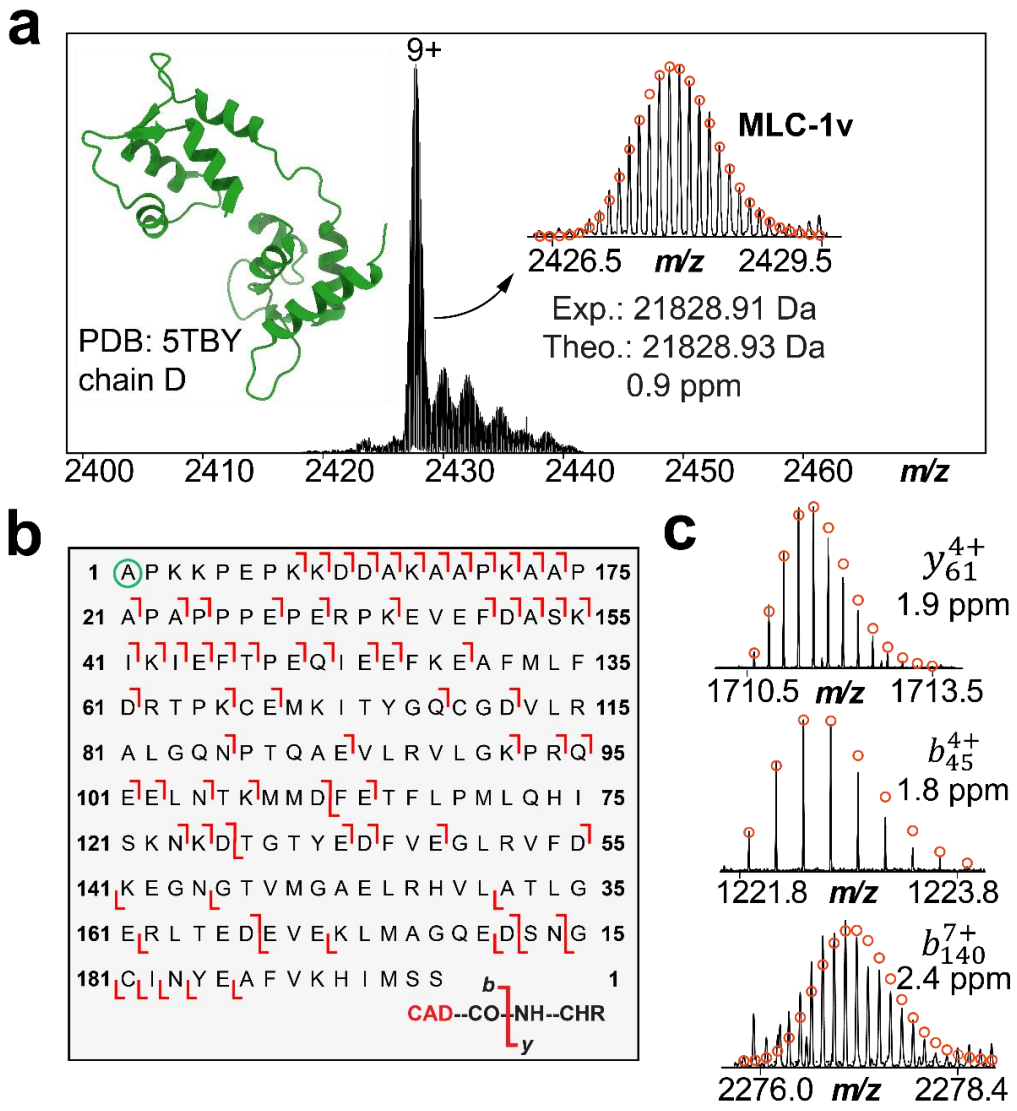


Figure 4.4. nTDMS characterization of endogenous MLC-1v. (a) MS1 native isolation spectra of endogenous MLC-1v (z = 9+). Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy proteoform characterization. PDB: 5TBY, chain D. **(b)** nTDMS CAD fragmentation map achieving 39% bond cleavage. The green circle is representative of N,N,N-trimethylation. **(c)** Representative fragment ions obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion.

Finally, we applied our nTDMS method to characterize α-Tpm (~32.7 kDa) directly from human cardiac tissue lysate. α-Tpm initially exists as a monomer in solution before undergoing

polymerization to form a dimeric α-helical coiled coil protein. During polymerization, α-Tpm molecules overlap head-to-tail, with the C-terminus of one monomer binding to the N-terminus of the adjacent monomer.[277] The precise alignment at the N- and C-termini is critical for the regulation of cardiac contraction. Specifically, only the dimeric α-helical coiled coil α-Tpm can serve as a binding scaffold for cardiac myosin and actin in response to an increase in intracellular Ca^{2+} during cardiac contraction.[277-279] In our native mass spectra, we detected the monomeric form α-Tpm which was lower abundant than MLC-2v and MLC-1v, displaying a charge state distribution from 11+ to 13+ between 2500 *m/z* to 2900 *m/z*. The native isolation spectra of α-Tpm (*z* = 12+) revealed two proteoforms with high mass accuracy, including unphosphorylated α-Tpm (32730.69 Da) and phosphorylated (32810.65 Da) α-Tpm (**Figure 4.5a, Table S4.1**). To obtain additional structural insights, CAD fragmentation of α-Tpm required a high collision voltage (75 V) to generate sequence informative fragment ions (**Figure S4.6**). We achieved 22% bond cleavage by CAD fragmentation with 47 *b* and 14 *y* fragment ions identified (**Figure 4.5b-c**), suggesting that the N-terminal region of α-Tpm is more solvent exposed than the C-terminal region. The cryo-EM structure of α-Tpm (PDB: 6X5Z) lacks resolution for both the N- and C-termini, as the central region forms the α-helical coiled coil. This explains the predominant observation of small terminal fragment ions, as opposed to larger fragment ions containing the central residues of α-Tpm. Notably, we detect α-Tpm as a monomer in our native mass spectra, rather than as a part of an α-Tpm homodimer or its reported supramolecular complex with actin and myosin.[280] Within this complex, the α-Tpm homodimer acts as a binding scaffold for actin and myosin, playing a crucial role in mediating cardiac contractility. Our nTDMS data on monomeric α-Tpm suggests a notable flexibility and lack of twisting at the N- and C-termini, likely attributed to the unstructured state of monomeric α-Tpm in the absence of its respective binding partners.

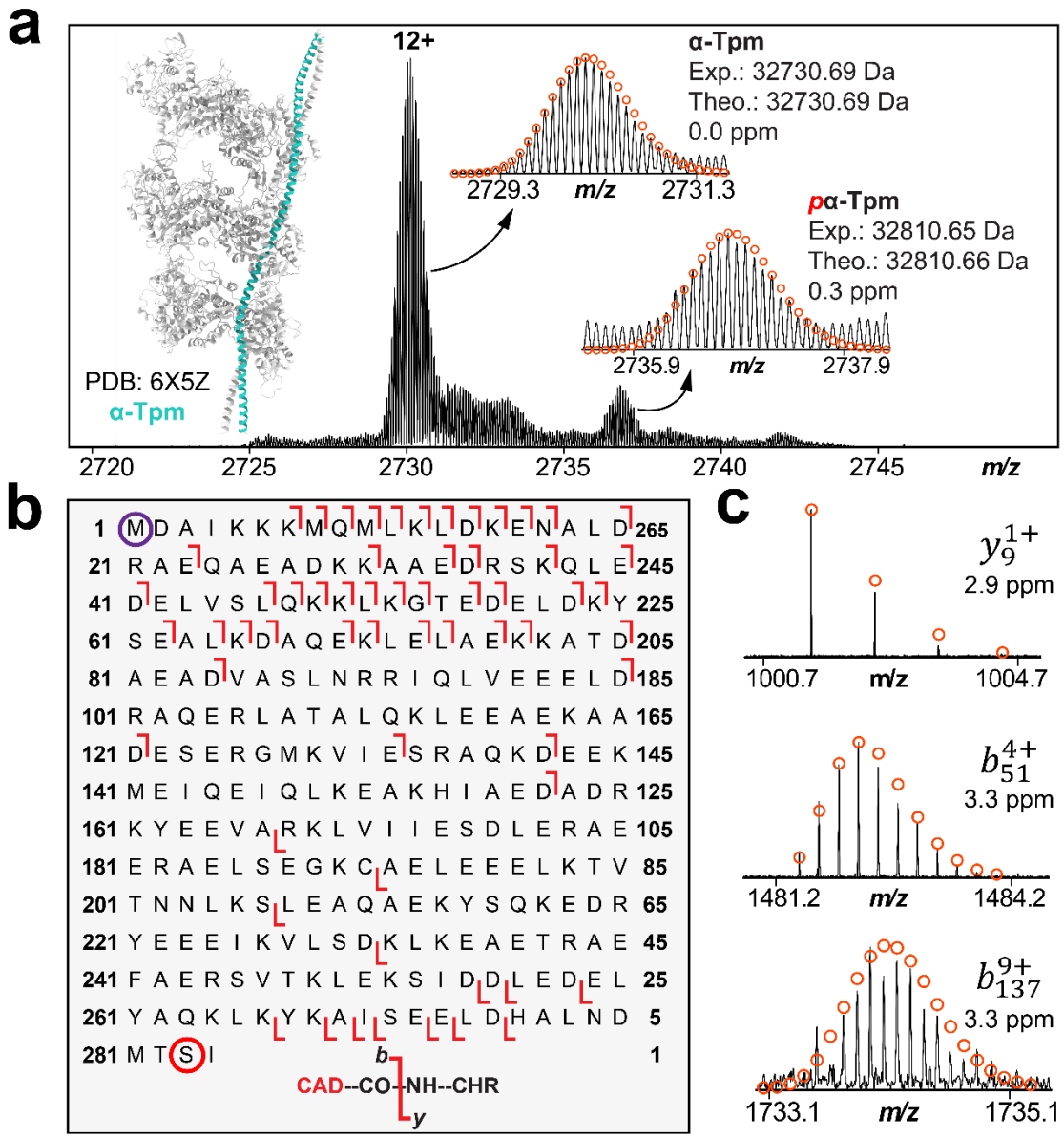


Figure 4.5. nTDMS characterization of endogenous α-Tpm. (a) MS1 native isolation spectra of endogenous α-Tpm (z = 12+). Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectrum to illustrate the high mass accuracy proteoform characterization. PDB: 6X5Z. **(b)** nTDMS CAD fragmentation map achieving 22% bond cleavage. The purple and red circles are representative of N-acetylation and phosphorylation, respectively. **(c)** Representative fragment ions obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion.

In turn, these findings may imply that monomeric α-Tpm lacks the ability to bind to other proteins, necessitating its operation as a dimer to function as a binding scaffold during cardiac contraction. Moreover, nTDMS-based detection and characterization of high molecular weight (MW) proteins and protein complexes from complex biological mixtures is challenging due to the exponential decrease in signal-to-noise ratio (S/N) with increasing MW and co-elution of low-MW proteins.[84, 113, 281] While, the presence of the dimeric form of α-Tpm in our tissue lysate is plausible; its detection may have been hindered by signal suppression from co-eluting low-MW proteins such as MLC-2v, MLC-1v, and α-Tpm monomer.

Overall, nTDMS characterization using FTICR-MS provided detailed and unique structural insights into the higher-order structure of MLC-2v, MLC-1v, and α-Tpm, establishing a direct link to functions in their endogenous environment. Our nTDMS approach, involving direct infusion FTICR-MS analysis of cardiac tissue lysates demonstrated simplicity and high- efficiency without the need for prior chromatographic separation methods. Notably, we have been able to detect and comprehensively characterize important sarcomeric proteins which are critical to cardiac function. Such straightforward direct infusion FTICR-based nTDMS workflows could be used for rapid tissue characterization or quality control purposes. Nevertheless, we recognize the significant challenge that remains in the direct analysis of complex heart tissue lysate due to the high dynamic range resulting in signal suppression through a concentration distribution spanning several magnitudes in endogenous matrices. This limitation restricts the characterization to only a few highly abundant proteins populating the mass spectrum. For example, the proteins identified in our study include some of the most abundant contractile proteins found in the sarcomere. This observation is consistent with previous native MS studies that directly analyzed proteins and protein complexes purified from biological samples, such as rhodopsin[48] and aquaporin-o[258] from

bovine eyes. As an alternative to conventional ESI, nanospray-desorption electrospray ionization (nano-DESI) for native mass spectrometry imaging (MSI) is a promising technique for top-down characterization directly from tissues. However, as noted by Cooper and co-workers, optimization of this method for top-down MS analysis is difficult due to the relatively low signal intensities observed.[257] Conceivably, multi-dimensional chromatographic strategies incorporating hydrophobic interaction chromatography (HIC)[87] and size exclusion chromatography (SEC)[79, 86] need to be developed for nTDMS to provide a deeper coverage into the human heart proteome. Moreover, high specificity enrichment methods could facilitate the identification of low abundant proteins and protein complexes from complex biological matrices.[30] Finally, we envision leveraging additional gas-phase filtering technologies, such as ion mobility[282] or quadrupole mass selection techniques, to select and enhance the signals of low abundant proteins in the native mass spectrum.

4.4 Conclusions

In summary, we have comprehensively characterized endogenous proteoforms directly from human cardiac tissue lysates using nTDMS and FTICR-MS, without prior separation strategies. nTDMS fragmentation with CAD and ECD either individually or in combination provided complimentary information on metal binding, PTMs, and protein structure. When directly infusing cardiac tissue lysate for nTDMS analysis, we identified six endogenous proteoforms for subsequent characterization of cardiac proteins MLC-2v, MLC-1v, and α-Tpm. Our data indicated that two rounds of buffer exchange into 150 mM ammonium acetate with spin filters was effective and provided clean, baseline-resolved native mass spectra. For MLC-2v, MLC-1v, and α-Tpm, the application of FTICR-MS for nTDMS characterization yielded

distinctive structural insights into their higher-order structure which was not represented by their respective cryo-EM structures. This analysis established a direct correlation between the structural features of these cardiac proteins and their functions within their native environment in the sarcomere. Moreover, we verified the Ca^{2+} binding domain in MLC-2v to EF-hand domain I between amino acid residues Lys45 and Asn59, which has not been previously reported. Overall, we demonstrate that direct injection of heart tissue lysates by nTDMS using an FTICR mass spectrometer provides unique structural insights, while offering a snapshot of the native proteome of human heart tissue.

4.5 Methods

4.5.1 Chemicals and Reagents

Ammonium acetate was purchased from Fisher Scientific (Fair Lawn, NJ, USA). Bio-Spin® columns with Bio-Gel® P-30 were purchased from Bio-Rad (Hercules, CA, USA).

4.5.2. Native Protein Extraction from Human Heart Tissues

Cardiac tissue (500 mg) was homogenized on ice using a Polytron homogenizer in 10 volumes (mL/g tissues) of native wash buffer (5 mM NaH_2PO_4, 5 mM Na_2HPO_4 (pH 7.0), 100 mM NaCl, 125 mM L-Methionine (pH 7.5), 1 mM PMSF and 1X Halt^{TM} protease and phosphatase inhibitor cocktail). The homogenate was centrifuged at 17,000 × ***g*** for 10 minutes at 4 ºC, and the supernatant was discarded. The washing and homogenization step was repeated, and the supernatant was discarded again. The protein pellet was then extracted using 5 volumes (mL/g tissues) of a native lithium chloride (LiCl) buffer at physiological pH (25 mM Tris (pH 7.5), 700 mM LiCl, 125 L-Methionine (pH 7.5), 1 mM PMSF and 1X Halt^{TM} protease and phosphatase inhibitor cocktail). The resulting homogenate was incubated at 4 ºC for 10 min and then centrifuged at 17,000 × ***g*** for 3 minutes at 4 ºC. The supernatant was transferred to new Eppendorf Protein Lo-Bind tubes and further centrifuged at 21,000 × ***g*** for 30 min at 4 ºC. The supernatants were once again transferred to new Eppendorf Protein Lo-Bind tubes, snap frozen in liquid nitrogen, and stored at -80 ºC.

4.5.3. Heart Tissue Lysate Preparation for nTDMS Analysis

Briefly, native protein extracts were thawed out on ice. 75 µL of thawed extract was then buffer-exchanged twice into 150 mM ammonium acetate using Bio-Spin® columns with Bio-Gel® P-30, following the protocol detailed from Bio-Rad resulting in a volume of 75 µL with a protein

concentration of approximately 0.7 μg/μL. The solution was then stored at 4 ºC until nTDMS analysis.

4.5.4. nTDMS Analysis using FTICR-MS

Samples were analyzed by nano-ESI via direct infusion using a TriVersa Nanomate system (Advion BioSciences, Ithaca, NY, USA) coupled to a 12-Tesla solariX XR FTICR mass spectrometer (Bruker Daltonics, Billerica, MA). 5 μL (3.5 μg) of sample was injected using the TriVersa autosampler. For the nano-ESI source using a TriVersa Nanomate, the desolvating gas pressure was maintained between 0.5-0.8 PSI and the voltage was set to 1.5-1.8 kV. The source dry gas flow rate was set to 4 L/min at 180 ℃. For the source optics, the capillary exit, deflector plate, funnel 1, skimmer voltage, funnel RF amplitude, octopole frequency, octopole RF amplitude, collision cell RF frequency, collision cell energy, and collision cell RF amplitude were set at 190.0 V, 200.0 V, 120.0 V, 100.0 V, 300.0 Vpp, 2 MHz, 600.0 Vpp, 1.4 MHz, 4-12 V, and 2000.0 Vpp, respectively. Mass spectra were acquired with an acquisition size of 2M-4M in the mass range between 200-8000 *m/z*. Ions were accumulated in the collision cell for 1 to 30 s and a time-of-flight of 2.000 ms was used for their transfer to the ICR cell. For broadband collisionally-activated dissociation (CAD) MS/MS experiments, the collisional energy was varied between 60-75 V. For electron-capture dissociation (ECD) experiments, ions were pulsed for 50 ms, with a lens bias of 1.0 V, and a lens voltage ranging from 12-15 V.

4.5.5. Human Cardiac Tissue Collection

Left ventricular (LV) myocardium from healthy donor hearts with no history of heart disease were obtained from the University of Wisconsin (UW)-Madison Organ and Tissue Donation-Surgical Recovery and Preservation Services. Heart tissues were kept in cardioplegic solution until

dissection, where they were immediately flash-frozen in liquid nitrogen and stored at -80 ºC. The procedures for the collection of human donor heart tissues were approved by the UW-Madison Institutional Review Board (IRB) (Protocol number 2013-1264).

4.5.6. Data Analysis

All mass spectrometry data were processed with Compass DataAnalysis v. 4.3 (Bruker Daltonics) software and Native MASH v. 1.0 software.[156] For MS1 peak picking, the Sophisticated Numerical Annotation Procedure (SNAP) peak-picking algorithm (quality factor: 0.4; signal-to-noise ratio (S/N): 3.0) was used. MS and MS/MS data were output from the DataAnalysis software and analyzed using MASH Native software using the TopFD deconvolution algorithm for sequence mapping and proteoform identification. All fragment ions were manually validated using a 20-ppm mass tolerance.

4.6 Supplemental Information

4.6.1. Supplemental Tables

Table S4.1. Comparison of theoretical and experimental monoisotopic masses for all proteins analyzed. Protein, modified forms, theoretical monoisotopic mass (Da), experimental monoisotopic mass (Da), and mass error (ppm) for the proteoforms we identified. Proteoforms were manually identified based on highly accurate intact mass measurements and MS/MS data together with prior knowledge of identified sarcomeric proteoforms from previous publications[39, 119, 144, 169] or UniProt.

Protein	Modified Forms	Theo Mass (Da)	Expt Mass (Da)	Mass Error (ppm)
MLC-2v	Met removal, N,N,N-trimethyl	18688.38	18688.37	0.5
MLC-2v	Met removal, N,N,N-trimethyl, Ca(II)	18726.33	18726.34	0.5
MLC-2v	Met removal, N,N,N-trimethyl, phospho	18768.35	18768.33	1.0
MLC-1v	Met removal, N,N,N-trimethyl	21828.93	21828.91	0.9
α-Tpm	N-acetyl	32730.69	32730.69	0
α-Tpm	N-acetyl, phospho	32810.66	32810.65	0.3

4.6.2. Supplemental Figures

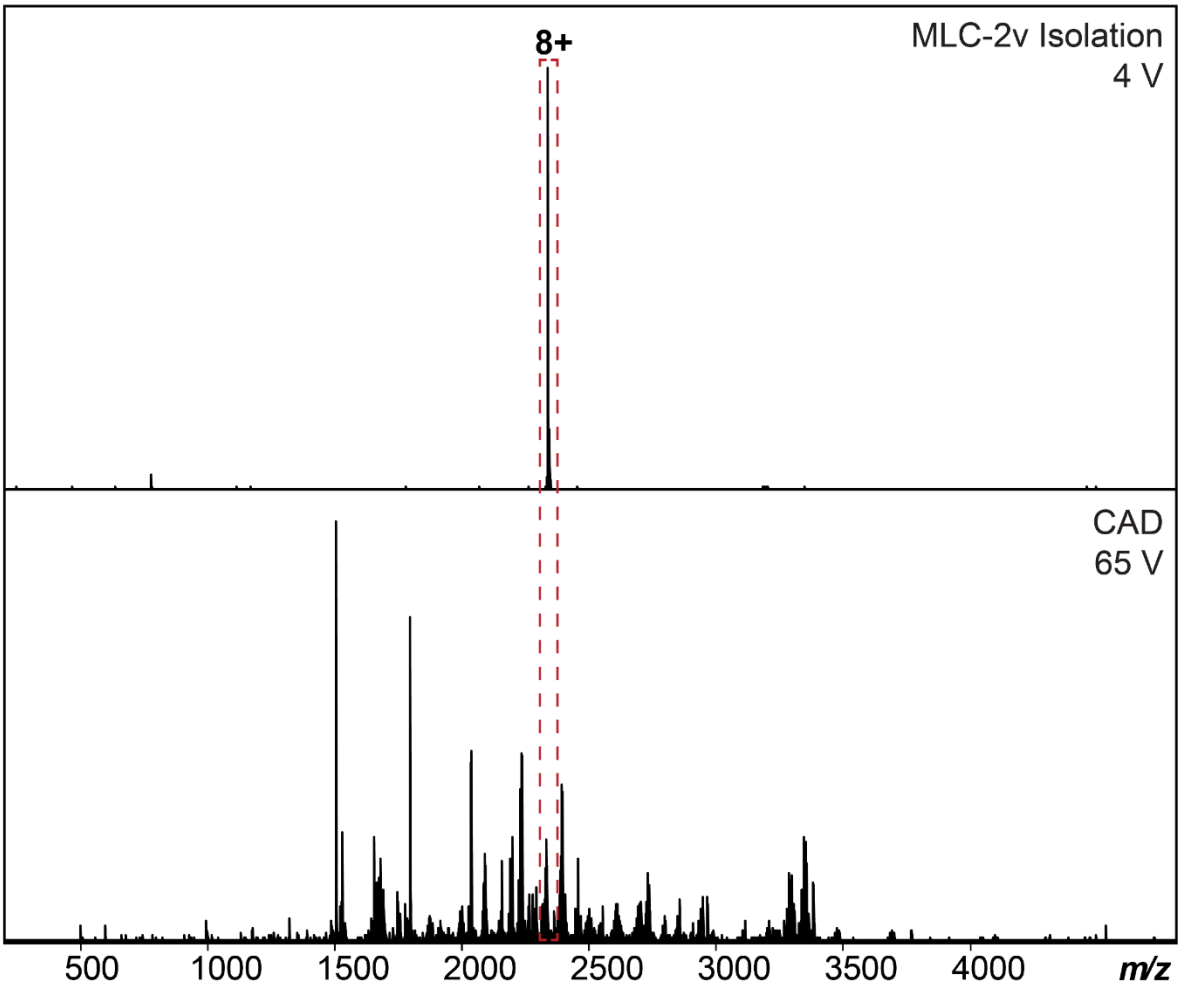


Figure S4.1. nTDMS characterization of endogenous MLC-2v using CAD. MS1 native isolation of MLC-2v (z = 8+, centered at 2343 m/z) followed by nTDMS analysis of the isolated precursor using a CAD energy of 65 V applied in the collision cell to generate sequence informative fragment ions.

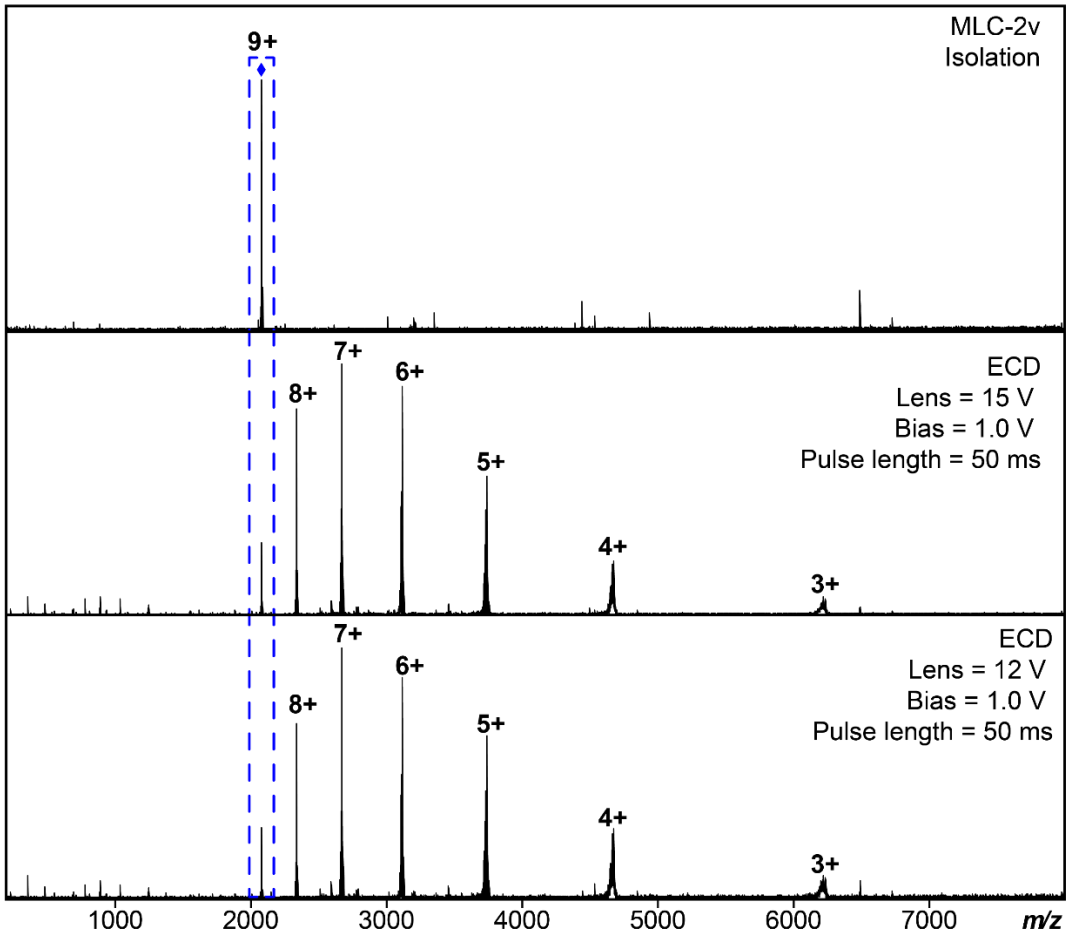


Figure S4.2. nTDMS characterization of MLC-2v using ECD. MS1 native isolation of MLC-2v (*z* = 9+, centered at 2082 *m/z*) followed by ECD fragmentation of the isolated precursor.

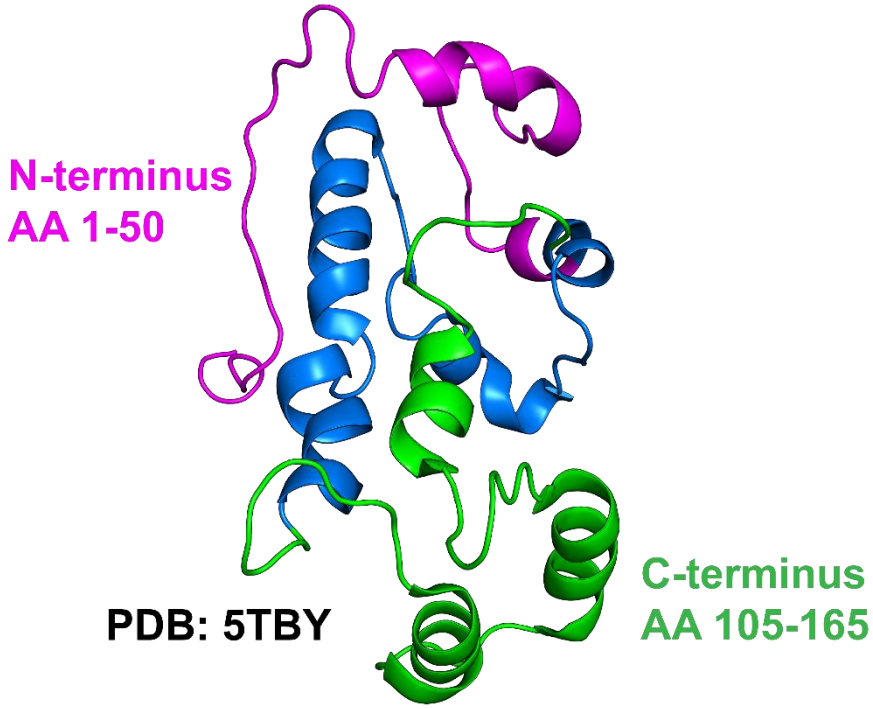


Figure S4.3. Structural representation of MLC-2v. MLC-2v crystal structure with the first 50 amino acid (AA) residues in the N-terminal domain highlighted in magenta and the last 50 AA residues in the C-terminal domain highlighted in green. The first 6 AA residues were not present in the PDB (5TBY, Chain F) structure.[283]

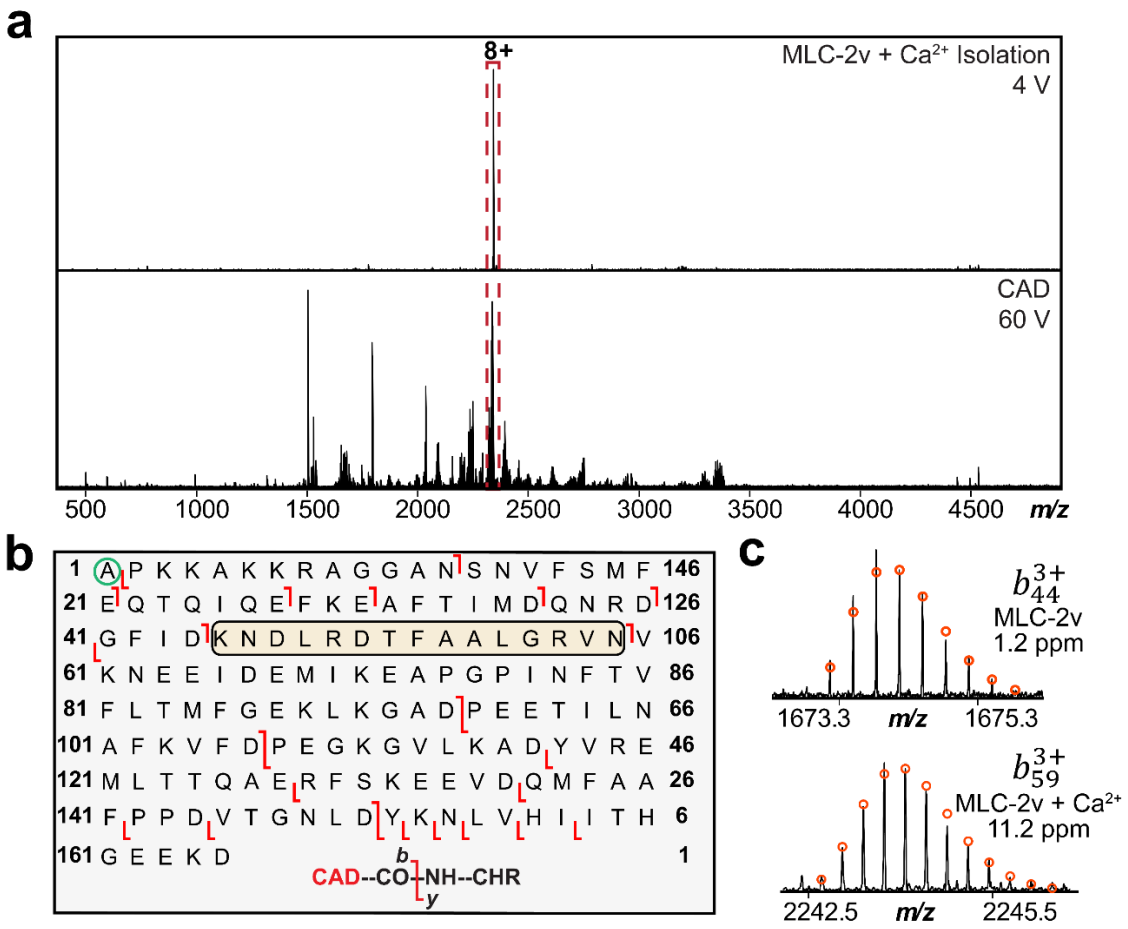


Figure S4.4. Localization of Ca^{2+} binding in MLC-2v by nTDMS. (a) MS1 native isolation of the endogenous MLC-2v proteoform with one Ca^{2+} bound (z = 8+, centered at 2343.5 *m/z*) followed by nTDMS characterization of the isolated precursor. A CAD energy of 60 V was applied in the collision cell to generate sequence informative fragment ions. **(b)** Sequence table of MLC-2v showing Ca^{2+} localized between Lys45 and Asp59. **(c)** Representative CAD fragment ions (b_{44}^{3+}, b_{59}^{3+}) obtained from the nTDMS analysis. Theoretical ion distributions are indicated by the red circles and mass accuracy errors are listed for each fragment ion.

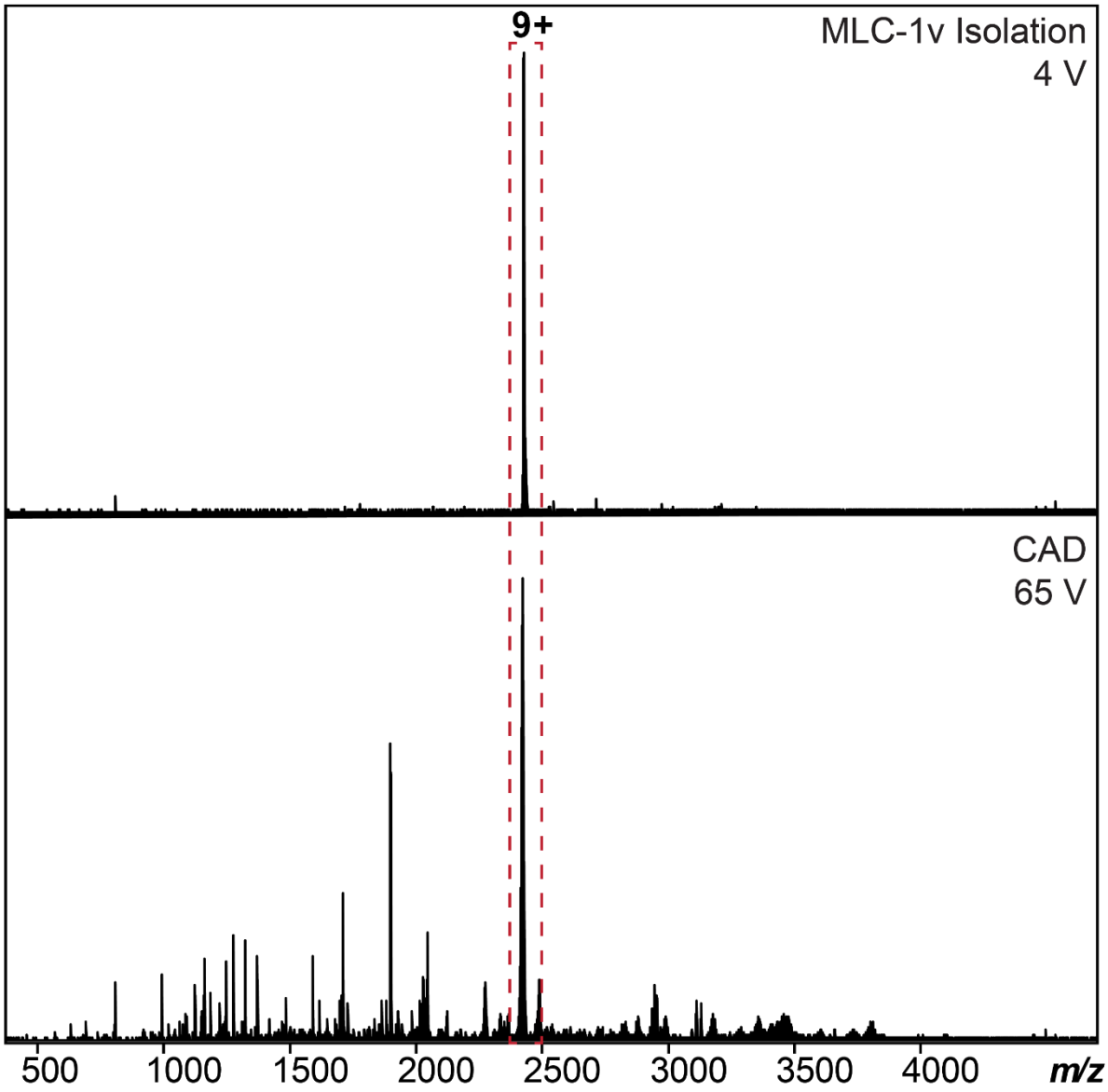


Figure S4.5. Characterization of endogenous MLC-1v using nTDMS. MS1 native isolation of endogenous MLC-1v (z = 9+, centered at 2433 m/z) followed by nTDMS analysis of the isolated precursor. A CAD energy of 65 V was applied in the collision cell to obtain sequence informative fragments from the MS/MS spectra.

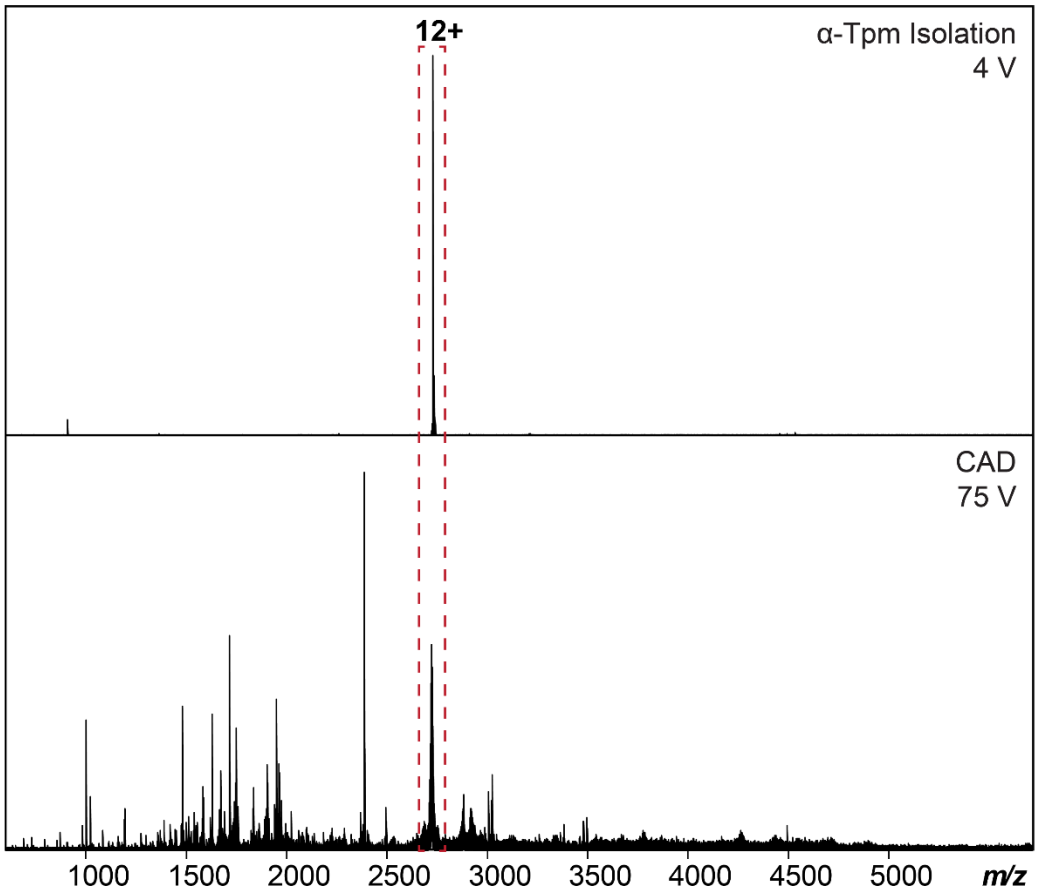
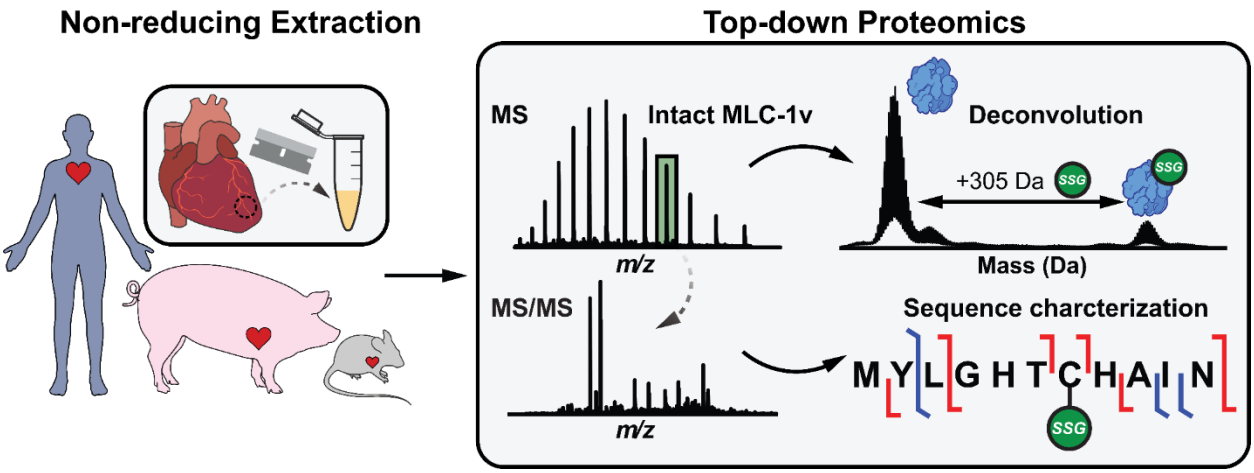


Figure S4.6. Characterization of endogenous α-Tpm using nTDMS. MS1 native isolation of α-Tpm (*z* = 12+, centered at 2737 *m/z*) followed by nTDMS characterization of the isolated precursor using CAD. A CAD energy of 75 V was applied in the collision cell to obtain sequence informative fragment ions.

CHAPTER 5:

In-Depth Characterization of S-Glutathionylation in Ventricular Myosin Light Chain 1 Across Species by Top-Down Proteomics



Portions of this chapter are adapted from the following pre-print:

Chapman, E.A.; Rogers, H.T.; Gao, Z.; Chan, H.; Alvarado, F.J.; Ge, Y. In-Depth Characterization of S-Glutathionylation in Ventricular Myosin Light Chain 1 Across Species by Top-Down Proteomics. *bioRxiv* **2024**. DOI: https://doi.org/10.1101/2024.12.11.628048.

5.1 Abstract

S-glutathionylation (SSG) is increasingly recognized as a critical signaling mechanism in the heart, yet SSG modifications in cardiac sarcomeric proteins remain understudied. Here we identified SSG of the ventricular isoform of myosin light chain 1 (MLC-1v) in human, swine, and mouse cardiac tissues using top-down mass spectrometry (MS)-based proteomics. Our results enabled the accurate identification, quantification, and site-specific localization of SSG in MLC-1v across different species. Notably, endogenous SSG of MLC-1v was observed in human and swine cardiac tissues but not in mice. Treating non-reduced cardiac tissue lysates with GSSG significantly increased SSG levels in MLC-1v across all three species.

5.2 Introduction

The balance between reactive oxygen species (ROS) and antioxidants is essential for maintaining redox homeostasis and regulating physiological function.[25] Disruption of this balance, often through increased ROS exposure, leads to the reversible oxidation of reactive cysteine (Cys) thiols through various post-translational modifications (PTMs), including S-glutathionylation (SSG).[284] SSG involves the conjugation of glutathione to a reactive protein Cys thiol, resulting in a 305.07 Da mass increase in the target protein.[285, 286] This modification can protect protein Cys thiols from irreversible oxidation and is increasingly recognized as a critical signaling mechanism in the heart, regulating key cellular processes such as oxidative phosphorylation, cellular metabolism, and protein synthesis.[287, 288] Cardiac sarcomeric proteins, most notably cardiac myosin-binding protein C (cMyPB-C), have been previously recognized as substrates of SSG.[287, 289, 290] However, SSG modifications remain understudied in endogenous tissues due to their low abundance and labile nature.[71, 287, 291] Additionally, it is unclear whether SSG is conserved across different species.

Top-down mass spectrometry (MS)-based proteomics analyzes intact proteins enabling the unambiguous identification and characterization of proteoforms - protein products from a single gene, including PTMs, alternative splice variants, and genetic mutations.[6, 34] In contrast to traditional bottom-up MS-based proteomics, which involves proteolytic digestion of proteins into small peptides, top-down proteomics allows for the precise identification and quantification of intact proteins and their associated proteoforms directly from complex biological samples without digestion.[39, 144] Additionally, top-down proteomics enables the fragmentation of intact proteoforms by tandem MS (MS/MS) for sequence characterization and site-localization of PTMs.[34] As such,

top-down proteomics provides an unbiased and detailed perspective on proteoform diversity and complexity.

Herein, we developed a top-down mass spectrometry (MS)-based proteomics platform to identify, quantify, and localize SSG modifications in sarcomeric proteins from human, swine, and mouse cardiac tissues. Our analysis unequivocally identified the ventricular isoform of myosin light chain 1 (MLC-1v) in human, swine, and mouse cardiac tissues as a substrate for protein SSG. Cardiac MLC-1v, also known as the essential light chain, is a key component of the hexameric myosin II complex and plays a crucial role in regulating cross-bridge cycling and cardiac contraction.[271, 292] Recently we have demonstrated that MLC-1v, rather than MLC-2v, is ventricle-specific in adult human hearts.[119] This study represents the first report of SSG in MLC-1v. Our results enabled the quantification and site-specific localization of SSG to Cys residues in MLC-1v across multiple species, demonstrating the power of top-down MS-based proteomics in uncovering novel PTMs within complex biological matrices.

5.3 Results & Discussion

A label-free top-down MS-based proteomics platform was used to investigate SSG modifications in sarcomeric proteins extracted directly from human (n = 12), swine (n = 6), and mouse (n = 8) left ventricular cardiac tissues **(Figure 5.1a, Figure S5.1)**. This platform involved tissue homogenization, extraction of proteins using a buffer without reducing agents to preserve protein SSG, and separation of intact proteins using reversed-phase liquid chromatography (RPLC). Online and offline MS/MS were then performed to comprehensively characterize protein SSG **(Figure 5.1a, Figure S5.1)**. The high reproducibility of our online LC-MS/MS method was

demonstrated between triplicate injections of the same amount of total protein **(Figure S5.2)**, enabling quantitative analysis of SSG proteoforms.

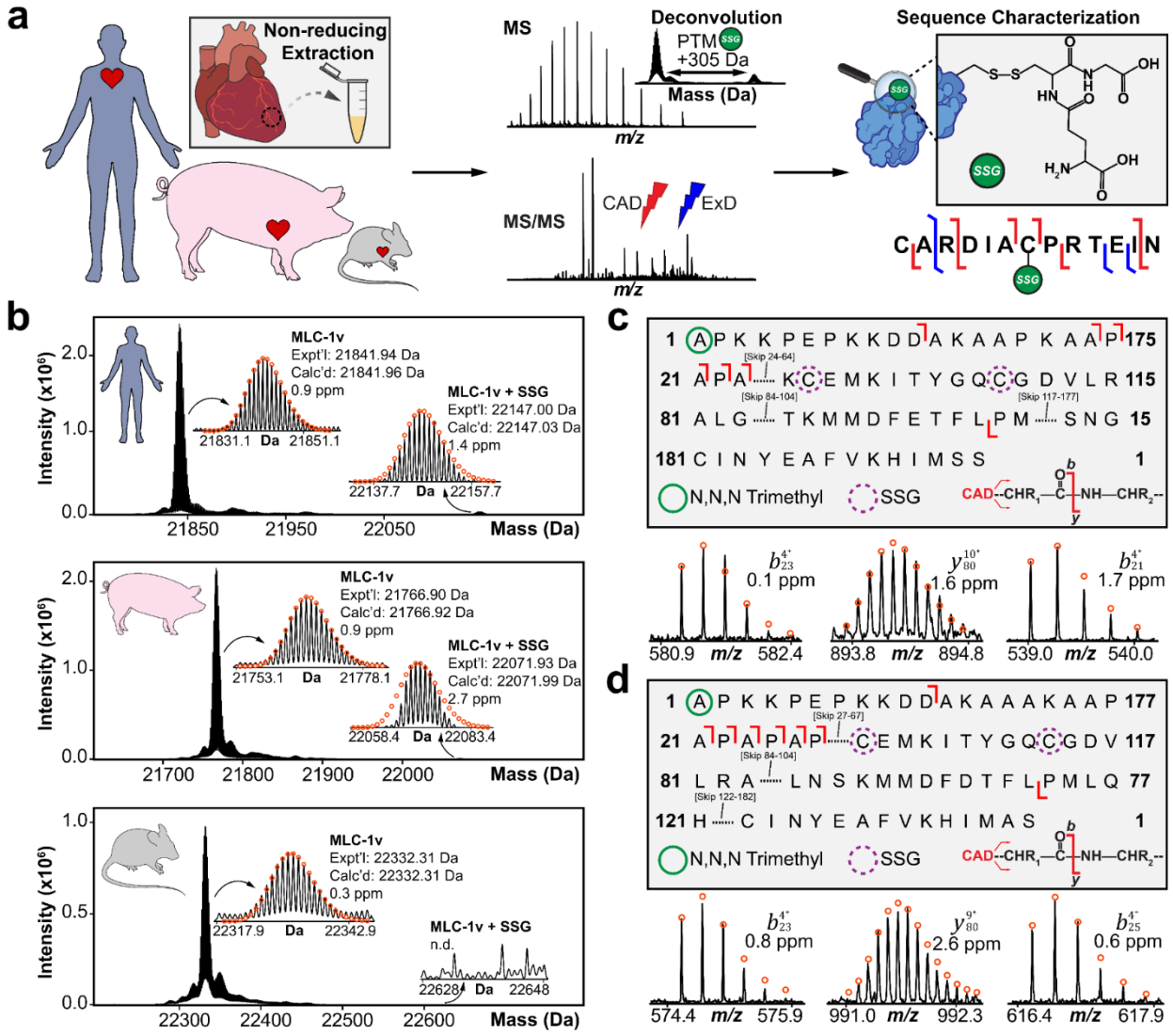


Figure 5.1. Endogenous SSG of MLC-1v is detected in human and swine cardiac tissues, but not in mouse tissues using top-down proteomics. (a) Top-down proteomics workflow for the identification of protein SSG (+305.07 Da) in human, swine, and mouse cardiac tissues. Intact sarcomeric proteins were extracted from cardiac tissue using a non-reducing extraction buffer. MS and MS/MS data were acquired for intact protein mass analysis, sequence characterization, and SSG localization. **(b)** Representative high-resolution deconvoluted mass spectra of MLC-1v in human, swine, and mouse cardiac tissues. Theoretical isotope distributions (red circles) are overlaid on the experimentally obtained mass spectra. The MLC-1v + SSG proteoform was isolated and fragmented in human and swine using collisionally activated dissociation

(CAD). The CAD fragmentation map for **(c)** human and **(d)** swine MLC-1v shows SSG is modified on either Cys66 or Cys75 in human and Cys68 or Cys77 in swine. The green circle represents N,N,N-trimethylation, while the dashed purple circle represents potential SSG sites. Representative CAD fragment ions are shown below the fragmentation maps. All individual ion assignments are within 10 ppm of the theoretical mass and the theoretical isotopic distributions are indicated by the red circles.

5.3.1. Top-Down Proteomics Reveals Endogenous SSG of MLC-1v

Our top-down proteomics data revealed low-abundant endogenous SSG on MLC-1v in human **(Figure 5.1b, Figure S5.3)** and swine **(Figure 5.1b, Figure S5.4)** cardiac tissues with high mass accuracy, indicated by a 305.07 Da mass increase from the most abundant MLC-1v proteoform **(Table S5.1)**. Interestingly, endogenous SSG was not detected in mouse cardiac tissues on MLC-1v **(Figure 5.1b, Figure S5.5)**. The total amount of SSG detected on MLC-1v (%SSG) was quantified by the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all MLC-1v proteoforms. Total SSG levels of MLC-1v in human tissues ranged from 0.2% to 1.5% across twelve biological replicates and from 0.1% to 0.4% across six biological replicates in swine tissues. No significant sex-based differences were detected in total SSG levels in our sex-balanced human cohort (6 females and 6 males) **(Figure S5.6)**. To ensure the reproducibility of our online LC-MS/MS method for quantification of endogenous MLC-1v SSG proteoforms, we performed three extraction replicates using human **(Figure S5.7)** and swine cardiac tissues **(Figure S5.8)**. Extraction reproducibility was demonstrated by the highly similar base peak chromatograms (BPC) and comparable %SSG values calculated for MLC-1v between extraction replicates. To localize the site of endogenous SSG in human and swine MLC-1v, we performed online LC-MS/MS with collisionally activated dissociation (CAD) to produce N- and C- terminal protein fragment ions of the isolated MLC-1v SSG proteoform **(Figure 5.1c-d, Figure S5.9-S5.10)**. Fragment ions confirmed N,N,N-trimethylation to Ala1 in both human and

swine cardiac tissues, as previously reported.[119, 193] However, due to the low abundance of the endogenous SSG proteoform, we could only narrow down the site of SSG to be on either Cys66 or Cys75 in human and Cys68 or Cys77 in swine, ruling out Cys181 and Cys183 as possible sites of SSG in human and swine MLC-1v, respectively. Due to the low signal-to-noise ratio of fragment ions in the online MS/MS spectra, which hindered site-localization of the SSG modification, we explored an alternative approach to further fragment and localize SSG in MLC-1v.

5.3.2. Treatment with Oxidized Glutathione (GSSG) Leads to Increased SSG Levels in MLC-1v

Human, swine, and mouse cardiac tissue lysates were treated with 1 mM of oxidized glutathione (GSSG) prior to top-down proteomics analysis. We observed a significant increase in the abundance of the MLC-1v SSG proteoform in human **(Figure 5.2a, Figure 5.2d, Figure S5.11)** and swine **(Figure 5.2b, Figure 5.2e, Figure S5.12)** cardiac tissues following GSSG treatment, compared to controls. Interestingly, SSG of MLC-1v was detected in mouse cardiac tissues after GSSG treatment, with total SSG levels ranging from 0.6% to 1.1% across three biological replicates **(Figure 5.2c, Figure 5.2f, Figure S5.13)**. To determine if GSSG treatment induced SSG of other major sarcomeric proteins detected in the LC-MS/MS analysis, we analyzed the deconvoluted mass spectra of cardiac troponin T (cTnT), cardiac troponin I (cTnI), alpha-tropomyosin (α-Tpm), troponin C (TnC), ventricular isoform of myosin light chain 2 (MLC-2v), and actin in human **(Figure S5.14)**, swine **(Figure S5.15)**, and mouse **(Figure S5.16)**. Evidently, we did not identify cTnT, cTnI, α-Tpm, TnC, or MLC-2v in human, swine, and mouse cardiac tissues to be modified with SSG following GSSG treatment, as we did not detect isotopically resolved proteoforms with a +305.07 Da mass shift from the most abundant proteoform with high mass accuracy. However, actin appears to be potentially modified with SSG in human, swine, and mouse cardiac tissues following incubation with 1 mM GSSG **(Figure S5.14-S5.16)**. To localize

the specific sites of SSG in MLC-1v in human, swine, and mouse cardiac tissues treated with GSSG, we employed a combination of offline and online MS/MS methods using CAD, electron transfer dissociation (ETD), and electron capture dissociation (ECD). For human MLC-1v, MS/MS analysis revealed key fragment ions containing the SSG modification, including b_{191}, y_{168}, and y_{171}, which closely matched the theoretical isotopic distribution with high mass accuracy **(Figure 5.2g, Figure S5.17)**. The detection of the y_{120} ion, which is not modified with SSG, confirmed Cys66 as the putative site of SSG in human MLC-1v. For swine MLC-1v, MS/MS analysis also revealed key fragment ions containing the SSG modification, including b_{193}, y_{169}, and y_{171} **(Figure 5.2h, Figure S5.18)**. The detection of the y_{127} and y_{128} ions, which are unmodified, confirmed Cys68 as the putative site of SSG in swine MLC-1v. Finally, MS/MS analysis of mouse MLC-1v confirmed N,N,N-trimethylation to Ala1 and ruled out Cys190 as the site of SSG modification, as C-terminal fragment ions, such as y_{80} and y_{61}, were unmodified **(Figure 5.2i, Figure S5.19)**. Based on the presence of the b_{125} ion containing the SSG modification with high mass accuracy, Cys84 was assigned as the putative site of SSG in mouse MLC-1v. Interestingly, sequence alignment of MLC-1v revealed that Cys66 in human and Cys68 swine are conserved, whereas this Cys residue is not conserved in mouse **(Figure S5.20)**. Instead, Cys84 in mouse, which can be modified with SSG, is conserved with Cys77 in human and Cys75 in swine; however, these residues were found to be unlikely substrates for SSG. Therefore, these findings highlight the species-specific nature of SSG and underscore the differences in how this modification may be regulated across species.

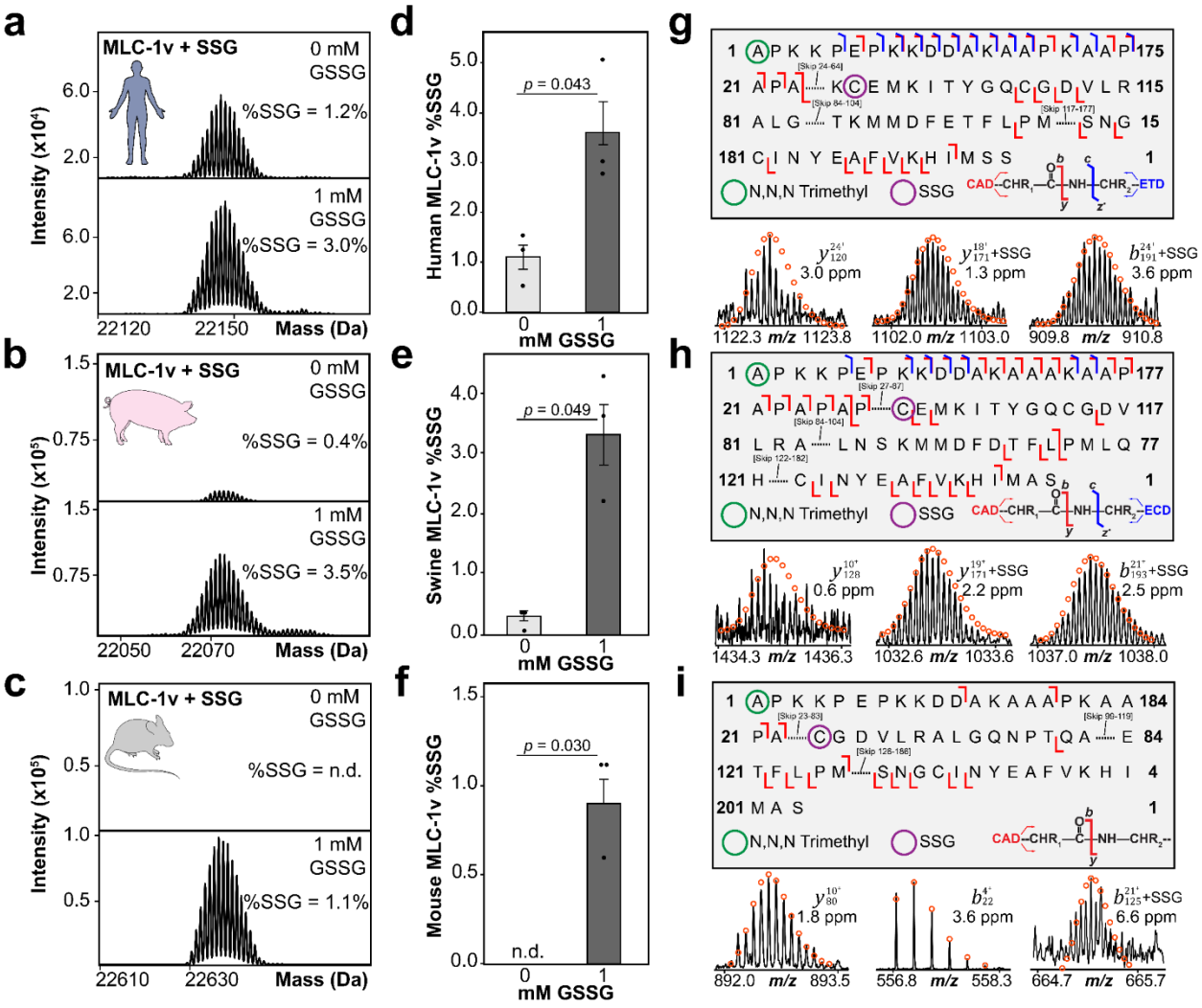


Figure 5.2. Top-down proteomics reveals treatment of non-reduced cardiac tissue lysates with GSSG increases MLC-1v SSG levels in human, swine, and mice. Representative deconvoluted mass spectra of the **(a)** human, **(b)** swine, and **(c)** mouse MLC-1v + SSG proteoform in control (0 mM GSSG) and treated (1 mM GSSG) samples. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v. Quantification of %SSG in **(d)** human, **(e)** swine, and **(f)** mouse MLC-1v for control (0 mM GSSG) versus treated (1 mM GSSG) samples ($n = 3$). Groups were considered significantly different by paired student's t-test with $p < 0.05$. Error bars represent the mean ± standard error of the mean (SEM). **(g)** The collisionally activated dissociation (CAD) and electron transfer dissociation (ETD) fragmentation map for human MLC-1v (1 mM GSSG) shows SSG is localized to Cys66. The green circle represents N,N,N-trimethylation, while the purple circle represents SSG. Representative CAD fragment ions are shown below the fragmentation map. All individual ion assignments are within 10 ppm of the theoretical mass, and the

theoretical isotopic distributions are indicated by the red circles. **(h)** The CAD fragmentation map for swine MLC-1v (1 mM GSSG) shows SSG is localized to Cys68. Representative CAD fragment ions are shown below the fragmentation map. **(i)** The CAD fragmentation map for mouse MLC-1v (1 mM GSSG) shows SSG is localized to Cys84. Representative CAD fragment ions are shown below the fragmentation map.

In this study, we employed top-down MS-based proteomics to identify, quantify, and localize SSG in MLC-1v from human, swine, and mouse cardiac tissues. To our knowledge, this is the first time SSG has been reported on MLC-1v. Previous studies have shown that key sarcomeric proteins, particularly cMyBP-C as well as titin, actin, and cTnI, may act as substrates for SSG, which have advanced our understanding of SSG in the sarcomere.[287, 289, 290, 293] Titin or cMyBP-C were not detected in the current top-down 1DLC-MS data due to their large molecular weights, as in top-down proteomics, the S/N decreases with increasing protein size.[56] Although we detected cTnI in our top-down MS data, SSG of cTnI was not observed in non-failing hearts using the non-reducing method employed in this study. A previous study has shown that SSG levels of cTnI are increased in end-stage human failing hearts compared to non-failing hearts.[287, 290] However, other studies have not detected SSG of cTnI.[294, 295] For example, in isolated rat control hearts, SSG was not reported on cTnI but was found to be specific to cMyBP-C.[294]-Interestingly, actin appears to be modified with SSG in human, swine, and mouse cardiac tissues upon incubation with 1 mM GSSG **(Figure S5.14-S5.16)**. Nevertheless, further effort needs to be allocated to isotopically resolve the presumable actin SSG proteoform and subject to MS/MS to confirm this modification. In this study, we focused on the in-depth characterization of SSG in MLC-1v at the basal level and after GSSG treatment in non-failing hearts from human, swine, and mouse cardiac tissues. MLC-1v is increasingly recognized for its essential role in modulating cardiac contractility by regulating the interaction between actin and myosin within the sarcomere.[271, 292] Recently we have demonstrated that MLC-1v, rather than MLC-2v, is exclusively expressed in the ventricles

in adult human hearts, making it a potential ventricular marker.[119] Note that SSG of MLC-1v was not detected in the previous studies[119,193] due to the presence of reducing agents added in the extraction procedures. Here, for the first time, our non-reducing extraction procedure enabled the accurate identification, quantification, and site-specific localization of SSG in MLC-1v across different species.

5.4 Conclusions

Overall, we identified low-abundance endogenous SSG on MLC-1v in human and swine cardiac tissues, but not in mouse tissues. Due to the low abundance of endogenous SSG on MLC-1v, samples were then treated with GSSG prior to MS analysis. Notably, SSG levels significantly increased in human and swine samples, and SSG was also detected in mouse samples following GSSG treatment. Additionally, MS/MS analysis confirmed the putative SSG sites in MLC-1v to be Cys66 in human, Cys68 in swine, and Cys84 in mouse. Overall, our results provide the first evidence that MLC-1v is a target of SSG in the sarcomere across different species, demonstrating the utility of top-down MS-based proteomics in identifying, quantifying, and localizing novel PTMs directly from cardiac tissues.

5.5 Methods

5.5.1. Chemicals and Reagents

All reagents were purchased from Sigma-Aldrich unless otherwise noted. Optima LC/MS grade acetonitrile and Titan3, 17mm PES membrane syringe filters were purchased from Thermo Fisher Scientific. Amicon, 0.5 mL cellulose ultra-centrifugal filters with a molecular weight cut-off (MWCO) of 10 kDa and HPLC grade LiChrosolv® Ethanol were purchased from MilliporeSigma.

5.5.2. Human Cardiac Tissue Collection

Left ventricular (LV) myocardium from healthy donor hearts with no history of heart disease were obtained from the University of Wisconsin (UW)-Madison Organ and Tissue Donation-Surgical Recovery and Preservation Services. The procedures for the collection of human donor heart tissues were approved by the UW-Madison Institutional Review Board (IRB). A summary of the clinical characteristics of human hearts used in this study can be found in **Table S5.2**.

5.5.3. Swine and Mouse Cardiac Tissue Collection

Swine heart tissue was obtained from adult Yorkshire domestic swine. Mouse heart tissue was obtained from C57Bl/6J wild-types, with an equal ratio of male to female mice. The hearts were excised from the organism, and the LV was isolated and immediately snap-frozen in liquid nitrogen and stored at -80 ºC. All experiments involving animals were conducted in accordance with the NIH Guide for the Care and Use of Laboratory Animals and using protocols approved by the University of Wisconsin Institutional Animal Care and Use Committee.

5.5.4. Non-Reducing Protein Extraction from Human, Swine, and Mouse Cardiac Tissues

Proteins were extracted from cardiac tissues as previously described.[79, 114] Cardiac left ventricular tissue (50-500 mg) was homogenized on ice using a Polytron homogenizer in 10 volumes (mL/g tissues) of phosphate wash buffer (5 mM NaH_2PO_4, 5 mM Na_2HPO_4 (pH 7.0), 100 mM NaCl, 125

mM L-Methionine (pH 7.5), 1 mM PMSF and 1X Halt™ protease and phosphatase inhibitor cocktail. The homogenate was centrifuged at 17,000 × *g* for 3 minutes at 4 ºC, and the supernatant was discarded. The washing and homogenization step was repeated, and the supernatant was once again discarded. The protein pellet was extracted using 5 volumes (mL/g tissues) of a lithium chloride (LiCl) buffer (25 mM Tris (pH 7.5), 700 mM LiCl, 125 L-Methionine (pH 7.5), 1 mM PMSF and 1X Halt™ protease and phosphatase inhibitor cocktail). The resulting homogenate was incubated at 4 ºC for 10 min and then centrifuged at 17,000 × *g* for 3 minutes at 4 ºC. The supernatant was transferred to new Eppendorf Protein Lo-Bind tubes and centrifuged at 21,000 × *g* for 30 min at 4 ºC to further clarify the tissue lysates. The resulting supernatant was passed through a Titan3, 17mm PES membrane syringe filter (pre-soaked with LiCl buffer) using a 5 mL Luer-lock syringe, into a new Eppendorf Protein Lo-Bind tube, snap-frozen in liquid nitrogen, and stored at -80 ºC.

5.5.5. Sample Preparation for Top-Down Proteomics Analysis

Cardiac tissue lysates were thawed out on ice for 15 minutes, desalted using a 10 kDa MWCO filter (Amicon, 0.5 mL, cellulose, MilliporeSigma), and buffer exchanged using 0.1% formic acid in nanopure water. The concentration of the desalted tissue lysates was determined by the Bradford protein assay. For direct infusion offline MS/MS experiments, non-reduced cardiac tissue lysates were buffer-exchanged twice using 0.1% formic acid in 10:10:80 IPA:ACN:water using Bio-Spin columns with Bio-Gel P-30, following the protocol detailed from Bio-Rad.

5.5.6. Online Top-Down LC-MS/MS Data Acquisition

LC-MS/MS analysis was carried out using either an Acquity ultra-high performance LC M-Class system (Waters) coupled to a maXis II quadrupole time-of-flight mass spectrometer (Bruker Daltonics) or a NanoAcquity ultra-high performance LC system (Waters) coupled to an Impact II

quadrupole time-of-flight mass spectrometer (Bruker Daltonics). 500 ng of total protein was injected onto a home-packed PLRP column (PLRP-S) (Agilent Technologies), 10 μm particle size, 250 or 500 μm inner diameter, 1,000 Å pore size using an organic gradient of 5% to 95% mobile phase B (mobile phase A: 0.1% formic acid in water; mobile phase B: 0.1% formic acid in 50:50 acetonitrile:ethanol) at a flow rate of 8-12 μL/min and temperature of 60 ℃. Mass spectra were acquired at a scan rate of 0.5 Hz over 200-3000 *m/z*. For the electrospray ionization source, the end plate offset, capillary, nebulizer, dry gas, and dry temp were set at 500 V, 4500 V, 0.5 bar, 4.0 L/min, and 200 °C, respectively. For tune settings, the Funnel 1 RF, isCID energy, multipole RF (for maXis II), Funnel 2 RF (for Impact II), Hexapole RF (for Impact II), quadrupole ion energy, collision energy, collision RF, transfer time, and pre-pulse storage were set at 400 Vpp, 10 eV, 400 Vpp, 400 Vpp, 400 Vpp, 5 eV, 5 eV, 2000 Vpp, 100 μs, and 15 μs, respectively. Proteoforms of interest were first isolated in the gas phase and fragmented by either collisionally activated dissociation (CAD) or electron transfer dissociation (ETD). For targeted CAD MS/MS experiments (Impact II), the precursor ion was set with a width of at least 5 *m*/*z* and mass spectra were acquired at a scan rate of 1.0 Hz over 200 – 2000 *m/z*. Collision energies were set to values ranging from 8 to 20 eV. For targeted ETD MS/MS experiments (maXis II), 3,4 hexanedione was used as the ETD reagent. The precursor ion was set with a width of at least 5 *m/z* and mass spectra were acquired at a scan rate of 1.0 Hz over 200-3000 *m/z*. The analyte accumulation time was set to 1000 ms with a 30 ms reagent injection time, and extended reaction time varied from 10 to 100 ms.

5.5.7. Offline Top-Down MS/MS Data Acquisition

Samples were either directly infused[114], or individual protein fractions were collected following reversed-phase LC separation for offline MS/MS analysis. Protein fractions were analyzed by

nano-ESI via direct infusion using a TriVersa Nanomate system (Advion BioSciences) coupled to a 12-T solariX XR FTICR mass spectrometer (Bruker Daltonics).[84, 172, 193] For the nano-ESI source using a TriVersa Nanomate, the desolvating gas pressure was maintained between 0.7 and 0.8 PSI and the voltage was set to 1.7–1.8 kV. The source dry gas flow rate was set to 3 L/min at 180 °C. For the source optics, the capillary exit, deflector plate, funnel 1, skimmer voltage, funnel RF amplitude, octopole frequency, octopole RF amplitude, collision cell RF frequency, collision cell energy, and collision cell RF amplitude were set at 240.0 V, 220.0 V, 100.0 V, 50.0 V, 300.0 Vpp, 5 MHz, 600.0 Vpp, 2.0 MHz, 4V, and 2000.0 Vpp, respectively. Mass spectra were acquired with an acquisition size of 1 to 2M in the mass range between 200 and 3000 *m*/*z*. Ions were accumulated in the collision cell for 1 to 15 s and a time-of-flight of 1.000 ms was used for their transfer to the ICR cell. For offline CAD MS/MS experiments, the collisional energy was varied between 10 and 25 V. For offline electron-capture dissociation (ECD) experiments, ions were pulsed for 20 ms with a lens bias of 1.5 V and a lens voltage of 15 V.

5.5.8. Data Analysis

All MS data were processed and analyzed using Compass DataAnalysis v. 4.3 (Bruker Daltonics) software. All chromatograms were smoothened by the Gauss algorithm with a smoothing width of 2.04 s. The Sophisticated Numerical Annotation Procedure (SNAP) peak-picking algorithm (quality factor: 0.4; signal-to-noise ratio (S/N): 3.0) was applied to determine the monoisotopic mass of all detected ions. Mass spectra were deconvoluted using the Maximum Entropy algorithm within the DataAnalysis v. 4.3 software. The resolving power for Maximum Entropy deconvolution was set to 50,000k (Impact II) and 60,000k (maXis II) for proteins that were isotopically resolved. Protein S-glutathionylation (SSG) was quantified based on the ratio of the peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of the

same protein using the deconvoluted mass spectrum. Online and offline MS/MS data were output from the DataAnalysis v. 4.3 software and analyzed using MASH Native v. 1.1 software[156] using the eThrash algorithm[147] for sequence mapping and proteoform identification. All fragment ions were manually validated using a 10-ppm mass tolerance.

5.5.9. Statistical Analysis

Paired Student's t-tests were used to determine the level of statistical significance for the quantification of SSG modifications. All *p*-values at $p < 0.05$ were considered significant. Error bars indicated in the figures represent the mean ± standard error of the mean (SEM).

5.6 Supplemental Information

5.6.1. Supplemental Tables

Table S5.1. MLC-1v proteoforms identified in top-down MS experiments from human, swine, and mouse cardiac tissues. Species, proteoform, modified forms, theoretical most abundant mass, experimental most abundant mass, and mass error for the proteoforms identified. Modified proteoforms were identified based on highly accurate intact mass measurements and MS/MS data. Abbreviations: ventricular isoform of myosin light chain 1 (MLC-1v), methionine (Met), N,N,N-trimethylation (N,N,N-trimethyl), and S-glutathionylation (SSG).

Species	Proteoform	Modified Forms	Theo Mass (Da)	Expt Mass (Da)	Mass Error (ppm)
Homo sapien	MLC-1v	Met removal, N,N,N-Trimethyl	21841.96	21841.94	0.9
Homo sapien	MLC-1v	Met removal, N,N,N-Trimethyl, SSG	22147.03	22147.00	1.4
Sus scrofa	MLC-1v	Met removal, N,N,N-Trimethyl	21766.92	21766.90	0.9
Sus scrofa	MLC-1v	Met removal, N,N,N-Trimethyl, SSG	22071.99	22071.93	2.7
Mus musculus	MLC-1v	Met removal, N,N,N-Trimethyl	22332.31	22332.31	0.3
Mus musculus	MLC-1v	Met removal, N,N,N-Trimethyl, SSG	22637.37	22637.35	0.9

Table S5.2. Clinical characteristics of human hearts used. Non-failing donor human heart tissues were obtained from the University of Wisconsin (UW) Organ and Tissue Donation, Surgical Recovery, and Preservation Services. Average and standard error of the mean (SEM) is reported for age. Number of donors (*n*) is indicated in parentheses after each reported percentage.

Clinical Characteristic	Precentage of Sample Cohort
Age	55 ± 11.5 years
	min: 25 years
	max: 65 years
Male	50% (*n* = 6)
Female	50% (*n* = 6)

5.6.2. Supplemental Figures

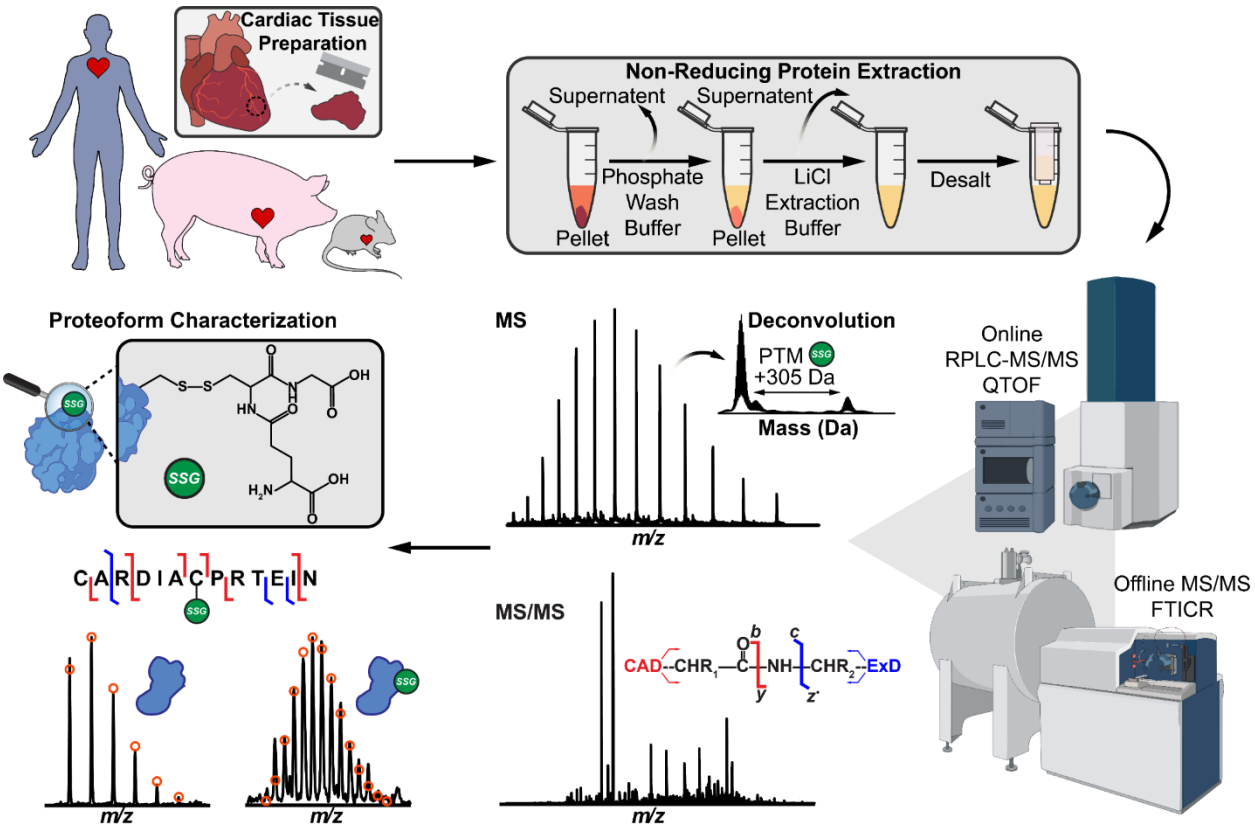


Figure S5.1. Top-down MS-based proteomics workflow for the identification of protein SSG in human, swine, and mouse cardiac tissues. Human (n = 12), swine (n = 6), and mouse (n = 8) cardiac tissues were first homogenized in phosphate wash buffer. Sarcomeric proteins were extracted from cardiac tissue using a non-reducing lithium chloride (LiCl) extraction buffer and then desalted using a 10 kDa molecular weight cut-off filter to prepare for MS analysis. Intact proteins (500 ng) were separated by reversed-phase liquid chromatography (RPLC) and online MS and tandem MS (MS/MS) data were acquired using a quadrupole-time-of-flight (QTOF) mass spectrometer. Offline MS/MS data were acquired using an ultrahigh-resolution Fourier-transform ion cyclotron resonance (FTICR) mass spectrometer. Protein SSG (+305.07 Da) was identified through intact mass measurements obtained from the deconvoluted mass spectra. To localize SSG, collisionally activated dissociation (CAD) and electron dissociation (ExD) fragmentation techniques were used. Proteoform characterization was performed by matching identified fragment ions to the target proteoform sequence using MASH Native v. 1.1.

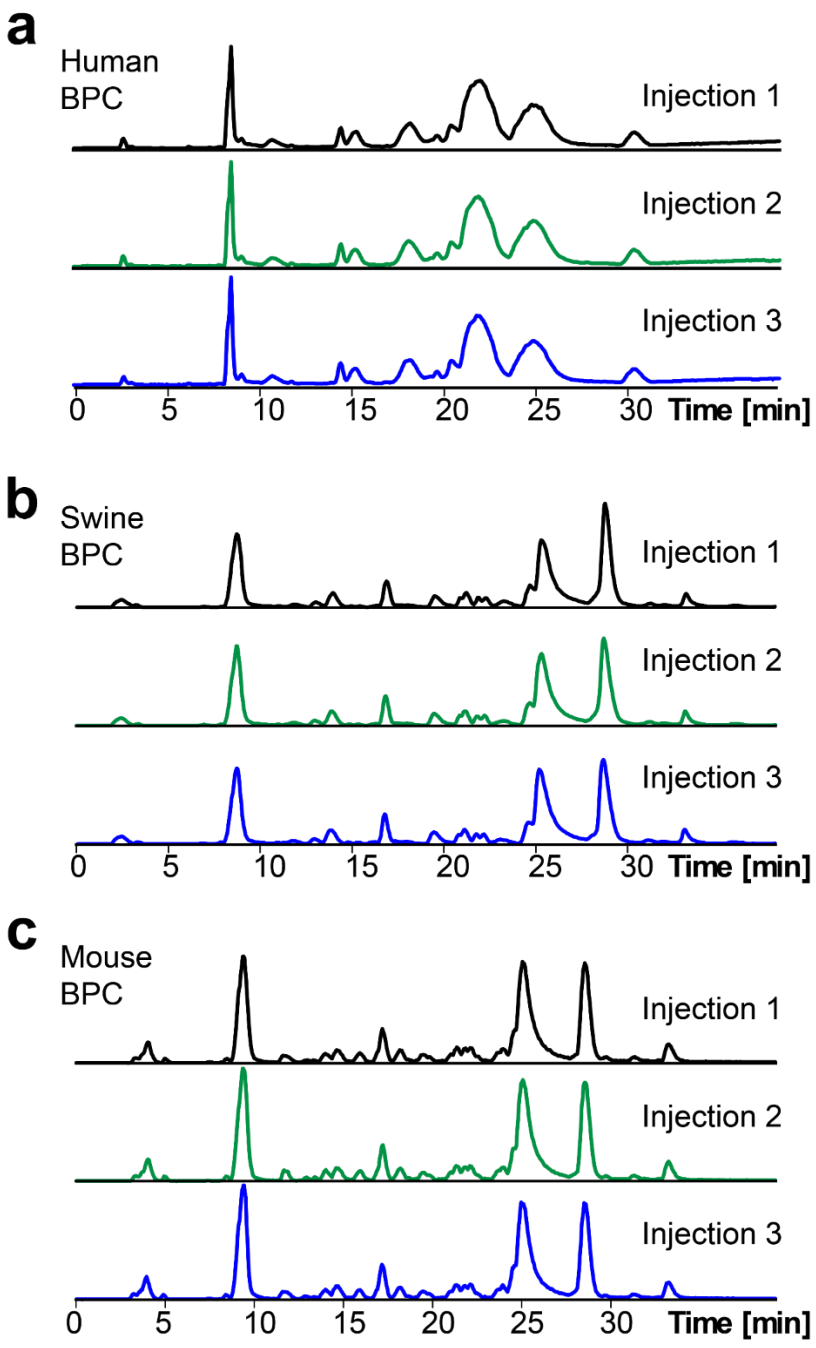


Figure S5.2. Highly reproducible top-down online LC-MS/MS method. Representative intensity normalized base peak chromatograms (BPC) for injection replicates (n = 3) of 500 ng of total protein from **(a)** human, **(b)** swine, and **(c)** mouse cardiac tissue lysates.

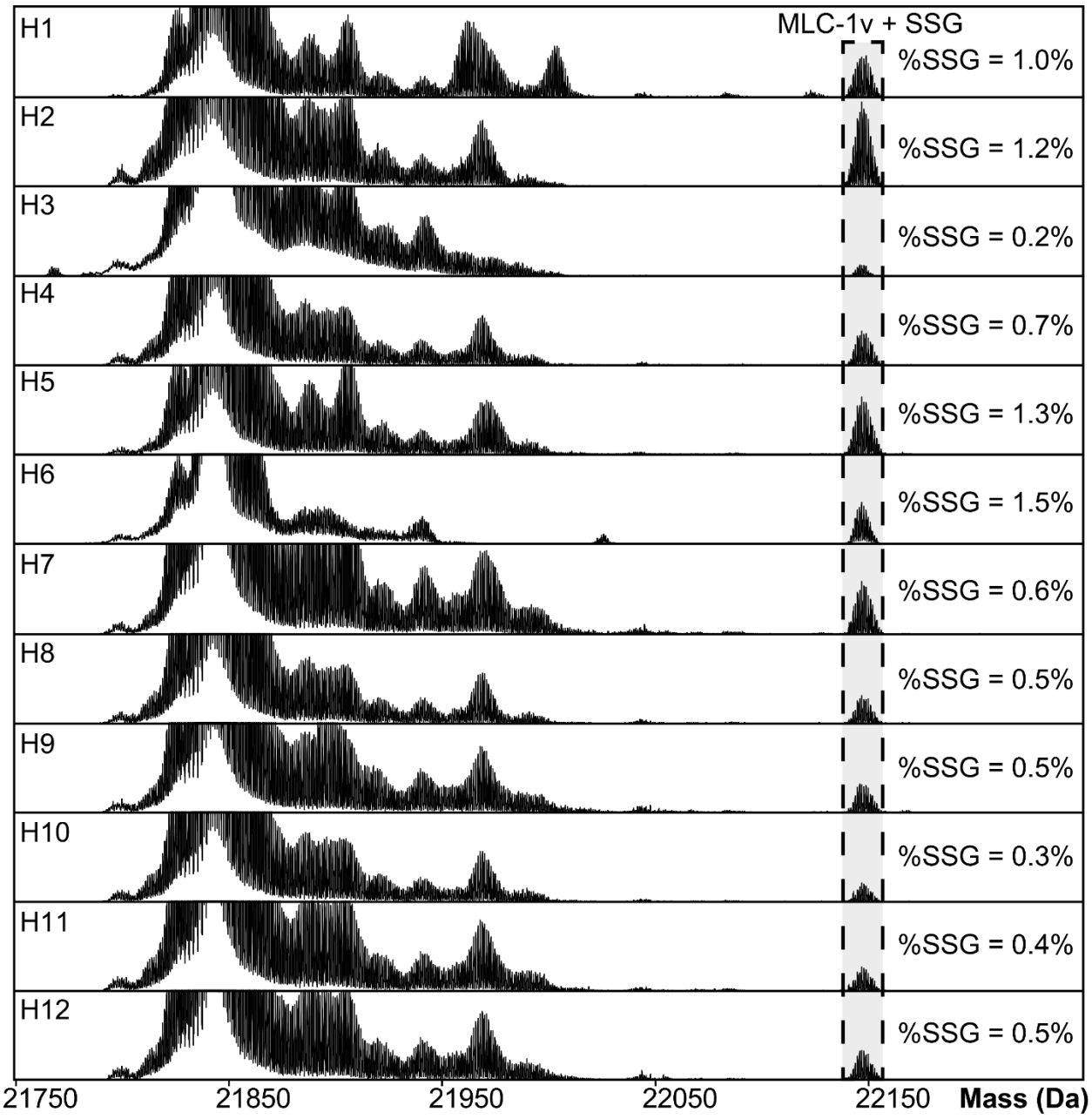


Figure S5.3. Top-down proteomics identifies endogenous SSG in human MLC-1v. Normalized deconvoluted mass spectra of MLC-1v + SSG from non-reduced human cardiac tissue lysates (*n* = 12; H1-H12) not incubated with GSSG. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

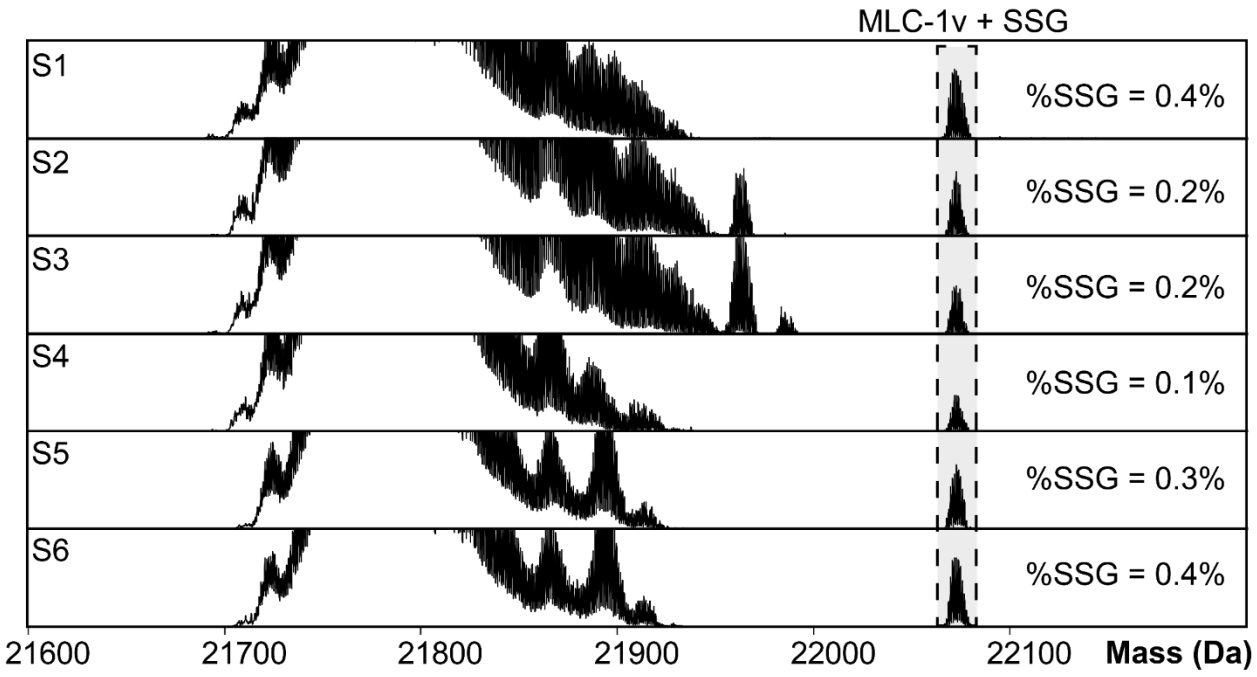


Figure S5.4. Top-down proteomics identifies endogenous SSG in swine MLC-1v. Normalized deconvoluted mass spectra of MLC-1v + SSG from non-reduced swine cardiac tissue lysates (*n* = 6; S1-S6) not incubated with GSSG. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

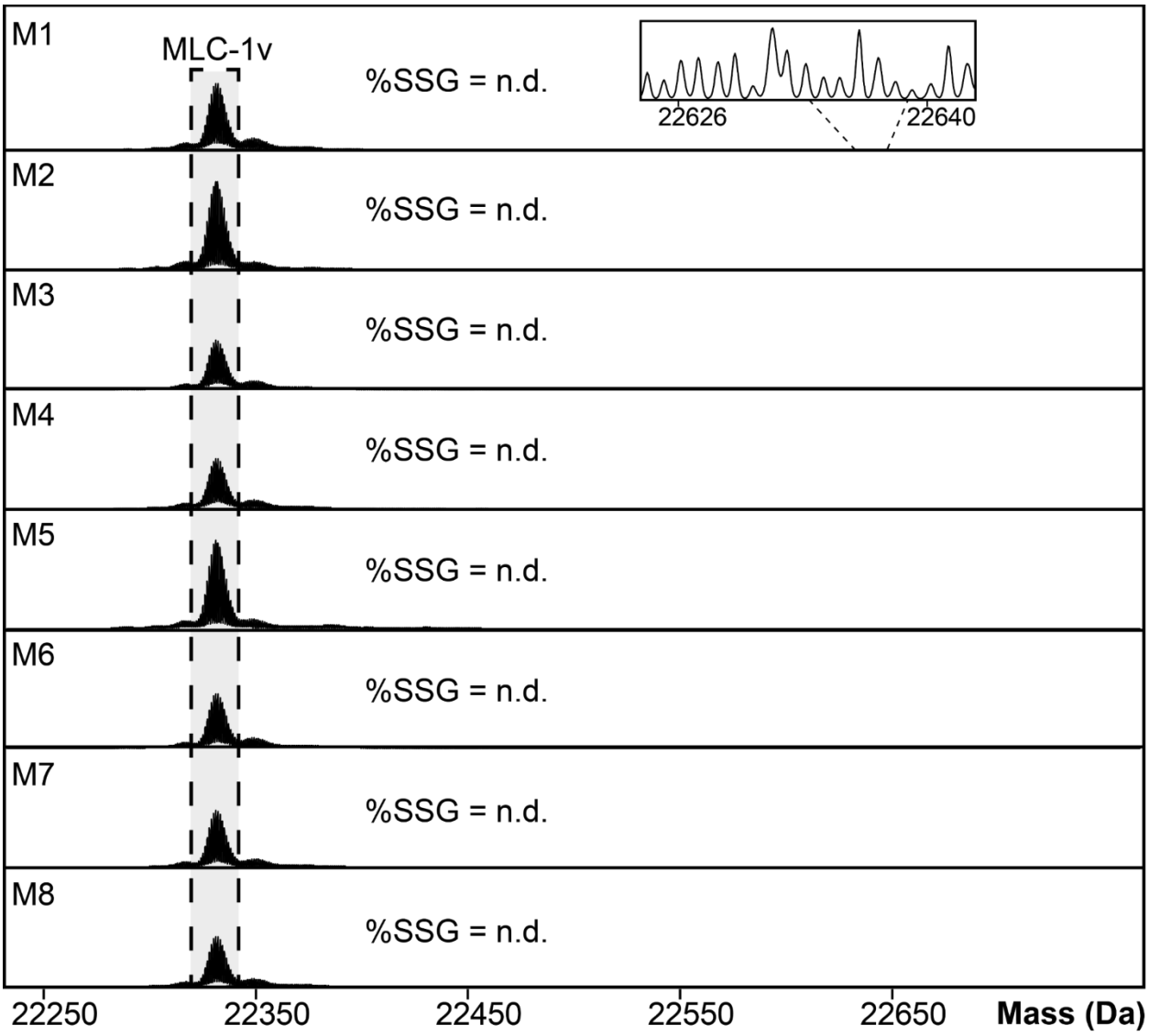


Figure S5.5. Endogenous SSG is not detected by top-down proteomics in mouse MLC-1v. Normalized deconvoluted mass spectra of MLC-1v from non-reduced mouse cardiac tissue lysates (*n* = 8; M1-M8) not incubated with GSSG. The SSG proteoform was not detected (n.d.) on MLC-1v.

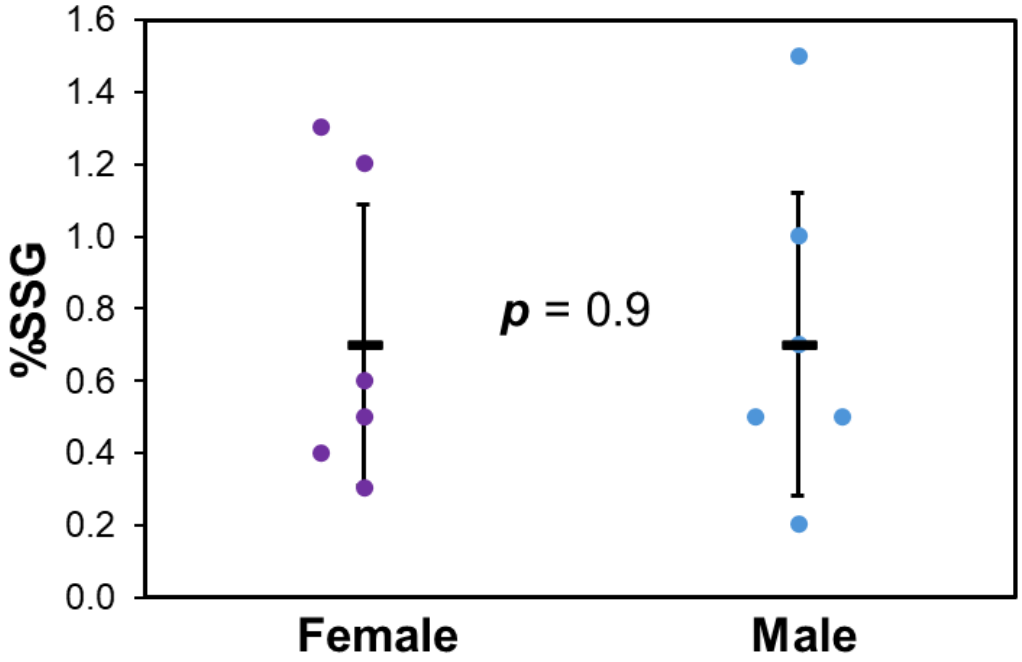


Figure S5.6. Sex-based differences of SSG in human MLC-1v. Comparison of the relative abundance of SSG (%SSG) in MLC-1v between female (purple; $n = 6$) and male (blue; $n = 6$) non-failing donor human hearts. Groups were considered significantly different by paired student t-tests with $p < 0.05$ and not significantly different with $p > 0.05$. Error bars represent the mean ± standard error of the mean (SEM).

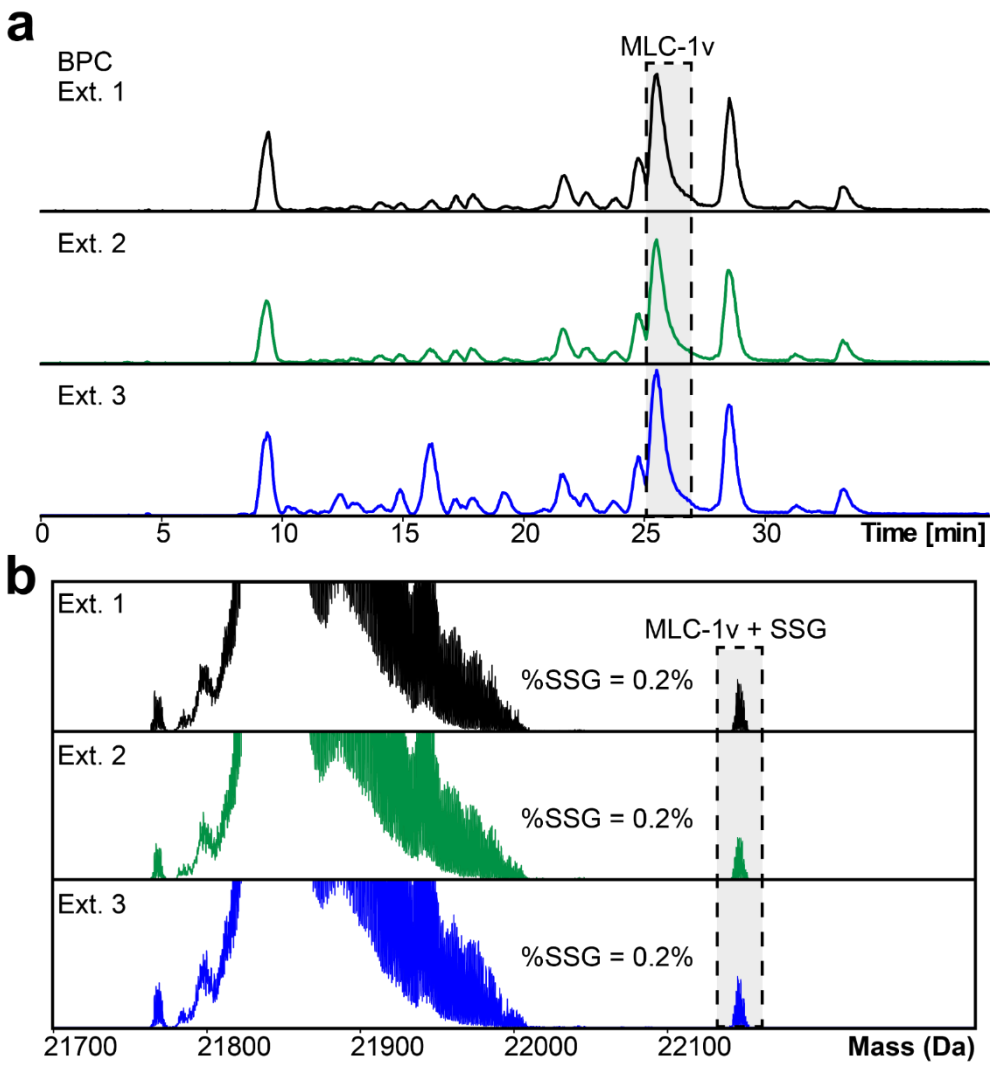


Figure S5.7. Reproducible online LC-MS/MS analysis of MLC-1v from human cardiac tissue. (a) Normalized base peak chromatograms (BPC) of individual extraction replicates (n = 3) from human cardiac tissue showing high reproducibility of the method. **(b)** Deconvoluted mass spectra of the MLC-1v + SSG proteoform in individual extraction replicates (n =3) from human cardiac tissue. Mass spectra were normalized to the same intensity. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

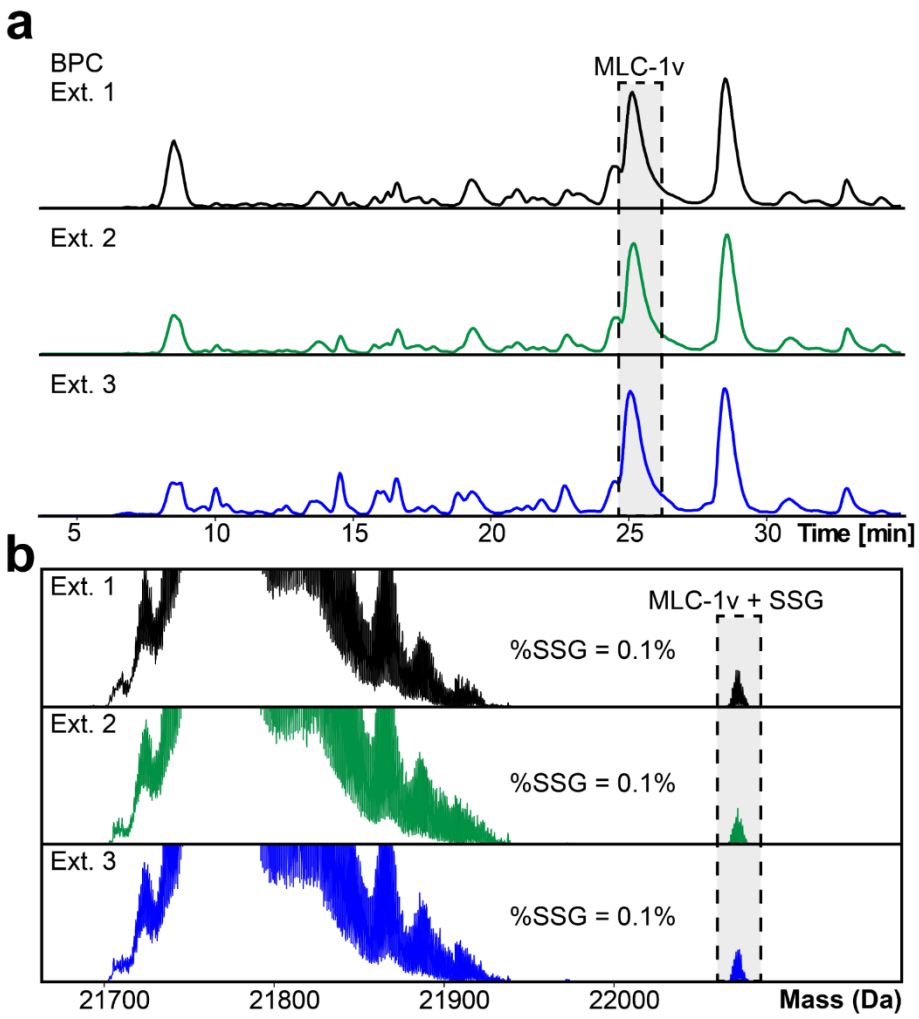


Figure S5.8. Reproducible online LC-MS/MS analysis of MLC-1v from swine cardiac tissue. (a) Normalized base peak chromatograms (BPC) of individual extraction replicates (n = 3) from swine cardiac tissue showing high reproducibility of the method. **(b)** Deconvoluted mass spectra of the MLC-1v + SSG proteoform in individual extraction replicates (n =3) from swine cardiac tissue. Mass spectra were normalized to the same intensity. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

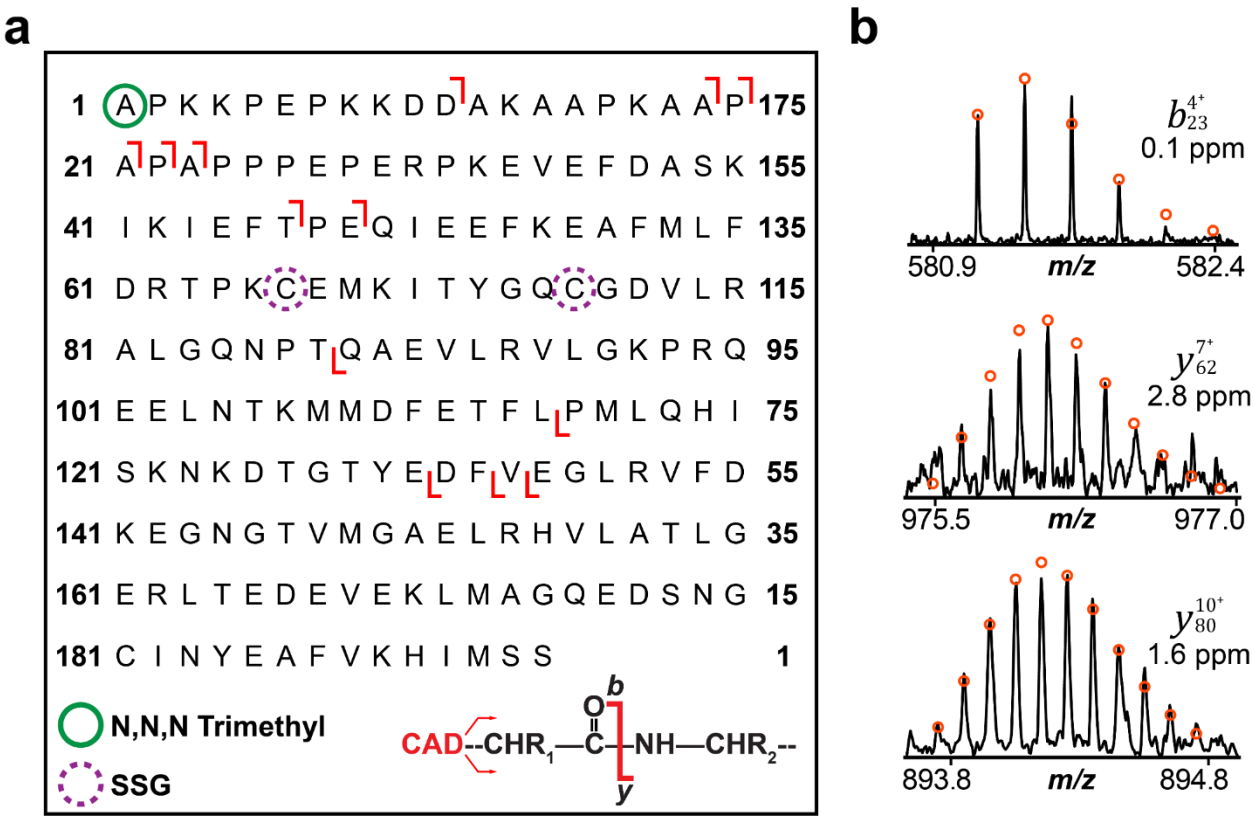


Figure S5.9. Characterization of endogenous human MLC-1v + SSG using top-down proteomics. Online LC-MS/MS analysis of the human MLC-1v + SSG proteoform (no incubation with GSSG). The MLC-1v + SSG proteoform was isolated and fragmented using collisionally activated dissociation (CAD). **(a)** The CAD fragmentation map for human MLC-1v shows SSG is potentially modified on either Cys66 or Cys75. The green circle represents N,N,N-trimethylation, while the dashed purple circle represents potential SSG modification sites. **(b)** Representative CAD fragment ions. All individual ion assignments are within 10 ppm of the theoretical mass, and the theoretical isotopic distributions are indicated by the red circles. For human MLC-1v + SSG there were 8 *b* ions and 5 *y* ions achieving 7% sequence coverage.

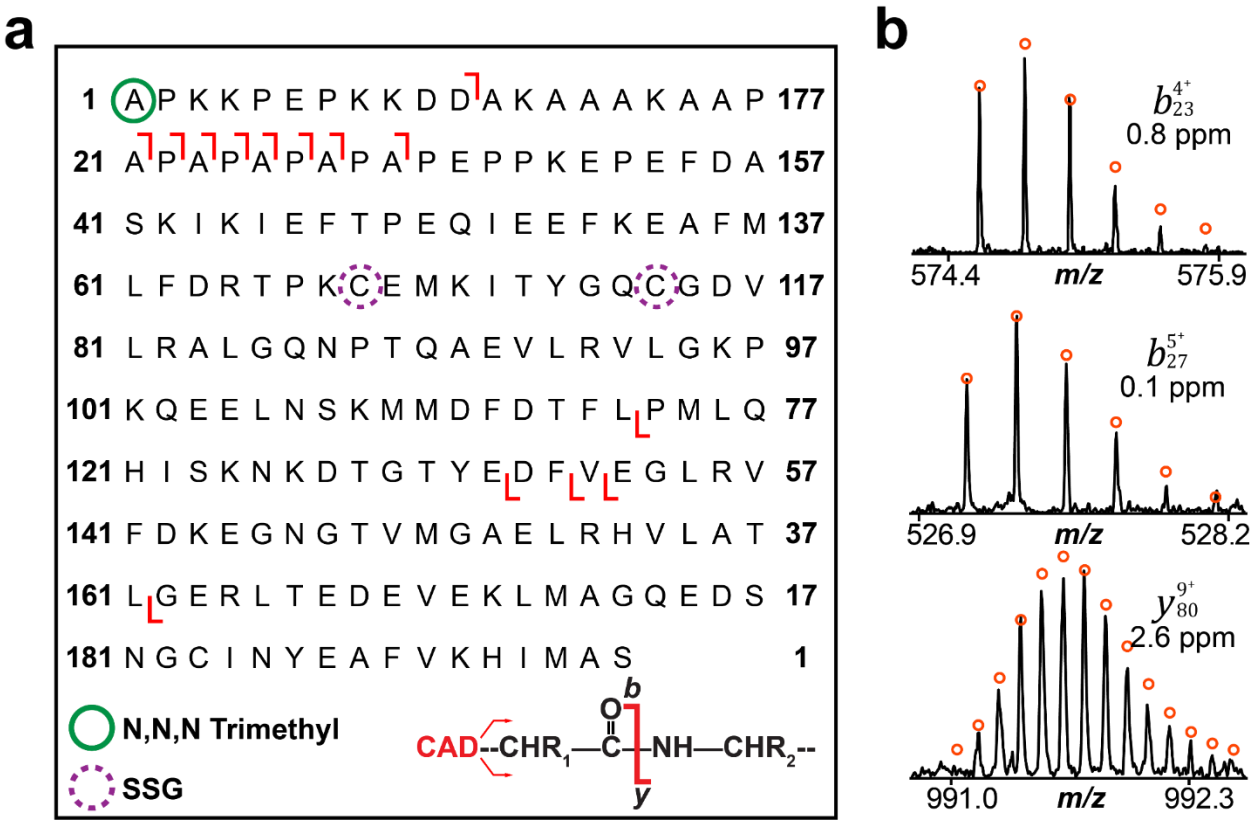


Figure S5.10. Characterization of endogenous swine MLC-1v + SSG using top-down proteomics. Online LC-MS/MS analysis of the swine MLC-1v + SSG proteoform (no incubation with GSSG). The MLC-1v + SSG proteoform was isolated and fragmented using collisionally activated dissociation (CAD). **(a)** The CAD fragmentation map for swine MLC-1v shows SSG is potentially modified on either Cys68 or Cys77. The green circle represents N,N,N-trimethylation, while the dashed purple circle represents potential SSG sites. **(b)** Representative CAD fragment ions. All individual ion assignments are within 10 ppm of the theoretical mass, and the theoretical isotopic distributions are indicated by the red circles. For swine MLC-1v + SSG there were 9 *b* ions and 5 *y* ions achieving 7% sequence coverage.

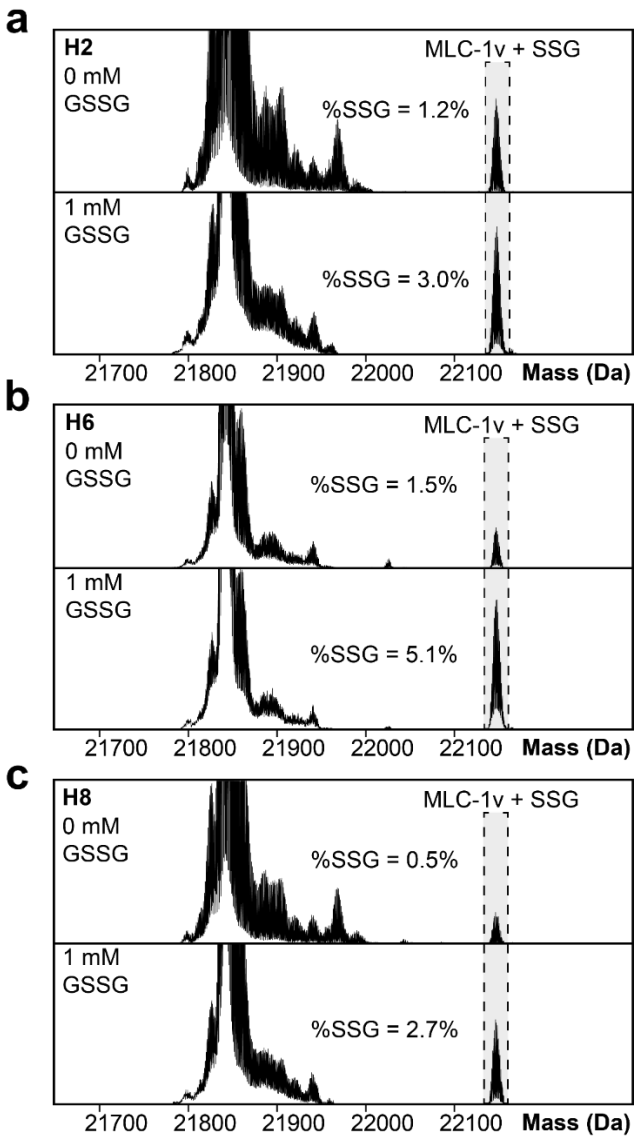


Figure S5.11. Top-down proteomics reveals treatment of non-reduced human cardiac tissue lysates with GSSG increases SSG levels in MLC-1v. Normalized deconvoluted mass spectra of the human MLC-1v + SSG proteoform in three biological replicates **(a)** H2 **(b)** H6 and **(c)** H8 that were incubated with 0 mM and 1 mM of GSSG at 4°C for 1 hour. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

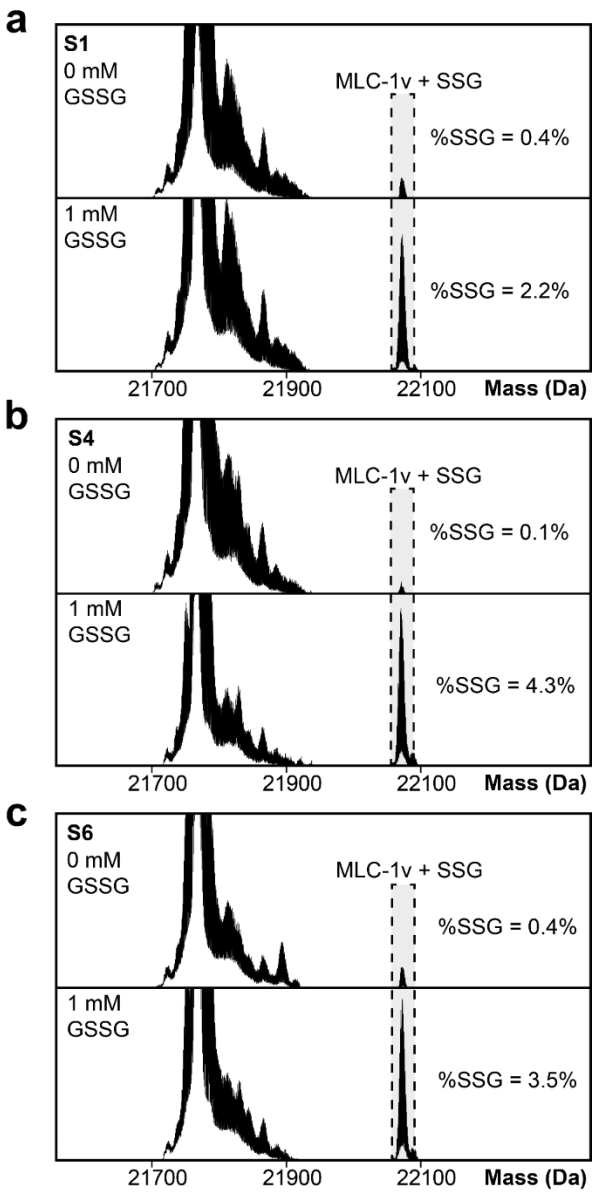


Figure S5.12. Top-down proteomics reveals treatment of non-reduced swine cardiac tissue lysates with GSSG increases SSG levels in MLC-1v. Normalized deconvoluted mass spectra of the swine MLC-1v + SSG proteoform in three biological replicates **(a)** S1 **(b)** S4 and **(c)** S6 that were incubated with 0 mM and 1 mM of GSSG at 4°C for 1 hour. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

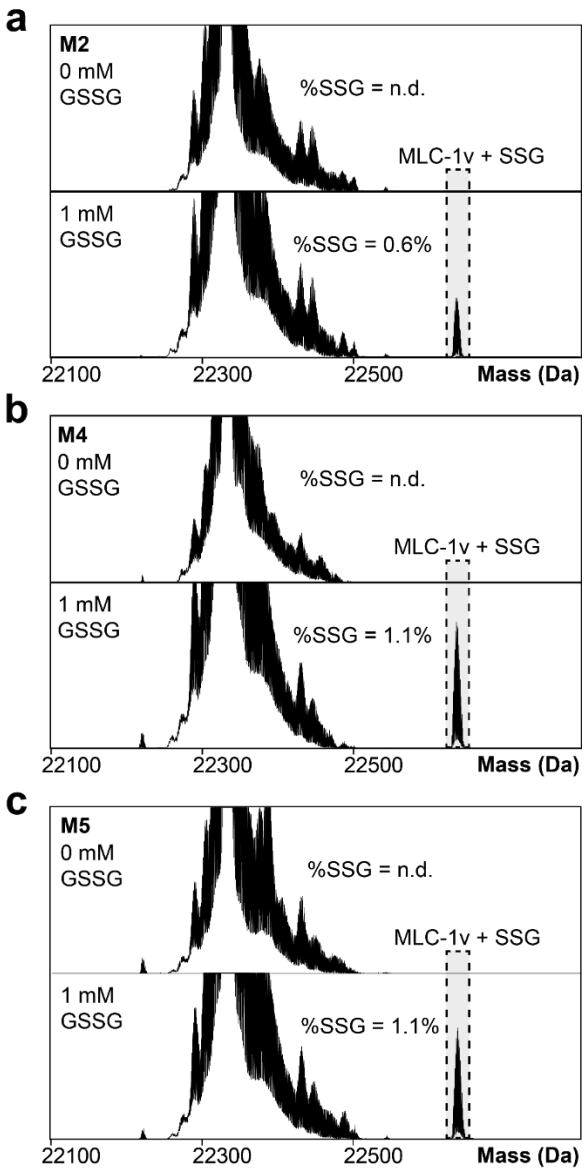


Figure S5.13. Top-down proteomics reveals treatment of non-reduced mouse cardiac tissue lysates with GSSG increases SSG levels in MLC-1v. Normalized deconvoluted mass spectra of the mouse MLC-1v + SSG proteoform in three biological replicates **(a)** M2 **(b)** M4 and **(c)** M5 that were incubated with 0 mM and 1 mM of GSSG at 4°C for 1 hour. Total protein SSG (%SSG) was quantified based on the ratio of the deconvoluted peak intensity of the SSG proteoform to the summed peak intensities of all proteoforms of MLC-1v.

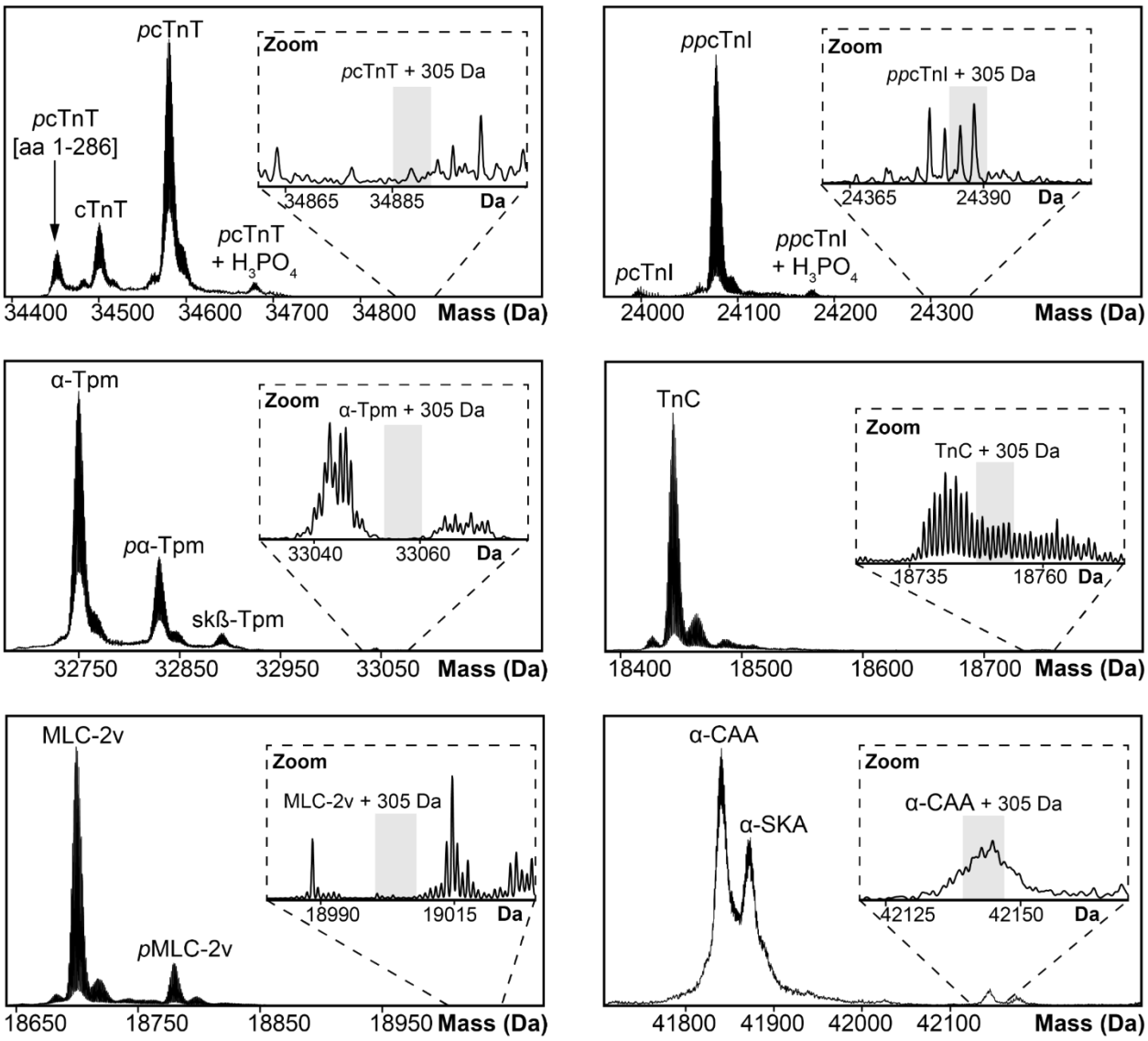


Figure S5.14. Overview of major cardiac proteoforms accessed by top-down proteomics when non-reduced human tissue lysate was incubated with 1 mM GSSG. Representative deconvoluted mass spectra showing proteoforms of major human cardiac proteins; cardiac troponin T (cTnT), cardiac troponin I (cTnI), alpha-tropomyosin (α-Tpm), troponin C (TnC), ventricular isoform of myosin light chain 2 (MLC-2v), alpha-cardiac actin (α-CAA), and alpha-skeletal actin (α-SKA). Mono- and bis-phosphorylated proteoforms are indicated with italicized "*p*" and "*pp*", respectively. Insets show a zoomed-in view of the most abundant proteoform with a +305 Da mass shift to represent the expected mass of the SSG proteoform for each protein.

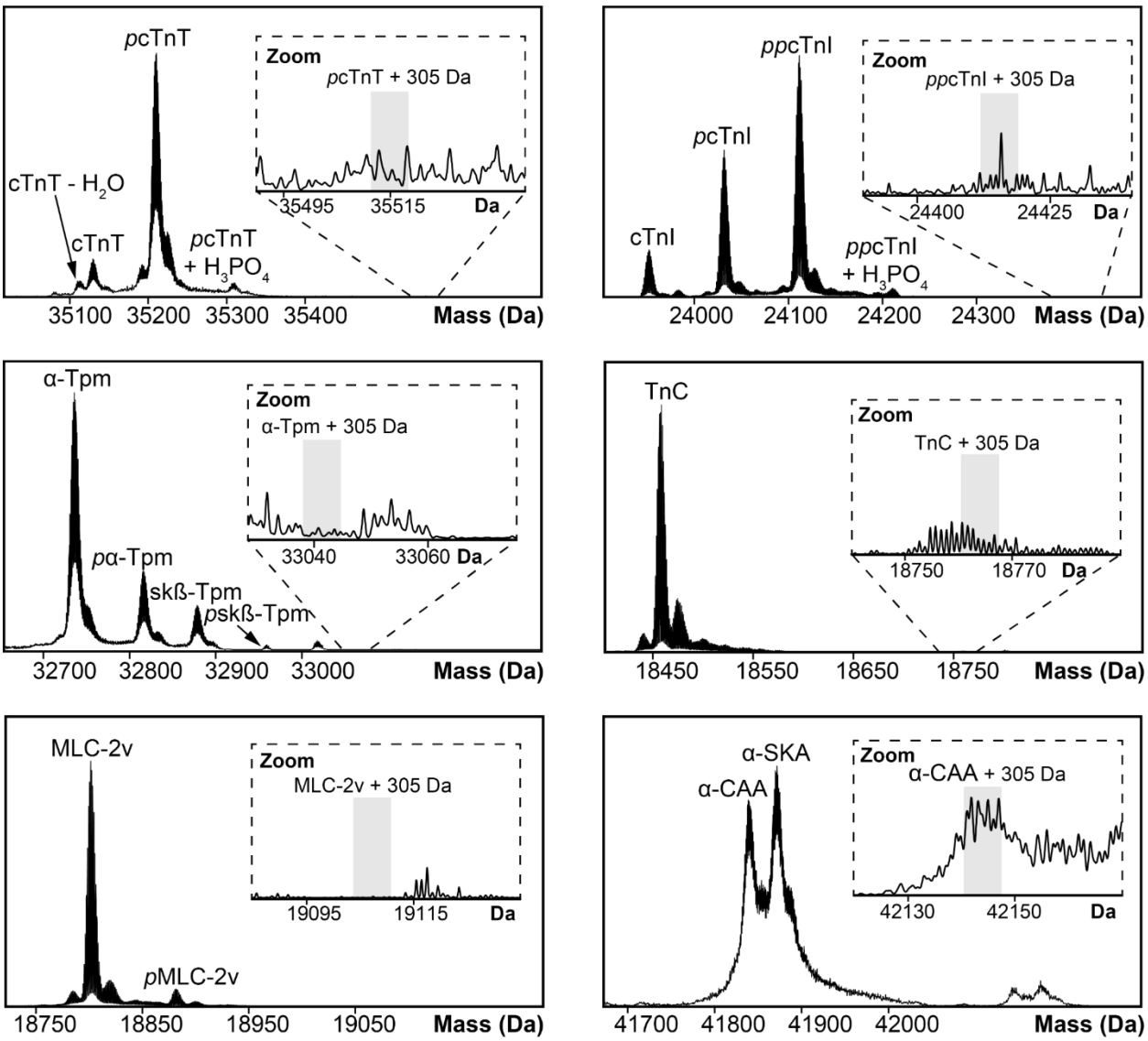


Figure S5.15. Overview of major cardiac proteoforms accessed by top-down proteomics when non-reduced swine tissue lysate was incubated with 1 mM GSSG. Representative deconvoluted mass spectra showing proteoforms of major swine cardiac proteins; cardiac troponin T (cTnT), cardiac troponin I (cTnI), alpha-tropomyosin (α-Tpm), troponin C (TnC), ventricular isoform of myosin light chain 2 (MLC-2v), alpha-cardiac actin (α-CAA), and alpha-skeletal actin (α-SKA). Mono- and bis-phosphorylated proteoforms are indicated with italicized "*p*" and "*pp*", respectively. Insets show a zoomed-in view of the most abundant proteoform with a +305 Da mass shift to represent the expected mass of the SSG proteoform for each protein.

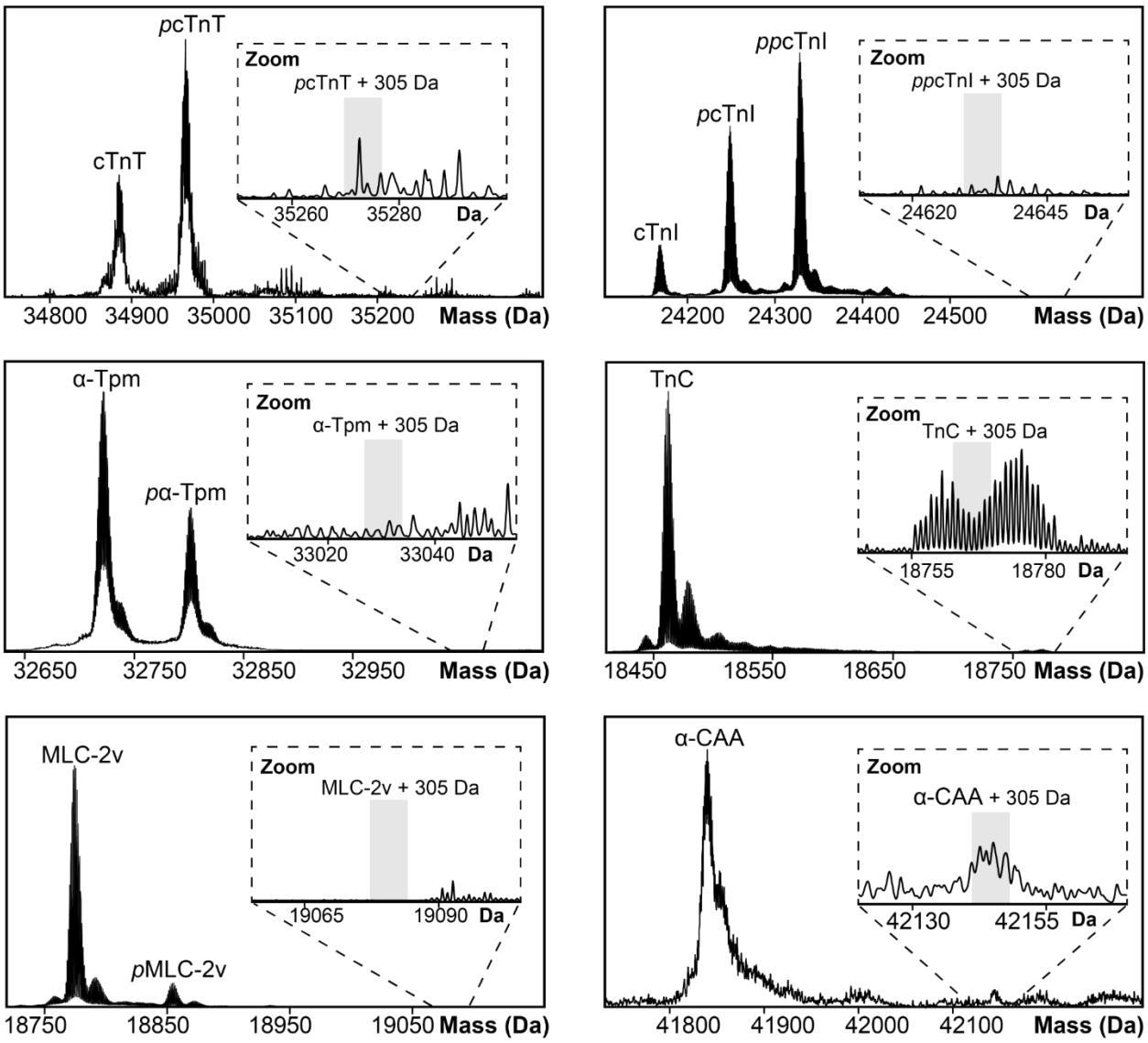


Figure S5.16. Overview of major cardiac proteoforms accessed by top-down proteomics when non-reduced mouse tissue lysate was incubated with 1 mM GSSG. Representative deconvoluted mass spectra showing proteoforms of major mouse cardiac proteins; cardiac troponin T (cTnT), cardiac troponin I (cTnI), alpha-tropomyosin (α-Tpm), troponin C (TnC), ventricular isoform of myosin light chain 2 (MLC-2v), alpha-cardiac actin (α-CAA), and alpha-skeletal actin (α-SKA). Mono- and bis-phosphorylated proteoforms are indicated with italicized "*p*" and "*pp*", respectively. Insets show a zoomed-in view of the most abundant proteoform with a +305 Da mass shift to represent the expected mass of the SSG proteoform for each protein.

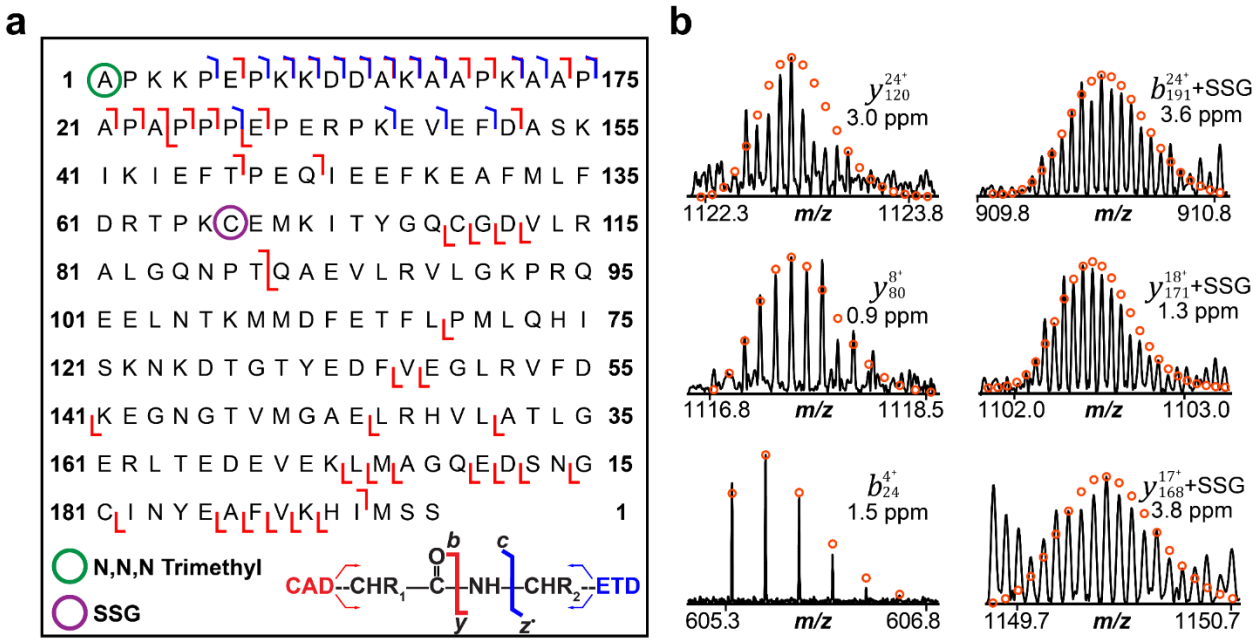


Figure S5.17. Characterization of the human MLC-1v + SSG proteoform following incubation with 1 mM GSSG using top-down proteomics. (a) Combined online and offline MS/MS fragmentation map for the MLC-1v + SSG proteoform in human cardiac tissue lysate incubated with 1 mM GSSG. MLC-1v + SSG was isolated and fragmented using collisionally activated dissociation (CAD) and electron transfer dissociation (ETD). The fragmentation map for human MLC-1v shows SSG is localized to Cys66. The green circle represents N,N,N-trimethylation, while the purple circle represents SSG. **(b)** Representative fragment ions. All individual ion assignments are within 10 ppm of the theoretical mass, and the theoretical isotopic distributions are indicated by the red circles. For human MLC-1v + SSG there were 24 *b* ions, 26 *y* ions, and 16 *c* ions achieving 29% sequence coverage.

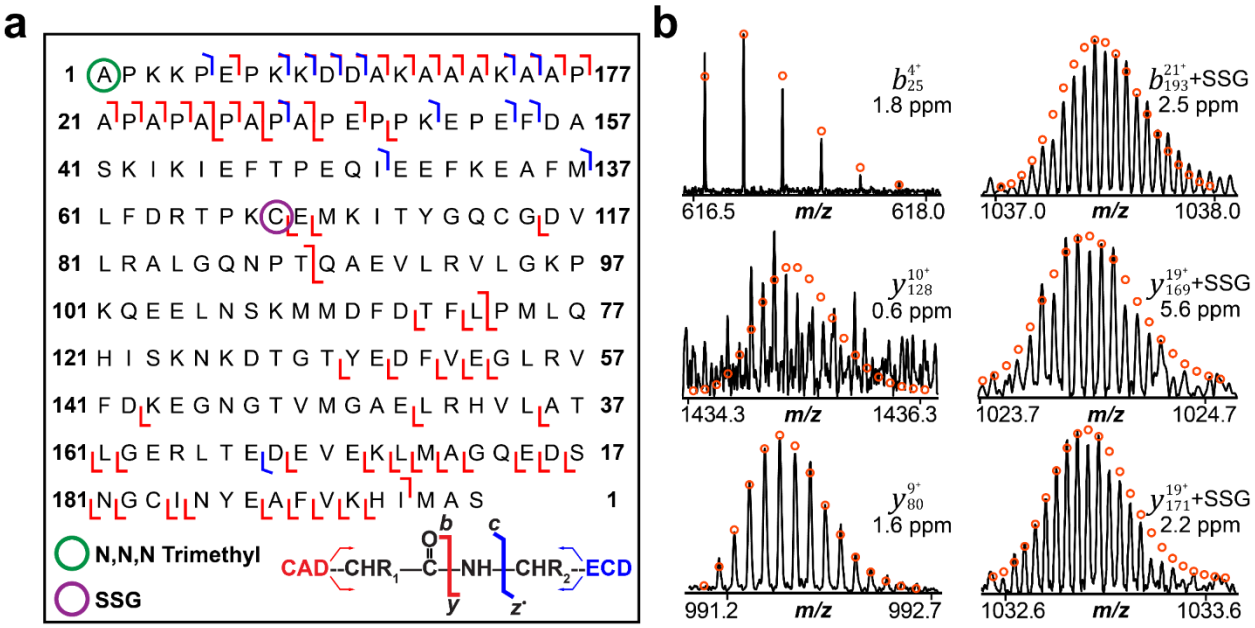


Figure S5.18. Characterization of the swine MLC-1v + SSG proteoform following incubation with 1 mM GSSG using top-down proteomics. (a) Combined online and offline MS/MS fragmentation map for the MLC-1v + SSG proteoform in swine cardiac tissue lysate incubated with 1 mM GSSG. MLC-1v + SSG was isolated and fragmented using collisionally activated dissociation (CAD) and electron capture dissociation (ECD). The fragmentation map for swine MLC-1v shows SSG is localized to Cys68. The green circle represents N,N,N-trimethylation, while the purple circle represents SSG. **(b)** Representative fragment ions. All individual ion assignments are within 10 ppm of the theoretical mass, and the theoretical isotopic distributions are indicated by the red circles. For swine MLC-1v + SSG there were 27 *b* ions, 39 *y* ions, 13 *c*, and 1 $z^{\bullet}$ ions achieving 37% sequence coverage.

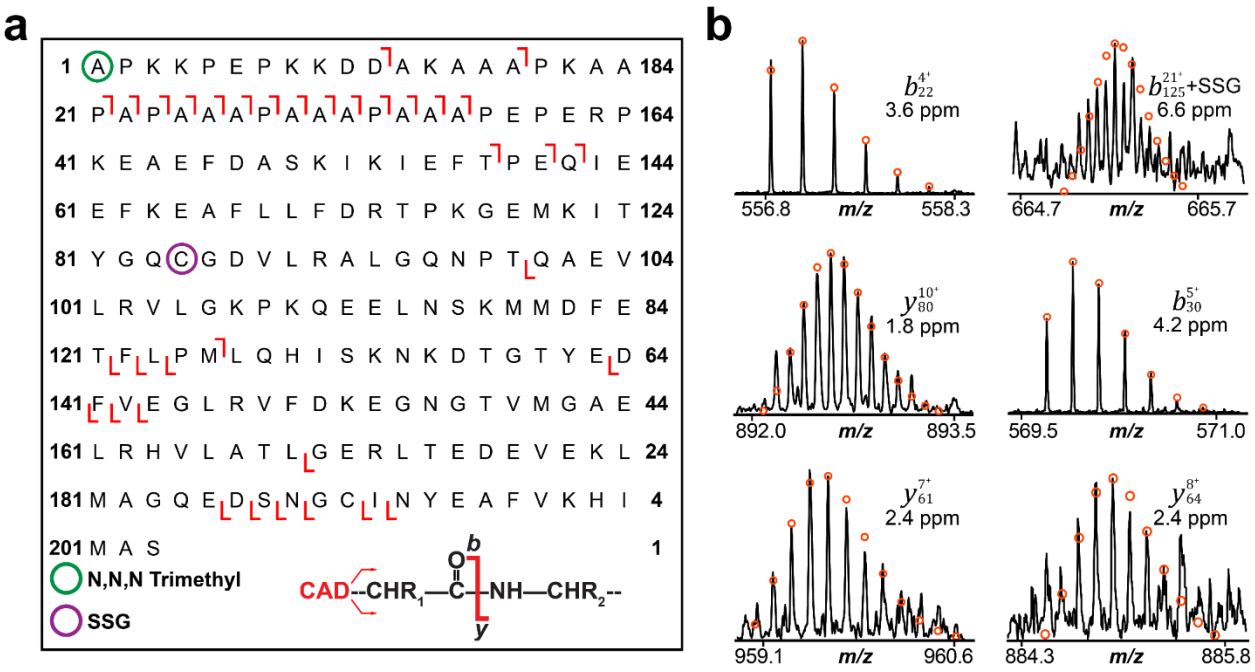


Figure S5.19. Characterization of the mouse MLC-1v + SSG proteoform following incubation with 1 mM GSSG using top-down proteomics. (a) Combined online MS/MS fragmentation map for the MLC-1v + SSG proteoform in mouse cardiac tissue lysate incubated with 1 mM GSSG. MLC-1v + SSG was isolated and fragmented using collisionally activated dissociation (CAD). The fragmentation map for mouse MLC-1v shows SSG is localized to Cys84. The green circle represents N,N,N-trimethylation, while the purple circle represents SSG. **(b)** Representative fragment ions. All individual ion assignments are within 10 ppm of the theoretical mass, and the theoretical isotopic distributions are indicated by the red circles. For mouse MLC-1v + SSG there were 20 *b* ions and 15 *y* ions achieving 17% sequence coverage.

Figure S5.20. Sequence alignment of human, swine, and mouse MLC-1v. MLC-1v in human (UniProt ID P08590; MYL3_HUMAN), mouse (UniProt ID P09542; MYL3_MOUSE), and swine (UniProt ID A0A8D1HG89; A0A8D1HG89_PIG)**.** Human and swine MLC-1v share 95.4% sequence similarity. Human and mouse share 95.9% sequence similarity. Mouse and swine share 95.4% sequence similarity.

CHAPTER 6:

Conclusions and Future Directions

Top-down mass spectrometry (MS)-based proteomics has emerged as a powerful approach for the identification, quantification, and characterization of proteoforms at the intact protein level under both denaturing and native conditions.[11, 30-34, 54] Significant advancements have been made in the field of top-down proteomics (TDP) over the last decade, particularly in instrumentation, sample preparation, and data analysis. However, several challenges remain, including the proteome's high dynamic range, the vast diversity of proteoforms, and the low abundance of heterogeneous protein complexes in their native environments.[30] These challenges limit the broader application of TDP strategies for directly characterizing endogenous proteins and protein complexes from biological samples.[178] In this dissertation, I presented new top-down MS-based proteomics strategies for the characterization of endogenous proteins and protein complexes directly from cardiac tissues, aiming to investigate biological heterogeneity at the proteoform level, elucidate protein structure-function relationships, and provide a comprehensive view of the endogenous heart proteoform landscape.

Chapter 2 described a denatured TDP method for the simultaneous quantification of sarcomeric protein expression and post-translational modifications (PTMs) from end-stage ischemic cardiomyopathy (ICM) and non-failing donor human cardiac tissues. In this study, we demonstrated the unique advantages of TDP for defining the diverse sarcomeric proteoform landscape and quantifying protein expression and PTMs in parallel. Additionally, we highlighted the importance of proteoform-level characterization to uncover molecular changes underlying human ICM. Despite revealing significant PTM and expression-level changes in sarcomeric proteins during ischemic heart failure, one major limitation of this study is the difficulty in detecting large proteins (>50 kDa), primarily due to the exponential decay in the signal-to-noise (S/N) ratio as protein molecular weight (MW) increases and coelution of proteins during the one

dimensional (1D) reversed phase liquid chromatography (RPLC) separation before tandem MS (MS/MS) analysis. Conceivably, to increase sarcomeric proteome coverage and to identify larger MW proteins in ICM, two dimensional (2D) LC methods employing serial size exclusion chromatography (sSEC) and RPLC prior to top-down MS analysis can be used. Moreover, single-omics approaches, primarily reflect only one aspect of the biological system and may not fully recapitulate the mechanisms, functional pathways, and molecular alterations underlying human ICM.[296, 297] Therefore, a MS-based multi-omics approach integrating proteomics, metabolomics, and lipidomics, could offer a more comprehensive understanding of the complex biological processes in ICM.

Chapter 3 utilized a native nanoproteomics platform using peptide-functionalized superparamagnetic nanoparticles (NPs) and native top-down MS (nTDMS) for the enrichment and analysis of the endogenous cardiac troponin (cTn) complex directly from human cardiac tissue. This native nanoproteomics platform enabled the direct characterization of structural features revealing the molecular composition, stoichiometry, and non-covalent interactions of the endogenous cTn complex. This study characterized the endogenous cTn complex from non-failing donor human cardiac tissue; however, it would be interesting to perform further experiments to enrich and characterize the cTn complex from cardiac tissues with different cardiomyopathies, including ICM, dilated cardiomyopathy (DCM), and hypertrophic cardiomyopathy (HCM). Alterations in the phosphorylation of the cardiac troponin I (cTnI) subunit have been associated with various cardiomyopathies and can influence calcium (Ca^{2+}) sensitivity within the cTn complex, ultimately affecting cardiac contractility.[232, 298-300] Thus, applying our native nanoproteomics platform to characterize the endogenous cTn complex across various cardiomyopathies could yield unique insights into Ca^{2+} binding, complex stoichiometry, and

PTMs, to enhance our understanding of how cardiac diseases influence heart contractility through disease-related changes in protein modifications and Ca^{2+} binding. Furthermore, this native nanoproteomics platform has clearly demonstrated that the enriched protein complex can retain its native state when suitable enrichment and elution conditions are developed. Therefore, this platform can be readily adaptable to other endogenous low-abundant protein targets such as kinases and receptors.

The nanoproteomics-based technology described in **Chapter 3** can also be applied under denaturing conditions for the enrichment and proteoform-resolved analysis of intact cTnI from human serum at clinically relevant concentrations.[78] cTnI, a gold-standard cardiac-specific biomarker for diagnosing acute myocardial infarction (AMI), is released into the bloodstream from cardiac tissue following myocardial injury.[221, 223, 301] Currently, antibody-based diagnostic assays such as enzyme-linked immunosorbent assays (ELISAs) are used in clinical settings to identify circulating cTnI concentrations and for the rapid rule-in/rule-out of AMI. Although ELISAs are high-throughput and automated, they lack standardization, cannot distinguish between cTnI proteoforms, and suffer from non-specific binding due to circulating autoantibodies, which may lead to misdiagnosis of AMI.[302, 303] Thus, a highly reproducible and sensitive MS-based diagnostic assay incorporating the denatured nanoproteomics platform to enrich cTnI proteoforms directly from human serum in patients with ST-elevation myocardial infarction (STEMI), typically considered a severe case of AMI, would be highly advantageous for the early detection and accurate diagnosis of AMI. This nanoproteomics MS-based diagnostic assay could be used to analyze circulating cTnI proteoforms enriched from serum samples of a large human cohort, including both healthy control and STEMI patients, to uncover correlations between cTnI proteoforms and underlying disease etiology. Patient-specific variables such as age, sex,

race/ethnicity, and co-morbidities will need to be considered as well. Of course, many challenges remain in incorporating this assay into the clinic, including the sensitivity of the method, as cTnI often circulates in serum at concentrations less than 0.04 ng/mL. Moreover, non-specific binding of highly abundant serum proteins, such as apolipoproteins, may occur on the NP surface. To address this, optimizing the NP surface chemistry and implementing additional depletion strategies, such as perchloric acid protein precipitation,[304] may be necessary. Furthermore, circulating troponin autoantibodies in serum may target the central region of cTnI to which the NP binds to. To prevent potential negative interference of circulating cTnI autoantibodies during enrichment of cTnI proteoforms from serum, we may need to consider adding blocking agents to make the central region of cTnI inaccessible for autoantibody binding. Overall, a nanoproteomics MS-based clinical assay could enable new precision medicine opportunities for more accurate diagnosis of AMI, personalized treatment, and improved patient outcomes by providing proteoform-resolved molecular fingerprints of circulating cTnI proteoforms.

Chapter 4 explores the application of direct infusion Fourier-transform ion cyclotron resonance (FTICR) for nTDMS analysis, enabling the characterization of endogenous sarcomeric proteins from human cardiac tissue lysate without the need for conventional separation strategies. This nTDMS approach illustrates the power of direct infusion FTICR MS in analyzing PTMs, Ca^{2+} binding, and higher-order structures of endogenous proteins from human cardiac tissues, revealing direct correlations between their structural features and functions within the sarcomere. However, only the most highly abundant monomeric proteins (>50 kDa) in the sarcomere were identified in this study. A significant challenge in the direct analysis of proteins from complex biological samples arises from the high dynamic range, which leads to signal suppression due to the concentration distribution spanning several magnitudes in endogenous matrices. Additionally,

analyzing proteins and protein complexes larger than 50 kDa using FTICR MS is inherently challenging, as the MS signal spreads across an increased number of isotopologues and charge states with increasing MW.[84] Thus, I envision that developing multidimensional chromatographic strategies incorporating SEC and hydrophobic interaction chromatography (HIC) prior to nTDMS analysis could reduce signal suppression of high MW proteins and protein complexes to provide deeper coverage of the native human heart proteome. Finally, utilizing additional gas-phase filtering technologies such as trapped ion mobility or quadrupole mass selection techniques could select and enhance the signals of low abundant ions in the native mass spectrum.

Chapter 5 describes the identification and characterization of endogenous S-glutathionylation (SSG) in the ventricular isoform of myosin light chain 1 (MLC-1v) from human, swine, and mouse cardiac tissues using a denatured TDP approach. This study is the first to report SSG on MLC-1v in non-failing control hearts across species. However, the functional consequences of this oxidative PTM in the sarcomere remain unclear. To better understand the role of SSG in the sarcomere, quantifying SSG levels in MLC-1v from failing hearts compared to non-failing donor hearts could reveal whether perturbations in SSG contribute to the onset and progression of various cardiomyopathies. Additionally, we could investigate other sarcomeric proteins that have been identified as potential substrates for SSG, including actin and cTnI. Overall, comprehensively characterizing SSG modifications in sarcomeric proteins across various cardiomyopathies could provide valuable insights into the functional role of SSG in the sarcomere and help identify new biomarker candidates.

REFERENCES

(1) Morris, R.; Black, Katrina A.; Stollar, Elliott J. Uncovering protein function: from classification to complexes. *Essays in Biochemistry* **2022**, *66* (3), 255-285. DOI: 10.1042/EBC20200108.

(2) Smith, L. M.; Agar, J. N.; Chamot-Rooke, J.; Danis, P. O.; Ge, Y.; Loo, J. A.; Paša-Tolić, L.; Tsybin, Y. O.; Kelleher, N. L.; The Consortium for Top-Down, P. The Human Proteoform Project: Defining the human proteome. *Science Advances 7* (46), eabk0734. DOI: 10.1126/sciadv.abk0734.

(3) Aebersold, R.; Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **2016**, *537* (7620), 347-355. DOI: 10.1038/nature19949.

(4) Gregorich, Z. R.; Chang, Y.-H.; Ge, Y. Proteomics in heart failure: top-down or bottom-up? *Pflügers Archiv - European Journal of Physiology* **2014**, *466* (6), 1199-1209. DOI: 10.1007/s00424-014-1471-9.

(5) Franklin, S.; Vondriska, T. M. Genomes, Proteomes, and the Central Dogma. *Circulation: Cardiovascular Genetics* **2011**, *4* (5), 576-576. DOI: 10.1161/CIRCGENETICS.110.957795.

(6) Smith, L. M.; Kelleher, N. L.; Linial, M.; Goodlett, D.; Langridge-Smith, P.; Ah Goo, Y.; Safford, G.; Bonilla*, L.; Kruppa, G.; Zubarev, R.; et al. Proteoform: a single term describing protein complexity. *Nature Methods* **2013**, *10* (3), 186-187. DOI: 10.1038/nmeth.2369.

(7) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; et al. How many human proteoforms are there? *Nature Chemical Biology* **2018**, *14* (3), 206-214. DOI: 10.1038/nchembio.2576.

(8) Smith, L. M.; Kelleher, N. L. Proteoforms as the next proteomics currency. *Science* **2018**, *359* (6380), 1106-1107. DOI: 10.1126/science.aat1884.

(9) Altelaar, A. F. M.; Munoz, J.; Heck, A. J. R. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nature Reviews Genetics* **2013**, *14* (1), 35-48. DOI: 10.1038/nrg3356.

(10) Marsh, Joseph A.; Hernández, H.; Hall, Z.; Ahnert, Sebastian E.; Perica, T.; Robinson, Carol V.; Teichmann, Sarah A. Protein Complexes Are under Evolutionary Selection to Assemble via Ordered Pathways. *Cell* **2013**, *153* (2), 461-470. DOI: 10.1016/j.cell.2013.02.044.

(11) Gregorich, Z. R.; Ge, Y. Top-down proteomics in health and disease: Challenges and opportunities. *PROTEOMICS* **2014**, *14* (10), 1195-1210. DOI: https://doi.org/10.1002/pmic.201300432.

(12) Birhanu, A. G. Mass spectrometry-based proteomics as an emerging tool in clinical laboratories. *Clinical Proteomics* **2023**, *20* (1), 32. DOI: 10.1186/s12014-023-09424-x.

(13) Keenan, E. K.; Zachman, D. K.; Hirschey, M. D. Discovering the landscape of protein modifications. *Molecular Cell* **2021**, *81* (9), 1868-1878. DOI: https://doi.org/10.1016/j.molcel.2021.03.015.

(14) Bagwan, N.; El Ali, H. H.; Lundby, A. Proteome-wide profiling and mapping of post translational modifications in human hearts. *Scientific Reports* **2021**, *11* (1), 2184. DOI: 10.1038/s41598-021-81986-y.

(15) Zhong, Q.; Xiao, X.; Qiu, Y.; Xu, Z.; Chen, C.; Chong, B.; Zhao, X.; Hai, S.; Li, S.; An, Z.; et al. Protein posttranslational modifications in health and diseases: Functions, regulatory mechanisms, and therapeutic implications. *MedComm* **2023**, *4* (3), e261. DOI: https://doi.org/10.1002/mco2.261.

(16) Wu, Z.; Tiambeng, T. N.; Cai, W.; Chen, B.; Lin, Z.; Gregorich, Z. R.; Ge, Y. Impact of Phosphorylation on the Mass Spectrometry Quantification of Intact Phosphoproteins. *Analytical Chemistry* **2018**, *90* (8), 4935-4939. DOI: 10.1021/acs.analchem.7b05246.

(17) Hunter, T. Protein kinases and phosphatases: The Yin and Yang of protein phosphorylation and signaling. *Cell* **1995**, *80* (2), 225-236. DOI: https://doi.org/10.1016/0092-8674(95)90405-0.

(18) Pang, K.; Wang, W.; Qin, J.-X.; Shi, Z.-D.; Hao, L.; Ma, Y.-Y.; Xu, H.; Wu, Z.-X.; Pan, D.; Chen, Z.-S.; et al. Role of protein phosphorylation in cell signaling, disease, and the intervention therapy. *MedComm* **2022**, *3* (4), e175. DOI: https://doi.org/10.1002/mco2.175.

(19) Henriques, A. G.; Müller, T.; Oliveira, J. M.; Cova, M.; da Cruz e Silva, C. B.; da Cruz e Silva, O. A. B. Altered protein phosphorylation as a resource for potential AD biomarkers. *Scientific Reports* **2016**, *6* (1), 30319. DOI: 10.1038/srep30319.

(20) Arnesen, T.; Van Damme, P.; Polevoda, B.; Helsens, K.; Evjenth, R.; Colaert, N.; Varhaug, J. E.; Vandekerckhove, J.; Lillehaug, J. R.; Sherman, F.; et al. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proceedings of the National Academy of Sciences* **2009**, *106* (20), 8157-8162. DOI: 10.1073/pnas.0901931106.

(21) Drazic, A.; Myklebust, L. M.; Ree, R.; Arnesen, T. The world of protein acetylation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2016**, *1864* (10), 1372-1401. DOI: https://doi.org/10.1016/j.bbapap.2016.06.007.

(22) Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M. Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* **2009**, *325* (5942), 834-840. DOI: 10.1126/science.1175371.

(23) Gujral, P.; Mahajan, V.; Lissaman, A. C.; Ponnampalam, A. P. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. *Reproductive Biology and Endocrinology* **2020**, *18* (1), 84. DOI: 10.1186/s12958-020-00637-5.

(24) Paulsen, C. E.; Carroll, K. S. Cysteine-Mediated Redox Signaling: Chemistry, Biology, and Tools for Discovery. *Chemical Reviews* **2013**, *113* (7), 4633-4679. DOI: 10.1021/cr300163e.

(25) Chung, H. S.; Wang, S.-B.; Venkatraman, V.; Murray, C. I.; Van Eyk, J. E. Cysteine Oxidative Posttranslational Modifications. *Circulation Research* **2013**, *112* (2), 382-392. DOI: 10.1161/CIRCRESAHA.112.268680.

(26) Gu, L.; Robinson, R. A. S. Proteomic approaches to quantify cysteine reversible modifications in aging and neurodegenerative diseases. *PROTEOMICS – Clinical Applications* **2016**, *10* (12), 1159-1177. DOI: https://doi.org/10.1002/prca.201600015.

(27) Mieyal, J. J.; Gallogly, M. M.; Qanungo, S.; Sabens, E. A.; Shelton, M. D. Molecular Mechanisms and Clinical Implications of Reversible Protein S-Glutathionylation. *Antioxidants & Redox Signaling* **2008**, *10* (11), 1941-1988. DOI: 10.1089/ars.2008.2089.

(28) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R., III. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chemical Reviews* **2013**, *113* (4), 2343-2394. DOI: 10.1021/cr3003533.

(29) Plubell, D. L.; Käll, L.; Webb-Robertson, B.-J.; Bramer, L. M.; Ives, A.; Kelleher, N. L.; Smith, L. M.; Montine, T. J.; Wu, C. C.; MacCoss, M. J. Putting Humpty Dumpty Back Together Again: What Does Protein Quantification Mean in Bottom-Up Proteomics? *Journal of Proteome Research* **2022**, *21* (4), 891-898. DOI: 10.1021/acs.jproteome.1c00894.

(30) Melby, J. A.; Roberts, D. S.; Larson, E. J.; Brown, K. A.; Bayne, E. F.; Jin, S.; Ge, Y. Novel Strategies to Address the Challenges in Top-Down Proteomics. *Journal of the American Society for Mass Spectrometry* **2021**, *32* (6), 1278-1294. DOI: 10.1021/jasms.1c00099.

(31) Li, H.; Nguyen, H. H.; Ogorzalek Loo, R. R.; Campuzano, I. D. G.; Loo, J. A. An integrated native mass spectrometry and top-down proteomics method that connects sequence to structure and function of macromolecular complexes. *Nature Chemistry* **2018**, *10* (2), 139-148. DOI: 10.1038/nchem.2908.

(32) Toby, T. K.; Fornelli, L.; Kelleher, N. L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. From 2016.

(33) Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Analytical Chemistry* **2018**, *90* (1), 110-127. DOI: 10.1021/acs.analchem.7b04747.

(34) Roberts, D. S.; Loo, J. A.; Tsybin, Y. O.; Liu, X.; Wu, S.; Chamot-Rooke, J.; Agar, J. N.; Paša-Tolić, L.; Smith, L. M.; Ge, Y. Top-down proteomics. *Nature Reviews Methods Primers* **2024**, *4* (1), 38. DOI: 10.1038/s43586-024-00318-2.

(35) Zhang, J.; Guy, M. J.; Norman, H. S.; Chen, Y.-C.; Xu, Q.; Dong, X.; Guner, H.; Wang, S.; Kohmoto, T.; Young, K. H.; et al. Top-Down Quantitative Proteomics Identified Phosphorylation

of Cardiac Troponin I as a Candidate Biomarker for Chronic Heart Failure. *Journal of Proteome Research* **2011**, *10* (9), 4054-4065. DOI: 10.1021/pr200258m.

(36) Siuti, N.; Kelleher, N. L. Decoding protein modifications using top-down mass spectrometry. *Nature Methods* **2007**, *4* (10), 817-821. DOI: 10.1038/nmeth1097.

(37) Zhang, H.; Ge, Y. Comprehensive Analysis of Protein Modifications by Top-Down Mass Spectrometry. *Circulation: Cardiovascular Genetics* **2011**, *4* (6), 711-711. DOI: 10.1161/CIRCGENETICS.110.957829.

(38) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; et al. Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature Methods* **2019**, *16* (7), 587-594. DOI: 10.1038/s41592-019-0457-0.

(39) Tucholski, T.; Cai, W.; Gregorich, Z. R.; Bayne, E. F.; Mitchell, S. D.; Mcilwain, S. J.; De Lange, W. J.; Wrobbel, M.; Karp, H.; Hite, Z.; et al. Distinct hypertrophic cardiomyopathy genotypes result in convergent sarcomeric proteoform profiles revealed by top-down proteomics. *Proceedings of the National Academy of Sciences* **2020**, *117* (40), 24691-24700. DOI: 10.1073/pnas.2006764117.

(40) Melby, J. A.; Brown, K. A.; Gregorich, Z. R.; Roberts, D. S.; Chapman, E. A.; Ehlers, L. E.; Gao, Z.; Larson, E. J.; Jin, Y.; Lopez, J. R.; et al. High sensitivity top–down proteomics captures single muscle cell heterogeneity in large proteoforms. *Proceedings of the National Academy of Sciences* **2023**, *120* (19), e2222081120. DOI: 10.1073/pnas.2222081120.

(41) Roberts, D. S.; Mann, M.; Melby, J. A.; Larson, E. J.; Zhu, Y.; Brasier, A. R.; Jin, S.; Ge, Y. Structural O-Glycoform Heterogeneity of the SARS-CoV-2 Spike Protein Receptor-Binding Domain Revealed by Top-Down Mass Spectrometry. *Journal of the American Chemical Society* **2021**, *143* (31), 12014-12024. DOI: 10.1021/jacs.1c02713.

(42) Rogers, H. T.; Roberts, D. S.; Larson, E. J.; Melby, J. A.; Rossler, K. J.; Carr, A. V.; Brown, K. A.; Ge, Y. Comprehensive Characterization of Endogenous Phospholamban Proteoforms

Enabled by Photocleavable Surfactant and Top-down Proteomics. *Analytical Chemistry* **2023**, *95* (35), 13091-13100. DOI: 10.1021/acs.analchem.3c01618.

(43) Habeck, T.; Brown, K. A.; Des Soye, B.; Lantz, C.; Zhou, M.; Alam, N.; Hossain, M. A.; Jung, W.; Keener, J. E.; Volny, M.; et al. Top-down mass spectrometry of native proteoforms and their complexes: a community study. *Nature Methods* **2024**, *21* (12), 2388-2396. DOI: 10.1038/s41592-024-02279-6.

(44) Tamara, S.; den Boer, M. A.; Heck, A. J. R. High-Resolution Native Mass Spectrometry. *Chemical Reviews* **2022**, *122* (8), 7269-7326. DOI: 10.1021/acs.chemrev.1c00212.

(45) Gault, J.; Liko, I.; Landreh, M.; Shutin, D.; Bolla, J. R.; Jefferies, D.; Agasid, M.; Yen, H.-Y.; Ladds, M. J. G. W.; Lane, D. P.; et al. Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nature Methods* **2020**, *17* (5), 505-508. DOI: 10.1038/s41592-020-0821-0.

(46) Yen, H.-Y.; Hopper, J. T. S.; Liko, I.; Allison, T. M.; Zhu, Y.; Wang, D.; Stegmann, M.; Mohammed, S.; Wu, B.; Robinson, C. V. Ligand binding to a G protein–coupled receptor captured in a mass spectrometer. *Science Advances 3* (6), e1701016. DOI: 10.1126/sciadv.1701016.

(47) Yen, H.-Y.; Liko, I.; Song, W.; Kapoor, P.; Almeida, F.; Toporowska, J.; Gherbi, K.; Hopper, J. T. S.; Charlton, S. J.; Politis, A.; et al. Mass spectrometry captures biased signalling and allosteric modulation of a G-protein-coupled receptor. *Nature Chemistry* **2022**, *14* (12), 1375-1382. DOI: 10.1038/s41557-022-01041-9.

(48) Chen, S.; Getter, T.; Salom, D.; Wu, D.; Quetschlich, D.; Chorev, D. S.; Palczewski, K.; Robinson, C. V. Capturing a rhodopsin receptor signalling cascade across a native membrane. *Nature* **2022**, *604* (7905), 384-390. DOI: 10.1038/s41586-022-04547-x.

(49) Heck, A. J. R. Native mass spectrometry: a bridge between interactomics and structural biology. *Nature Methods* **2008**, *5* (11), 927-933. DOI: 10.1038/nmeth.1265.

(50) Skinner, O. S.; Haverland, N. A.; Fornelli, L.; Melani, R. D.; Do Vale, L. H. F.; Seckler, H. S.; Doubleday, P. F.; Schachner, L. F.; Srzentić, K.; Kelleher, N. L.; et al. Top-down

characterization of endogenous protein complexes with native proteomics. *Nature Chemical Biology* **2018**, *14* (1), 36-41. DOI: 10.1038/nchembio.2515.

(51) Zhou, M.; Lantz, C.; Brown, K. A.; Ge, Y.; Paša-Tolić, L.; Loo, J. A.; Lermyte, F. Higher-order structural characterisation of native proteins and complexes by top-down mass spectrometry. *Chemical Science* **2020**, *11* (48), 12918-12936, 10.1039/D0SC04392C. DOI: 10.1039/D0SC04392C.

(52) Liu, R.; Xia, S.; Li, H. Native top-down mass spectrometry for higher-order structural characterization of proteins and complexes. *Mass Spectrometry Reviews* **2023**, *42* (5), 1876-1926. DOI: https://doi.org/10.1002/mas.21793.

(53) Po, A.; Eyers, C. E. Top-Down Proteomics and the Challenges of True Proteoform Characterization. *Journal of Proteome Research* **2023**, *22* (12), 3663-3675. DOI: 10.1021/acs.jproteome.3c00416.

(54) Brown, K. A.; Melby, J. A.; Roberts, D. S.; Ge, Y. Top-down proteomics: challenges, innovations, and applications in basic and clinical research. *Expert Review of Proteomics* **2020**, *17* (10), 719-733. DOI: 10.1080/14789450.2020.1855982.

(55) Anderson, N. L.; Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects*. *Molecular & Cellular Proteomics* **2002**, *1* (11), 845-867. DOI: https://doi.org/10.1074/mcp.R200007-MCP200.

(56) Cai, W.; Tucholski, T.; Chen, B.; Alpert, A. J.; McIlwain, S.; Kohmoto, T.; Jin, S.; Ge, Y. Top-Down Proteomics of Large Proteins up to 223 kDa Enabled by Serial Size Exclusion Chromatography Strategy. *Analytical Chemistry* **2017**, *89* (10), 5467-5475. DOI: 10.1021/acs.analchem.7b00380.

(57) Zhang, Z.; Wu, S.; Stenoien, D. L.; Paša-Tolić, L. High-Throughput Proteomics. From 2014.

(58) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; et al. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* **2011**, *480* (7376), 254-258. DOI: 10.1038/nature10575.

(59) Loo, R. R. O.; Dales, N.; Andrews, P. C. Surfactant effects on protein structure examined by electrospray ionization mass spectrometry. *Protein Science* **1994**, *3* (11), 1975-1983. DOI: https://doi.org/10.1002/pro.5560031109.

(60) Botelho, D.; Wall, M. J.; Vieira, D. B.; Fitzsimmons, S.; Liu, F.; Doucette, A. Top-Down and Bottom-Up Proteomics of SDS-Containing Solutions Following Mass-Based Separation. *Journal of Proteome Research* **2010**, *9* (6), 2863-2870. DOI: 10.1021/pr900949p.

(61) Puchades, M.; Westman, A.; Blennow, K.; Davidsson, P. Removal of sodium dodecyl sulfate from protein samples prior to matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1999**, *13* (5), 344-349.

(62) Wessel, D.; Flügge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry* **1984**, *138* (1), 141-143. DOI: https://doi.org/10.1016/0003-2697(84)90782-6.

(63) Kachuk, C.; Doucette, A. A. The benefits (and misfortunes) of SDS in top-down proteomics. *Journal of Proteomics* **2018**, *175*, 75-86. DOI: https://doi.org/10.1016/j.jprot.2017.03.002.

(64) Brown, K. A.; Chen, B.; Guardado-Alvarez, T. M.; Lin, Z.; Hwang, L.; Ayaz-Guner, S.; Jin, S.; Ge, Y. A photocleavable surfactant for top-down proteomics. *Nature Methods* **2019**, *16* (5), 417-420. DOI: 10.1038/s41592-019-0391-1.

(65) Brown, K. A.; Gugger, M. K.; Yu, Z.; Moreno, D.; Jin, S.; Ge, Y. Nonionic, Cleavable Surfactant for Top-Down Proteomics. *Analytical Chemistry* **2023**, *95* (3), 1801-1804. DOI: 10.1021/acs.analchem.2c03916.

(66) Cai, W.; Hite, Z. L.; Lyu, B.; Wu, Z.; Lin, Z.; Gregorich, Z. R.; Messer, A. E.; McIlwain, S. J.; Marston, S. B.; Kohmoto, T.; et al. Temperature-sensitive sarcomeric protein post-translational modifications revealed by top-down proteomics. *Journal of Molecular and Cellular Cardiology* **2018**, *122*, 11-22. DOI: https://doi.org/10.1016/j.yjmcc.2018.07.247.

(67) Lee, K. J.; Jordan, J. S.; Williams, E. R. Is Native Mass Spectrometry in Ammonium Acetate Really Native? Protein Stability Differences in Biochemically Relevant Salt Solutions. *Analytical Chemistry* **2024**, *96* (44), 17586-17593. DOI: 10.1021/acs.analchem.4c03415.

(68) Urner, L. H.; Liko, I.; Yen, H.-Y.; Hoi, K.-K.; Bolla, J. R.; Gault, J.; Almeida, F. G.; Schweder, M.-P.; Shutin, D.; Ehrmann, S.; et al. Modular detergents tailor the purification and structural analysis of membrane proteins including G-protein coupled receptors. *Nature Communications* **2020**, *11* (1), 564. DOI: 10.1038/s41467-020-14424-8.

(69) Yen, H.-Y.; Abramsson, M. L.; Agasid, M. T.; Lama, D.; Gault, J.; Liko, I.; Kaldmäe, M.; Saluri, M.; Qureshi, A. A.; Suades, A.; et al. Electrospray ionization of native membrane proteins proceeds via a charge equilibration step. *RSC Advances* **2022**, *12* (16), 9671-9680.

(70) Cox, B.; Emili, A. Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nature Protocols* **2006**, *1* (4), 1872-1878. DOI: 10.1038/nprot.2006.273.

(71) Karpov, O. A.; Stotland, A.; Raedschelders, K.; Chazarin, B.; Ai, L.; Murray, C. I.; Van Eyk, J. E. Proteomics of the heart. *Physiological Reviews* **2024**, *104* (3), 931-982. DOI: 10.1152/physrev.00026.2023.

(72) Catherman, A. D.; Durbin, K. R.; Ahlf, D. R.; Early, B. P.; Fellers, R. T.; Tran, J. C.; Thomas, P. M.; Kelleher, N. L. Large-scale Top-down Proteomics of the Human Proteome: Membrane Proteins, Mitochondria, and Senescence*. *Molecular & Cellular Proteomics* **2013**, *12* (12), 3465-3473. DOI: https://doi.org/10.1074/mcp.M113.030114.

(73) Bauer, A.; Kuster, B. Affinity purification-mass spectrometry. *European Journal of Biochemistry* **2003**, *270* (4), 570-578. DOI: https://doi.org/10.1046/j.1432-1033.2003.03428.x.

(74) Ge, Y.; Rybakova, I. N.; Xu, Q.; Moss, R. L. Top-down high-resolution mass spectrometry of cardiac myosin binding protein C revealed that truncation alters protein phosphorylation state. *Proceedings of the National Academy of Sciences* **2009**, *106* (31), 12658-12663. DOI: 10.1073/pnas.0813369106.

(75) Shao, X.; Tian, M.; Yin, J.; Duan, H.; Tian, Y.; Wang, H.; Xia, C.; Wang, Z.; Zhu, Y.; Wang, Y.; et al. Biofunctionalized dissolvable hydrogel microbeads enable efficient characterization of native protein complexes. *Nature Communications* **2024**, *15* (1), 8633. DOI: 10.1038/s41467-024-52948-5.

(76) Baker, M. Reproducibility crisis: Blame it on the antibodies. *Nature* **2015**, *521* (7552), 274-276. DOI: 10.1038/521274a.

(77) Janes, K. A. Fragile epitopes—Antibody's guess is as good as yours. *Science Signaling* **2020**, *13* (616), eaaz8130. DOI: 10.1126/scisignal.aaz8130.

(78) Tiambeng, T. N.; Roberts, D. S.; Brown, K. A.; Zhu, Y.; Chen, B.; Wu, Z.; Mitchell, S. D.; Guardado-Alvarez, T. M.; Jin, S.; Ge, Y. Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum. *Nature Communications* **2020**, *11* (1), 3903. DOI: 10.1038/s41467-020-17643-1.

(79) Chapman, E. A.; Roberts, D. S.; Tiambeng, T. N.; Andrews, J.; Wang, M.-D.; Reasoner, E. A.; Melby, J. A.; Li, B. H.; Kim, D.; Alpert, A. J.; et al. Structure and dynamics of endogenous cardiac troponin complex in human heart tissue captured by native nanoproteomics. *Nature Communications* **2023**, *14* (1), 8400. DOI: 10.1038/s41467-023-43321-z.

(80) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Analytical Chemistry* **2011**, *83* (17), 6868-6874. DOI: 10.1021/ac2010795.

(81) Cupp-Sutton, K. A.; Wang, Z.; Yu, D.; Wu, S. RPLC-RPLC-MS/MS for Proteoform Identification. In *Proteoform Identification: Methods and Protocols*, Sun, L., Liu, X. Eds.; Springer US, 2022; pp 31-42.

(82) Shen, Y.; Tolić, N.; Piehowski, P. D.; Shukla, A. K.; Kim, S.; Zhao, R.; Qu, Y.; Robinson, E.; Smith, R. D.; Paša-Tolić, L. High-resolution ultrahigh-pressure long column reversed-phase liquid chromatography for top-down proteomics. *Journal of Chromatography A* **2017**, *1498*, 99-110. DOI: https://doi.org/10.1016/j.chroma.2017.01.008.

(83) Chen, X.; Ge, Y. Ultrahigh pressure fast size exclusion chromatography for top-down proteomics. *PROTEOMICS* **2013**, *13* (17), 2563-2566. DOI: https://doi.org/10.1002/pmic.201200594.

(84) Tucholski, T.; Knott, S. J.; Chen, B.; Pistono, P.; Lin, Z.; Ge, Y. A Top-Down Proteomics Platform Coupling Serial Size Exclusion Chromatography and Fourier Transform Ion Cyclotron

Resonance Mass Spectrometry. *Analytical Chemistry* **2019**, *91* (6), 3835-3844. DOI: 10.1021/acs.analchem.8b04082.

(85) Rogers, H. T.; Melby, J. A.; Ehlers, L. E.; Fischer, M. S.; Larson, E. J.; Gao, Z.; Rossler, K. J.; Wang, D.; Alpert, A. J.; Ge, Y. Small-Scale Serial Size Exclusion Chromatography (s3SEC) for High Sensitivity Top-Down Proteomics of Large Proteoforms. *Analytical Chemistry* **2024**, *96* (7), 2748-2753. DOI: 10.1021/acs.analchem.3c05733.

(86) VanAernum, Z. L.; Busch, F.; Jones, B. J.; Jia, M.; Chen, Z.; Boyken, S. E.; Sahasrabuddhe, A.; Baker, D.; Wysocki, V. H. Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry. *Nature Protocols* **2020**, *15* (3), 1132-1157. DOI: 10.1038/s41596-019-0281-0.

(87) Chen, B.; Peng, Y.; Valeja, S. G.; Xiu, L.; Alpert, A. J.; Ge, Y. Online Hydrophobic Interaction Chromatography–Mass Spectrometry for Top-Down Proteomics. *Analytical Chemistry* **2016**, *88* (3), 1885-1891. DOI: 10.1021/acs.analchem.5b04285.

(88) Queiroz, J. A.; Tomaz, C. T.; Cabral, J. M. S. Hydrophobic interaction chromatography of proteins. *Journal of Biotechnology* **2001**, *87* (2), 143-159. DOI: https://doi.org/10.1016/S0168-1656(01)00237-1.

(89) Chen, B.; Lin, Z.; Alpert, A. J.; Fu, C.; Zhang, Q.; Pritts, W. A.; Ge, Y. Online Hydrophobic Interaction Chromatography–Mass Spectrometry for the Analysis of Intact Monoclonal Antibodies. *Analytical Chemistry* **2018**, *90* (12), 7135-7138. DOI: 10.1021/acs.analchem.8b01865.

(90) Kopaciewicz, W.; Rounds, M. A.; Fausnaugh, J.; Regnier, F. E. Retention model for high-performance ion-exchange chromatography. *Journal of Chromatography A* **1983**, *266*, 3-21. DOI: https://doi.org/10.1016/S0021-9673(01)90875-1.

(91) Baek, J.; Schwahn, A. B.; Lin, S.; Pohl, C. A.; De Pra, M.; Tremintin, S. M.; Cook, K. New Insights into the Chromatography Mechanisms of Ion-Exchange Charge Variant Analysis: Dispelling Myths and Providing Guidance for Robust Method Optimization. *Analytical Chemistry* **2020**, *92* (19), 13411-13419. DOI: 10.1021/acs.analchem.0c02775.

(92) Fischer, M. S.; Rogers, H. T.; Chapman, E. A.; Chan, H.-J.; Krichel, B.; Gao, Z.; Larson, E. J.; Ge, Y. Online Mixed-Bed Ion Exchange Chromatography for Native Top-Down Proteomics of Complex Mixtures. *Journal of Proteome Research* **2024**, *23* (7), 2315-2322. DOI: 10.1021/acs.jproteome.4c00430.

(93) Guo, Y.; Cupp-Sutton, K. A.; Zhao, Z.; Anjum, S.; Wu, S. Multidimensional separations in top–down proteomics. *Analytical Science Advances* **2023**, *4* (5-6), 181-203. DOI: https://doi.org/10.1002/ansa.202300016.

(94) Brown, K. A.; Tucholski, T.; Alpert, A. J.; Eken, C.; Wesemann, L.; Kyrvasilis, A.; Jin, S.; Ge, Y. Top-Down Proteomics of Endogenous Membrane Proteins Enabled by Cloud Point Enrichment and Multidimensional Liquid Chromatography–Mass Spectrometry. *Analytical Chemistry* **2020**, *92* (24), 15726-15735. DOI: 10.1021/acs.analchem.0c02533.

(95) Xiu, L.; Valeja, S. G.; Alpert, A. J.; Jin, S.; Ge, Y. Effective Protein Separation by Coupling Hydrophobic Interaction and Reverse Phase Chromatography for Top-down Proteomics. *Analytical Chemistry* **2014**, *86* (15), 7899-7906. DOI: 10.1021/ac501836k.

(96) Valeja, S. G.; Xiu, L.; Gregorich, Z. R.; Guner, H.; Jin, S.; Ge, Y. Three Dimensional Liquid Chromatography Coupling Ion Exchange Chromatography/Hydrophobic Interaction Chromatography/Reverse Phase Chromatography for Effective Protein Separation in Top-Down Proteomics. *Analytical Chemistry* **2015**, *87* (10), 5363-5371. DOI: 10.1021/acs.analchem.5b00657.

(97) Henley, W. H.; He, Y.; Mellors, J. S.; Batz, N. G.; Ramsey, J. M.; Jorgenson, J. W. High resolution separations of charge variants and disulfide isomers of monoclonal antibodies and antibody drug conjugates using ultra-high voltage capillary electrophoresis with high electric field strength. *Journal of Chromatography A* **2017**, *1523*, 72-79. DOI: https://doi.org/10.1016/j.chroma.2017.07.059.

(98) Sorensen, M. J.; Miller, K. E.; Jorgenson, J. W.; Kennedy, R. T. Two-dimensional liquid chromatography-mass spectrometry for lipidomics using off-line coupling of hydrophilic interaction liquid chromatography with 50 cm long reversed phase capillary columns. *Journal of Chromatography A* **2023**, *1687*, 463707. DOI: https://doi.org/10.1016/j.chroma.2022.463707.

(99) Pursch, M.; Buckenmaier, S. Loop-Based Multiple Heart-Cutting Two-Dimensional Liquid Chromatography for Target Analysis in Complex Matrices. *Analytical Chemistry* **2015**, *87* (10), 5310-5317. DOI: 10.1021/acs.analchem.5b00492.

(100) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246* (4926), 64-71. DOI: 10.1126/science.2675315.

(101) Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry* **1988**, *60* (20), 2299-2301. DOI: 10.1021/ac00171a028.

(102) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988**, *2* (8), 151-153. DOI: https://doi.org/10.1002/rcm.1290020802.

(103) Banerjee, S.; Mazumdar, S. Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry* **2012**, *2012* (1), 282574. DOI: https://doi.org/10.1155/2012/282574.

(104) Wilm, M. S.; Mann, M. Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last? *International Journal of Mass Spectrometry and Ion Processes* **1994**, *136* (2), 167-180. DOI: https://doi.org/10.1016/0168-1176(94)04024-9.

(105) Karas, M.; Bahr, U.; Dülcks, T. Nano-electrospray ionization mass spectrometry: addressing analytical problems beyond routine. *Fresenius' Journal of Analytical Chemistry* **2000**, *366* (6), 669-676. DOI: 10.1007/s002160051561.

(106) El-Faramawy, A.; Siu, K. W. M.; Thomson, B. A. Efficiency of Nano-Electrospray Ionization. *Journal of the American Society for Mass Spectrometry* **2005**, *16* (10), 1702-1707. DOI: https://doi.org/10.1016/j.jasms.2005.06.011.

(107) Mamyrin, B. A. Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry* **2001**, *206* (3), 251-266. DOI: https://doi.org/10.1016/S1387-3806(00)00392-4.

(108) Beck, S.; Michalski, A.; Raether, O.; Lubeck, M.; Kaspar, S.; Goedecke, N.; Baessmann, C.; Hornburg, D.; Meier, F.; Paron, I.; et al. The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics *. *Molecular & Cellular Proteomics* **2015**, *14* (7), 2014-2029. DOI: https://doi.org/10.1074/mcp.M114.047407.

(109) Mallis, C. S.; Zheng, X.; Qiu, X.; McCabe, J. W.; Shirzadeh, M.; Lyu, J.; Laganowsky, A.; Russell, D. H. Development of native MS capabilities on an extended mass range Q-TOF MS. *International Journal of Mass Spectrometry* **2020**, *458*, 116451. DOI: https://doi.org/10.1016/j.ijms.2020.116451.

(110) Fernandez-Lima, F.; Kaplan, D. A.; Suetering, J.; Park, M. A. Gas-phase separation using a trapped ion mobility spectrometer. *International Journal for Ion Mobility Spectrometry* **2011**, *14* (2), 93-98. DOI: 10.1007/s12127-011-0067-8.

(111) Panczyk, E. M.; Lin, Y.-F.; Harvey, S. R.; Snyder, D. T.; Liu, F. C.; Ridgeway, M. E.; Park, M. A.; Bleiholder, C.; Wysocki, V. H. Evaluation of a Commercial TIMS-Q-TOF Platform for Native Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2024**, *35* (7), 1394-1402. DOI: 10.1021/jasms.3c00320.

(112) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrometry Reviews* **1998**, *17* (1), 1-35.

(113) Tucholski, T.; Ge, Y. Fourier-transform ion cyclotron resonance mass spectrometry for characterizing proteoforms. *Mass Spectrometry Reviews* **2022**, *41* (2), 158-177, https://doi.org/10.1002/mas.21653. DOI: https://doi.org/10.1002/mas.21653.

(114) Chapman, E. A.; Li, B. H.; Krichel, B.; Chan, H.-J.; Buck, K. M.; Roberts, D. S.; Ge, Y. Native Top-Down Mass Spectrometry for Characterizing Sarcomeric Proteins Directly from Cardiac Tissue Lysate. *Journal of the American Society for Mass Spectrometry* **2024**, *35* (4), 738-745. DOI: 10.1021/jasms.3c00430.

(115) Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nature Methods* **2012**, *9* (11), 1084-1086. DOI: 10.1038/nmeth.2208.

(116) Fornelli, L.; Durbin, K. R.; Fellers, R. T.; Early, B. P.; Greer, J. B.; LeDuc, R. D.; Compton, P. D.; Kelleher, N. L. Advancing Top-down Analysis of the Human Proteome Using a Benchtop Quadrupole-Orbitrap Mass Spectrometer. *Journal of Proteome Research* **2017**, *16* (2), 609-618. DOI: 10.1021/acs.jproteome.6b00698.

(117) Gault, J.; Donlan, J. A. C.; Liko, I.; Hopper, J. T. S.; Gupta, K.; Housden, N. G.; Struwe, W. B.; Marty, M. T.; Mize, T.; Bechara, C.; et al. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature Methods* **2016**, *13* (4), 333-336. DOI: 10.1038/nmeth.3771.

(118) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* **2007**, *4* (9), 709-712. DOI: 10.1038/nmeth1060.

(119) Bayne, E. F.; Rossler, K. J.; Gregorich, Z. R.; Aballo, T. J.; Roberts, D. S.; Chapman, E. A.; Guo, W.; Palecek, S. P.; Ralphe, J. C.; Kamp, T. J.; et al. Top-down proteomics of myosin light chain isoforms define chamber-specific expression in the human heart. *Journal of Molecular and Cellular Cardiology* **2023**, *181*, 89-97. DOI: 10.1016/j.yjmcc.2023.06.003.

(120) Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Analytical Chemistry* **2016**, *88* (1), 30-51. DOI: 10.1021/acs.analchem.5b04563.

(121) Lantz, C.; Wei, B.; Zhao, B.; Jung, W.; Goring, A. K.; Le, J.; Miller, J.; Loo, R. R. O.; Loo, J. A. Native Top-Down Mass Spectrometry with Collisionally Activated Dissociation Yields Higher-Order Structure Information for Protein Complexes. *Journal of the American Chemical Society* **2022**, *144* (48), 21826-21830. DOI: 10.1021/jacs.2c06726.

(122) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *Journal of the American Chemical Society* **1998**, *120* (13), 3265-3266. DOI: 10.1021/ja973478k.

(123) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences* **2004**, *101* (26), 9528-9533. DOI: 10.1073/pnas.0402700101.

(124) Xu, F.; Xu, Q.; Dong, X.; Guy, M.; Guner, H.; Hacker, T. A.; Ge, Y. Top-down high-resolution electron capture dissociation mass spectrometry for comprehensive characterization of post-translational modifications in Rhesus monkey cardiac troponin I. *International Journal of Mass Spectrometry* **2011**, *305* (2), 95-102. DOI: https://doi.org/10.1016/j.ijms.2010.09.007.

(125) Jin, Y.; Peng, Y.; Lin, Z.; Chen, Y.-C.; Wei, L.; Hacker, T. A.; Larsson, L.; Ge, Y. Comprehensive analysis of tropomyosin isoforms in skeletal muscles by top-down proteomics. *Journal of Muscle Research and Cell Motility* **2016**, *37* (1), 41-52. DOI: 10.1007/s10974-016-9443-7.

(126) Peng, Y.; Chen, X.; Zhang, H.; Xu, Q.; Hacker, T. A.; Ge, Y. Top-down Targeted Proteomics for Deep Sequencing of Tropomyosin Isoforms. *Journal of Proteome Research* **2013**, *12* (1), 187-198. DOI: 10.1021/pr301054n.

(127) Li, H.; Wolff, J. J.; Van Orden, S. L.; Loo, J. A. Native Top-Down Electrospray Ionization-Mass Spectrometry of 158 kDa Protein Complex by High-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry* **2014**, *86* (1), 317-320. DOI: 10.1021/ac4033214.

(128) Zhang, J.; Malmirchegini, G. R.; Clubb, R. T.; Loo, J. A. Native Top-down Mass Spectrometry for the Structural Characterization of Human Hemoglobin. *European Journal of Mass Spectrometry* **2015**, *21* (3), 221-231. DOI: 10.1255/ejms.1340.

(129) Wongkongkathep, P.; Han, J. Y.; Choi, T. S.; Yin, S.; Kim, H. I.; Loo, J. A. Native Top-Down Mass Spectrometry and Ion Mobility MS for Characterizing the Cobalt and Manganese Metal Binding of α-Synuclein Protein. *Journal of The American Society for Mass Spectrometry* **2018**, *29* (9), 1870-1880. DOI: 10.1007/s13361-018-2002-2.

(130) Greisch, J.-F.; den Boer, M. A.; Lai, S.-H.; Gallagher, K.; Bondt, A.; Commandeur, J.; Heck, A. J. R. Extending Native Top-Down Electron Capture Dissociation to MDa Immunoglobulin

Complexes Provides Useful Sequence Tags Covering Their Critical Variable Complementarity-Determining Regions. *Analytical Chemistry* **2021**, *93* (48), 16068-16075. DOI: 10.1021/acs.analchem.1c03740.

(131) Iavarone, A. T.; Paech, K.; Williams, E. R. Effects of Charge State and Cationizing Agent on the Electron Capture Dissociation of a Peptide. *Analytical Chemistry* **2004**, *76* (8), 2231-2238. DOI: 10.1021/ac035431p.

(132) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *Journal of the American Chemical Society* **2013**, *135* (34), 12646-12651. DOI: 10.1021/ja4029654.

(133) O'Brien, J. P.; Li, W.; Zhang, Y.; Brodbelt, J. S. Characterization of Native Protein Complexes Using Ultraviolet Photodissociation Mass Spectrometry. *Journal of the American Chemical Society* **2014**, *136* (37), 12920-12928. DOI: 10.1021/ja505217w.

(134) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chemical Reviews* **2020**, *120* (7), 3328-3380. DOI: 10.1021/acs.chemrev.9b00440.

(135) Cai, W.; Tucholski, T. M.; Gregorich, Z. R.; Ge, Y. Top-down Proteomics: Technology Advancements and Applications to Heart Diseases. *Expert Review of Proteomics* **2016**, *13* (8), 717-730. DOI: 10.1080/14789450.2016.1209414.

(136) Cupp-Sutton, K. A.; Wu, S. High-throughput quantitative top-down proteomics. *Molecular Omics* **2020**, *16* (2), 91-99. DOI: 10.1039/C9MO00154A.

(137) Rhoads, T. W.; Rose, C. M.; Bailey, D. J.; Riley, N. M.; Molden, R. C.; Nestler, A. J.; Merrill, A. E.; Smith, L. M.; Hebert, A. S.; Westphall, M. S.; et al. Neutron-Encoded Mass Signatures for Quantitative Top-Down Proteomics. *Analytical Chemistry* **2014**, *86* (5), 2314-2319. DOI: 10.1021/ac403579s.

(138) Shortreed, M. R.; Frey, B. L.; Scalf, M.; Knoener, R. A.; Cesnik, A. J.; Smith, L. M. Elucidating Proteoform Families from Proteoform Intact-Mass and Lysine-Count Measurements. *Journal of Proteome Research* **2016**, *15* (4), 1213-1221. DOI: 10.1021/acs.jproteome.5b01090.

(139) Waanders, L. F.; Hanke, S.; Mann, M. Top-down quantitation and characterization of SILAC-labeled proteins. *Journal of the American Society for Mass Spectrometry* **2007**, *18* (11), 2058-2064. DOI: 10.1016/j.jasms.2007.09.001.

(140) Ntai, I.; Kim, K.; Fellers, R. T.; Skinner, O. S.; Smith, A. D. I. V.; Early, B. P.; Savaryn, J. P.; LeDuc, R. D.; Thomas, P. M.; Kelleher, N. L. Applying Label-Free Quantitation to Top Down Proteomics. *Analytical Chemistry* **2014**, *86* (10), 4961-4968. DOI: 10.1021/ac500395k.

(141) Wang, T.; Holt, M. V.; Young, N. L. The histone H4 proteoform dynamics in response to SUV4-20 inhibition reveals single molecule mechanisms of inhibitor resistance. *Epigenetics & Chromatin* **2018**, *11* (1), 29. DOI: 10.1186/s13072-018-0198-9.

(142) Durbin, K. R.; Fornelli, L.; Fellers, R. T.; Doubleday, P. F.; Narita, M.; Kelleher, N. L. Quantitation and Identification of Thousands of Human Proteoforms below 30 kDa. *Journal of Proteome Research* **2016**, *15* (3), 976-982. DOI: 10.1021/acs.jproteome.5b00997.

(143) Lin, Z.; Wei, L.; Cai, W.; Zhu, Y.; Tucholski, T.; Mitchell, S. D.; Guo, W.; Ford, S. P.; Diffee, G. M.; Ge, Y. Simultaneous Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome*[S]. *Molecular & Cellular Proteomics* **2019**, *18* (3), 594-605. DOI: https://doi.org/10.1074/mcp.TIR118.001086.

(144) Chapman, E. A.; Aballo, T. J.; Melby, J. A.; Zhou, T.; Price, S. J.; Rossler, K. J.; Lei, I.; Tang, P. C.; Ge, Y. Defining the Sarcomeric Proteoform Landscape in Ischemic Cardiomyopathy by Top-Down Proteomics. *Journal of Proteome Research* **2023**, *22* (3), 931-941. DOI: 10.1021/acs.jproteome.2c00729.

(145) Liu, X.; Inbar, Y.; Dorrestein, P. C.; Wynne, C.; Edwards, N.; Souda, P.; Whitelegge, J. P.; Bafna, V.; Pevzner, P. A. Deconvolution and Database Search of Complex Tandem Mass Spectra of Intact Proteins*. *Molecular & Cellular Proteomics* **2010**, *9* (12), 2772-2782. DOI: https://doi.org/10.1074/mcp.M110.002766.

(146) Tabb, D. L.; Jeong, K.; Druart, K.; Gant, M. S.; Brown, K. A.; Nicora, C.; Zhou, M.; Couvillion, S.; Nakayasu, E.; Williams, J. E.; et al. Comparing Top-Down Proteoform Identification: Deconvolution, PrSM Overlap, and PTM Detection. *Journal of Proteome Research* **2023**, *22* (7), 2199-2217. DOI: 10.1021/acs.jproteome.2c00673.

(147) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. Automated reduction and interpretation of. *Journal of the American Society for Mass Spectrometry* **2000**, *11* (4), 320-332. DOI: 10.1016/S1044-0305(99)00157-9.

(148) Kou, Q.; Xun, L.; Liu, X. TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics* **2016**, *32* (22), 3495-3497. DOI: 10.1093/bioinformatics/btw398.

(149) Yuan, Z.-F. e.; Liu, C.; Wang, H.-P.; Sun, R.-X.; Fu, Y.; Zhang, J.-F.; Wang, L.-H.; Chi, H.; Li, Y.; Xiu, L.-Y.; et al. pParse: A method for accurate determination of monoisotopic peaks in high-resolution mass spectra. *PROTEOMICS* **2012**, *12* (2), 226-235. DOI: https://doi.org/10.1002/pmic.201100081.

(150) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Analytical Chemistry* **2015**, *87* (8), 4370-4376. DOI: 10.1021/acs.analchem.5b00140.

(151) Kostelic, M. M.; Marty, M. T. Deconvolving Native and Intact Protein Mass Spectra with UniDec. In *Proteoform Identification: Methods and Protocols*, Sun, L., Liu, X. Eds.; Springer US, 2022; pp 159-180.

(152) Liu, X.; Hengel, S.; Wu, S.; Tolić, N.; Pasa-Tolić, L.; Pevzner, P. A. Identification of Ultramodified Proteins Using Top-Down Tandem Mass Spectra. *Journal of Proteome Research* **2013**, *12* (12), 5830-5838. DOI: 10.1021/pr400849y.

(153) Cesnik, A. J.; Shortreed, M. R.; Schaffer, L. V.; Knoener, R. A.; Frey, B. L.; Scalf, M.; Solntsev, S. K.; Dai, Y.; Gasch, A. P.; Smith, L. M. Proteoform Suite: Software for Constructing,

Quantifying, and Visualizing Proteoform Families. *Journal of Proteome Research* **2018**, *17* (1), 568-578. DOI: 10.1021/acs.jproteome.7b00685.

(154) LeDuc, R. D.; Taylor, G. K.; Kim, Y.-B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Research* **2004**, *32* (suppl_2), W340-W345. DOI: 10.1093/nar/gkh447.

(155) Karabacak, N. M.; Li, L.; Tiwari, A.; Hayward, L. J.; Hong, P.; Easterling, M. L.; Agar, J. N. Sensitive and Specific Identification of Wild Type and Variant Proteins from 8 to 669 kDa Using Top-down Mass Spectrometry*S. *Molecular & Cellular Proteomics* **2009**, *8* (4), 846-856. DOI: https://doi.org/10.1074/mcp.M800099-MCP200.

(156) Larson, E. J.; Pergande, M. R.; Moss, M. E.; Rossler, K. J.; Wenger, R. K.; Krichel, B.; Josyer, H.; Melby, J. A.; Roberts, D. S.; Pike, K.; et al. MASH Native: A Unified Solution for Native Top-Down Proteomics Data Processing. *Bioinformatics* **2023**, btad359. DOI: 10.1093/bioinformatics/btad359.

(157) Martin, S. S.; Aday, A. W.; Almarzooq, Z. I.; Anderson, C. A. M.; Arora, P.; Avery, C. L.; Baker-Smith, C. M.; Barone Gibbs, B.; Beaton, A. Z.; Boehme, A. K.; et al. 2024 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation* **2024**, *149* (8), e347-e913. DOI: 10.1161/CIR.0000000000001209.

(158) Simonetto, C.; Rospleszcz, S.; Kaiser, J. C.; Furukawa, K. Heterogeneity in coronary heart disease risk. *Scientific Reports* **2022**, *12* (1), 10131. DOI: 10.1038/s41598-022-14013-3.

(159) Chan, A. S.; Wu, S.; Vernon, S. T.; Tang, O.; Figtree, G. A.; Liu, T.; Yang, J. Y. H.; Patrick, E. Overcoming cohort heterogeneity for the prediction of subclinical cardiovascular disease risk. *iScience* **2023**, *26* (5), 106633. DOI: https://doi.org/10.1016/j.isci.2023.106633.

(160) Yuan, C.; Solaro, R. J. Myofilament proteins: From cardiac disorders to proteomic changes. *PROTEOMICS – Clinical Applications* **2008**, *2* (6), 788-799. DOI: https://doi.org/10.1002/prca.200780076.

(161) van der Velden, J.; Stienen, G. J. M. Cardiac Disorders and Pathophysiology of Sarcomeric Proteins. *Physiological Reviews* **2018**, *99* (1), 381-426. DOI: 10.1152/physrev.00040.2017.

(162) Hwang, P. M.; Sykes, B. D. Targeting the sarcomere to correct muscle function. *Nature Reviews Drug Discovery* **2015**, *14* (5), 313-328. DOI: 10.1038/nrd4554.

(163) Solís, C.; Solaro, R. J. Novel insights into sarcomere regulatory systems control of cardiac thin filament activation. *Journal of General Physiology* **2021**, *153* (7), e202012777. DOI: 10.1085/jgp.202012777.

(164) Solaro, R. J.; Henze, M.; Kobayashi, T. Integration of Troponin I Phosphorylation With Cardiac Regulatory Networks. *Circulation Research* **2013**, *112* (2), 355-366. DOI: 10.1161/CIRCRESAHA.112.268672.

(165) Pyle, W. G.; Solaro, R. J. At the Crossroads of Myocardial Signaling. *Circulation Research* **2004**, *94* (3), 296-305. DOI: 10.1161/01.RES.0000116143.74830.A9.

(166) Wang, X.; Su, H. Unraveling Enigma in the Z-Disks. *Circulation Research* **2010**, *107* (3), 321-323. DOI: 10.1161/CIRCRESAHA.110.225615.

(167) Fert-Bober, J.; Murray, C. I.; Parker, S. J.; Van Eyk, J. E. Precision Profiling of the Cardiovascular Post-Translationally Modified Proteome. *Circulation Research* **2018**, *122* (9), 1221-1237. DOI: 10.1161/CIRCRESAHA.118.310966.

(168) Solaro, R. J.; Warren, C. M.; Scruggs, S. B. Why is it important to analyze the cardiac sarcomere subproteome? *Expert Review of Proteomics* **2010**, *7* (3), 311-314. DOI: 10.1586/epr.10.15.

(169) Peng, Y.; Yu, D.; Gregorich, Z.; Chen, X.; Beyer, A. M.; Gutterman, D. D.; Ge, Y. In-depth proteomic analysis of human tropomyosin by top-down mass spectrometry. *Journal of Muscle Research and Cell Motility* **2013**, *34* (3), 199-210. DOI: 10.1007/s10974-013-9352-y.

(170) Chen, Y.-C.; Ayaz-Guner, S.; Peng, Y.; Lane, N. M.; Locher, M. R.; Kohmoto, T.; Larsson, L.; Moss, R. L.; Ge, Y. Effective Top-Down LC/MS+ Method for Assessing Actin Isoforms as a

Potential Cardiac Disease Marker. *Analytical Chemistry* **2015**, *87* (16), 8399-8406. DOI: 10.1021/acs.analchem.5b01745.

(171) Zhang, J.; Dong, X.; Hacker, T. A.; Ge, Y. Deciphering modifications in swine cardiac troponin I by top-down high-resolution tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* **2010**, *21* (6), 940-948. DOI: 10.1016/j.jasms.2010.02.005.

(172) Peng, Y.; Gregorich, Z. R.; Valeja, S. G.; Zhang, H.; Cai, W.; Chen, Y.-C.; Guner, H.; Chen, A. J.; Schwahn, D. J.; Hacker, T. A.; et al. Top-down Proteomics Reveals Concerted Reductions in Myofilament and Z-disc Protein Phosphorylation after Acute Myocardial Infarction*. *Molecular & Cellular Proteomics* **2014**, *13* (10), 2752-2764. DOI: https://doi.org/10.1074/mcp.M114.040675.

(173) Ayaz-Guner, S.; Zhang, J.; Li, L.; Walker, J. W.; Ge, Y. In Vivo Phosphorylation Site Mapping in Mouse Cardiac Troponin I by High Resolution Top-Down Electron Capture Dissociation Mass Spectrometry: Ser22/23 Are the Only Sites Basally Phosphorylated. *Biochemistry* **2009**, *48* (34), 8161-8170. DOI: 10.1021/bi900739f.

(174) Dong, X.; Sumandea, C. A.; Chen, Y.-C.; Garcia-Cazarin, M. L.; Zhang, J.; Balke, C. W.; Sumandea, M. P.; Ge, Y. Augmented Phosphorylation of Cardiac Troponin I in Hypertensive Heart Failure*. *Journal of Biological Chemistry* **2012**, *287* (2), 848-857. DOI: https://doi.org/10.1074/jbc.M111.293258.

(175) Sancho Solis, R.; Ge, Y.; Walker, J. W. Single amino acid sequence polymorphisms in rat cardiac troponin revealed by top–down tandem mass spectrometry. *Journal of Muscle Research and Cell Motility* **2008**, *29* (6), 203-212. DOI: 10.1007/s10974-009-9168-y.

(176) Chapman, E. A.; Rogers, H. T.; Gao, Z.; Chan, H.-J.; Alvarado, F. J.; Ge, Y. In-depth Characterization of S-Glutathionylation in Ventricular Myosin Light Chain 1 Across Species by Top-Down Proteomics. *bioRxiv* **2024**, 2024.2012.2011.628048. DOI: 10.1101/2024.12.11.628048.

(177) Gregorich, Z. R.; Peng, Y.; Lane, N. M.; Wolff, J. J.; Wang, S.; Guo, W.; Guner, H.; Doop, J.; Hacker, T. A.; Ge, Y. Comprehensive assessment of chamber-specific and transmural heterogeneity in myofilament protein phosphorylation by top-down mass spectrometry. *Journal of*

Molecular and Cellular Cardiology **2015**, *87*, 102-112. DOI: https://doi.org/10.1016/j.yjmcc.2015.08.007.

(178) Rogawski, R.; Sharon, M. Characterizing Endogenous Protein Complexes with Biological Mass Spectrometry. *Chemical Reviews* **2022**, *122* (8), 7386-7414. DOI: 10.1021/acs.chemrev.1c00217.

(179) Virani, S. S.; Alonso, A.; Aparicio, H. J.; Benjamin, E. J.; Bittencourt, M. S.; Callaway, C. W.; Carson, A. P.; Chamberlain, A. M.; Cheng, S.; Delling, F. N.; et al. Heart Disease and Stroke Statistics—2021 Update. *Circulation* **2021**, *143* (8), e254-e743. DOI: 10.1161/CIR.0000000000000950.

(180) Mudd, J. O.; Kass, D. A. Tackling heart failure in the twenty-first century. *Nature* **2008**, *451* (7181), 919-928. DOI: 10.1038/nature06798.

(181) Arrigo, M.; Jessup, M.; Mullens, W.; Reza, N.; Shah, A. M.; Sliwa, K.; Mebazaa, A. Acute heart failure. *Nature Reviews Disease Primers* **2020**, *6* (1), 16. DOI: 10.1038/s41572-020-0151-7.

(182) Ziaeian, B.; Fonarow, G. C. Epidemiology and aetiology of heart failure. *Nature Reviews Cardiology* **2016**, *13* (6), 368-378. DOI: 10.1038/nrcardio.2016.25.

(183) Briceno, N.; Schuster, A.; Lumley, M.; Perera, D. Ischaemic cardiomyopathy: pathophysiology, assessment and the role of revascularisation. *Heart* **2016**, *102* (5), 397. DOI: 10.1136/heartjnl-2015-308037.

(184) Felker, G. M.; Shaw, L. K.; O'Connor, C. M. A standardized definition of ischemic cardiomyopathy for use in clinical research. *Journal of the American College of Cardiology* **2002**, *39* (2), 210-218. DOI: https://doi.org/10.1016/S0735-1097(01)01738-7.

(185) Bloom, M. W.; Greenberg, B.; Jaarsma, T.; Januzzi, J. L.; Lam, C. S. P.; Maggioni, A. P.; Trochu, J.-N.; Butler, J. Heart failure with reduced ejection fraction. *Nature Reviews Disease Primers* **2017**, *3* (1), 17058. DOI: 10.1038/nrdp.2017.58.

(186) Marston, S. B.; Walker, J. W. Back to the future: new techniques show that forgotten phosphorylation sites are present in contractile proteins of the heart whilst intensively studied sites

appear to be absent. *Journal of Muscle Research and Cell Motility* **2009**, *30* (3), 93-95. DOI: 10.1007/s10974-009-9184-y.

(187) Ramirez-Correa, G. A.; Jin, W.; Wang, Z.; Zhong, X.; Gao, W. D.; Dias, W. B.; Vecoli, C.; Hart, G. W.; Murphy, A. M. O-Linked GlcNAc Modification of Cardiac Myofilament Proteins. *Circulation Research* **2008**, *103* (12), 1354-1358. DOI: 10.1161/CIRCRESAHA.108.184978.

(188) Lau, E.; Han, Y.; Williams, D. R.; Thomas, C. T.; Shrestha, R.; Wu, J. C.; Lam, M. P. Y. Splice-Junction-Based Mapping of Alternative Isoforms in the Human Proteome. *Cell Reports* **2019**, *29* (11), 3751-3765.e3755. DOI: https://doi.org/10.1016/j.celrep.2019.11.026.

(189) Drown, B. S.; Jooß, K.; Melani, R. D.; Lloyd-Jones, C.; Camarillo, J. M.; Kelleher, N. L. Mapping the Proteoform Landscape of Five Human Tissues. *Journal of Proteome Research* **2022**, *21* (5), 1299-1310. DOI: 10.1021/acs.jproteome.2c00034.

(190) Tiambeng, T. N.; Tucholski, T.; Wu, Z.; Zhu, Y.; Mitchell, S. D.; Roberts, D. S.; Jin, Y.; Ge, Y. Chapter Fifteen - Analysis of cardiac troponin proteoforms by top-down mass spectrometry. In *Methods in Enzymology*, Garcia, B. A. Ed.; Vol. 626; Academic Press, 2019; pp 347-374.

(191) Li, A.; Ponten, F.; dos Remedios, C. G. The interactome of LIM domain proteins: The contributions of LIM domain proteins to heart failure and heart development. *PROTEOMICS* **2012**, *12* (2), 203-225. DOI: https://doi.org/10.1002/pmic.201100492.

(192) Wadmore, K.; Azad, A. J.; Gehmlich, K. The Role of Z-disc Proteins in Myopathy and Cardiomyopathy. In *International Journal of Molecular Sciences*, 2021; Vol. 22.

(193) Gregorich, Z. R.; Cai, W.; Lin, Z.; Chen, A. J.; Peng, Y.; Kohmoto, T.; Ge, Y. Distinct sequences and post-translational modifications in cardiac atrial and ventricular myosin light chains revealed by top-down mass spectrometry. *Journal of Molecular and Cellular Cardiology* **2017**, *107*, 13-21. DOI: https://doi.org/10.1016/j.yjmcc.2017.04.002.

(194) Scruggs, S. B.; Reisdorph, R.; Armstrong, M. L.; Warren, C. M.; Reisdorph, N.; Solaro, R. J.; Buttrick, P. M. A Novel, In-solution Separation of Endogenous Cardiac Sarcomeric Proteins and Identification of Distinct Charged Variants of Regulatory Light Chain*. *Molecular & Cellular Proteomics* **2010**, *9* (9), 1804-1818. DOI: https://doi.org/10.1074/mcp.M110.000075.

(195) Chang, A. N.; Chen, G.; Gerard, R. D.; Kamm, K. E.; Stull, J. T. Cardiac Myosin Is a Substrate for Zipper-interacting Protein Kinase (ZIPK)*. *Journal of Biological Chemistry* **2010**, *285* (8), 5122-5126. DOI: https://doi.org/10.1074/jbc.C109.076489.

(196) Wang, C.; Taskinen, J. H.; Segersvärd, H.; Immonen, K.; Kosonen, R.; Tolva, J. M.; Mäyränpää, M. I.; Kovanen, P. T.; Olkkonen, V. M.; Sinisalo, J.; et al. Alterations of Cardiac Protein Kinases in Cyclic Nucleotide-Dependent Signaling Pathways in Human Ischemic Heart Failure. *Frontiers in Cardiovascular Medicine* **2022**, *9*.

(197) Liu, Y.; Chen, J.; Fontes, S. K.; Bautista, E. N.; Cheng, Z. Physiological and pathological roles of protein kinase A in the heart. *Cardiovascular Research* **2022**, *118* (2), 386-398. DOI: 10.1093/cvr/cvab008.

(198) Grandi, E.; Ripplinger, C. M. Antiarrhythmic mechanisms of beta blocker therapy. *Pharmacological Research* **2019**, *146*, 104274. DOI: https://doi.org/10.1016/j.phrs.2019.104274.

(199) Ong, H. T. β blockers in hypertension and cardiovascular disease. *BMJ* **2007**, *334* (7600), 946. DOI: 10.1136/bmj.39185.440382.47.

(200) Peng, Y.; Ayaz-Guner, S.; Yu, D.; Ge, Y. Top-down mass spectrometry of cardiac myofilament proteins in health and disease. *PROTEOMICS – Clinical Applications* **2014**, *8* (7-8), 554-568. DOI: https://doi.org/10.1002/prca.201400043.

(201) Geeves, M. A.; Hitchcock-DeGregori, S. E.; Gunning, P. W. A systematic nomenclature for mammalian tropomyosin isoforms. *Journal of Muscle Research and Cell Motility* **2015**, *36* (2), 147-153. DOI: 10.1007/s10974-014-9389-6.

(202) Garg, P.; Morris, P.; Fazlanie, A. L.; Vijayan, S.; Dancso, B.; Dastidar, A. G.; Plein, S.; Mueller, C.; Haaf, P. Cardiac biomarkers of acute coronary syndrome: from history to high-sensitivity cardiac troponin. *Internal and Emergency Medicine* **2017**, *12* (2), 147-155. DOI: 10.1007/s11739-017-1612-1.

(203) Zhang, J.; Zhang, H.; Ayaz-Guner, S.; Chen, Y.-C.; Dong, X.; Xu, Q.; Ge, Y. Phosphorylation, but Not Alternative Splicing or Proteolytic Degradation, Is Conserved in Human and Mouse Cardiac Troponin T. *Biochemistry* **2011**, *50* (27), 6081-6092. DOI: 10.1021/bi2006256.

(204) Perry, S. V. Troponin T: genetics, properties and function. *Journal of Muscle Research & Cell Motility* **1998**, *19* (6), 575-602. DOI: 10.1023/A:1005397501968.

(205) Wei, B.; Jin, J. P. Troponin T isoforms and posttranscriptional modifications: Evolution, regulation and function. *Archives of Biochemistry and Biophysics* **2011**, *505* (2), 144-154. DOI: https://doi.org/10.1016/j.abb.2010.10.013.

(206) Millat, G.; Bouvagnet, P.; Chevalier, P.; Sebbag, L.; Dulac, A.; Dauphin, C.; Jouk, P.-S.; Delrue, M.-A.; Thambo, J.-B.; Le Metayer, P.; et al. Clinical and mutational spectrum in a cohort of 105 unrelated patients with dilated cardiomyopathy. *European Journal of Medical Genetics* **2011**, *54* (6), e570-e575. DOI: https://doi.org/10.1016/j.ejmg.2011.07.005.

(207) Nakajima-Taniguchi, C.; Matsui, H.; Fujio, Y.; Nagata, S.; Kishimoto, T.; Yamauchi-Takihara, K. Novel Missense Mutation in Cardiac Troponin T Gene Found in Japanese Patient with Hypertrophic Cardiomyopathy. *Journal of Molecular and Cellular Cardiology* **1997**, *29* (2), 839-843. DOI: https://doi.org/10.1006/jmcc.1996.0322.

(208) Vafiadaki, E.; Arvanitis, D. A.; Sanoudou, D. Muscle LIM Protein: Master regulator of cardiac and skeletal muscle functions. *Gene* **2015**, *566* (1), 1-7. DOI: https://doi.org/10.1016/j.gene.2015.04.077.

(209) Frey, N.; Richardson, J. A.; Olson, E. N. Calsarcins, a novel family of sarcomeric calcineurin-binding proteins. *Proceedings of the National Academy of Sciences* **2000**, *97* (26), 14632-14637. DOI: 10.1073/pnas.260501097.

(210) Frank, D.; Kuhn, C.; van Eickels, M.; Gehring, D.; Hanselmann, C.; Lippl, S.; Will, R.; Katus, H. A.; Frey, N. Calsarcin-1 Protects Against Angiotensin-II–Induced Cardiac Hypertrophy. *Circulation* **2007**, *116* (22), 2587-2596. DOI: 10.1161/CIRCULATIONAHA.107.711317.

(211) Buyandelger, B.; Ng, K.-E.; Miocic, S.; Piotrowska, I.; Gunkel, S.; Ku, C.-H.; Knöll, R. MLP (muscle LIM protein) as a stress sensor in the heart. *Pflügers Archiv - European Journal of Physiology* **2011**, *462* (1), 135-142. DOI: 10.1007/s00424-011-0961-2.

(212) Roselló-Lletí, E.; Alonso, J.; Cortés, R.; Almenar, L.; Martínez-Dolz, L.; Sánchez-Lázaro, I.; Lago, F.; Azorín, I.; Juanatey, J. R. G.; Portolés, M.; et al. Cardiac protein changes in ischaemic

and dilated cardiomyopathy: a proteomic study of human left ventricular tissue. *Journal of Cellular and Molecular Medicine* **2012**, *16* (10), 2471-2486. DOI: https://doi.org/10.1111/j.1582-4934.2012.01565.x.

(213) Yi, X.; Jiang, D.-S.; Feng, G.; Jiang, X.-j.; Zeng, H.-L. An altered left ventricle protein profile in human ischemic cardiomyopathy revealed in comparative quantitative proteomics. *Polish Heart Journal (Kardiologia Polska)* **2019**, *77* (10), 951-959.

(214) Han, J. J.; Acker, M. A.; Atluri, P. Left Ventricular Assist Devices. *Circulation* **2018**, *138* (24), 2841-2851. DOI: 10.1161/CIRCULATIONAHA.118.035566.

(215) Nakayasu, E. S.; Gritsenko, M.; Piehowski, P. D.; Gao, Y.; Orton, D. J.; Schepmoes, A. A.; Fillmore, T. L.; Frohnert, B. I.; Rewers, M.; Krischer, J. P.; et al. Tutorial: best practices and considerations for mass-spectrometry-based protein biomarker discovery and validation. *Nature Protocols* **2021**, *16* (8), 3737-3760. DOI: 10.1038/s41596-021-00566-6.

(216) Laganowsky, A.; Reading, E.; Hopper, J. T. S.; Robinson, C. V. Mass spectrometry of intact membrane protein complexes. *Nature Protocols* **2013**, *8* (4), 639-651. DOI: 10.1038/nprot.2013.024.

(217) Yen, H.-Y.; Hoi, K. K.; Liko, I.; Hedger, G.; Horrell, M. R.; Song, W.; Wu, D.; Heine, P.; Warne, T.; Lee, Y.; et al. PtdIns(4,5)P2 stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature* **2018**, *559* (7714), 423-427. DOI: 10.1038/s41586-018-0325-6.

(218) Wu, D.; Robinson, C. V. Native Top-Down Mass Spectrometry Reveals a Role for Interfacial Glycans on Therapeutic Cytokine and Hormone Assemblies. *Angewandte Chemie International Edition* **2022**, *61* (49), e202213170, https://doi.org/10.1002/anie.202213170. DOI: https://doi.org/10.1002/anie.202213170.

(219) Bowman, J. D.; Lindert, S. Computational Studies of Cardiac and Skeletal Troponin. *Frontiers in Molecular Biosciences* **2019**, *6*.

(220) Jayasundar, J. J.; Xing, J.; Robinson, J. M.; Cheung, H. C.; Dong, W.-J. Molecular Dynamics Simulations of the Cardiac Troponin Complex Performed with FRET Distances as Restraints. *PLOS ONE* **2014**, *9* (2), e87135. DOI: 10.1371/journal.pone.0087135.

(221) Park, K. C.; Gaze, D. C.; Collinson, P. O.; Marber, M. S. Cardiac troponins: from myocardial infarction to chronic disease. *Cardiovascular Research* **2017**, *113* (14), 1708-1718. DOI: 10.1093/cvr/cvx183.

(222) Farah, C. S.; Reinach, F. C. The troponin complex and regulation of muscle contraction. *The FASEB Journal* **1995**, *9* (9), 755-767, https://doi.org/10.1096/fasebj.9.9.7601340. DOI: https://doi.org/10.1096/fasebj.9.9.7601340.

(223) van der Linden, N.; Wildi, K.; Twerenbold, R.; Pickering, J. W.; Than, M.; Cullen, L.; Greenslade, J.; Parsonage, W.; Nestelberger, T.; Boeddinghaus, J.; et al. Combining High-Sensitivity Cardiac Troponin I and Cardiac Troponin T in the Early Diagnosis of Acute Myocardial Infarction. *Circulation* **2018**, *138* (10), 989-999. DOI: 10.1161/CIRCULATIONAHA.117.032003.

(224) Layland, J.; Solaro, R. J.; Shah, A. M. Regulation of cardiac contractile function by troponin I phosphorylation. *Cardiovascular Research* **2005**, *66* (1), 12-21. DOI: 10.1016/j.cardiores.2004.12.022.

(225) Gomes, A. V.; Potter, J. D.; Szczesna-Cordary, D. The Role of Troponins in Muscle Contraction. *IUBMB Life* **2002**, *54* (6), 323-333, https://doi.org/10.1080/15216540216037. DOI: https://doi.org/10.1080/15216540216037.

(226) Takeda, S.; Yamashita, A.; Maeda, K.; Maéda, Y. Structure of the core domain of human cardiac troponin in the Ca2+-saturated form. *Nature* **2003**, *424* (6944), 35-41. DOI: 10.1038/nature01780.

(227) Davis, J. P.; Norman, C.; Kobayashi, T.; Solaro, R. J.; Swartz, D. R.; Tikunova, S. B. Effects of Thin and Thick Filament Proteins on Calcium Binding and Exchange with Cardiac Troponin C. *Biophysical Journal* **2007**, *92* (9), 3195-3206. DOI: https://doi.org/10.1529/biophysj.106.095406.

(228) Li, M. X.; Hwang, P. M. Structure and function of cardiac troponin C (TNNC1): Implications for heart failure, cardiomyopathies, and troponin modulating drugs. *Gene* **2015**, *571* (2), 153-166. DOI: https://doi.org/10.1016/j.gene.2015.07.074.

(229) Soetkamp, D.; Raedschelders, K.; Mastali, M.; Sobhani, K.; Bairey Merz, C. N.; Van Eyk, J. The continuing evolution of cardiac troponin I biomarker analysis: from protein to proteoform. *Expert Review of Proteomics* **2017**, *14* (11), 973-986. DOI: 10.1080/14789450.2017.1387054.

(230) Marston, S.; Zamora, J. E. Troponin structure and function: a view of recent progress. *Journal of Muscle Research and Cell Motility* **2020**, *41* (1), 71-89. DOI: 10.1007/s10974-019-09513-1.

(231) Metskas, L. A.; Rhoades, E. Order–Disorder Transitions in the Cardiac Troponin Complex. *Journal of Molecular Biology* **2016**, *428* (15), 2965-2977. DOI: https://doi.org/10.1016/j.jmb.2016.06.022.

(232) Liu, B.; Tikunova, S. B.; Kline, K. P.; Siddiqui, J. K.; Davis, J. P. Disease-Related Cardiac Troponins Alter Thin Filament Ca2+ Association and Dissociation Rates. *PLOS ONE* **2012**, *7* (6), e38259. DOI: 10.1371/journal.pone.0038259.

(233) Albury, A. N. J.; Swindle, N.; Swartz, D. R.; Tikunova, S. B. Effect of Hypertrophic Cardiomyopathy-Linked Troponin C Mutations on the Response of Reconstituted Thin Filaments to Calcium upon Troponin I Phosphorylation. *Biochemistry* **2012**, *51* (17), 3614-3621. DOI: 10.1021/bi300187k.

(234) Johnson, R. A.; Fulcher, L. M.; Vang, K.; Palmer, C. D.; Grossoehme, N. E.; Spuches, A. M. In depth, thermodynamic analysis of Ca2+ binding to human cardiac troponin C: Extracting buffer-independent binding parameters. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2019**, *1867* (4), 359-366. DOI: https://doi.org/10.1016/j.bbapap.2019.01.004.

(235) Rayani, K.; Seffernick, J.; Li, A. Y.; Davis, J. P.; Spuches, A. M.; Van Petegem, F.; Solaro, R. J.; Lindert, S.; Tibbits, G. F. Binding of calcium and magnesium to human cardiac troponin C. *Journal of Biological Chemistry* **2021**, *296*, 100350. DOI: https://doi.org/10.1016/j.jbc.2021.100350.

(236) Xiao, X.; Kuang, Z.; Slocik, J. M.; Tadepalli, S.; Brothers, M.; Kim, S.; Mirau, P. A.; Butkus, C.; Farmer, B. L.; Singamaneni, S.; et al. Advancing Peptide-Based Biorecognition Elements for

Biosensors Using in-Silico Evolution. *ACS Sensors* **2018**, *3* (5), 1024-1031. DOI: 10.1021/acssensors.8b00159.

(237) Townsend, J. A.; Keener, J. E.; Miller, Z. M.; Prell, J. S.; Marty, M. T. Imidazole Derivatives Improve Charge Reduction and Stabilization for Native Mass Spectrometry. *Analytical Chemistry* **2019**, *91* (22), 14765-14772. DOI: 10.1021/acs.analchem.9b04263.

(238) Shukla, D.; Trout, B. L. Understanding the Synergistic Effect of Arginine and Glutamic Acid Mixtures on Protein Solubility. *The Journal of Physical Chemistry B* **2011**, *115* (41), 11831-11839. DOI: 10.1021/jp204462t.

(239) Golovanov, A. P.; Hautbergue, G. M.; Wilson, S. A.; Lian, L.-Y. A Simple Method for Improving Protein Solubility and Long-Term Stability. *Journal of the American Chemical Society* **2004**, *126* (29), 8933-8939. DOI: 10.1021/ja049297h.

(240) Brgles, M.; Sviben, D.; Forčić, D.; Halassy, B. Nonspecific native elution of proteins and mumps virus in immunoaffinity chromatography. *Journal of Chromatography A* **2016**, *1447*, 107-114. DOI: https://doi.org/10.1016/j.chroma.2016.04.022.

(241) Cavanagh, J.; Benson, L. M.; Thompson, R.; Naylor, S. In-Line Desalting Mass Spectrometry for the Study of Noncovalent Biological Complexes. *Analytical Chemistry* **2003**, *75* (14), 3281-3286. DOI: 10.1021/ac030182q.

(242) Lermyte, F.; Tsybin, Y. O.; O'Connor, P. B.; Loo, J. A. Top or Middle? Up or Down? Toward a Standard Lexicon for Protein Top-Down and Allied Mass Spectrometry Approaches. *Journal of the American Society for Mass Spectrometry* **2019**, *30* (7), 1149-1157. DOI: 10.1007/s13361-019-02201-x.

(243) Hwang, P. M.; Cai, F.; Pineda-Sanabria, S. E.; Corson, D. C.; Sykes, B. D. The cardiac-specific N-terminal region of troponin I positions the regulatory domain of troponin C. *Proceedings of the National Academy of Sciences* **2014**, *111* (40), 14412-14417. DOI: 10.1073/pnas.1410775111.

(244) Ferrières, G.; Pugnière, M.; Mani, J.-C.; Villard, S.; Laprade, M.; Doutre, P.; Pau, B.; Granier, C. Systematic mapping of regions of human cardiac troponin I involved in binding to

cardiac troponin C: N- and C-terminal low affinity contributing regions. *FEBS Letters* **2000**, *479* (3), 99-105. DOI: https://doi.org/10.1016/S0014-5793(00)01881-0.

(245) Cheung, J. Y.; Tillotson, D. L.; Yelamarty, R. V.; Scaduto, R. C. Cytosolic free calcium concentration in individual cardiac myocytes in primary culture. *American Journal of Physiology-Cell Physiology* **1989**, *256* (6), C1120-C1130. DOI: 10.1152/ajpcell.1989.256.6.C1120.

(246) Schober, T.; Huke, S.; Venkataraman, R.; Gryshchenko, O.; Kryshtal, D.; Hwang, H. S.; Baudenbacher, F. J.; Knollmann, B. C. Myofilament Ca Sensitization Increases Cytosolic Ca Binding Affinity, Alters Intracellular Ca Homeostasis, and Causes Pause-Dependent Ca-Triggered Arrhythmia. *Circulation Research* **2012**, *111* (2), 170-179. DOI: 10.1161/CIRCRESAHA.112.270041.

(247) Bers, D. M. Calcium Fluxes Involved in Control of Cardiac Myocyte Contraction. *Circulation Research* **2000**, *87* (4), 275-281. DOI: 10.1161/01.RES.87.4.275.

(248) Johnson, J. D.; Nakkula, R. J.; Vasulka, C.; Smillie, L. B. Modulation of Ca2+ exchange with the Ca(2+)-specific regulatory sites of troponin C. *Journal of Biological Chemistry* **1994**, *269* (12), 8919-8923. DOI: https://doi.org/10.1016/S0021-9258(17)37055-2.

(249) Tang, N.; Skibsted, L. H. Calcium Binding to Amino Acids and Small Glycine Peptides in Aqueous Solution: Toward Peptide Design for Better Calcium Bioavailability. *Journal of Agricultural and Food Chemistry* **2016**, *64* (21), 4376-4389. DOI: 10.1021/acs.jafc.6b01534.

(250) Vinogradova, M. V.; Stone, D. B.; Malanina, G. G.; Karatzaferi, C.; Cooke, R.; Mendelson, R. A.; Fletterick, R. J. Ca2+-regulated structural changes in troponin. *Proceedings of the National Academy of Sciences* **2005**, *102* (14), 5038-5043. DOI: 10.1073/pnas.0408882102.

(251) Jeanne Dit Fouque, K.; Garabedian, A.; Leng, F.; Tse-Dinh, Y.-C.; Ridgeway, M. E.; Park, M. A.; Fernandez-Lima, F. Trapped Ion Mobility Spectrometry of Native Macromolecular Assemblies. *Analytical Chemistry* **2021**, *93* (5), 2933-2941. DOI: 10.1021/acs.analchem.0c04556.

(252) Dodds, J. N.; Baker, E. S. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *Journal of the American Society for Mass Spectrometry* **2019**, *30* (11), 2185-2195. DOI: 10.1007/s13361-019-02288-2.

(253) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion mobility–mass spectrometry analysis of large protein complexes. *Nature Protocols* **2008**, *3* (7), 1139-1152. DOI: 10.1038/nprot.2008.78.

(254) Marklund, Erik G.; Degiacomi, Matteo T.; Robinson, Carol V.; Baldwin, Andrew J.; Benesch, Justin L. P. Collision Cross Sections for Structural Proteomics. *Structure* **2015**, *23* (4), 791-799. DOI: https://doi.org/10.1016/j.str.2015.02.010.

(255) Kleerekoper, Q.; Howarth, J. W.; Guo, X.; Solaro, R. J.; Rosevear, P. R. Cardiac Troponin I Induced Conformational Changes in Cardiac Troponin C As Monitored by NMR Using Site-Directed Spin and Isotope Labeling. *Biochemistry* **1995**, *34* (41), 13343-13352. DOI: 10.1021/bi00041a010.

(256) Lindhout, D. A.; Sykes, B. D. Structure and Dynamics of the C-domain of Human Cardiac Troponin C in Complex with the Inhibitory Region of Human Cardiac Troponin I*. *Journal of Biological Chemistry* **2003**, *278* (29), 27024-27034. DOI: https://doi.org/10.1074/jbc.M302497200.

(257) Hale, O. J.; Cooper, H. J. Native Mass Spectrometry Imaging of Proteins and Protein Complexes by Nano-DESI. *Analytical Chemistry* **2021**, *93* (10), 4619-4627. DOI: 10.1021/acs.analchem.0c05277.

(258) Harvey, S. R.; O'Neale, C.; Schey, K. L.; Wysocki, V. H. Native Mass Spectrometry and Surface Induced Dissociation Provide Insight into the Post-Translational Modifications of Tetrameric AQP0 Isolated from Bovine Eye Lens. *Analytical Chemistry* **2022**, *94* (3), 1515-1519. DOI: 10.1021/acs.analchem.1c04322.

(259) van der Velden, J. Diastolic myofilament dysfunction in the failing human heart. *Pflügers Archiv - European Journal of Physiology* **2011**, *462* (1), 155-163. DOI: 10.1007/s00424-011-0960-3.

(260) Reiffert, S.; Maytum, R.; Geeves, M.; Lohmann, K.; Greis, T.; Blüggel, M.; Meyer, H. E.; Heilmeyer, L. M. G.; Jaquet, K. Characterization of the cardiac holotroponin complex reconstituted from native cardiac troponin T and recombinant I and C. *European Journal of*

Biochemistry **1999**, *261* (1), 40-47, https://doi.org/10.1046/j.1432-1327.1999.00261.x. DOI: https://doi.org/10.1046/j.1432-1327.1999.00261.x.

(261) Swindle, N.; Tikunova, S. B. Hypertrophic Cardiomyopathy-Linked Mutation D145E Drastically Alters Calcium Binding by the C-Domain of Cardiac Troponin C. *Biochemistry* **2010**, *49* (23), 4813-4820. DOI: 10.1021/bi100400h.

(262) Negele, J. C.; Dotson, D. G.; Liu, W.; Sweeney, H. L.; Putkey, J. A. Mutation of the high affinity calcium binding sites in cardiac troponin C. *Journal of Biological Chemistry* **1992**, *267* (2), 825-831. DOI: https://doi.org/10.1016/S0021-9258(18)48358-5.

(263) Britt, H. M.; Cragnolini, T.; Thalassinos, K. Integration of Mass Spectrometry Data for Structural Biology. *Chemical Reviews* **2022**, *122* (8), 7952-7986. DOI: 10.1021/acs.chemrev.1c00356.

(264) Vimer, S.; Ben-Nissan, G.; Marty, M.; Fleishman, S. J.; Sharon, M. Direct-MS analysis of antibody-antigen complexes. *PROTEOMICS* **2021**, *21* (21-22), 2000300. DOI: https://doi.org/10.1002/pmic.202000300.

(265) Karch, K. R.; Snyder, D. T.; Harvey, S. R.; Wysocki, V. H. Native Mass Spectrometry: Recent Progress and Remaining Challenges. *Annual Review of Biophysics* **2022**, *51* (1), 157-179. DOI: 10.1146/annurev-biophys-092721-085421.

(266) Zhang, H.; Cui, W.; Wen, J.; Blankenship, R. E.; Gross, M. L. Native electrospray and electron-capture dissociation in FTICR mass spectrometry provide top-down sequencing of a protein component in an intact protein assembly. *Journal of the American Society for Mass Spectrometry* **2010**, *21* (12), 1966-1968. DOI: 10.1016/j.jasms.2010.08.006.

(267) Zhang, H.; Cui, W.; Wen, J.; Blankenship, R. E.; Gross, M. L. Native Electrospray and Electron-Capture Dissociation FTICR Mass Spectrometry for Top-Down Studies of Protein Assemblies. *Analytical Chemistry* **2011**, *83* (14), 5598-5606. DOI: 10.1021/ac200695d.

(268) Zhang, H.; Cui, W.; Gross, M. L. Native electrospray ionization and electron-capture dissociation for comparison of protein structure in solution and the gas phase. *International*

Journal of Mass Spectrometry **2013**, *354-355*, 288-291. DOI: https://doi.org/10.1016/j.ijms.2013.06.019.

(269) Jacob, E.; Unger, R. A tale of two tails: why are terminal residues of proteins exposed? *Bioinformatics* **2007**, *23* (2), e225-e230. DOI: 10.1093/bioinformatics/btl318.

(270) Gregorich, Z. R.; Peng, Y.; Cai, W.; Jin, Y.; Wei, L.; Chen, A. J.; McKiernan, S. H.; Aiken, J. M.; Moss, R. L.; Diffee, G. M.; et al. Top-Down Targeted Proteomics Reveals Decrease in Myosin Regulatory Light-Chain Phosphorylation That Contributes to Sarcopenic Muscle Dysfunction. *Journal of Proteome Research* **2016**, *15* (8), 2706-2716. DOI: 10.1021/acs.jproteome.6b00244.

(271) Sitbon, Y. H.; Yadav, S.; Kazmierczak, K.; Szczesna-Cordary, D. Insights into myosin regulatory and essential light chains: a focus on their roles in cardiac and skeletal muscle function, development and disease. *Journal of Muscle Research and Cell Motility* **2020**, *41* (4), 313-327. DOI: 10.1007/s10974-019-09517-x.

(272) Sheikh, F.; Lyon, R. C.; Chen, J. Functions of myosin light chain-2 (MYL2) in cardiac muscle and disease. *Gene* **2015**, *569* (1), 14-20. DOI: https://doi.org/10.1016/j.gene.2015.06.027.

(273) Kannan, K.; Babu, G. J.; Veluraja, K.; Rajamanickam, C. Molecular modelling of N-terminal region of human cardiac myosin light chain 2 induced during cardiac hypertrophy. *Current Science* **1995**, *68* (8), 819-825. (acccessed 2023/11/15/).JSTOR.

(274) Muthu, P.; Wang, L.; Yuan, C.-C.; Kazmierczak, K.; Huang, W.; Hernandez, O. M.; Kawai, M.; Irving, T. C.; Szczesna-Cordary, D. Structural and functional aspects of the myosin essential light chain in cardiac muscle contraction. *The FASEB Journal* **2011**, *25* (12), 4394-4405. DOI: https://doi.org/10.1096/fj.11-191973.

(275) Collins, J. H. Myosin light chains and troponin C: Structural and evolutionary relationships revealed by amino acid sequence comparisons. *Journal of Muscle Research & Cell Motility* **1991**, *12* (1), 3-25. DOI: 10.1007/BF01781170.

(276) Wang, Y.; Ajtai, K.; Kazmierczak, K.; Szczesna-Cordary, D.; Burghardt, T. P. N-Terminus of Cardiac Myosin Essential Light Chain Modulates Myosin Step-Size. *Biochemistry* **2016**, *55* (1), 186-198. DOI: 10.1021/acs.biochem.5b00817.

(277) Murakami, K.; Stewart, M.; Nozawa, K.; Tomii, K.; Kudou, N.; Igarashi, N.; Shirakihara, Y.; Wakatsuki, S.; Yasunaga, T.; Wakabayashi, T. Structural basis for tropomyosin overlap in thin (actin) filaments and the generation of a molecular swivel by troponin-T. *Proceedings of the National Academy of Sciences* **2008**, *105* (20), 7200-7205. DOI: 10.1073/pnas.0801950105.

(278) Li, X.; Holmes, K. C.; Lehman, W.; Jung, H.; Fischer, S. The Shape and Flexibility of Tropomyosin Coiled Coils: Implications for Actin Filament Assembly and Regulation. *Journal of Molecular Biology* **2010**, *395* (2), 327-339. DOI: https://doi.org/10.1016/j.jmb.2009.10.060.

(279) Brown, J. H.; Kim, K.-H.; Jun, G.; Greenfield, N. J.; Dominguez, R.; Volkmann, N.; Hitchcock-DeGregori, S. E.; Cohen, C. Deciphering the design of the tropomyosin molecule. *Proceedings of the National Academy of Sciences* **2001**, *98* (15), 8496-8501. DOI: 10.1073/pnas.131219198.

(280) Doran, M. H.; Pavadai, E.; Rynkiewicz, M. J.; Walklate, J.; Bullitt, E.; Moore, J. R.; Regnier, M.; Geeves, M. A.; Lehman, W. Cryo-EM and Molecular Docking Shows Myosin Loop 4 Contacts Actin and Tropomyosin on Thin Filaments. *Biophysical Journal* **2020**, *119* (4), 821-830. DOI: https://doi.org/10.1016/j.bpj.2020.07.006.

(281) McCool, E. N.; Xu, T.; Chen, W.; Beller, N. C.; Nolan, S. M.; Hummon, A. B.; Liu, X.; Sun, L. Deep top-down proteomics revealed significant proteoform-level differences between metastatic and nonmetastatic colorectal cancer cells. *Science Advances 8* (51), eabq6348. DOI: 10.1126/sciadv.abq6348.

(282) Christofi, E.; Barran, P. Ion Mobility Mass Spectrometry (IM-MS) for Structural Biology: Insights Gained by Measuring Mass, Charge, and Collision Cross Section. *Chemical Reviews* **2023**, *123* (6), 2902-2949. DOI: 10.1021/acs.chemrev.2c00600.

(283) Alamo, L.; Ware, J. S.; Pinto, A.; Gillilan, R. E.; Seidman, J. G.; Seidman, C. E.; Padrón, R. Effects of myosin variants on interacting-heads motif explain distinct hypertrophic and dilated cardiomyopathy phenotypes. *eLife* **2017**, *6*, e24634. DOI: 10.7554/eLife.24634.

(284) Rashdan, N. A.; Shrestha, B.; Pattillo, C. B. S-glutathionylation, friend or foe in cardiovascular health and disease. *Redox Biology* **2020**, *37*, 101693. DOI: https://doi.org/10.1016/j.redox.2020.101693.

(285) Li, X.; Zhang, T.; Day, N. J.; Feng, S.; Gaffrey, M. J.; Qian, W.-J. Defining the S-Glutathionylation Proteome by Biochemical and Mass Spectrometric Approaches. In *Antioxidants*, 2022; Vol. 11.

(286) Kukulage, D. S. K.; Matarage Don, N. N. J.; Ahn, Y.-H. Emerging chemistry and biology in protein glutathionylation. *Current Opinion in Chemical Biology* **2022**, *71*, 102221. DOI: https://doi.org/10.1016/j.cbpa.2022.102221.

(287) Rosas, P. C.; Solaro, R. J. Implications of S-glutathionylation of sarcomere proteins in cardiac disorders, therapies, and diagnosis. *Frontiers in Cardiovascular Medicine* **2023**, *9*.

(288) Pastore, A.; Piemonte, F. Protein Glutathionylation in Cardiovascular Diseases. In *International Journal of Molecular Sciences*, 2013; Vol. 14, pp 20845-20876.

(289) Patel, B. G.; Wilder, T.; Solaro, R. J. Novel control of cardiac myofilament response to calcium by S-glutathionylation at specific sites of myosin binding protein C. *Frontiers in Physiology* **2013**, *4*, Original Research.

(290) Budde, H.; Hassoun, R.; Tangos, M.; Zhazykbayeva, S.; Herwig, M.; Varatnitskaya, M.; Sieme, M.; Delalat, S.; Sultana, I.; Kolijn, D.; et al. The Interplay between S-Glutathionylation and Phosphorylation of Cardiac Troponin I and Myosin Binding Protein C in End-Stage Human Failing Hearts. In *Antioxidants*, 2021; Vol. 10.

(291) Li, X.; Gluth, A.; Zhang, T.; Qian, W.-J. Thiol redox proteomics: Characterization of thiol-based post-translational modifications. *PROTEOMICS* **2023**, *23* (13-14), 2200194. DOI: https://doi.org/10.1002/pmic.202200194.

(292) Hernandez, O. M.; Jones, M.; Guzman, G.; Szczesna-Cordary, D. Myosin essential light chain in health and disease. *American Journal of Physiology-Heart and Circulatory Physiology* **2007**, *292* (4), H1643-H1654. DOI: 10.1152/ajpheart.00931.2006.

(293) Alegre-Cebollada, J.; Kosuri, P.; Giganti, D.; Eckels, E.; Rivas-Pardo, Jaime A.; Hamdani, N.; Warren, Chad M.; Solaro, R. J.; Linke, Wolfgang A.; Fernández, Julio M. S-Glutathionylation of Cryptic Cysteines Enhances Titin Elasticity by Blocking Protein Folding. *Cell* **2014**, *156* (6), 1235-1246. DOI: https://doi.org/10.1016/j.cell.2014.01.056.

(294) Chakouri, N.; Reboul, C.; Boulghobra, D.; Kleindienst, A.; Nottin, S.; Gayrard, S.; Roubille, F.; Matecki, S.; Lacampagne, A.; Cazorla, O. Stress-induced protein S-glutathionylation and phosphorylation crosstalk in cardiac sarcomeric proteins - Impact on heart function. *International Journal of Cardiology* **2018**, *258*, 207-216. DOI: https://doi.org/10.1016/j.ijcard.2017.12.004.

(295) Wilder, T.; Ryba, D. M.; Wieczorek, D. F.; Wolska, B. M.; Solaro, R. J. N-acetylcysteine reverses diastolic dysfunction and hypertrophy in familial hypertrophic cardiomyopathy. *American Journal of Physiology-Heart and Circulatory Physiology* **2015**, *309* (10), H1720-H1730. DOI: 10.1152/ajpheart.00339.2015.

(296) Karczewski, K. J.; Snyder, M. P. Integrative omics for health and disease. *Nature Reviews Genetics* **2018**, *19* (5), 299-310. DOI: 10.1038/nrg.2018.4.

(297) Coman, C.; Solari, F. A.; Hentschel, A.; Sickmann, A.; Zahedi, R. P.; Ahrends, R. Simultaneous Metabolite, Protein, Lipid Extraction (SIMPLEX): A Combinatorial Multimolecular Omics Approach for Systems Biology*. *Molecular & Cellular Proteomics* **2016**, *15* (4), 1435-1466. DOI: https://doi.org/10.1074/mcp.M115.053702.

(298) Davis, J. P.; Tikunova, S. B. Ca2+ exchange with troponin C and cardiac muscle dynamics. *Cardiovascular Research* **2008**, *77* (4), 619-626. DOI: 10.1093/cvr/cvm098.

(299) Robertson, S. P.; Johnson, J. D.; Holroyde, M. J.; Kranias, E. G.; Potter, J. D.; Solaro, R. J. The effect of troponin I phosphorylation on the Ca2+-binding properties of the Ca2+-regulatory site of bovine cardiac troponin. *Journal of Biological Chemistry* **1982**, *257* (1), 260-263. DOI: https://doi.org/10.1016/S0021-9258(19)68355-9.

(300) Kooij, V.; Saes, M.; Jaquet, K.; Zaremba, R.; Foster, D. B.; Murphy, A. M.; dos Remedios, C.; van der Velden, J.; Stienen, G. J. M. Effect of troponin I Ser23/24 phosphorylation on Ca2+-sensitivity in human myocardium depends on the phosphorylation background. *Journal of Molecular and Cellular Cardiology* **2010**, *48* (5), 954-963. DOI: https://doi.org/10.1016/j.yjmcc.2010.01.002.

(301) Katrukha, I. A.; Katrukha, A. G. Myocardial Injury and the Release of Troponins I and T in the Blood of Patients. *Clinical Chemistry* **2021**, *67* (1), 124-130. DOI: 10.1093/clinchem/hvaa281.

(302) Labugger, R.; Organ, L.; Collier, C.; Atar, D.; Van Eyk, J. E. Extensive Troponin I and T Modification Detected in Serum From Patients With Acute Myocardial Infarction. *Circulation* **2000**, *102* (11), 1221-1226. DOI: 10.1161/01.CIR.102.11.1221.

(303) O'Donohoe, T. J.; Ketheesan, N.; Schrale, R. G. Anti-troponin antibodies following myocardial infarction. *Journal of Cardiology* **2017**, *69* (1), 38-45. DOI: https://doi.org/10.1016/j.jjcc.2016.07.018.

(304) Viode, A.; van Zalm, P.; Smolen, K. K.; Fatou, B.; Stevenson, D.; Jha, M.; Levy, O.; Steen, J.; Steen, H.; On behalf of the, I. N. A simple, time- and cost-effective, high-throughput depletion strategy for deep plasma proteomics. *Science Advances 9* (13), eadf9717. DOI: 10.1126/sciadv.adf9717.