NANOPROTEOMICS-ENABLED STRATEGIES FOR THE ANALYSIS OF CARDIAC TROPONIN

By

Timothy-Neil Trinidad Tiambeng

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

Doctor of Philosophy

(Chemistry)

At the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: January 21st, 2022

Ying Ge, Professor, Chemistry and Cell and Regenerative Biology

Song Jin, Professor, Chemistry

Andrew R. Buller, Professor, Chemistry

Helen E. Blackwell, Professor, Chemistry

Acknowledgements

I would first like to acknowledge and thank my advisors and mentors, Professor Ying Ge and Professor Song Jin, who have guided me on this journey to become a better scientist and person. They have been my "red pen editors" in a literal and figurative sense—not only for scientific writing or grants, but also to highlight the areas where my "story" can improve. Over the past 5 and a half years, I remember countless stories that Ying shared, of the way her late Ph.D advisor, Prof. Fred McLafferty, positively changed and impacted her life. I am grateful that Ying demonstrates her gratitude to Prof. McLafferty by "paying it forward" to the students she mentors. She will frequently remind us that her students are her top priority, and I am thankful for her tireless advocacy and unwavering support for her students—even with the many responsibilities she juggles. She has stimulated an environment and group culture where students are warm-hearted, collaborative, and can creatively pursue new ideas and research directions. It is an environment where students can learn, explore, and grow beyond the safety of their comfort zone, and I am forever grateful to have learned from Ying, and to have learned mass spectrometry in Ying's lab. I am both grateful and inspired by Song's seemingly boundless scientific curiosity, and his profound perspectives on science. The questions that Song asks, both in literature reviews and in research presentations, have helped me to better understand what elements can elevate research to great and significant science. His wisdom through stories, humor, vision, encouragement and empathy—especially when science became difficult and sometimes frustrating—is something I will deeply miss. I would like to thank the other members on my committee who I have been fortunate to interact with these past few years: Prof. Andrew Buller, and Prof. Helen Blackwell. Andrew, thank you so much for being so kind to me! I am so grateful for the wisdom and advice

you provided at each of my yearly committee meetings together—particularly the wisdom you provided me at my 2nd year TBE—"Fail fast." It remains a lingering memory and voice in the back of my mind, helping me when I approach the design of experiments or trying to systematically troubleshoot scientific problems. Helen, thank you for leading, along with Dr. Cara Jenkins, an amazing and stimulating Chemistry-Biology Interface (CBI) Training program. Learning how to broadly communicate to a diverse audience of scientists, as well as engaging in colloquiums with my fellow trainees, has been so enriching for me and I consider myself so fortunate to have participated in the CBI program under your leadership.

I would like to thank my high school chemistry teachers, and undergraduate research mentors at UC Irvine for the passion they displayed for their craft, and for showing me how fun and exciting both chemistry and the research endeavor can be. Mr. Groenboer, thank you for bringing chemistry to life with your interactive demos, and for sparking my interest in chemistry after my initial difficulties with freshman year science courses. Professor Zhibin Guan, thank you for allowing me to participate in chemistry research in your lab, and for instilling in your members that research questions which are significant, yet challenging, are often the most rewarding. Dr. Hanxiang Zeng, thank you for being such a kind and fantastic graduate student mentor to me, for teaching me how to do organic chemistry, for attending and listening to all my undergraduate research talks (and treating me to hotpot afterwards), and encouraging me to attend graduate school.

I would like to acknowledge and thank my former lab members in both the Jin and Ge groups who came before me: Matt Stolt, Lichen Xiu, Yongping Fu, Melinda Shearer, Lianna Dang, Wenjie Li, Nitish Mathur, Yuzhou Zhao, Matt Hautzinger, Zach Gregorich, Vege Cai, Ziqing Lin,

Bifan Chen, Yutong Jin, Kyle Brown, Trisha Tucholski, Zhijie (Abe) Wu, Austin Carr and Sam Knott. Specifically, I would like to thank Matt Stolt and Lianna Dang for their words of advice, encouragement, and for helping me with TEM experiments in my early years in the Jin group, especially as they were already on their way out. I would like to deeply thank the "big three": Kyle, Trisha, and Abe, who were one year ahead of me, and with whom I have interacted frequently, even after their graduation. Kyle, your technical prowess and scientific insights have been phenomenal, and have helped direct me and others in the right direction on multiple occasions. Trisha, you deserve nomination for sainthood for how helpful and patient you've been to me—thank you for the many, many late-night edits on my first protocol paper. Abe, thank you for being such a fun roommate for two years, and a great labmate and friend for 5+, especially now that we're working together in Regeneron! Austin, thank you for being a fantastic deskmate, an fun conversationalist, and for your hospitality over the years. Sam, thank you for help and support, particularly so in the earlier years of graduate school together.

I would like to acknowledge and thank current lab members in both the Jin and Ge groups, for making the lab such an incredible and unforgettable place to do research—I will miss this environment very much: Chih-Jung Chen, Chris Roy, David Roberts, Dom Ross, Daniel Kim, Dongxu Pan, Emily Reasoner, Erika Jackson, Hongyuan Shen, Jaan Andrews, Jaeyong Park, Katelyn Michael, Kristel Forlano, Rui Wang, Willa Mihalyi-Koch, Ben Wancewicz, David Moreno, Eli Larson, Elizabeth Bayne, Emily Chapman, Holden Rogers, Jake Melby, Kalina Reese, Kent Wenger, Kevin Buck, Timothy Aballo, and Yanlong Zhu. Specifically, I would like to deeply thank my good friend David Roberts, who is an amazing human being and an unbelievable force of nature in the lab—I will be keeping an eye on your future lab and research with great interest!

David Roberts, I'm grateful for your curiosity, your energy, your humor, and your passion and high-standards for research. Working with you, I always found myself laughing or joking, and I greatly appreciated your presence when we worked together on the weekends or on a late-night experiment. You've encouraged and inspired me when I didn't think I could do this, and I would not have enjoyed graduate school as much without you. Daniel, I am fortunate to have mentored you, and am so thankful for how hard-working and dedicated you are, especially as an undergraduate researcher. Emily Reasoner, it was a pleasure to work together in the nanoproteomics subgroup. David Moreno, thank you for being my foodie buddy and for our explorations together finding the best taquerias in Madison! Eli and Jake, thank you so much for allowing me to pick your brains with LC-MS questions, and for your warm energy and inviting presence, both in and out of the lab. Yanlong, thank you so much for your patience with instrument troubleshooting and showing me the ropes—I have learned so much practical knowledge of MS from you. I would like to thank Emily Chapman and Timothy Aballo—I've gotten to know both of you so much better over the last year and a half, and I'm so excited to see what the future holds for you both. It has been too much fun sitting next to both of you, and I'm grateful for our lunchtime walks around WIMR, exploring Philadelphia together, and our ability to endlessly laugh and share whatever is on our minds.

I would like to thank my friends here at UW-Madison, particularly my classmates of the 2016 Chemistry cohort: Marshall Padilla, Adam Kleman, Steven Chapman, Lei Liu, Aspen Herman, Keywan Johnson—thank you for being a fun and hilarious bunch to be around—I am grateful to have gotten to know you during our time together. Marshall, thanks for being a fantastic, fun, and understanding roommate, and encouraging me to take care of my physical and mental

health, especially during the height of the pandemic. Adam, thanks for your witty sense of humor, for guiding me on my five-month stint as a vegan during grad school, and your general thoughtfulness. Steven, thank you for being my tennis buddy and for your southern hospitality and charm. Lei, thank you letting me be your travel buddy to Milwaukee and Chicago, and for having me taste-test your amazing dumplings, matcha cake, and ox tail soup. Aspen, thank you for your many letters and cards of encouragement. Keywan, thank you for infectious laughter and for opening your house for so many board game nights—it was a warm, friendly environment where I was able to meet so many other people I wouldn't otherwise have met in the department.

I would like to thank my friends from my time at UC Irvine, who have been such a wonderful source of support for me in graduate school—especially since we all entered graduate school or professional school around the same time! I am grateful that our friendship together has only continued to grow and become stronger, even after our graduation. I would like to deeply thank David Cao, Joseph Kim, Bryan Xie, Greg Chan, Jenna Hung, Ashley Eastman, and Hannah Nguyen. Although we are scattered across the US now, I am so grateful to feel that you all are never that far away.

I would like to thank my partner, Alex, for his overwhelming support, encouragement, patience, especially during the last few months of graduate school and writing up this thesis. I'm grateful that you continue to remind me that I am valued in many more ways as a person than merely what I can do or provide. You have been an incredible bright spot for me during the pandemic, and I am excited to see what the future holds for us together.

Lastly, I would like to thank my family—particularly my parents—for being one of my primary sources of inspiration, and for their unconditional love and encouragement for me. They have taught and demonstrated to me through their love and actions that sometimes those with very little, can be the most generous of all. I am so lucky and blessed to have you as parents, and to be your son.

Table of Contents

Acknowledgements i

Table of Contents vii

Abstract ix

Chapter 1- Introduction 1

1.1 Mass spectrometry-based proteomics in health and disease applications 2

1.2 Cardiac biomarkers for diagnosis of acute myocardial infarction 4

1.3 The human blood proteome dynamic range is exceptionally challenging for the analysis of low-abundance proteins 8

1.4 LC-MS and LC-MS/MS analysis of proteins 13

References 21

Chapter 2 - Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum 28

Abstract 29

Introduction 30

Results 32

Conclusion 43

Material and Methods 45

References 60

Figures and Supplemental Information 65

Chapter 3 - Native Mass Spectrometry of Human Cardiac Troponin Complexes Enabled by Functionalized Nanoparticle Enrichment and Online Size Exclusion Chromatography.....113

Abstract 114

Introduction.....115

Materials and methods 119

Results and discussion 130

Conclusion 137

References.....139

Figures and Supplemental Information.....143

Chapter 4 – Conclusion and Future Directions.....156

Appendix A. Analysis of Cardiac Troponin Proteoforms by Top-Down Mass Spectrometry ...161

Appendix B. Size Exclusion Chromatography Strategies and MASH Explorer for Targeted Proteoform Characterization.....205

Abstract

Cardiac troponin I (cTnI) is a universally-accepted, clinical biomarker for the diagnosis of acute myocardial infarction (a.k.a., heart attack). cTnI is released into the bloodstream following cardiovascular injury where it circulates with low abundance (ng/ml) and in myriad proteoforms (e.g., phosphorylated, acetylated, oxidized, and truncated). Top-down mass spectrometry (MS)-based proteomics has emerged as a powerful technology for characterizing and quantifying proteoforms arising from post-translational modifications, alternative splicing, and sequence variations at the intact protein level. However, the large dynamic range of the blood proteome makes the MS-detection of low-abundance biomarker proteins extremely challenging. Thus, protein enrichment strategies are necessary to specifically capture and enrich low-abundance proteins from complex mixtures prior to MS-analysis.

In chapter 1, an introduction is given regarding the use of top-down mass spectrometry in health and disease cardiac protein biomarkers, and the need for front-end enrichment methods. In chapter 2, I describe the development of an integrated "nanoproteomics" method coupling peptide-functionalized superparamagnetic nanoparticles (NPs) with top-down MS for the enrichment and comprehensive analysis of cTnI proteoform monomers, directly from serum for the first time. Chapter 3 utilizes the NP platform developed in chapter 2 to purify and enable the native MS analysis of intact, non-recombinant cardiac troponin complexes (cTnT-cTnI-cTnC; 77 kDa) from human heart tissues, with native MS coupled online with size exclusion chromatography (SEC), thereby enabling the native MS analysis of cardiac troponin complexes for the first time. Chapter 4 describes future directions with using nanoproteomics for the analysis of clinical acute myocardial infarction samples, and the structural analysis of cardiac troponin complexes. Finally,

the thesis is concluded with two Appendices containing step-by-step protocols and methods for the analysis of cardiac troponin proteoforms by top-down mass spectrometry, and size exclusion chromatography strategies combined with MASH Explorer for targeted proteoform characterization. Overall, the work described here demonstrates how nanoproteomics can provide proteoform- and protein complex- resolved details of cardiac troponin to elucidate the molecular mechanisms underlying health and disease.

Chapter 1

Introduction

1.1 Mass spectrometry-based proteomics in health and disease applications

The central dogma of molecular biology describes the process of the flow of genetic information that is encoded in a biological system[1]. Sections of DNA molecules containing genetic information called genes, are transcribed into RNA transcript molecules which carry the information encoded in the DNA and are then translated into amino-acid chains that form proteins. However, after the completion of the Human Genome Project, which sought to identify all of the genes in human DNA, scientists discovered that there are only ~20,500 protein-coding genes[2]. From this finding, it became increasingly recognized that much of the complexity encountered in biological systems is present at the level of protein variation, rather than due to a high number of genes[3]. In the post-genomics era, the field of proteomics, or the large-scale analysis of proteins, has come to the forefront to better understand complex phenotypes present in health and disease. The proteome is highly complex and exhibits considerable plasticity owing to proteoforms—a term encompassing all the protein species produced by a single gene through alternative splicing of mRNA transcripts and protein post-translational modifications (PTMs)[3, 4], and highly homologous protein isoforms produced by multiple homologous genes[5]. In particular, protein PTMs play critical roles in the regulation of protein localization, activity, degradation, and can be altered in disease[6]. Thus, proteomics and the analysis of proteoforms will be indispensable for evaluation of disease states and the discovery of new protein biomarkers and drug targets.

Mass spectrometry (MS)-based proteomics has emerged as the most versatile and comprehensive strategy for protein characterization. In the “bottom-up” MS-based approach, proteins are extracted from a biological sample of interest, and are digested into peptides enzymatically. The peptides are then separated by liquid chromatography (LC) coupled with online

MS and fragmented by tandem MS (MS/MS) to obtain their amino acid sequences. Using bioinformatics software, peptides can be identified and quantified to infer their respective protein identity and quantity. Bottom-up proteomics is highly efficient for sequencing small, trypsin-digested peptides and valuable for large scale, high throughput protein identification[7]. However, the 'peptide-to-protein inference- problem in bottom-up proteomics can lead to a loss of proteoform information because homologous sequences used to identify proteins can be shared among similar protein isoforms and protein variants[8].

In contrast, the "top-down" MS-based approach analyzes intact proteins rather than proteolytically digested peptides, enabling all PTMs and sequence variations to be resolved at the molecular ion level without a priori knowledge[4, 5, 7, 9-11]. In the top-down approach, proteins are separated and analyzed by MS. Proteoforms of interest can be specifically isolated in the gas phase for MS/MS analysis to characterize protein PTMs and sequence variants. Moreover, because the ionization and gas phase physicochemical properties of larger protein molecules are less affected by PTMs than smaller peptide molecules are, top-down MS enables the relative quantification of proteoforms present within the same mass spectrum[12-14]. Thus, top-down MS is a powerful tool to understand molecular differences that arise in biological samples, to discover pathways that are altered in disease, and to identify proteoform-specific biomarkers[15].

Here, I will summarize our understanding of cardiac proteoform biomarkers for acute myocardial infarction, and the need for affinity-purification based approaches for these important biomarkers. Finally, I will discuss top-down LC/MS and LC/MS/MS-based strategies to analyze proteins.

1.2 Cardiac biomarkers for diagnosis of acute myocardial infarction

Significance of myocardial infarction and cardiac biomarker testing

Acute myocardial infarction (AMI, aka heart attack) is a major cause of morbidity and mortality worldwide. There are two types of heart attack: 1) ST-segment elevation MI (STEMI), which physicians can identify by the atypical cardiac rhythm patterns (ST-segment elevation) measured by electrocardiography alone and also by blood testing, and 2) Non-ST-segment elevation MI (NSTEMI), which produces typical cardiac rhythm patterns, but is only identifiable by blood testing. More than 3 million people each year are estimated to have an acute ST-elevation myocardial infarction (STEMI), with more than 4 million having a non-ST-elevation myocardial infarction (NSTEMI) [16, 17]. While NSTEMI is normally a less severe form of heart attack than STEMI, the pathophysiology resulting in myocardial infarction arises from occlusion or blockage of the coronary artery, which ultimately reduces the concentration of oxygen available to cardiomyocytes, leading to cardiomyocyte necrosis and the release of cellular contents circulating into the blood. Thus, myocardial infarction is defined pathologically as myocardial cell death due to prolonged ischemia, or lack of oxygen to cells. As early as 15 minutes after the onset of ischemia, changes in cellular glycogen, relaxed myofibrils, and disruption of the sarcomere can be noted[18]. Mitochondrial abnormalities are also seen as early as 10 minutes after occlusion by electron microscopy[19]. Timely diagnosis and oxygen reperfusion therapy, reduces ischemic injury of the heart, and is paramount for improved patient outcomes following AMI.[17]

In the United States alone, almost 30 million visits per year to the emergency department results in cardiac biomarker testing, accounting for over 1 in every 6 emergency department visits[20]. The detection of cardiomyocyte necrosis from the heart has historically relied upon the measurement of circulating cardiac markers in the blood, such as cardiac enzymes (e.g., creatine kinase-MB isoenzyme [CK-MB]), and cell contents (e.g., cardiac troponin I (cTnI), cardiac troponin T (cTnT), myoglobin)[21, 22]. Both the sensitivity and specificity for cardiovascular injury vary significantly among these markers, but cTnI and cTnT have emerged as the most sensitive and specific, as these cardiac protein isoforms are not also expressed in skeletal muscle tissue and have not been reported to increase following injury to non-cardiac tissues. As a result, an essential criterion for the diagnosis of AMI is the presence of myocardial injury detected by a rise and/or fall of cardiac troponins with at least one value above the 99th percentile upper reference limit of a healthy population, in the setting of evidence of myocardial ischemia[17].

Although elevated circulating cTn values reflect that injury has occurred in myocardial cells, there are many causes of elevated troponin values in clinical settings other than AMI. Higher concentrations of troponins do not indicate the underlying pathophysiological mechanisms and can be detected following physiological stresses in otherwise normal hearts, such as after participating in a marathon race[23, 24]. Various causes have been suggested for the release and detection of troponins, including normal turnover of myocardial cells, apoptosis, cellular release of cTn degradation products, and increased cellular wall permeability[25, 26]. However, it is not yet clinically possible to distinguish which increases in cTn are caused by a particular mechanism or pathway[27]. We hypothesize that altered PTM profiles on myocardial proteins such as the cardiac troponins

can function as molecular fingerprints of cellular signaling pathway activity with the potential to serve as the next-generation biomarkers beyond just measurement of concentration alone[10, 28, 29].

The cardiac troponin complex

Contraction of the cardiomyocyte occurs through the calcium-dependent interaction of the thick filament (primarily composed of myosin) and the thin filament (primarily composed of actin and tropomyosin)[30]. The ternary cardiac troponin complex, composed of cTnI, cTnT, and cTnC, serves as the molecular switch that links the rise and fall of intracellular Ca^{2+} concentration with cardiac muscle contraction and relaxation. Thus, the troponin complex functions as a key regulator of contractility in the cardiac muscle[31-33]. Each protein within the troponin complex fulfills a particular role in regulating interactions and cooperativity of the proteins found in the thick and thin filaments[34]. cTnT anchors tropomyosin onto the actin-tropomyosin thin filament, while organizing the other subunits in the troponin complex and fixing the complex itself onto the thin filament[35, 36 37]. cTnC acts as the Ca^{2+} sensor of the myofilaments, which binds intracellular calcium, inducing a change in its structure to propagate the signal to the other members of the troponin complex[38]. cTnI serves as the inhibitory subunit and interacts with actin in the absence of Ca^{2+}, causing a shift in tropomyosin to sterically impede the interactions between the thick and thin filaments, thus turning contraction on and off in the presence of intracellular calcium[39, 40].

Post-translational modifications of troponin complex subunits and regulation of muscle contraction

Like most important cellular regulators, the cardiac troponin subunits are under tight regulatory control, primarily through signaling pathways that ultimately lead to phosphorylation of specific serine (Ser) residues. cTnI and cTnT can be phosphorylated by protein kinases, which

have a considerable effect on the structure and properties of the troponin complex. Phosphorylation of cTnI by protein kinase A (PKA) is particularly interesting because it phosphorylates two residues in the N-terminus of cTnI that are not present in the skeletal muscle isoforms of the protein[41, 42]. As a result, this pathway is a specific regulatory mechanism expressed only in cardiac muscle. Using NMR, bisphosphorylation of cTnI at Ser-22 and Ser-23 leads to significant conformational changes—the formation of an alpha-helical structure by residues 20-29, the dissociation of the N-terminal domain of cTnI from cTnC, and a shift of the N-terminal domain to the inhibitory domain of cTnI[43]. It is argued that the interactions of the N-terminal domain with the inhibitory domain of cTnI mediate the physiological effects of phosphorylation by PKA[43]. PKA-dependent phosphorylation of TnI also plays a significant role in the development of cardiovascular diseases. Phosphorylation at Ser-22 and Ser-23 ultimately lead to a decrease in cardiac muscle sensitivity to Ca^{2+} , promoting an increase in the relaxation rate of cardiomyocytes[44]. It has been repeatedly observed that the rate of phosphorylation at Ser-22 and Ser-23 is decreased in the hearts of patients with heart failure and cardiomyopathies, leading to higher sensitivity to Ca^{2+} than normal hearts and a substantial increase in the contraction rate in the heart[10, 42, 45, 46]. Unlike cTnI, cTnT is not phosphorylated by PKA but rather by troponin T kinase on Ser-2[47, 48]. From MS-based experiments performed on human, rat, and mouse hearts[49, 50], cTnT is only found to be phosphorylated on Ser-2, despite the many putative phosphorylation sites in vitro. However, the physiological role of phosphorylation on Ser-2 is not yet fully understood because it does not affect Ca^{2+} sensitivity or the interaction between actin and myosin[36], although in vitro phosphorylation sites have demonstrated that cTnT phosphorylation inhibits the Ca^{2+} activation of the thin filament, and decreases the maximum force generated by

cardiomyocytes[51, 52]. Because of the substantial body of evidence that suggests the PTM status of the cardiac troponins can modulate heart function, the detection and quantification of released modified forms can provide a window into cardiac disease status if we are able to develop analytical technologies with sufficient sensitivity and specificity[28, 53].

1.3 **The human blood proteome dynamic range is exceptionally challenging for the analysis of low-abundance proteins**

While proteomics has achieved remarkable sensitivity, MS is biased towards high abundance proteins, which often suppress low abundance protein signal of biomarkers of interest in various clinical tissue samples or biofluids[54, 55]. Whole blood is an exceptionally complex fluid, comprised of both cells (i.e., red and white blood cells, platelets) and protein[56]. Depending on the preparation of this liquid (whether prepared with clotting factors, or without clotting factors), the liquid supernatant is considered plasma or serum, respectively. Such preparation also is known to potentially affect the availability of certain biomarkers, so care must be taken to ensure that analytes of interest are not being misrepresented during analysis[57]. Human serum and plasma have remained clinical biofluids of interest because they are easily accessible and less invasive to access compared to tissues, and contain a diversity of proteins that are secreted or leaked from damaged cells or tissues especially under disease conditions, or released from infectious organisms[58]. The concentration dynamic range in the human serum/plasma proteome spans at least 10 orders of magnitude, from the most abundant albumin in mg/ml to cytokines, interleukins, and troponins in ng/ml to pg/ml in normal individuals[55]. The high dynamic range is a significant challenge, which limits the analysis and depth possible with just conventional liquid chromatography-tandem mass

spectrometry alone. To address these challenges, the main approaches to address the dynamic range have been protein depletion or affinity enrichment strategies.

Depletion of highly abundant plasma proteins is often achieved through using commercialized depletion kits[59] or by precipitation of large molecular weight globular proteins using organic, acetonitrile-based extraction methods[60]. Commercialized depletion kits featuring solid phase extraction columns are widely used for the depletion of high abundance proteins as a first step in biomarker-related MS-based proteomic studies[61]. Bacterial Protein A and Protein G antibody-based columns, which specifically bind to the Fc region of immunoglobulin G (IgG), a protein comprising about 10-20% of the serum/plasma proteome[62], have been successfully used for specific separation of IgGs from serum and plasma[63, 64]. In another example featured acetonitrile-based extraction, in a study targeting 57 high to medium abundance serum proteins, 29 were able to be detected by liquid-chromatography/mass spectrometry (LC/MS) following this extraction, with albumin, the most abundant protein in serum and plasma, identified as the 20th most abundant protein[60]. While some success has been seen, in general, these preparations have several caveats[63]. A major disadvantage of commercial depletion kits is the relatively high cost and low sample capacity when working with antibody-based extraction columns[61]. Another problem associated with protein depletion is that low abundance proteins may be removed along with abundant species[65]. Albumin serves as a transport protein that binds a large variety of compounds, including hormones, lipoproteins, and amino acids. Thus, removal of albumin through depletion strategies may also result in the non-specific removal of low-abundance biomarkers of interest[63, 66]. In fact, removal of albumin from serum using depletion strategies has been previously shown to result in the loss of cardiac biomarkers from serum/plasma, such as cTnI[67]. Thus, while the ability to

potentially identify more, lower-abundance proteins using depletion methods is a step in the right direction, the nonselective loss of proteins can hinder downstream approaches for targeted biomarker analysis.

In contrast, enrichment strategies such as affinity purification can be coupled with mass spectrometry workflows to examine a targeted, specific subset of proteins that are present within complex biological samples[68]. In general, affinity purification refers to the capture of proteins of interest via specific interaction and enrichment with a ligand coupled to a solid support[69]. Many types of ligands have been used in affinity purification, such as DNA and RNA oligonucleotides, small molecule inhibitors, lipids, peptides, or proteins, with antibodies being the most widely used ligands[70-73]. After immobilizing a ligand onto a solid support (traditionally agarose beads functionalized with a reactive electrophile), this coupled ligand can be used to capture the proteins of interest from solution. After washing unbound proteins from the solid support, the captured proteins are eluted by disrupting the ligand-antigen interactions. In particular, drastic changes in pH are commonly used to disrupt antibody-antigen interactions in immunoaffinity purification using small molecule buffers such as glycine-HCl, pH 2.0, or triethylamine, pH 11.5[74]. These small molecules can then be readily removed via buffer exchange or ion exchange prior to downstream MS applications. The Ge group has previously utilized immunoaffinity purification featuring agarose-functionalized monoclonal antibodies for the enrichment of cardiac troponin I (cTnI) from a variety of animal and human cardiac tissues for the study of disease-associated cTnI proteoform alterations using the top-down approach. However, in the MS analysis of disease-modified cTnI proteoforms from human serum/plasma, progress has been severely hampered by limitations inherent to the conventional immunoaffinity material platform. A long-standing problem with

agarose beads is that despite blocking electrophilic reactive sites, affinity-purified proteins are often contaminated by proteins that are non-specifically adsorbed to the largely porous matrix during the incubation step and are later eluted during the elution step as undesired contaminants[75]. Poor colloidal dispersion of the materials caused by their large sizes (>1 μm), can also cause incomplete interactions with the proteins of interest[76]. Even despite the use of monoclonal antibodies, due to the incredibly high non-specific binding in the agarose platform, large quantities of albumin and IgG are reported to co-elute and overwhelm the small amount of cTnI and cTnT that are commonly present in serum after AMI[53]. Therefore, to address and improve upon the considerable drawbacks with conventional agarose-based platforms, developing nanomaterials as scaffolds for applications such as protein enrichment and high-sensitivity detection remains an active area of research.

Gold Nanoparticles

Gold nanoparticles (AuNPs) are attractive choices for biological applications because of their unique physical and chemical properties[77, 78]. AuNPs can be easily functionalized via simple thiolate chemistry and show significant size-dependent color differences[79]. Thus, AuNPs have been used for the specific detection of DNAs and proteins present in extremely low abundance[80]. Wang and co-workers demonstrated that AuNPs could be used to enrich protein present in human urine, where the total protein concentration is less than 1.4 ppm[81]. Gold nanoparticles have been shown to enrich protein present in human urine, which contains 1.4 ppm amount of protein. This relied on the use of the affinity of gold-thiol chemistry, in which proteins containing high amounts of cysteine residues were enriched from the human urine samples. Similarly, Tan and coworkers used aptamer-conjugated gold nanorods to enrich and detect low-abundance proteins from human

plasma.[76] 15-mer thiol-terminated aptamers with high affinity for thrombin, and thiol-terminated PEG were conjugated onto <100 nm gold nanorods. Using this strategy, they were able to enrich 1 ng (10 ppb) of α-thrombin that was spiked into human plasma, and visualize protein with SDS-PAGE, demonstrating substantial enrichment. The authors noted that by having a material where binding can only occur on the surface, it eliminates the necessity of pore penetration, thereby decreasing nonspecific binding.

Iron oxide nanoparticles

Iron oxide nanoparticles (magnetite, Fe_3O_4 NPs) are another major class of nanomaterials that have received considerable attention because they offer unique advantages over other materials for biological applications. They are relatively inexpensive, physically and chemically stable, biocompatible, and are superparamagnetic (below 20 nm) [82-84]. The synthesis of Fe_3O_4 NPs is highly scalable, with researchers reproducibly reporting synthetic procedures generating batches as large as 40 g with yields over > 95%[85]. As iron oxide NPs are typically synthesized in the presence of hydrophobic surfactants, they require an additional surface modification step to make them hydrophilic for biological applications[86]. A common method is to use hydrophilic alkoxysilane derivatives to perform ligand exchange reactions, producing a covalently attached siloxane network (-Fe-O-Si-) on the surface of Fe_3O_4 NPs. Roberts et. al., reported a method for the reproducible, large-scale production (> 200 mg) of silanized NPs for downstream protein enrichment applications. This method is broadly applicable for the surface-modification of Fe_3O_4 NPs with a variety of hydrophilic alkoxysilane derivatives for the attachment of diverse affinity ligands[87]. The small sizes of Fe_3O_4 NPs and their superparamagnetism are also attractive qualities for intact protein enrichment applications prior to MS applications. Hwang et. al., reported the

specific enrichment of phosphoproteins using 6 nm Fe_3O_4 NPs functionalized with a Zn^{2+} affinity ligand that specifically binds to phosphate groups[88]. The small size of the NPs is comparable to the size of proteins, enabling NPs to (i) penetrate better in protein mixtures, leading to a higher binding rate; (ii) reduce the probability of protein denaturation. As Fe_3O_4 NPs below 20 nm are superparamagnetic, they avoid aggregating by themselves without the application of an external magnetic field. This is a useful property with respect to protein enrichment applications, as poor colloidal dispersion of materials can lead to poorer interactions with proteins in aqueous solutions[76]. Using these functionalized materials coupled with top-down MS, several low abundance phosphoproteins (> 8000 Da) that were nearly undetectable in the original swine tissue lysate before enrichment were significantly enriched and detected in the elution after enrichment—enabling a comprehensive characterization of the phosphoproteome by top-down LC/MS.

1.4 LC-MS and LC-MS/MS analysis of proteins

Chromatography-based methods for online protein separation coupled to MS

While enrichment applications largely alleviate the sample complexity and enable analysis of low abundance proteoforms from complex biological mixtures, online LC separations are often utilized to facilitate subsequent comprehensive proteoform characterization by LC/MS and MS/MS. The most common online chromatography separations techniques for intact proteins include reverse phase chromatography (RPC), hydrophobic interaction chromatography (HIC), hydrophilic interaction chromatography (HILIC), ion-exchange chromatography (IEC), and size-exclusion chromatography (SEC).

In RPC, denatured proteins adsorb onto a hydrophobic stationary phase in the initial presence of aqueous solvent and are eluted into the mass spectrometer using increasing amounts of organic solvent. As a result, hydrophilic proteins are eluted first, followed by hydrophobic proteins. The stationary phase consists of silica particles which are modified with different alkyl substituents such as n-butyl (C4), octyl (C8), and phenyl groups, which have different retention characteristics and can be used to optimize protein separations. For positive mode MS, formic acid is often included as a mobile phase additive in both aqueous and organic solvents because it improves chromatographic performance, is volatile and does not cause protein adducts, and assists in ionizing proteins in the gas phase.

Like RPC, HIC separates proteins based on their hydrophobicity. In contrast to RPC, which is a denaturing chromatography mode, HIC can be used to separate proteins while maintaining their tertiary and quaternary structures[89]. In HIC, non-denatured proteins adsorb onto a moderately hydrophobic (propyl and phenyl) stationary phase in the presence of aqueous mobile phases containing high concentrations of salt. Due to the high presence of salt, the solvation of proteins with water is marginal, and proteins adsorb to the stationary phase. As the salt concentration decreases, proteins are rehydrated and allow hydrophilic proteins to elute first, followed by hydrophobic proteins—with high recovery and unusually high sensitivity to conformational variations[90]. HIC requires the use of MS compatible salts such as ammonium acetate that are volatile and do not suppress MS signal[91].

HILIC is most akin to ‘normal-phase’ chromatography, where denatured proteins adsorb onto a polar, hydrophilic stationary phase (ethyl or cyano) in the presence of organic solvents and are eluted into the mass spectrometer using increasing amounts of aqueous solvent. HILIC is useful

for analyzing polar compounds that have poor retention in RPC[92]. However, the use of HILIC for intact proteins has been more limited, as many proteins tend to precipitate using the high concentrations of polar, organic solvents used for binding to HILIC. Many applications have involved hydrophobic membrane proteins[93] or apolipoproteins commonly found in serum/plasma[94].

IEC separates proteins based on their net charge through electrostatic interactions to a stationary phase with the opposite charge[95]. At a pH higher than that corresponding to a protein's isoelectric point (pI), the analyte has a net (-) charge, and is retained by an anion-exchange resin (cationic stationary phase). At a pH lower than that corresponding to a protein's pI, the analyte has a net (+) charge, and is retained by a cation-exchange resin (anionic stationary phase). Elution in ion-exchange involves an increasing gradient in salt concentration, to analyte binding to the stationary phase through electrostatic interactions. Because this separation is based on surface interactions, IEC can have sensitivity towards proteoforms that differ in charge, because it is possible for a protein to have non-uniform distribution of charged residues on its surface[95].

SEC separates proteins based on their size or hydrodynamic diameter, while reducing sample loss due to minimal interactions of proteins with the SEC stationary phase. SEC uses hydrophilic stationary phases with well-defined pore diameters, where smaller proteins are more likely to diffuse into the pores, and larger proteins are less likely to enter. Because proteins do not interact with the SEC stationary phase, elution gradients in SEC are isocratic and can be performed with either denaturing or non-denaturing mobile phases such as 1% formic acid in water (pH 2.0), or 200 mM ammonium acetate (pH 7.0). Effective separations using SEC has enabled MS-detection of high MW proteins up to 223 kDa[96].

Ionization techniques

A mass spectrometer consists of an ionization source, a mass analyzer, and a detector which are operated under high vacuum conditions. Mass spectrometers are designed to generate ions from neutral molecules, accelerating and guiding them in a controlled direction to determine the mass to charge (m/z) ratio of these ions. To ionize large biomacromolecules such as proteins for MS-analysis, matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization have emerged as the most useful and widely used techniques because the ionization process does not fragment proteins and preserves their structural integrity. MALDI normally relies on preparing and mixing a sample in a solid matrix and applying energy from a pulsed laser to ultimately ionize the analyte, generally yielding singly charged protein ions[97]. In contrast, ESI transfers analytes from solution to the gas phase, generating an electrostatically charged aerosol at ambient pressure that is transferred into the high vacuum of a mass analyzer, producing multiply-charged protein ions[98]. Thus, ESI is ideal for mass spectrometry coupled to liquid chromatography applications, and for the characterization of intact proteins.

Mass analyzers

MALDI/ESI ionization sources are often coupled to several types of commercially available, high-resolution mass analyzers, with contemporary hybrid instruments featuring a combination of two or more mass analyzers. Ion trap analyzers are composed of four rod-shaped electrodes, to which direct current (DC) and alternating current (AC) are applied to produce electric fields that confine ions in space. When ions enter the trap, they adopt a secular motion that

is proportional to the amplitude of the AC applied, and the mass of the ion. Increasing the amplitude of the AC current increases secular motion of an ion of a given m/z until it is no longer stable, ejecting it and allowed it to be detected by the mass detector[99]. Quadrupole mass filters also consist of four rod-shaped electrodes, but has quadrupolar DC applied, meaning that rods opposite of each other in space have the same sign (both positive), while the other set of rods opposite of each other has the opposite sign (both negative). This difference in sign affects how ions move in a quadrupole mass filter vs. a linear ion trap—ions in a quadrupole mass filter move from one end to the other partially based on the AC current applied and can be used to 'filter' ions. Thus, m/z is measured as a function of ion stability within the electromagnetic field. A time-of-flight (TOF) mass analyzers measure ions' time of flight, which is intrinsically related to their m/z. As ions are accelerated under an electric field, ions carry the same kinetic energy as those with the same charge; however, the velocity resulting in the different arrival times is dependent on the ion's mass, where heavier ions have lower velocity, and lighter ions have higher velocity[100]. Finally, Fourier-Transform (FT)-based instruments such as the Orbitrap and Ion Cyclotron Resonance (ICR) mass spectrometers, measure m/z as a function of ion frequency during oscillatory motion. The Orbitrap instruments use an electromagnetic field to place the ions in an oscillatory motion, while FTICR instruments use superconducting magnets. The oscillation of ions in the time domain is converted to frequency using the Fourier transform, which enables acquisition of ions' m/z[101, 102]. FTMS offers the highest resolving power and mass accuracy over all the other mass analyzers.

Fragmentation techniques

For proteoform characterization, it is necessary to perform tandem MS (MS/MS) experiments to isolate and fragment the proteoforms of interest to localize PTMs and characterize sequence variants. Common fragmentation methods for MS/MS include collisional activated dissociated (CAD), electron capture dissociation (ECD), and electron transfer dissociation (ETD).

For CAD, protein fragmentation is based on excitation of vibrational modes of the molecules[103]. During this process, a neutral gas (commonly argon, helium, or nitrogen) is present in the collision cell of the mass spectrometer. Inelastic collisions between neutral gas molecules and the proteoform ions of interest, converting the collisional energy to the ions' vibrational energy. The vibrational energy is transferred throughout the entirety of the molecule, resulting in the cleavage of the most labile bonds. In proteins, this leads to the cleavage of the backbone amide bond (-CONHR), resulting in *b* and *y* ions. However, some post-translational modifications, such as phosphorylation or glycosylation, consist of bonds that are more labile than that of the amide backbone, and can potentially be lost during CAD fragmentation—resulting in a loss of PTM information.

To address the shortcomings of CAD techniques, electron-based dissociation methods such as ETD and ECD are useful for preserving labile bonds from PTMs, which enable PTM site characterization. ECD is usually implemented in FTICR instruments, where electrons can be trapped and are stable in the superconducting magnet of the ICR cell. During ECD, multiply charged precursor protein cations capture electrons, forming a charge-reduced species and initiating ion-radical fragmentation reactions[104, 105]. The predominant cleavage is the backbone N-

Cα bond, resulting in a c' and z · ion fragment pair. Thus, electron-based fragmentation techniques are complementary to CID and are useful for proteoform characterization. Similar to ECD, ETD is an ion-ion reaction involving multiply charged precursor protein cations, and singly charged radical reagent anions[106]. An electron from the reagent anion is transferred to the protein cation, resulting in an odd-number electron cation that also results in cleavage of the N-Cα bond in proteins. ETD is accessible on a wide variety of hybrid mass spectrometers, such as quadrupole linear ion traps, quadrupole time-of-flight, and Orbitraps.

Data Analysis

Following acquisition of top-down LC/MS and LC/MS/MS spectra, data analysis tools are required for the deconvolution of protein MS and MS/MS spectra. Deconvolution is the process that combines m/z peaks belonging to the same mass value formed from different charge states, by taking advantage of the fact that charge can only be an integer value (e.g., 17+, 16+, 15+ etc.)[107]. However, spectral deconvolution of protein mass spectra to generate simplified mass lists is challenging, because elemental isotopes (e.g., $^{13}C/^{12}C$) must be considered for intact proteins (the most abundant protein mass observed in a mass spectrum no longer corresponds to the monoisotopic mass of a protein). Thus, a group of isotopic peaks is necessary to calculate the charge state for ions corresponding to a higher molecular weight protein. In top-down proteomics, the Averagine model[108] developed by McLafferty and co-workers, is used to predict the chemical composition of an ion, to estimate how many "heavy" isotopes are present. The peaks in a theoretical envelope are matched to the peaks found in a spectrum to find an experimental isotopic envelope, and a scoring function is applied to assess the fitting. If the theoretical and experimental isotopic envelop fit, a monoisotopic mass can be computed[109]. Following MS1 acquisition of the

monoisotopic mass, MS/MS fragment ions are also deconvoluted to generate mass lists that can be used to match against theoretically generated mass lists calculated from amino acid sequences during protein database searching.

1. Crick, F., Central dogma of molecular biology. *Nature* **1970,** *227* (5258), 561-563.
2. Clamp, M.; Fry, B.; Kamal, M.; Xie, X.; Cuff, J.; Lin, M. F.; Kellis, M.; Lindblad-Toh, K.; Lander, E. S., Distinguishing protein-coding and noncoding genes in the human genome. *Proceedings of the National Academy of Sciences* **2007,** *104* (49), 19428-19433.
3. Smith, L. M.; Kelleher, N. L., Proteoform: a single term describing protein complexity. *Nature methods* **2013,** *10* (3), 186-187.
4. 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.
5. Gregorich, Z. R.; Ge, Y., Top-down proteomics in health and disease: Challenges and opportunities. *Proteomics* **2014,** *14* (10), 1195-1210.
6. Mann, M.; Jensen, O. N., Proteomic analysis of post-translational modifications. *Nature biotechnology* **2003,** *21* (3), 255-261.
7. 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.
8. 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.
9. Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nature methods* **2007,** *4* (10), 817-821.
10. Zhang, J.; Guy, M. J.; Norman, H. S.; Chen, Y.-C.; Xu, Q.; Dong, X.; Guner, H.; Wang, S.; Kohmoto, T.; Young, K. H., 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.
11. 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.
12. 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., 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.
13. Jiang, L.; Smith, J. N.; Anderson, S. L.; Ma, P.; Mizzen, C. A.; Kelleher, N. L., Global assessment of combinatorial post-translational modification of core histones in yeast using contemporary mass spectrometry: LYS4 trimethylation correlates with degree of acetylation on the same H3 tail. *Journal of biological chemistry* **2007,** *282* (38), 27923-27934.
14. Chamot-Rooke, J.; Mikaty, G.; Malosse, C.; Soyer, M.; Dumont, A.; Gault, J.; Imhaus, A.-F.; Martin, P.; Trellet, M.; Clary, G., Posttranslational modification of pili upon cell contact triggers N. meningitidis dissemination. *Science* **2011,** *331* (6018), 778-782.
15. Schaffer, L. V.; Millikin, R. J.; Miller, R. M.; Anderson, L. C.; Fellers, R. T.; Ge, Y.; Kelleher, N. L.; LeDuc, R. D.; Liu, X.; Payne, S. H., Identification and quantification of proteoforms by mass spectrometry. *Proteomics* **2019,** *19* (10), 1800361.

16. White, H. D.; Chew, D. P., Acute myocardial infarction. *The Lancet* **2008,** *372* (9638), 570-584.
17. Thygesen, K.; Alpert, J. S.; Jaffe, A. S.; Chaitman, B. R.; Bax, J. J.; Morrow, D. A.; White, H. D.; null, n., Fourth Universal Definition of Myocardial Infarction (2018). *Circulation* **2018,** *138* (20), e618-e651.
18. Ooi, D. S.; Isotalo, P. A.; Veinot, J. P., Correlation of antemortem serum creatine kinase, creatine kinase-MB, troponin I, and troponin T with cardiac pathology. *Clinical chemistry* **2000,** *46* (3), 338-344.
19. Virmani, R.; Forman, M.; Kolodgie, F., Myocardial reperfusion injury. Histopathological effects of perfluorochemical. *Circulation* **1990,** *81* (3 Suppl), IV57-68.
20. Makam, A. N.; Nguyen, O. K., Use of cardiac biomarker testing in the emergency department. *JAMA internal medicine* **2015,** *175* (1), 67-75.
21. Al-Hadi, H. A.; Fox, K. A., Cardiac markers in the early diagnosis and management of patients with acute coronary syndrome. *Sultan Qaboos University Medical Journal* **2009,** *9* (3), 231.
22. Alvin, M. D.; Jaffe, A. S.; Ziegelstein, R. C.; Trost, J. C., Eliminating creatine kinase–myocardial band testing in suspected acute coronary syndrome: A value-based quality improvement. *JAMA internal medicine* **2017,** *177* (10), 1508-1512.
23. Paana, T.; Jaakkola, S.; Bamberg, K.; Saraste, A.; Tuunainen, E.; Wittfooth, S.; Kallio, P.; Heinonen, O. J.; Knuuti, J.; Pettersson, K., Cardiac troponin elevations in marathon runners. Role of coronary atherosclerosis and skeletal muscle injury. The MaraCat Study. *International journal of cardiology* **2019,** *295*, 25-28.
24. Weil, B. R.; Suzuki, G.; Young, R. F.; Iyer, V.; Canty, J. M., Troponin release and reversible left ventricular dysfunction after transient pressure overload. *Journal of the American college of cardiology* **2018,** *71* (25), 2906-2916.
25. White, H. D., Pathobiology of troponin elevations: do elevations occur with myocardial ischemia as well as necrosis? American College of Cardiology Foundation Washington, DC: 2011; Vol. 57, pp 2406-2408.
26. Mair, J.; Lindahl, B.; Hammarsten, O.; Müller, C.; Giannitsis, E.; Huber, K.; Möckel, M.; Plebani, M.; Thygesen, K.; Jaffe, A. S., How is cardiac troponin released from injured myocardium? *European heart journal: acute cardiovascular care* **2018,** *7* (6), 553-560.
27. Jaffe, A. S.; Wu, A. H., Troponin release—reversible or irreversible injury? Should we care? *Clinical chemistry* **2012,** *58* (1), 148-150.
28. 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.
29. McDonough, J. L.; Van Eyk, J. E., Developing the next generation of cardiac markers: disease-induced modifications of troponin I. *Progress in cardiovascular diseases* **2004,** *47* (3), 207-216.
30. Gordon, A. M.; Homsher, E.; Regnier, M., Regulation of contraction in striated muscle. *Physiological reviews* **2000,** *80* (2), 853-924.
31. Greaser, M. L.; Gergely, J., Purification and properties of the components from troponin. *Journal of Biological Chemistry* **1973,** *248* (6), 2125-2133.

32. Ebashi, S., Separation of troponin into its three components. *The Journal of Biochemistry* **1972,** *72* (3), 787-790.
33. Schaub, M.; Perry, S., The relaxing protein system of striated muscle. Resolution of the troponin complex into inhibitory and calcium ion-sensitizing factors and their relationship to tropomyosin. *Biochemical Journal* **1969,** *115* (5), 993-1004.
34. Metzger, J. M.; Westfall, M. V., Covalent and noncovalent modification of thin filament action: the essential role of troponin in cardiac muscle regulation. *Circulation research* **2004,** *94* (2), 146-158.
35. Perry, S., Troponin T: genetics, properties and function. *Journal of Muscle Research & Cell Motility* **1998,** *19* (6), 575-602.
36. Tobacman, L. S., Thin filament-mediated regulation of cardiac contraction. *Annual review of physiology* **1996,** *58* (1), 447-481.
37. Katrukha, I., Human cardiac troponin complex. Structure and functions. *Biochemistry (Moscow)* **2013,** *78* (13), 1447-1465.
38. 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.
39. Solaro, R. J., The special structure and function of troponin I in regulation of cardiac contraction and relaxation. *Molecular and Cellular Aspects of Muscle Contraction* **2003**, 389-402.
40. Kobayashi, T.; Patrick, S. E.; Kobayashi, M., Ala scanning of the inhibitory region of cardiac troponin I. *Journal of Biological Chemistry* **2009,** *284* (30), 20052-20060.
41. Mittmann, K.; Jaquet, K.; Heilmeyer, L. M., A common motif of two adjacent phosphoserines in bovine, rabbit and human cardiac troponin I. *FEBS letters* **1990,** *273* (1-2), 41-45.
42. van der Velden, J.; Papp, Z.; Zaremba, R.; Boontje, N.; de Jong, J. W.; Owen, V.; Burton, P.; Goldmann, P.; Jaquet, K.; Stienen, G., Increased Ca2+-sensitivity of the contractile apparatus in end-stage human heart failure results from altered phosphorylation of contractile proteins. *Cardiovascular research* **2003,** *57* (1), 37-47.
43. Howarth, J. W.; Meller, J.; Solaro, R. J.; Trewhella, J.; Rosevear, P. R., Phosphorylation-dependent conformational transition of the cardiac specific N-extension of troponin I in cardiac troponin. *Journal of molecular biology* **2007,** *373* (3), 706-722.
44. Zhang, R.; Zhao, J.; Mandveno, A.; Potter, J. D., Cardiac troponin I phosphorylation increases the rate of cardiac muscle relaxation. *Circulation research* **1995,** *76* (6), 1028-1035.
45. Messer, A. E.; Jacques, A. M.; Marston, S. B., Troponin phosphorylation and regulatory function in human heart muscle: dephosphorylation of Ser23/24 on troponin I could account for the contractile defect in end-stage heart failure. *Journal of molecular and cellular cardiology* **2007,** *42* (1), 247-259.
46. Zhang, P.; Kirk, J. A.; Ji, W.; Dos Remedios, C. G.; Kass, D. A.; Van Eyk, J. E.; Murphy, A. M., Multiple reaction monitoring to identify site-specific troponin I phosphorylated residues in the failing human heart. *Circulation* **2012,** *126* (15), 1828-1837.
47. Risnik, V. V.; Gusev, N. B., Some properties of the nucleotide-binding site of troponin T kinase-casein kinase type II from skeletal muscle. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* **1984,** *790* (2), 108-116.

48. Gusev, N. B.; Dobrovolskii, A. B.; Severin, S. E., Isolation and some properties of troponin T kinase from rabbit skeletal muscle. *Biochemical Journal* **1980,** *189* (2), 219-226.
49. Solis, R. S.; 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* (6-8), 203-212.
50. 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.
51. Noland Jr, T. A.; Kuo, J., Protein kinase C phosphorylation of cardiac troponin I and troponin T inhibits Ca2+-stimulated MgATPase activity in reconstituted actomyosin and isolated myofibrils, and decreases actin-myosin interactions. *Journal of molecular and cellular cardiology* **1993,** *25* (1), 53-65.
52. Sumandea, M. P.; Burkart, E. M.; Kobayashi, T.; De Tombe, P. P.; SOlARO, R. J., Molecular and integrated biology of thin filament protein phosphorylation in heart muscle. *Annals of the New York Academy of Sciences* **2004,** *1015* (1), 39-52.
53. 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.
54. Zubarev, R. A., The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **2013,** *13* (5), 723-726.
55. Hortin, G. L.; Sviridov, D., The dynamic range problem in the analysis of the plasma proteome. *Journal of proteomics* **2010,** *73* (3), 629-636.
56. Anderson, N. L.; Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics* **2002,** *1* (11), 845-867.
57. Miles, R. R.; Roberts, R. F.; Putnam, A. R.; Roberts, W. L., Comparison of serum and heparinized plasma samples for measurement of chemistry analytes. *Clinical chemistry* **2004,** *50* (9), 1704-1706.
58. Veenstra, T. D.; Conrads, T. P.; Hood, B. L.; Avellino, A. M.; Ellenbogen, R. G.; Morrison, R. S., Biomarkers: mining the biofluid proteome. *Molecular & Cellular Proteomics* **2005,** *4* (4), 409-418.
59. Echan, L. A.; Tang, H. Y.; Ali-Khan, N.; Lee, K.; Speicher, D. W., Depletion of multiple high-abundance proteins improves protein profiling capacities of human serum and plasma. *Proteomics* **2005,** *5* (13), 3292-3303.
60. Kay, R.; Barton, C.; Ratcliffe, L.; Matharoo-Ball, B.; Brown, P.; Roberts, J.; Teale, P.; Creaser, C., Enrichment of low molecular weight serum proteins using acetonitrile precipitation for mass spectrometry based proteomic analysis. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up-to-the-Minute Research in Mass Spectrometry* **2008,** *22* (20), 3255-3260.
61. Luque-Garcia, J. L.; Neubert, T. A., Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. *Journal of Chromatography A* **2007,** *1153* (1-2), 259-276.
62. Vidarsson, G.; Dekkers, G.; Rispens, T., IgG subclasses and allotypes: from structure to effector functions. *Frontiers in immunology* **2014,** *5*, 520.

63. Adkins, J. N.; Varnum, S. M.; Auberry, K. J.; Moore, R. J.; Angell, N. H.; Smith, R. D.; Springer, D. L.; Pounds, J. G., Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Molecular & Cellular Proteomics* **2002,** *1* (12), 947-955.
64. Björhall, K.; Miliotis, T.; Davidsson, P., Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. *Proteomics* **2005,** *5* (1), 307-317.
65. Scopes, R. K., *Protein purification: principles and practice*. Springer Science & Business Media: 1993.
66. Yocum, A. K.; Yu, K.; Oe, T.; Blair, I. A., Effect of immunoaffinity depletion of human serum during proteomic investigations. *Journal of proteome research* **2005,** *4* (5), 1722-1731.
67. Kuhn, E.; Addona, T.; Keshishian, H.; Burgess, M.; Mani, D.; Lee, R. T.; Sabatine, M. S.; Gerszten, R. E.; Carr, S. A., Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry. *Clinical chemistry* **2009,** *55* (6), 1108-1117.
68. Dunham, W. H.; Mullin, M.; Gingras, A. C., Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *Proteomics* **2012,** *12* (10), 1576-1590.
69. Roque, A. C. A.; Lowe, C. R., Affinity Chromatography. *Affinity chromatography* **2008**, 1-23.
70. Bantscheff, M.; Eberhard, D.; Abraham, Y.; Bastuck, S.; Boesche, M.; Hobson, S.; Mathieson, T.; Perrin, J.; Raida, M.; Rau, C., Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nature biotechnology* **2007,** *25* (9), 1035-1044.
71. Kool, J.; Jonker, N.; Irth, H.; Niessen, W. M., Studying protein–protein affinity and immobilized ligand–protein affinity interactions using MS-based methods. *Analytical and bioanalytical chemistry* **2011,** *401* (4), 1109-1125.
72. Slobodin, B.; Gerst, J. E., RaPID: an aptamer-based mRNA affinity purification technique for the identification of RNA and protein factors present in ribonucleoprotein complexes. In *RNA Detection and Visualization*, Springer: 2011; pp 387-406.
73. Lambert, J. P.; Pawson, T.; Gingras, A. C., Mapping physical interactions within chromatin by proteomic approaches. *Proteomics* **2012,** *12* (10), 1609-1622.
74. Harlow, E.; Lane, D., A laboratory manual. *New York: Cold Spring Harbor Laboratory* **1988,** *579*.
75. Fleminger, G.; Solomon, B.; Wolf, T.; Hadas, E., Effect of polyethylene glycol on the non-specific adsorption of proteins to Eupergit C and agarose. *Journal of Chromatography A* **1990,** *510*, 271-279.
76. Yasun, E.; Gulbakan, B.; Ocsoy, I.; Yuan, Q.; Shukoor, M. I.; Li, C.; Tan, W., Enrichment and Detection of Rare Proteins with Aptamer-Conjugated Gold Nanorods. *Analytical Chemistry* **2012,** *84* (14), 6008-6015.
77. Ocsoy, I.; Gulbakan, B.; Shukoor, M. I.; Xiong, X.; Chen, T.; Powell, D. H.; Tan, W., Aptamer-conjugated multifunctional nanoflowers as a platform for targeting, capture, and detection in laser desorption ionization mass spectrometry. *ACS nano* **2013,** *7* (1), 417-427.
78. Dreaden, E. C.; Alkilany, A. M.; Huang, X.; Murphy, C. J.; El-Sayed, M. A., The golden age: gold nanoparticles for biomedicine. *Chemical Society Reviews* **2012,** *41* (7), 2740-2779.

79. Teng, C.-H.; Ho, K.-C.; Lin, Y.-S.; Chen, Y.-C., Gold nanoparticles as selective and concentrating probes for samples in MALDI MS analysis. *Analytical chemistry* **2004,** *76* (15), 4337-4342.
80. Elghanian, R.; Storhoff, J. J.; Mucic, R. C.; Letsinger, R. L.; Mirkin, C. A., Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. *Science* **1997,** *277* (5329), 1078-1081.
81. Wang, A.; Wu, C.-J.; Chen, S.-H., Gold Nanoparticle-Assisted Protein Enrichment and Electroelution for Biological Samples Containing Low Protein Concentration A Prelude of Gel Electrophoresis. *Journal of proteome research* **2006,** *5* (6), 1488-1492.
82. Arias, L. S.; Pessan, J. P.; Vieira, A. P. M.; Lima, T. M. T. d.; Delbem, A. C. B.; Monteiro, D. R., Iron oxide nanoparticles for biomedical applications: a perspective on synthesis, drugs, antimicrobial activity, and toxicity. *Antibiotics* **2018,** *7* (2), 46.
83. Wu, W.; Wu, Z.; Yu, T.; Jiang, C.; Kim, W.-S., Recent progress on magnetic iron oxide nanoparticles: synthesis, surface functional strategies and biomedical applications. *Science and technology of advanced materials* **2015**.
84. Sun, S., Recent advances in chemical synthesis, self-assembly, and applications of FePt nanoparticles. *Advanced Materials* **2006,** *18* (4), 393-403.
85. Park, J.; An, K.; Hwang, Y.; Park, J.-G.; Noh, H.-J.; Kim, J.-Y.; Park, J.-H.; Hwang, N.-M.; Hyeon, T., Ultra-large-scale syntheses of monodisperse nanocrystals. *Nature materials* **2004,** *3* (12), 891-895.
86. De Palma, R.; Peeters, S.; Van Bael, M. J.; Van den Rul, H.; Bonroy, K.; Laureyn, W.; Mullens, J.; Borghs, G.; Maes, G., Silane ligand exchange to make hydrophobic superparamagnetic nanoparticles water-dispersible. *Chemistry of Materials* **2007,** *19* (7), 1821-1831.
87. Roberts, D. S.; Chen, B.; Tiambeng, T. N.; Wu, Z.; Ge, Y.; Jin, S., Reproducible large-scale synthesis of surface silanized nanoparticles as an enabling nanoproteomics platform: Enrichment of the human heart phosphoproteome. *Nano research* **2019,** *12* (6), 1473.
88. Hwang, L.; Ayaz-Guner, S.; Gregorich, Z. R.; Cai, W.; Valeja, S. G.; Jin, S.; Ge, Y., Specific enrichment of phosphoproteins using functionalized multivalent nanoparticles. *Journal of the American Chemical Society* **2015,** *137* (7), 2432-2435.
89. 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.
90. Queiroz, J.; Tomaz, C.; Cabral, J., Hydrophobic interaction chromatography of proteins. *Journal of biotechnology* **2001,** *87* (2), 143-159.
91. Heck, A. J., Native mass spectrometry: a bridge between interactomics and structural biology. *Nature methods* **2008,** *5* (11), 927-933.
92. Buszewski, B.; Noga, S., Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. *Analytical and bioanalytical chemistry* **2012,** *402* (1), 231-247.
93. Carroll, J.; Fearnley, I. M.; Walker, J. E., Definition of the mitochondrial proteome by measurement of molecular masses of membrane proteins. *Proceedings of the National Academy of Sciences* **2006,** *103* (44), 16170-16175.

94. Tetaz, T.; Detzner, S.; Friedlein, A.; Molitor, B.; Mary, J.-L., Hydrophilic interaction chromatography of intact, soluble proteins. *Journal of Chromatography A* **2011,** *1218* (35), 5892-5896.
95. Alpert, A. J., Protein fractionation and enrichment prior to proteomics sample preparation. *Modern Proteomics–Sample Preparation, Analysis and Practical Applications* **2016**, 23-41.
96. 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.
97. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **1988,** *60* (20), 2299-2301.
98. 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.
99. Savaryn, J. P.; Toby, T. K.; Kelleher, N. L., A researcher's guide to mass spectrometry-based proteomics. *Proteomics* **2016,** *16* (18), 2435-2443.
100. Guilhaus, M.; Selby, D.; Mlynski, V., Orthogonal acceleration time-of-flight mass spectrometry. *Mass spectrometry reviews* **2000,** *19* (2), 65-107.
101. 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.
102. Amster, I. J., Fourier transform mass spectrometry. *Journal of mass spectrometry* **1996,** *31* (12), 1325-1337.
103. McLafferty, F.; Bente, P.; Kornfeld, R.; Tsai, S.-C.; Howe, I., Metastable ion characteristics. XXII. Collisional activation spectra of organic ions. *Journal of the American Chemical Society* **1973,** *95* (7), 2120-2129.
104. Qi, Y.; Volmer, D. A., Electron-based fragmentation methods in mass spectrometry: an overview. *Mass spectrometry reviews* **2017,** *36* (1), 4-15.
105. Zubarev, R. A., Electron-capture dissociation tandem mass spectrometry. *Current opinion in biotechnology* **2004,** *15* (1), 12-16.
106. Riley, N. M.; Coon, J. J., The role of electron transfer dissociation in modern proteomics. *Analytical chemistry* **2018,** *90* (1), 40-64.
107. 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.
108. Senko, M. W.; Beu, S. C.; McLaffertycor, F. W., Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. *Journal of the American Society for Mass Spectrometry* **1995,** *6* (4), 229-233.
109. Basharat, A. R.; Ning, X.; Liu, X., EnvCNN: A Convolutional Neural Network Model for Evaluating Isotopic Envelopes in Top-Down Mass-Spectral Deconvolution. *Analytical chemistry* **2020,** *92* (11), 7778-7785.

Chapter 2

Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum

Adapted from: Timothy N. Tiambeng[1,†], David S. Roberts[1,†], Kyle A. Brown[1], Yanlong Zhu[2,3], Bifan Chen[1], Zhijie Wu[1], Stanford D. Mitchell[3,4], Tania M. Guardado-Alvarez[1], Song Jin[1]*, Ying Ge[1,2,3]*, *Nat. Commun*. **2020**, 11, 3903.

[†]These authors contributed equally to this work.

Abstract

Top-down mass spectrometry (MS)-based proteomics provides a comprehensive analysis of proteoforms to achieve a proteome-wide understanding of protein functions. However, the MS detection of low-abundance proteins from blood remains an unsolved challenge due to the extraordinary dynamic range of the blood proteome. Here we develop an integrated nanoproteomics method coupling peptide-functionalized superparamagnetic nanoparticles (NPs) with top-down MS for the enrichment and comprehensive analysis of cardiac troponin I (cTnI), a gold-standard cardiac biomarker, directly from serum. These NPs enable the sensitive enrichment of cTnI (< 1 ng/mL) with high specificity and reproducibility, while simultaneously depleting highly abundant proteins such as human serum albumin (>10^{10} more abundant than cTnI). We demonstrate that top-down nanoproteomics can provide high-resolution proteoform-resolved molecular fingerprints of diverse cTnI proteoforms to establish proteoform-pathophysiology relationships. This scalable and reproducible antibody-free strategy can generally enable the proteoform-resolved analysis of low-abundance proteins directly from serum to reveal previously unachievable molecular details.

Introduction

The structure, activity, and function of proteins are modulated by post-translational modifications (PTMs), making the comprehensive analysis of proteoforms[1] — the various protein products arising from myriad PTMs and splicing isoforms from a single gene — crucial for understanding all biological systems at a functional level[2-4]. Mapping the proteoform landscape with absolute molecular specificity represents the next stage of proteomics to achieve proteome-wide characterization and understanding of protein function[3-6]. Top-down MS is the most powerful technology to analyze proteoforms to understand the biological variations that underlie complex phenotypes[7]. However, the MS detection of low-abundance proteins remains an unsolved challenge due to high dynamic range of the human proteome[8, 9].

In particular, the human blood proteome, a circulating representation of all pathophysiological processes, is immensely complex and the huge dynamic range of the blood proteome (10^{12}) dominated by the presence of numerous highly abundant proteins, such as human serum albumin (HSA) makes the detection of low-abundance proteins extremely difficult[9, 10]. Front-end protein enrichment strategies are therefore required to specifically capture and enrich low-abundance proteins from the complex biological milieu prior to MS analysis[9, 11]. Antibodies have so far been the dominating affinity reagents for protein capture and quantification in biological research[12, 13]. However, antibody-based platforms suffer from significant limitations including the batch-to-batch variability, high cost of the antibody production, and relatively low chemical stability[14, 15]. Moreover, the intrinsic challenge of designing and evaluating epitope-specific antibodies for PTMs makes it difficult to develop robust antibodies suitable for proteoform

analysis[16]. Hence, there is an urgent need to develop the next generation of affinity platforms capable of global proteoform capture[12, 13, 17]. Nanoparticles (NPs) are highly effective for such sensitive and specific proteoform enrichment because: (1) they are commensurate in size and diffusion kinetics with proteins allowing effective penetration through complex biological mixtures; (2) they have high surface-to-volume ratios to enhance protein interaction; (3) they are versatile scaffolds to couple diverse affinity ligands for protein binding and capture[18-21].

Cardiac troponin I (cTnI) is a gold-standard biomarker for cardiovascular diseases (the leading cause of death worldwide)[22-25]. cTnI forms the inhibitory subunit of the cTn complex (cTnI-cTnT-TnC; Supplementary Fig. 2.1) and is released into the bloodstream following cardiac injury[23-25] where it circulates with low-abundance (typically < 50 ng/mL)[26] and in myriad proteoforms (e.g., phosphorylated, acetylated, oxidized, truncated), making detection and analysis extremely challenging. Recent studies have shown that cTnI is heavily modified and its proteoforms arising from various PTMs can provide new insights to the molecular mechanisms underlying cardiovascular diseases[27-31]. The PTM profiles of cTnI can function as molecular fingerprints of cellular signaling pathway activity with the potential to serve as the next generation biomarkers[27, 32, 33]. Therefore, a comprehensive proteoform-resolved cTnI analysis that can detect cTnI in blood with detailed PTM information is urgently needed[27].

Here we report an antibody-free top-down nanoproteomics approach enabled by the integration of nanotechnology and top-down proteomics to specifically enrich low-abundance proteins directly from human serum using surface functionalized NPs and comprehensively analyze the enriched proteoforms by top-down MS. We design an organosilane surface-

functionalization molecule bearing a cysteine-thiol reactive handle for the robust synthesis of peptide-functionalized superparamagnetic NPs to capture and specifically enrich low-abundance proteins from human serum. We focus on the development of a top-down nanoproteomics strategy for comprehensive proteoform-resolved analysis of low-abundance cTnI in the blood. We demonstrate that top-down nanoproteomics can provide a high-resolution proteoform-resolved landscape of diverse cTnI molecular fingerprints arising from various PTMs to establish proteoform-pathophysiology relationships.

Results

Synthesis and characterization of peptide-functionalized NPs. We first rationally designed surface-functionalized superparamagnetic iron oxide (magnetite, Fe_3O_4) NPs to specifically enrich cTnI from complex mixtures (Fig. 2.1). We synthesized an organosilane linker molecule N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide (Fig. 2.1a, Supplementary Fig. 2.2-2.6), hereinafter referred to as BAPTES, which we use to silanize the oleic acid coated Fe_3O_4 NPs[34] following a method for reproducible surface silanization[35]. The terminal allene carboxamide functional group of the BAPTES ligand possesses high chemoselectivity towards the thiol sidechain of cysteine (Cys), and forms a stable and irreversible conjugate not prone to hydrolysis[36]. Importantly, instead of using antibodies to target cTnI, we chose a short, linear peptide evolved for high cTnI affinity by phage display and *in silico* evolution[37]. The use of peptides offers significant advantages for protein enrichment in comparison with antibodies, such as improved chemical stability to changes in pH and reducing environments, thermal stability, scalability using solid-phase peptide synthesis, and batch-to-batch reproducibility. This specific peptide (HWQIAYNEHQWQ) not only exhibits

an impressive binding affinity (K_d) of 270 pM comparable to that of antibodies[37], but also targets the central portion of cTnI (amino acid residues 114-144) with less susceptibility to proteolysis and is postulated to be an optimal targeting epitope to detect all forms of cTnI present in blood circulation[38-40]. To evaluate this peptide functionalization approach, we first analyzed the reaction of BAPTES with a C-terminal Cys-modified derivative of the high affinity peptide (HWQIAYNEHQWQC) using high-resolution tandem MS (MS/MS). The peptide-Cys reaction with BAPTES occurred exclusively even in the presence of other biologically relevant nucleophiles, such as hydroxyls, amines, and carboxylates (Supplementary Fig. 2.7-2.8). It should be noted that such allene carboxamide chemistry relies on the presence of a Cys nonessential to the peptide binding sequence. In the case where such a Cys is not available, alternative bioorthogonal coupling approaches, such as azides and cyclooctynes, can be employed[41, 42].

After confirming such allene carboxamide coupling chemistry, we functionalized the BAPTES silanized NPs (NP-BAPTES) with the high affinity cTnI binding-peptide onto the NP surface (Fig. 2.1b), which are hereafter referred to as NP-Pep. Transmission electron microscopy (TEM) revealed the uniformity and monodispersity of as-synthesized Fe_3O_4-oleic acid (Fe_3O_4-OA) NPs with an average diameter of 8.0 ± 0.3 nm (Fig. 2.1c) and confirmed the morphology did not significantly change after BAPTES silanization (Fig. 2.1d) and further peptide surface functionalization (Fig. 2.1e). Physicochemical properties of the surface functionalized NPs measured at each reaction step confirmed the proper functionalization and elucidated the surface properties of the NPs. Fourier transform infrared spectroscopy (FTIR) analysis of the NP-BAPTES revealed strong peak intensities at 1970 and 1947 cm^{-1} characteristic of the allene-containing molecules (C=C=C)[43] (Fig. 2.1f). Moreover, such IR signatures of as-synthesized NP-BAPTES

were consistent and reproducible across different synthesis batches (Supplementary Fig. 2.9). High-resolution magic angle spinning nuclear magnetic spectroscopy (HRMAS-NMR) confirmed the successful silanization of the NPs with BAPTES (Supplementary Fig. 2.10). After peptide coupling, the relative intensities of the characteristic allene IR peaks were reduced (Fig. 2.1f), indicating the successful consumption of the allene groups upon peptide conjugation. Thermogravimetric analysis (TGA) of the NP-BAPTES revealed a 26% weight loss, accounting for a thin silane coating on the surface functionalized NPs (Fig. 2.1g). Comparatively, an increased measured weight loss percentage (32%) was observed after peptide coupling, which suggests the successful attachment of the 13-mer cTnI affinity peptide. From the difference in weight loss (~ 6%) between the NP-BAPTES and the final NP-Pep, a surface density of ~ 0.034 μmol peptide/mg NP was inferred (Supplementary Tables 2.1-2.3). To demonstrate the colloidal stability of the NP-Pep, we determined the zeta potential (□-potential) of the NP-Pep suspended in 0.1x PBS buffer (pH 7.4) to be ~ -38 mV, which has been previously shown to be both ideal for serum protein applications and sufficient for electrostatic repulsive forces to dominate over the van der Waals force, such that agglomeration is suppressed (Supplementary Fig. 2.11)[44, 45]. Furthermore, photographs of the NPs dispersed in a biphasic mixture of dichloromethane (CH_2Cl_2) and water illustrate the drastic change in NP-solvent compatibility before and after successful conjugation with the hydrophilic cTnI binding peptide (Fig. 2.1h).

Reproducible and faithful recapitulation of cTnI proteoforms. We then evaluated the cTnI enrichment performance of the NP-Pep using sarcomeric protein extracts prepared from human heart tissues (Supplementary Fig. 2.1, detailed conditions shown in Supplementary Table 2.4), which were found to contain approximately 0.3% cTnI by ELISA. We first used sodium dodecyl

sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to visualize the protein components in the sarcomeric extracts by SYPRO Ruby staining to confirm the reproducibility of the protein extraction procedure (Supplementary Fig. 2.12). The cTnI enrichment experiments using the functionalized NP-Pep (Fig. 2.2a) proceed as following: (1) incubating the NP-Pep in the protein mixture, (2) magnetically isolating the NP-Pep to wash and remove unbound nonspecific proteins, and (3) eluting bound cTnI off of the NP-Pep using an acidic buffer to disrupt the intermolecular interactions between NP-Pep and the bound cTnI (see experimental details in the Methods). After NP enrichment, the protein bands corresponding to cTnI and cTnT were far more prominent in the elution solutions (E) compared to the initial loading mixtures (L) which contained abundant sarcomeric proteins such as actin (lower panel, Fig. 2.2b). We found that salt (NaCl) concentration of the wash buffer was a critical parameter and could be tuned to promote effective cTnI enrichment using the NP-Pep (Supplementary Fig. 2.13). Subsequently, we demonstrated the highly effective and reproducible enrichment of cTnI from sarcomeric extracts by the NP-Pep from both three different intra- and inter-batch syntheses (Fig. 2.2b, Supplementary Figs. 2.14, 2.15). We further investigated whether the high affinity cTnI-binding peptide functionalization onto the NPs was a critical factor for the high specificity enrichment by using the following NP controls: (1) NP-BAPTES with no peptide, denoted as NP-Ctrl; (2) NP-BAPTES functionalized with peptide at an acidic pH to inhibit formation of the covalent Cys-thiol conjugate; (3) NP-BAPTES functionalized with a negative-control peptide[37] to reduce cTnI-binding affinity. All of these control experiments showed no appreciable cTnI enrichment (Supplementary Fig. 2.16). This confirms only NPs properly functionalized with the high affinity cTnI-binding peptide (NP-Pep) allow for successful cTnI enrichment.

We next evaluated the cTnI enrichment performance of the NP-Pep by top-down liquid chromatography (LC)/MS analysis of the initial protein loading mixture, the resulting flow-through, and the final elution mixtures after cTnI enrichment. Top-down MS provided a bird's eye view of all cTnI proteoforms present for direct quantification of the relative abundance of individual cTnI proteoform normalized to the total cTnI proteoforms[32, 46]. Importantly, the NP-Pep preserved all endogenous cTnI proteoform distributions and faithfully retained the endogenous cTnI PTM profiles at every step of the enrichment process with no artifactual modifications (Fig. 2.2c). Moreover, the NP-Pep enriched cTnI over 3-fold ($p < 0.0001$) relative to the loading mixture (Fig. 2.2d, Supplementary Fig. 2.14b-c). High-resolution top-down MS analysis of the intact cTnI proteoforms demonstrated the high reproducibility of this process (Fig. 2.2e, Supplementary Figs. 2.14d, 2.15) and directly revealed all endogenous cTnI proteoforms present in the donor heart sample: bisphosphorylated cTnI with C-terminal truncation (ppcTnI[1-207]), unphosphorylated cTnI, monophosphorylated cTnI (pcTnI), and bisphosphorylated cTnI (ppcTnI) (Supplementary Table 2.5).

To illustrate the unique advantages of the NP-Pep system for cTnI enrichment over the existing immunoaffinity purification systems, we next compared the cTnI enrichment performance of the NPs against a conventional agarose-based solid support functionalized with the same high affinity cTnI-binding peptide (Agarose-Pep) and an anti-cTnI monoclonal antibody (M46; Agarose-mAb) possessing a similar cTnI-binding epitope (amino acids 130 to 145) to the high affinity peptide (Fig. 2.2f-h, Supplementary Fig. 2.17, details in the Methods). This cTnI enrichment comparison corrects for the bias due to affinity ligand and allows direct assessment of cTnI enrichment performance as a result of the choice of materials system. We evaluated the cTnI

enrichment performance of the NP and Agarose platforms from sarcomeric extracts containing low amounts of cTnI (< 700 ng) obtained from three different human heart samples: a donor heart with normal cardiac function (Donor), a heart with dilated cardiomyopathy (Diseased), and a post-mortem heart with no known history of cardiac dysfunction (Post-mortem). SDS-PAGE analysis revealed an intense cTnI band present in the NP-Pep elution mixtures (upper panels of Fig. 2.2f-h), visually demonstrating effective cTnI enrichment by the NP-Pep across each of the three human heart samples. Comparatively, the Agarose-Pep and Agarose-mAb showed reduced cTnI enrichment in the same SDS-PAGE (upper panels of Fig. 2.2f-h, Supplementary Fig. 2.17). These results highlight the unique advantages that these functionalized NPs possess for protein enrichment over antibody-based platforms: a) comparable size and diffusion kinetics to proteins, promoting effective capture of low abundance proteins within complex mixtures[20, 47, 48]; b) high surface area per volume and high number of flexible binding sites, contributing to high ligand density and binding efficacy;[49, 50] c) inexpensive, simple, fast, and scalable synthesis with minimal batch-to-batch variability,[35] in contrast to antibodies[14, 15, 51-53].

Furthermore, top-down LC/MS coupling reversed-phase liquid chromatography (RPLC) to high-resolution MS (see Methods) revealed the molecular details of the endogenous cTnI proteoforms found in the loading mixture (L), flow-through (F), and elution (E) for each heart sample (lower panels of Fig. 2.2f-h). We were able to unambiguously identify and characterize the various cTnI proteoforms inherent to each of the heart samples by accurate mass measurements (lower panels of Fig, 2.2f-h**,** Supplementary Table 2.6). We found ppcTnI to be the predominant cTnI proteoform in the non-diseased donor heart (Fig. 2.2f) whereas non-phosphorylated cTnI was found to be the most abundant cTnI proteoform in the dilated cardiomyopathy heart (Fig. 2.2g),

consistent with previous reports suggesting phosphorylation of cTnI as a potential biomarker for heart diseases[27, 31]. We also identified and characterized C-terminal truncated cTnI proteoforms found in the post-mortem samples (Fig. 2.2h), which is consistent with our previous report[32]. Despite the rich diversity of cTnI proteoforms between the various heart samples, the NP-Pep was capable of faithfully retaining the endogenous cTnI proteoform distribution initially present in the donor (Fig. 2.2f), dilated cardiomyopathy (Fig. 2.2g), and post-mortem (Fig. 2.2h) heart samples both before and after enrichment (Supplementary Fig. 2.18). Collectively, these results suggest NP-Pep can be used as an effective antibody-replacement for cTnI enrichment.

Enrichment of low-abundance proteoforms from human serum. Critically, MS-based detection of low-abundance cTnI proteoforms directly in the blood remains an unsolved challenge[27] due to the exceedingly complex and large dynamic range (~10^{12}) of the blood proteome[10] combined with the presence of highly abundant blood proteins such as HSA (~60% of total blood mass and >10 orders of magnitude in abundance than circulating cTnI)[9, 10]. To this end, we performed cTnI enrichment from human serum spiked-in with the cTnI extracted from the same donor, dilated cardiomyopathy, and post-mortem heart samples (*vide supra*). The endogenous cTnI obtained from different human heart samples were used as the reference cTnI sources and served to mimic circulating cTnI found in the blood. SDS-PAGE analysis revealed the striking contrast in cTnI enrichment performance and nonspecific blood protein resistance between the NP-Pep *versus* the conventional agarose platform (Fig. 2.3a). Importantly, the NP-Pep demonstrated impressive resistance to nonspecific adsorption from the highly abundant HSA in human serum, in contrast to the Agarose-Pep and Agarose-mAb. Although cTnI was undetectable in the original serum mixture due to its low abundance compared to the higher-

abundance blood proteins, the NP-Pep elution fraction showed significant cTnI enrichment compared to the original serum mixture (Fig. 2.3a) and nearly complete depletion of high-abundance proteins such as HSA (Supplementary Fig. 2.19). SDS-PAGE analysis of cTnI enrichment from human serum spiked with dilated cardiomyopathy and post-mortem cTnI-containing sarcomeric extracts also revealed consistent HSA depletion by the NP-Pep, in contrast with the Agarose-Pep and Agarose-mAb (Supplementary Fig. 2.20). This is further demonstrated by analysis of extracted ion chromatograms (EICs) corresponding to HSA and cTnI from the same serum samples in both the NP-Pep and the agarose platforms (Fig. 2.3b, Supplementary Figs. 2.21**a, 2.2**2a).

Unlike the NP-Pep, the agarose platforms retained a large amount of HSA in the elution fraction, which likely limited its enrichment efficacy for cTnI. Importantly, the deconvoluted top-down mass spectra of the serum spike-in cTnI samples illustrate not only did the NP-Pep enrich bisphosphorylated cTnI (ppcTnI), but also it was able to capture unphosphorylated cTnI and monophosphorylated cTnI (pcTnI) present in lower abundance from the non-diseased donor heart (Fig. 2.3c top panel). In contrast, minimal cTnI proteoforms were detected after enrichment by Agarose-Pep and Agarose-mAb, respectively (Fig. 3c bottom panels, Supplementary Figs. 2.21b, 2.22b). Meanwhile, the deconvoluted top-down mass spectra for HSA (Fig. 2.3d, Supplementary Figs. 2.21c, 2.22c) further confirmed the significant depletion of HSA by the NP-Pep compared to the Agarose-Pep and Agarose-mAb, where HSA remains in high abundance. These serum cTnI enrichment results illustrate the major benefits of functionalized NPs compared to current antibody-based approaches: (1) NPs are capable of reproducibly and effectively enriching cTnI (Supplementary Fig. 2.23) across all tested serum samples without the use of antibody, thus

overcoming the intrinsic antibody-related limitations (*vide supra*)[14, 15, 51-53]. (2) NPs retain all endogenous cTnI proteoforms from serum without introducing artifactual modification. (3) NPs show highly reproducible and effective depletion of HSA concurrently with cTnI enrichment. Although the NP-Pep demonstrates highly effective HSA depletion, there are some nonspecific proteins still retained from the serum enrichment (Fig. 2.3a and Supplementary Fig. 2.20). To investigate these coeluted proteins, we performed a detailed analysis of all top-down LC/MS proteins identified in the NP-Pep serum elution mixtures (Supplementary Fig. 2.24). cTnI was demonstrated to be confidently identified and consistently captured in all NP-Pep serum enrichment trials (Supplementary Fig. 2.24).

Following the determination of the absolute cTnI concentration using an enzyme-linked immunosorbent assay (ELISA) (Supplementary Fig. 2.25), we calculated the limit of detection (LOD) for cTnI using top-down MS by injecting controlled amounts of purified cTnI (3.1 ng to 0.006 ng) from the donor heart sample used in Fig. 2.2 and Fig. 2.3a-d. Top-down RPLC/MS with a CaptiveSpray (CS) ionization source fitted to a maXis II ETD mass spectrometer was sufficiently sensitive to detect cTnI with a LOD (3.3 σ/s) as low as 0.06 ng/mL (Fig. 2.3e, Supplementary Fig. 2.26). After demonstrating the capability of top-down MS for accurate detection and analysis of cTnI at low concentrations, we next evaluated the detection sensitivity of the current NP-Pep platform for cTnI enrichment from human serum. Notably, this integrated nanoproteomics method with specific enrichment of cTnI using the NP-Pep followed by top-down MS analysis enabled the detection of cTnI proteoforms from human serum at with a low LOD of 0.75 ng/mL (Fig. 2.3f, g, Supplementary Fig. 2.27). Note that the estimated levels of cTnI released after cardiac injury events such as AMI are typically in the range of 0.5 – 50 ng/mL[25, 27]. Hence, this nanoproteomics

approach enables analysis of cTnI proteoforms from human serum at clinically relevant concentrations. Additionally, we evaluated the serum cTnI enrichment performance of the NP-Pep compared to the Agarose-mAb by ELISA detection of the cTnI amount before and after enrichment (Supplementary Fig. 2.28). The NP-Pep demonstrated a high cTnI enrichment factor (115-fold) and the serum cTnI percent recovery of the NP-Pep is ~ 51% (3-fold higher than the Agarose-mAb, ~ 17%). We believe that the cTnI percent recovery can be further improved with future optimizations in automating the NP-Pep enrichment workflow to reduce sample handling and transfer steps which may result in unnecessary protein loss. Furthermore, additional instrumentation improvements in top-down MS[8] will further improve the LOD of the platform toward the diagnostic cutoff value used by contemporary cTnI ELISA (≤ 0.04 ng/mL).

Protoform-resolved analysis of cTnI from human serum. Finally, we applied this nanoproteomics method to human serum with a minimal amount of spike-in cTnI originating from various non-diseased donor, diseased, or post-mortem heart samples (Supplementary Table 2.7) to mimic the diversity of circulating cTnI proteoforms present in patients' blood (Fig. 2.4). The effective cTnI enrichment by the NP-Pep allowed for the subsequent comprehensive analysis of cTnI proteoforms using top-down LC/MS (Fig. 2.4a) to reveal unique cTnI proteoform fingerprints in each human serum sample representing distinct pathophysiology (Fig. 2.4b, Supplementary Fig. 2.29). The deconvoluted mass spectra corresponding to cTnI proteoforms enriched from human serum samples revealed that the endogenous cTnI was primarily in its phosphorylated (mono- and bis-phosphorylated) state in the two donor samples (Fig. 2.4b; panels i and ii), whereas drastic decrease of phosphorylated cTnI was detected in the two diseased samples with dilated cardiomyopathy (Fig. 2.4b; panels iii and iv). Moreover, cTnI degradation due to c-terminal

proteolysis was clearly seen in the two post-mortem samples (Fig. 2.4b; v, vi). Extensive modifications including phosphorylation and degradation of cTnI have previously been observed in serum samples from patients with AMI[28]. Tandem MS/MS analysis of the detected serum cTnI proteoforms were used to validate proteoform assignments across the various heart pathologies (Supplementary Fig. 2.30). Notably, cTnI phosphorylation is considered as an integral mechanism in cardiac homeostasis and has potential to reveal the functional status of the heart[54], whereas selective cTnI proteolytic degradation contributes to cardiac dysfunction in ischemia/reperfusion injury[29]. Altered PTM profiles of cTnI are associated with dysregulated cellular signaling during the onset and progression of diseases, thus disease-induced cTnI proteoforms are believed to have the potential to serve as the next generation cardiac biomarkers for diagnosis of specific cardiovascular syndromes[27, 32, 33].

Here, we demonstrated that NP-Pep enables highly effective and unbiased cTnI proteoform enrichment across all tested serum samples with the faithful recapitulation of cTnI PTM profiles (Supplementary Fig. 2.31) despite the rich diversity of cTnI proteoforms present in each of these cardiac samples (Supplementary Table 2.8). The relative abundance of cTnI proteoforms present in the initial sarcomeric extracts did not significantly change after serum spike-in and incubation in pooled human serum followed by subsequent NP-Pep enrichment (Supplementary Fig. 2.31). The concurrent enrichment of cTnI and depletion of HSA together with the reliable preservation of the cTnI proteoform fingerprints during the NP enrichment solves the major challenge in the MS detection of cTnI proteoforms directly from the blood[27] and will enable the application of this top-down nanoproteomics strategy to clinical samples to eventually establish a proteoform-resolved cTnI assay.

Conclusion

Here we develop a proteoform-resolved method for the analysis of low-abundance proteins directly from serum enabled by top-down nanoproteomics. We demonstrate that carefully designed peptide-functionalized NPs can directly capture and enrich cTnI from human serum with high specificity and reproducibility, while simultaneously depleting highly abundant blood proteins, such as HSA. These NPs not only outperform conventional monoclonal antibody platforms for serum cTnI enrichment, but also faithfully and holistically preserve all endogenous cTnI proteoforms. Thus, these NPs can serve as replacements to conventional immuno-based techniques to overcome the antibody-related limitations in cTnI immunoassays and potentially address the current reproducibility crisis caused by antibodies in general[14-16, 51, 53]. This antibody-free approach can be leveraged in future clinical cTnI diagnostic assays. By further applying to a large human cohort, patient blood samples can be analyzed to comprehensively detect all cTnI proteoforms and establish the relationships between cTnI proteoforms and underlying disease etiology[27]. Ultimately, this nanoproteomics strategy could enable next-generation precision medicine approaches for comprehensive cTnI analysis toward accurate diagnosis, better risk stratification, and improved outcome assessment of patients presenting with various cardiovascular syndromes.

Beyond cTnI, we expect that this scalable and reproducible top-down nanoproteomics approach can be generally applied to other low-abundance plasma/serum proteins of interest, such as cytokines or growth factors, if the NPs can be functionalized with a suitable affinity reagent[55, 56]. With the exciting recent advances in phage display libraries[57], *in silico* techniques[37], and designing aptamers/affimers largely owing to advancements in high-throughput methods for systemic evolution of ligands by exponential enrichment (SELEX)[58, 59], the design and selection

of effective affinity reagents for targeted protein analysis is more accessible than ever before. This nanoproteomics strategy is capable of providing previously unachievable molecular details of low-abundance serum proteins in general and can serve as an enabling technology to comprehensively map the proteoform landscape.

Methods

Materials and Reagents. All chemicals and reagents were purchased from MilliporeSigma (St. Louis, MO, USA) and used as received without further purification unless otherwise noted. Sodium oleate (97%) was purchased from Tokyo Chemical Industry (TCI) America (Portland, OR, USA). The high affinity cTnI peptide (95%) with *C*-terminal Cysteine residue (HWQIAYNEHQWQC) and the negative peptide (95%) with *C*-terminal Cysteine residue (HWNMAANEHMQWC) were purchased from GenScript USA Inc. (Pescataway, NJ, USA). (3-aminopropyl)triethoxysilane (APTES) was purchased from Gelest (Morrisville, PA, USA). Human male AB serum (H4522) was purchased from MilliporeSigma. Human cardiac troponin I-C monoclonal antibody (M46, cat. # sc-52277) and phosphatase inhibitor cocktail A (cat.# sc-45044) were purchased from Santa Cruz Biotechnology, Inc. (Dallas, TX, USA). N-hydroxysuccinimide (NHS) activated agarose slurry (cat.# 26200) and Halt™ Protease inhibitor cocktail were purchased from ThermoFisher Scientific (Rockford, IL, USA). Human cardiac troponin I ELISA kit (AccuBind® ELISA, cat. # 3825-300) was purchased from Monobind Inc. (Lake Forest, CA, USA). Extraction solutions were made in nanopure deionized water (H_2O) from Milli-Q® water (MilliporeSigma). Bradford protein assay reagent was purchased from Bio-Rad (Hercules, CA, USA). 12.5% gel (10 or 15 comb well, 10.0 cm × 10.0 cm, 1.0 mm thick) for SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) was home-made. DynaMag™-2 Magnet and Thermo Scientific™ Cimarec+™ stirring hotplate were purchased from ThermoFisher Scientific. Amicon, 0.5 mL cellulose centrifugal filters with a molecular-weight cutoff (MWCO) of 10 kDa were purchased from MilliporeSigma.

Synthesis of 3-Butynoic acid. 3-Butynoic acid was synthesized following a previously reported procedure[36], with modification. Deionized water (135 mL) was added to a 500 mL single neck round bottom flask fitted with magnetic stirrer bar. 70% HNO_3 (0.170 mL, 0.05 eq., 2.5 mmol), $Na_2Cr_2O_7 \cdot 2\ H_2O$ (0.15 g, 0.01 eq., 0.5 mmol) and $NaIO_4$ (34.2 g, 3.2 eq., 160 mmol) were subsequently added to the flask and the mixture was stirred vigorously on an ice bath for 15 min. 3-Butyn-1-ol (3.50 g, 1 eq., 50 mmol) was diluted with chilled deionized water (135 mL), added to the flask slowly, and stirred overnight. After this time, the reaction mixture was filtered to remove insoluble salts, and the product was extracted from the aqueous phase using diethyl ether (80 mL x 5). The organic fractions were combined, dried over anhydrous magnesium sulfate, and concentrated under rotary evaporation to give an orange viscous liquid. Subsequent washing with pentane (200 mL x 3) and further concentration under rotary evaporation times yielded 3-butynoic acid as an off-white solid (2.19 g, 52% yield). ^{1}H NMR (500 MHz, Chloroform-*d*) δ (ppm) 3.38 (d, 2H, J = 2.7 Hz), 2.25 (t, 1H, J = 2.7 Hz); ^{13}C NMR (125 MHz, Chloroform-*d*) δ (ppm) 173.8, 74.8, 72.4, 25.6.

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 (6.39 g, 1 eq., 25 mmol), and dichloromethane (250 mL) were added to an oven dried 500 mL three-neck flask fitted with a reflux condenser. The solution was heated to reflux under N_2. In a separate flask, (3-aminopropyl)triethoxysilane (5.53 g, 1 eq., 25 mmol) and N,N-diisopropylethylamine (6.46 g, 2 eq., 50 mmol) were diluted with dichloromethane (125 mL), and added to the refluxing three-neck flask by a syringe. The reaction mixture was allowed to reflux for 1 hour, and then N,N-diisopropylethylamine (6.46 mL, 2 eq., 50

mmol) was additionally added to effectively isomerize the propargylic isomer to N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide. The reaction was allowed to reflux overnight, and the product was concentrated by rotary evaporation. The crude product was redispersed in ethyl acetate, centrifuged at 5,500 x *g* for 5 min, and the supernatant was concentrated by rotary evaporation. The product was purified with flash column chromatography using a gradient of 50:50 ethyl acetate: n-hexane to 100:0 ethyl acetate: n-hexane to yield BAPTES as a clear, orange oil (4.80 g, 67% yield). R_f (Ethyl Acetate)= 0.67; ^{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*.

Synthesis of iron-oleate precursor. Iron oleate was synthesized using a previously established method[34]. 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-hour 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 by rotary evaporation.

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.

Synthesis of 8 nm magnetite (Fe_3O_4) nanocrystals. Iron oleate (10 mmol, 9.0 g) and oleic acid (5.5 mmol, 1.56 g) were added into a three-neck round bottom flask (250 mL) with a solvent mixture of 1-octadecene:1-tetradecene (40 g: 10 g). The mixture was stirred and degassed on a Schlenk line at 110 °C for 3 h. The mixture was then placed under nitrogen flow and the reaction solution was heated to 300 °C at a heating rate of 3.3 °C/min using a temperature controller. Reaction time was counted starting from when 300 °C was reached. After 1 h, the reaction solution was quickly cooled to room temperature by blowing air across the reaction flask. The resulting iron-oxide nanocrystals were precipitated using ethanol, isolated via centrifugation (5,500 xg, 20 min), and then washed with three precipitation/redispersion cycles (5,500 x *g*, 20 min) using ethanol. The resulting nanocrystals were then dried under vacuum, weighed, and redispersed in n-hexane at a concentration of 20 mg/mL for further use.

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 reaction under stirring 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 (5,500 x *g*, 10 min) to remove excess silane molecules and surfactants. The NPs were then dried under vacuum for later use.

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) 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 hour and later collected into Eppendorf tubes for washing. The peptide-functionalized NPs were washed three times with nanopure water via centrifugation (15,000 x *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.

Material characterization. Transmission electron microscopy (TEM) samples were prepared by pipetting a 10 μL drop of as-synthesized NPs at a concentration of 0.125 mg/mL onto a copper TEM grid with lacey carbon film. TEM was conducted on a FEI T12 microscope (FEI Company, Hillsboro, OR, USA) operated at 120 kV, equipped with a Gatan CCD image system with digital

micrograph software program. Transmission Fourier transform infrared (FTIR) spectroscopy measurements were recorded on a Bruker Equinox 55 FT-IR spectrometer (Bruker Optik GmbH, Ettlingen, Germany) in the range of 4,000 cm^{-1} to 400 cm^{-1} at 2 cm^{-1} resolution on NPs in a potassium bromide (KBr) pellet, at a sample mass loading of 0.33 wt.%. Thermogravimetric analysis (TGA) was carried out using a TA Instruments Q500 (TA Instruments, New Castle, DE, USA) thermal analysis system under a N_2 atmosphere and at a constant heating rate of 10 °C/min from 100 °C to 600 °C. All samples were first heated to 100 °C and held at that temperature for 3 min to remove adsorbed water. Zeta-potential (ζ-potential) measurements was carried out using a Malvern Zetasizer Nano.

Small molecule analysis by FTICR-MS. Small molecule samples (< 3 kDa) were diluted 500-fold in 50:50:0.1 (acetonitrile: water: formic acid) MS-grade solvent for positive electrospray ionization mode analysis. Samples were analyzed by direct infusion using a TriVersa Nanomate system (Advion BioSciences, Ithaca, NY, USA) coupled to a solariX XR 12-Tesla Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS, Bruker Daltonics, Bremen, Germany). For the nano-electrospray ionization source using a TriVersa Nanomate, the desolvating gas pressure was set at 0.5 PSI and the voltage was set to 1.2 to 1.6 kV versus the inlet of the mass spectrometer. Mass spectra were acquired with an acquisition size of 1M, in the mass range between 150-2000 *m/z* (with a resolution of 270,000 at 400 *m/z*), and 50 scans were accumulated for each sample. Ions were accumulated in the collision cell for 0.05 s, and a time of flight of 0.500 ms was used prior to their transfer to the ICR cell. For collisionally activated dissociation tandem MS (MS/MS) experiments, the collision energy was varied from 10 to 20V.

Tandem mass spectra were output from the DataAnalysis software and analyzed using MASH software[60]. The methods described here correspond to the data presented in Supplementary Figs. 3, 7, and 8.

Sarcomeric protein extraction from human cardiac tissue. All protein extraction procedures were performed in a cold room (4 °C) using freshly prepared buffers. All human cardiac tissue was voluntarily obtained from the University of Wisconsin Hospital and Clinic with informed consent. Discarded myocardial tissues from patients who undergo cardiac surgery as well as donor heart tissue from the Organ Procurement Organization (OPO), were used in this study. The procedure for the collection and de-identification of human cardiac tissue was approved by the Institutional Review Board of the University of Wisconsin-Madison. 500 mg of tissue was homogenized in wash buffer (5 mM NaH_2PO_4, 5 mM Na_2HPO_4, 5 mM $MgCl_2$, 0.5 mM EGTA, 0.1 M NaCl, 1% Triton X-100, 5 mM DTT, 1 mM PMSF, 1x HALT protease inhibitor cocktail, and 1x phosphatase inhibitor cocktail A, pH 7.4) using a Polytron electric homogenizer (Model PRO200; PRO Scientific, Oxford, CT, USA) on ice. The resulting homogenate was centrifuged at 10,000 x *g* (Avanti J-25i; Beckman Coulter, Fullerton, CA, USA) for 10 min at 4 ºC. After centrifugation, the supernatant was removed, and the pellet was washed once more with wash buffer. The wash supernatant was removed, and the pellet was re-suspended in protein extraction buffer (0.7 M LiCl, 25 mM Tris, 5 mM EGTA, 0.1 mM $CaCl_2$, 5 mM DTT, 1 mM PMSF, 1x HALT protease inhibitor cocktail, and 1x phosphatase inhibitor cocktail A, pH 7.5), and the suspension was agitated on a nutating mixer (Thermo Scientific, Boston, MA, USA) at 4 ºC for 45 min. The sarcomeric protein extract was centrifuged at 16,000 x *g* (Centrifuge 5415R, Eppendorf,

Hamburg, Germany) for 10 min at 4 ºC and the resulting supernatant was again centrifuged at 21,000 x *g* for 30 min to remove all tissue debris. The concentration of the tissue lysate was determined by Bradford protein assay. Samples were stored at -80 ºC for later study. The buffers and reagent preparation for the methods described here correspond to the data presented in Supplementary Table 4.

cTnI enrichment using NP-Pep from sarcomere extracts. NP-Pep (5 mg) was redispersed in a 2 mL Eppendorf Protein Lo-Bind tube with equilibration buffer (50 mM Tris, pH 7.5, 150 mM LiCl). The NPs were then centrifuged at 15,000 x *g* for 2 min at 4 °C, isolated from the solution using the DynaMag, and the supernatant was removed. Equilibration buffer was then added to the NPs, the mixture was sonicated and vortexed to prepare for protein loading. Protein loading mixture (L), from tissue extract was diluted to a final volume of 1 mL with a buffered solution (50 mM Tris, pH 7.4, 150 mM LiCl) to a total protein loading of 0.3 mg/mL and was added to the NP-Pep mixture, at a NP concentration of 5 mg/mL. After this mixture was agitated on a nutating mixer at 4 °C for 40 min, the NPs were centrifuged at 15,000 x *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 a wash buffer (50 mM Tris, pH 7.5, 300 mM NaCl; 0.20 mL/mg NP) following the same centrifugation and magnetic isolation steps to remove unbound, non-specific proteins. To elute the bound cTnI, 500 µL of 200 mM glycine hydrochloride buffer (pH 2.2) was added. After centrifugation and magnetic isolation, the resulting supernatant was collected as the elution fraction (E). We found that the majority of proteins were eluted in the first elution fraction and that subsequent elutions using the same 200

mM glycine hydrochloride buffer or 1% SDS yielded minimal protein. Typically, only the first elution mixture was used for further analysis. All protein fractions (L, F, and E) were desalted prior to MS analysis using a 10 kDa MWCO filter (Amicon, 0.5 mL, cellulose, MilliporeSigma) and buffered exchanged using 0.2% formic acid in nanopure water. To evaluate the enrichment performance of the functionalized NPs, all collected fractions (L, F, and E) were equally loaded (500 ng), separated using a polyacrylamide gel (12.5%), and stained with SYPRO Ruby (ThermoFisher Scientific) fluorescent dye. The gel was imaged using a ChemiDoc™ MP Imaging System (170-8280; Bio-Rad, Hercules, CA, USA). Uncropped and unprocessed scans of all blots and gels are available in the Source Data Excel file.

cTnI enrichment using agarose beads from sarcomere extracts. mAb or peptide was conjugated to the agarose beads following the manufacturer's recommendations and blocked with a 1.0 M ethanolamine solution, pH 7.4. Protein loading mixture (L) was incubated with 500 µL of mAb-conjugated or peptide-conjugated agarose beads in a disposable affinity column for 40 min on a nutating mixer at 4 ºC. After incubation, the supernatant was collected and saved as the flow-through (F) fraction. The agarose beads were then washed three times with 1 mL of wash buffer (50 mM Tris pH 7.4, 300 mM NaCl) to elute unbound proteins. Subsequently, the bound proteins were eluted using four equal fractions of 500 µL of 200 mM glycine hydrochloride (pH 2.2). The four elution fractions (E) were pooled and all protein fractions (L, F, and E) were desalted prior to MS analysis using a 10 kDa molecular weight cutoff (MWCO) filter (Amicon, 0.5 mL, cellulose, MilliporeSigma) and buffered exchanged using 0.2% formic acid in nanopure water.

cTnI enrichment using the NP-Pep from human serum. Serum loading mixtures were prepared by serially spiking in human sarcomeric protein extracts into 200 µL of serum (10 mg protein) from human male AB plasma and diluting the serum mixture to a final volume of 1 mL using a buffered solution (50 mM Tris pH 7.4, 1.0 M LiCl). The concentrations of cTnI spiked into serum samples were measured using a cTnI AccuBind ELISA kit (Monobind Inc.). The cTnI spiked human serum loading mixture (L) was incubated with NP-Pep (5 mg) in a 2 mL Eppendorf Lo-Bind tube. Typically, a relative loading of 0.5 mg NP-Pep per 1 mg of human serum was determined to be sufficient for effective serum cTnI enrichment. After this mixture was agitated on a nutating mixer at 4 °C for 40 min, the NPs were centrifuged at 15,000 x *g* for 2 min at 4 °C, and then isolated from the solution using a DynaMag. The supernatant was collected and saved as the flow-through (F) fraction. The NPs were then washed three times with 1 mL of wash buffer (50 mM Tris pH 7.4, 700 mM NaCl) following the same centrifugation and magnetic isolation steps to remove unbound, non-specific proteins. To elute the bound cTnI, 500 µL of 200 mM glycine hydrochloride buffer (pH 2.2) was added. We found that the majority of proteins were eluted in the first elution fraction and that subsequent elutions using the same 200 mM glycine hydrochloride buffer or 1% SDS yielded minimal protein. Typically, only the first elution mixture was used for further analysis. After centrifugation and magnetic isolation, the resulting supernatant was collected as the elution fraction (E). All protein fractions (L, F, and E) were desalted prior to MS analysis using a 10 kDa MWCO filter (Amicon, 0.5 mL, cellulose, MilliporeSigma) and buffer exchanged using 0.2% formic acid in nanopure water.

cTnI enrichment using agarose beads from human serum. Serum loading mixtures were prepared by serially spiking in human sarcomeric protein extracts into 200 µL of serum (10 mg protein) from human male AB plasma and diluting the serum mixture to a final volume of 1 mL using a buffered solution (50 mM Tris pH 7.4, 1.0 M LiCl). mAb or peptide was conjugated to the agarose beads following the manufacturer's recommendations and blocked with a 1.0 M ethanolamine solution, pH 7.4. The serum loading mixture (L) was added to 500 µL of mAb-conjugated or peptide-conjugated agarose beads in a disposable affinity column and agitated for 40 min on a nutating mixer at 4 ºC. After incubation, the supernatant was collected and saved as the flow-through (F) fraction. The agarose beads were then washed three times with 1 mL of wash buffer (50 mM Tris pH 7.4, 700 mM NaCl) to elute unbound proteins. Subsequently, the bound proteins were eluted using four equal fractions of 500 µL of 200 mM glycine hydrochloride (pH 2.2). The four elution fractions (E) were pooled and all protein fractions (L, F, and E) were desalted prior to MS analysis using a 10 kDa MWCO filter (Amicon, 0.5 mL, cellulose, MilliporeSigma) and buffered exchanged using 0.2% formic acid in nanopure water.

Comparison of cTnI enrichment from sarcomere extracts. The enrichment workflow was performed in a cold room held at 4 °C to minimize possible artifactual protein modifications, such as oxidation. NP-Control (NP-BAPTES), NP-Pep (Fe_3O_4-BAPTES-Peptide), Agarose-Control beads (no coupling), Agarose-Pep beads, and Agarose-mAb beads were incubated with sarcomeric extract obtained from completely de-identified healthy donor heart tissue, diseased heart tissue of dilated cardiomyopathy, and post-mortem heart tissue containing 306 ng cTnI, 482 ng cTnI, and 465 ng cTnI in the protein extract, respectively. cTnI values were determined by ELISA

quantification. The loading mixture (L), flow through (F), and elution mixture (E) were collected, desalted using a 10 kDa MWCO filter, and buffer exchanged prior to SDS-PAGE gel analysis or liquid chromatography-mass spectrometry (LC-MS) analysis as described in the tissue enrichment protocol. The methods described here correspond to the data presented in Fig. 2f-k and Supplementary Figs. 17-18.

Comparison of cTnI enrichment from serum spiked with cTnI. The enrichment workflow was performed in a cold room held at 4 °C to minimize possible artifactual protein modifications, such as oxidation. NP-Control (NP-BAPTES), NP-Pep (Fe_3O_4-BAPTES-Peptide), Agarose-Control beads (no coupling), Agarose-Pep beads, and Agarose-mAb beads were incubated with in a protein loading mixture containing human male AB serum (10 mg) and sarcomeric tissue extract obtained from healthy donor heart tissue, diseased heart tissue of dilated cardiomyopathy, and post-mortem heart tissue containing 306 ng cTnI, 482 ng cTnI, and 465 ng cTnI in the protein extract, respectively. cTnI values were determined by ELISA quantification. The loading mixture (L), flow through (F), and elution mixture (E) were collected, desalted using a 10 kDa MWCO, and buffer exchanged prior to SDS-PAGE gel analysis or LC-MS analysis as described in the tissue enrichment protocol. The methods described here correspond to the data presented in Fig. 2a-d and Supplementary Figs. 20-22.

Typical reverse phase chromatography procedure. Reverse phase chromatography (RPC) was performed with a nanoACQUITY UPLC system (Waters Corporation; Milford, MA, USA). Mobile phase A (MPA) contained 0.2 % formic acid in nanopure water, and mobile phase B (MPB)

contained 0.2% formic acid in 50:50 acetonitrile: isopropanol. Prior to injection, protein samples were desalted by washing through 10 kDa MWCO filters using 0.2% formic acid in nanopure water six times. For each injection, 500 ng of desalted protein sample was loaded onto a BIOshell A400 C4, 3.4 µm, 15 cm x 200 µm capillary column (Supelco, Bellafonte, PA, USA). The column was placed in a column heater set at 60 °C with a constant 3 µL/min flow rate. The RPC gradient consisted of the following concentrations of MPB: 10% MPB at 0 min, 10% at 5 min, 65% at 45 min, 90% at 50 min, held at 90% until 55 min, adjusted back to 10% at 55.1 min, and held at 10% until 60 min. Each run was 60 min long.

Top-down MS analysis. Samples eluted from RPC separation were ionized using a CaptiveSpray source (Bruker Daltonics) into a maXis II Q-TOF mass spectrometer (Bruker Daltonics) for online LC-MS and LC-MS/MS experiments. End plate offset and capillary voltage were set at 500 and 4000 V, respectively. The nebulizer was set to 0.3 bar, and the dry gas flow rate was 4.0 L/min at 200 °C. The quadruple low mass cutoff was set to 500 *m/z* for MS and 200 *m/z* for MS/MS. Mass range was set to 200-3000 *m/z* and spectra were acquired at 1 Hz for LC-MS runs. All data were collected with otofControl 3.4 (Bruker Daltonics), analyzed and processed in DataAnalysis 4.3 (Bruker Daltonics). Maximum Entropy algorithm (Bruker Daltonics) was used to deconvolute all mass spectra with the resolution set to 80,000. 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 the monoisotopic mass of all detected ions. Chromatograms in Fig. 3 and Supplementary Figs. 21-22 shown were smoothened by Gauss algorithm with a smoothing width of 1.67 s. To quantify protein expression across samples, the

top 5 most abundant charge states' ions (average ± 0.2 *m/z*) of all major proteoforms from the same protein were retrieved collectively as one extracted peak in the extracted ion chromatogram[61]. The area under curve was manually determined for each protein isoform using DataAnalysis. To quantify protein modifications, the relative abundances of specific modifications were calculated as their corresponding percentages among all the detected protein forms in the deconvoluted averaged mass. Tandem mass spectra were output from the DataAnalysis software and analyzed using MASH software[60]. For the targeted collision-induced dissociation (CID) LC-MS/MS analysis shown in Supplementary Fig. 30, the collision energy was varied from 18V to 30V and the quadruple low mass was set to 500 *m/z* with a scan range of 200 to 3000 *m/z*. The total ion current (TIC) corresponding to all obtained MS/MS signal was averaged across the LC retention window corresponding to cTnI, and the averaged MS/MS data was then directly imported into MASH Explorer for proteoform identification and sequence mapping. All the program-processed data were manually validated. Peak extraction was performed using a signal-to-noise ratio of 3 and a minimum fit of 60%, and all peaks were subjected to manual validation. Alignment-based MS-Align+ algorithm (TopPIC)[62] for intact protein identification based on protein spectrum matches was used to search against the UniprotSwissprot Human database, which was released on June 26th 2019. Fragment mass tolerance was set to 15 ppm. All identifications were validated with statistically significant P and E values (<0.01) and satisfactory numbers of assigned fragment (>10). A 15-ppm mass tolerance was used to match the experimental fragment ions to the calculated fragment ions based on amino acid sequence.

Statistical analysis

All statistical data were presented as the mean ± standard error of the mean (SEM) unless stated otherwise. Two-tailed student's t-test was used for comparison of two groups to evaluate the statistical significance of variance for the validation of the simultaneous quantification of protein expression and modification changes. Differences among means were considered significant at $p < 0.05$. All error bars shown in the figures were based on SEMs unless stated otherwise.

Data availability

All data generated or analyzed during this study are presented in this manuscript or in the supplementary information. All the raw data files or spectra are available upon request. All source data underlying Figs. 2.1f, 2.1g, 2.2b-h, 2.3a-g, 2.4b, Supplementary Fig. 2.3, 2.7-9, 2.11-24, 2.25b-c, 2.26c, 2.28, 2.28b-c, 2.29, 2.30a-c are available as individual data files. Additionally, proteomics data has been uploaded to PRIDE repository via ProteomeXchange with identifier PXD019712.

References

1. Smith, L.M. et al. Proteoform: a single term describing protein complexity. *Nature Methods* **10**, 186 (2013).
2. 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* **10**, 139-148 (2018).
3. Smith, L.M. & Kelleher, N.L. Proteoforms as the next proteomics currency. *Science* **359**, 1106-1107 (2018).
4. Aebersold, R. et al. How many human proteoforms are there? *Nature Chemical Biology* **14**, 206-214 (2018).
5. Cravatt, B.F., Simon, G.M. & Yates Iii, J.R. The biological impact of mass-spectrometry-based proteomics. *Nature* **450**, 991-1000 (2007).
6. Toby, T.K., Fornelli, L. & Kelleher, N.L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annual Review of Analytical Chemistry* **9**, 499-519 (2016).
7. Kelleher, N.L., Thomas, P.M., Ntai, I., Compton, P.D. & LeDuc, R.D. Deep and quantitative top-down proteomics in clinical and translational research. *Expert Review of Proteomics* **11**, 649-651 (2014).
8. Chen, B., Brown, K.A., Lin, Z. & Ge, Y. Top-Down Proteomics: Ready for Prime Time? *Analytical Chemistry* **90**, 110-127 (2018).
9. Rifai, N., Gillette, M.A. & Carr, S.A. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nature Biotechnology* **24**, 971-983 (2006).
10. Anderson, N.L. & Anderson, N.G. The Human Plasma Proteome. *History, Character, and Diagnostic Prospects* **1**, 845-867 (2002).
11. Xie, S., Moya, C., Bilgin, B., Jayaraman, A. & Walton, S.P. Emerging affinity-based techniques in proteomics. *Expert review of proteomics* **6**, 573-583 (2009).
12. Lollo, B., Steele, F. & Gold, L. Beyond antibodies: New affinity reagents to unlock the proteome. *PROTEOMICS* **14**, 638-644 (2014).
13. Smith, J.G. & Gerszten Robert, E. Emerging Affinity-Based Proteomic Technologies for Large-Scale Plasma Profiling in Cardiovascular Disease. *Circulation* **135**, 1651-1664 (2017).
14. Baker, M. Reproducibility crisis: Blame it on the antibodies. *Nature* **521**, 274-276 (2015).
15. Bradbury, A. & Plückthun, A. Reproducibility: Standardize antibodies used in research. *Nature* **518**, 27-29 (2015).

16. Janes, K.A. Fragile epitopes—Antibody's guess is as good as yours. *Science Signaling* **13**, eaaz8130 (2020).

17. Marx, V. Calling the next generation of affinity reagents. *Nature Methods* **10**, 829 (2013).

18. You, C.-C. et al. Detection and identification of proteins using nanoparticle–fluorescent polymer 'chemical nose' sensors. *Nature Nanotechnology* **2**, 318-323 (2007).

19. Qian, E.A. et al. Atomically precise organomimetic cluster nanomolecules assembled via perfluoroaryl-thiol SNAr chemistry. *Nature Chemistry* **9**, 333-340 (2017).

20. Kelley, S.O. et al. Advancing the speed, sensitivity and accuracy of biomolecular detection using multi-length-scale engineering. *Nature Nanotechnology* **9**, 969-980 (2014).

21. Labib, M. et al. Single-cell mRNA cytometry via sequence-specific nanoparticle clustering and trapping. *Nature Chemistry* **10**, 489-495 (2018).

22. Benjamin, E.J. et al. Heart Disease and Stroke Statistics—2018 Update: A Report From the American Heart Association. *Circulation* **137**, e67-e492 (2018).

23. Westermann, D., Neumann, J.T., Sörensen, N.A. & Blankenberg, S. High-sensitivity assays for troponin in patients with cardiac disease. *Nature Reviews Cardiology* **14**, 472 (2017).

24. Antman, E.M. et al. Cardiac-Specific Troponin I Levels to Predict the Risk of Mortality in Patients with Acute Coronary Syndromes. *New England Journal of Medicine* **335**, 1342-1349 (1996).

25. Park, K.C., Gaze, D.C., Collinson, P.O. & Marber, M.S. Cardiac troponins: from myocardial infarction to chronic disease. *Cardiovascular Research* **113**, 1708-1718 (2017).

26. Twerenbold, R. et al. Clinical Use of High-Sensitivity Cardiac Troponin in Patients With Suspected Myocardial Infarction. *Journal of the American College of Cardiology* **70**, 996-1012 (2017).

27. Soetkamp, D. et al. The continuing evolution of cardiac troponin I biomarker analysis: from protein to proteoform. *Expert Review of Proteomics* **14**, 973-986 (2017).

28. Labugger, R., Organ, L., Collier, C., Atar, D. & Van Eyk Jennifer, E. Extensive Troponin I and T Modification Detected in Serum From Patients With Acute Myocardial Infarction. *Circulation* **102**, 1221-1226 (2000).

29. Murphy, A.M. et al. Transgenic Mouse Model of Stunned Myocardium. *Science* **287**, 488-491 (2000).

30. Peng, Y. et al. Top-down Proteomics Reveals Concerted Reductions in Myofilament and Z-disc Protein Phosphorylation after Acute Myocardial Infarction. *Mol Cell Proteomics* **13**, 2752-2764 (2014).

31. Dong, X. et al. Augmented Phosphorylation of Cardiac Troponin I in Hypertensive Heart Failure. *Journal of Biological Chemistry* **287**, 848-857 (2012).

32. Zhang, J. et al. Top-Down Quantitative Proteomics Identified Phosphorylation of Cardiac Troponin I as a Candidate Biomarker for Chronic Heart Failure. *Journal of Proteome Research* **10**, 4054-4065 (2011).

33. McDonough, J.L. & Van Eyk, J.E. Developing the next generation of cardiac markers: Disease-induced modifications of troponin I. *Progress in Cardiovascular Diseases* **47**, 207-216 (2004).

34. Park, J. et al. Ultra-large-scale syntheses of monodisperse nanocrystals. *Nature Materials* **3**, 891-895 (2004).

35. Roberts, D.S. et al. Reproducible large-scale synthesis of surface silanized nanoparticles as an enabling nanoproteomics platform: Enrichment of the human heart phosphoproteome. *Nano Research* **12**, 1473-1481 (2019).

36. Abbas, A., Xing, B. & Loh, T.-P. Allenamides as Orthogonal Handles for Selective Modification of Cysteine in Peptides and Proteins. *Angewandte Chemie International Edition* **53**, 7491-7494 (2014).

37. Xiao, X. et al. Advancing Peptide-Based Biorecognition Elements for Biosensors Using in-Silico Evolution. *ACS Sensors* **3**, 1024-1031 (2018).

38. Katrukha, A.G. et al. Degradation of cardiac troponin I: implication for reliable immunodetection. *Clinical Chemistry* **44**, 2433 (1998).

39. Bates, K.J. et al. Circulating Immunoreactive Cardiac Troponin Forms Determined by Gel Filtration Chromatography after Acute Myocardial Infarction. *Clinical Chemistry* **56**, 952 (2010).

40. Apple, F.S. & Collinson, P.O. Analytical Characteristics of High-Sensitivity Cardiac Troponin Assays. *Clinical Chemistry* **58**, 54 (2012).

41. Agard, N.J., Prescher, J.A. & Bertozzi, C.R. A Strain-Promoted [3 + 2] Azide−Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society* **126**, 15046-15047 (2004).

42. McKay, C.S. & Finn, M.G. Click Chemistry in Complex Mixtures: Bioorthogonal Bioconjugation. *Chemistry & Biology* **21**, 1075-1101 (2014).

43. Es-sebbar, E., Jolly, A., Benilan, Y. & Farooq, A. Quantitative mid-infrared spectra of allene and propyne from room to high temperatures. *Journal of Molecular Spectroscopy* **305**, 10-16 (2014).

44. Jiang, J., Oberdörster, G. & Biswas, P. Characterization of size, surface charge, and agglomeration state of nanoparticle dispersions for toxicological studies. *Journal of Nanoparticle Research* **11**, 77-89 (2009).

45. Blanco, E., Shen, H. & Ferrari, M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nature Biotechnology* **33**, 941-951 (2015).

46. Zabrouskov, V., Ge, Y., Schwartz, J. & Walker, J.W. Unraveling Molecular Complexity of Phosphorylated Human Cardiac Troponin I by Top Down Electron Capture Dissociation/Electron Transfer Dissociation Mass Spectrometry. *Mol Cell Proteomics* **7**, 1838-1849 (2008).

47. Walkey, C.D. & Chan, W.C.W. Understanding and controlling the interaction of nanomaterials with proteins in a physiological environment. *Chemical Society Reviews* **41**, 2780-2799 (2012).

48. De, M., Ghosh, P.S. & Rotello, V.M. Applications of Nanoparticles in Biology. *Advanced Materials* **20**, 4225-4241 (2008).

49. Gao, J., Gu, H. & Xu, B. Multifunctional Magnetic Nanoparticles: Design, Synthesis, and Biomedical Applications. *Accounts of Chemical Research* **42**, 1097-1107 (2009).

50. Rosi, N.L. & Mirkin, C.A. Nanostructures in Biodiagnostics. *Chemical Reviews* **105**, 1547-1562 (2005).

51. Apple, F.S. Counterpoint: Standardization of Cardiac Troponin I Assays Will Not Occur in My Lifetime. *Clinical Chemistry* **58**, 169-171 (2012).

52. Apple, F.S., Sandoval, Y., Jaffe, A.S. & Ordonez-Llanos, J. Cardiac Troponin Assays: Guide to Understanding Analytical Characteristics and Their Impact on Clinical Care. *Clinical Chemistry* **63**, 73 (2017).

53. Herman, D.S., Kavsak, P.A. & Greene, D.N. Variability and Error in Cardiac Troponin Testing: An ACLPS Critical Review. *American Journal of Clinical Pathology* **148**, 281-295 (2017).

54. Solaro, R.J., Henze, M. & Kobayashi, T. Integration of Troponin I Phosphorylation With Cardiac Regulatory Networks. *Circulation Research* **112**, 355-366 (2013).

55. Dunn, M.R., Jimenez, R.M. & Chaput, J.C. Analysis of aptamer discovery and technology. *Nature Reviews Chemistry* **1**, 0076 (2017).

56. Gupta, S. et al. Chemically Modified DNA Aptamers Bind Interleukin-6 with High Affinity and Inhibit Signaling by Blocking Its Interaction with Interleukin-6 Receptor. *Journal of Biological Chemistry* **289**, 8706-8719 (2014).

57. Wu, C.-H., Liu, I.J., Lu, R.-M. & Wu, H.-C. Advancement and applications of peptide phage display technology in biomedical science. *Journal of Biomedical Science* **23**, 8 (2016).

58. Wang, T., Chen, C., Larcher, L.M., Barrero, R.A. & Veedu, R.N. Three decades of nucleic acid aptamer technologies: Lessons learned, progress and opportunities on aptamer development. *Biotechnology Advances* **37**, 28-50 (2019).

59. Kyle, S. Affimer Proteins: Theranostics of the Future? *Trends in Biochemical Sciences* **43**, 230-232 (2018).

60. Cai, W.X. et al. MASH Suite Pro: A Comprehensive Software Tool for Top-Down Proteomics. *Molecular & Cellular Proteomics* **15**, 703-714 (2016).

61. Lin, Z. et al. Simultaneous Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome. *Molecular & Cellular Proteomics* **18**, 594-605 (2019).

62. Kou, Q., Xun, L. & Liu, X. TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics* **32**, 3495-3497 (2016).

Acknowledgement

This research is supported by NIH R01 GM117058 (to S.J. and Y.G.). Y.G. would like to acknowledge NIH R01 GM125085, R01 HL096971, and S10 OD018475. T.N.T. also thanks the fellowship support from the NIH Chemistry-Biology Interface Training Program NIH T32GM008505. K.A.B. would like to acknowledge support from the Training Program in Translational Cardiovascular Science, T32 HL007936-19. We thank Prof. Timothy Kamp for critical reading of this manuscript and insightful discussions. We also thank Prof. Amish Raval and William Swain for the helpful discussions. Moreover, we would like to thank Dr. Takushi Kohmoto, Dr. Ken Young, Elizabeth Bayne, James Anderson, and Carrie Sparks for helping collect the human heart samples and providing the clinical information.

Author contributions

T.N.T. & D.S.R. contributed equally to this work. Y.G. and S.J. conceived the study and supervised the research; T.N.T., D.S.R., K.A.B., Y.Z., T.M.G.-A., S.J., and Y.G. designed the research; T.N.T. and D.S.R. performed the research; T.N.T., D.S.R., K.A.B., Y.Z, B.C., Z.W., and S.D.M. analyzed the data; T.N.T., D.S.R., S.J., and Y.G. wrote the manuscript with input from all co-authors.

Competing interests

The University of Wisconsin-Madison has filed a provisional patent application serial no. 62/949,869 (December 18, 2019) on the basis of this work. Y.G., S.J., T.N.T., and D.S.R. are named as the inventors on the provisional patent application.

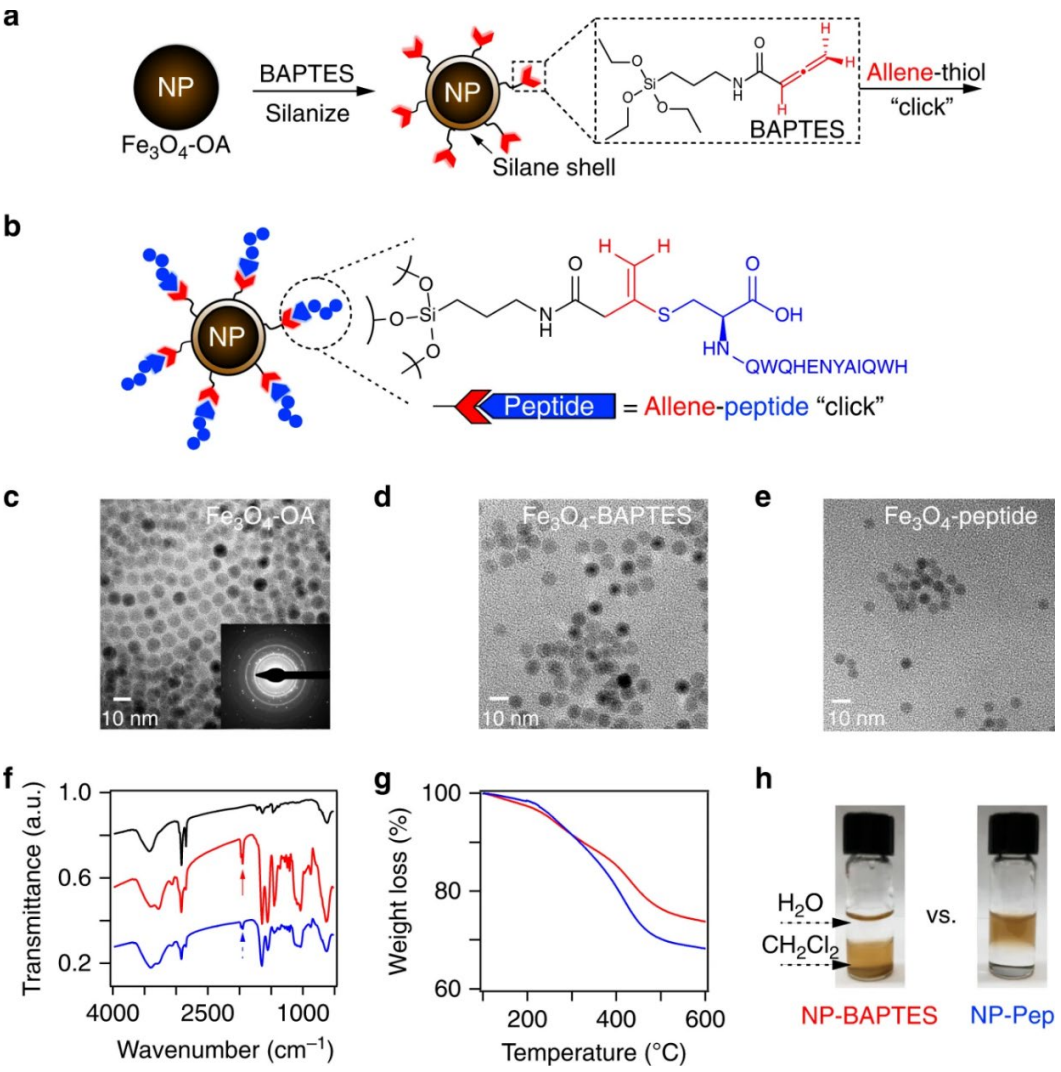


Fig. 2.1 | Design and characterization of surface functionalized nanoparticles (NPs) for capturing cTnI. a, Silanization of Fe_3O_4 NPs using an allene carboxamide-based organosilane monomer (BAPTES) for cysteine thiol-specific bioconjugation. **b**, Illustration of the rationally designed NPs that are surface functionalized with a 13-mer peptide that has a high affinity for cTnI (NP-Pep) for cTnI enrichment. The 13-mer peptide possesses a C-terminal cysteine that selectively reacts with the allene carboxamide moiety on the silanized NPs. **c-e**, Representative TEM images of surface functionalized NPs: Fe_3O_4-OA NPs (**c**) (inset shows the selected area electron diffraction pattern), Fe_3O_4-BAPTES NPs (**d**), and Fe_3O_4-Peptide NPs (**e**). **f**, **g,** Representative FTIR spectra (**f**) and TGA analysis (**g**) of various NPs: Fe_3O_4-OA NPs (black), Fe_3O_4-BAPTES NPs (red), and Fe_3O_4-peptide NPs (blue). FTIR spectra are offset along the y-axis for clarity, with arrows denoting the characteristic allene (C=C=C) vibrational modes (1971, 1947 cm^{-1}) corresponding to the silanized NPs. TGA analysis reveals increased organic content on NPs after peptide coupling. **h**, Photographs of functionalized NPs in a biphasic mixture of dichloromethane (CH_2Cl_2) and water (H_2O), comparing the solvent compatibility of the NP-BAPTES and the NP-Pep. The NP-BAPTES are dispersible in dicholoromethane but the NP-Pep are stable and dispersible in water. The displayed NP-Pep and NP-BAPTES originated from the same synthetic batch. NP-BAPTES is referred to as NP-Ctrl in the subsequent figures and text. Data are representative of n = 3 independent experiments. Source data are provided as a Source Data file.

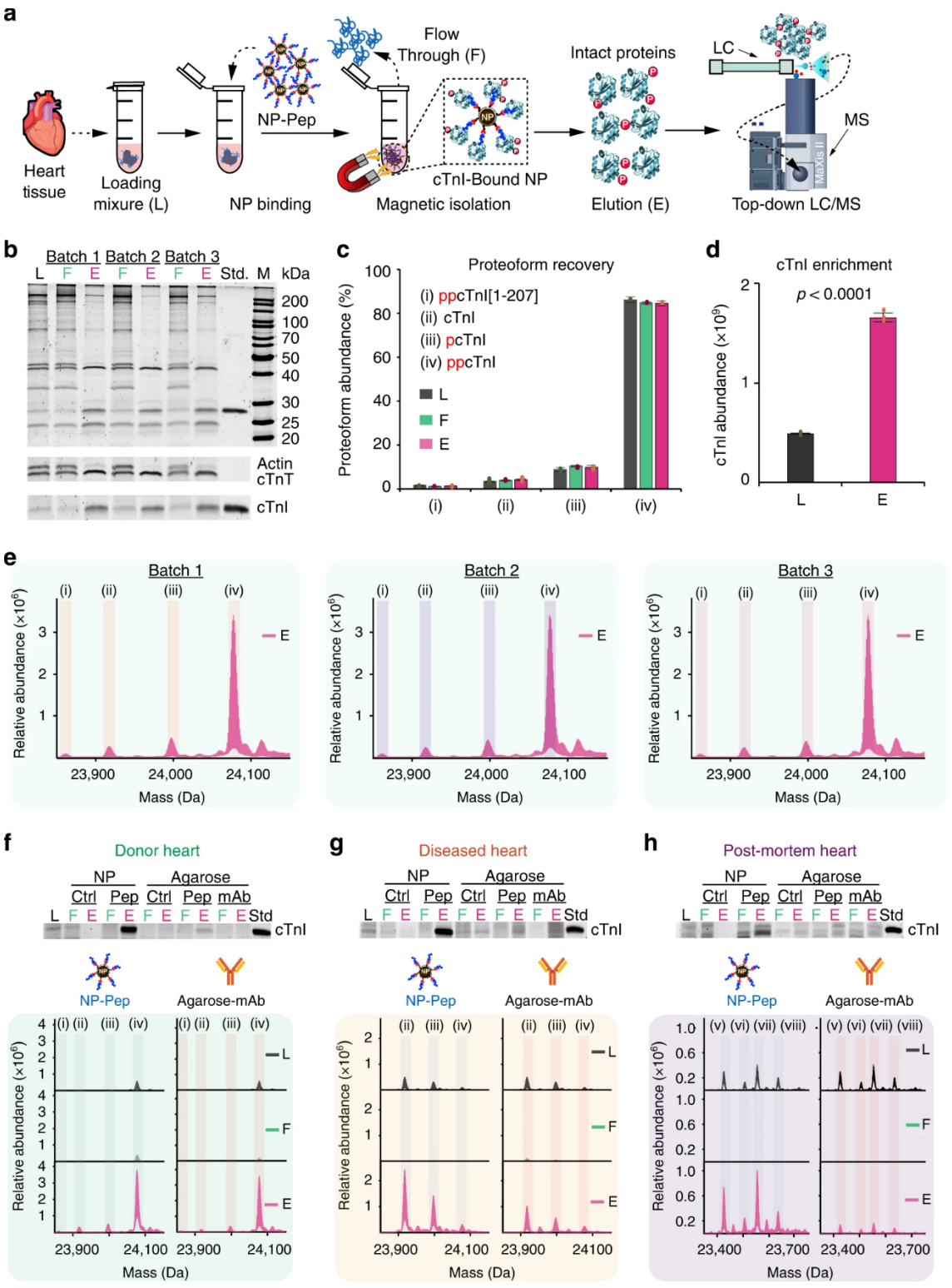

a
Heart tissue
Loading mixure (L)
NP-Pep
NP binding
Flow Through (F)
Magnetic isolation
cTnI-Bound NP
Intact proteins
Elution (E)
LC
MS
Top-down LC/MS
b
Batch 1 Batch 2 Batch 3
L F E F E F E Std. M kDa
200
100
70
50
40
30
25
20
Actin
cTnT
cTnI
c
Proteoform recovery
(i) ppcTnI[1-207]
(ii) cTnI
(iii) pcTnI
(iv) ppcTnI
L
F
E
Proteoform abundance (%)
d
cTnI enrichment
p < 0.0001
cTnI abundance (×10⁹)
e
Batch 1
Batch 2
Batch 3
Relative abundance (×10⁶)
Mass (Da)
f
Donor heart
g
Diseased heart
h
Post-mortem heart
NP
Agarose
Ctrl Pep Ctrl Pep mAb
cTnI
NP-Pep
Agarose-mAb

Fig. 2.2| Evaluation of enrichment performance and reproducibility of NP-Pep for capturing cTnI. a, Schematic illustration of the nanoproteomics strategy for cTnI enrichment and top-down MS analysis of cTnI proteoforms. Heart tissue extract (Loading mixture, L) is first incubated with the NP-Pep. Following magnetic isolation, the nonspecific proteins are removed as flow-through (F). The NPs are washed and the NP-bound proteins of interest are eluted (E) and analyzed by top-down LC/MS. b**,** SDS-PAGE demonstrates the high reproducibility obtained from three different NP-Pep syntheses (inter-batch). The lower panels highlight the specific enrichment of cTnI and cTnT in E as compared to L and F. **c, d,** Evaluation of total cTnI proteoform recovery (**c**), and relative abundance of cTnI (**d**), before and after NP-Pep enrichment. The same L, F, and E shown in the SDS-PAGE from (**b**) were equally loaded (500 ng) for LC/MS analysis. The deconvoluted top-down mass spectra corresponding to cTnI proteoforms were used to calculate the relative abundance of each cTnI proteoform when normalizing for total protein amount injected. Proteoform abundance data (**c**) are representative of n = 6 independent experiments with error bars indicating standard error of the mean. cTnI relative abundance data (**d**) are representative of n = 3 independent experiments with error bars indicating standard error of the mean. **e,** MS-based evaluation of cTnI enrichment using three different synthetic batches of NP-Pep showing the reproducible enrichment performance. **f-h**, Enrichment from three different human heart samples using NP-Pep in comparison with agarose-mAb. Upper panels feature SDS-PAGE strips that visualize cTnI enrichment performance of NP-Pep and agarose-mAb. Equal protein amount (500 ng) of the L, F, and E was loaded on the gel in **b, f-h.** Roman numerals correspond to N-terminally acetylated cTnI proteoforms following Met exclusion: (i) ppcTnI[1-207]; (ii) cTnI; (iii) pcTnI; (iv) ppcTnI; (v) cTnI[1-205]; (vi) pcTnI[1-205]; (vii) cTnI[1-206]; (viii) pcTnI[1-206]. M., protein marker; Std., endogenous cTnI protein standard. p*,* phosphorylation. pp*,* bisphosphorylation. A summary of the enriched proteoforms and their respective mass measurements by top-down MS are listed in Supplementary Tables 5 and 6. Source data are provided as a Source Data file.

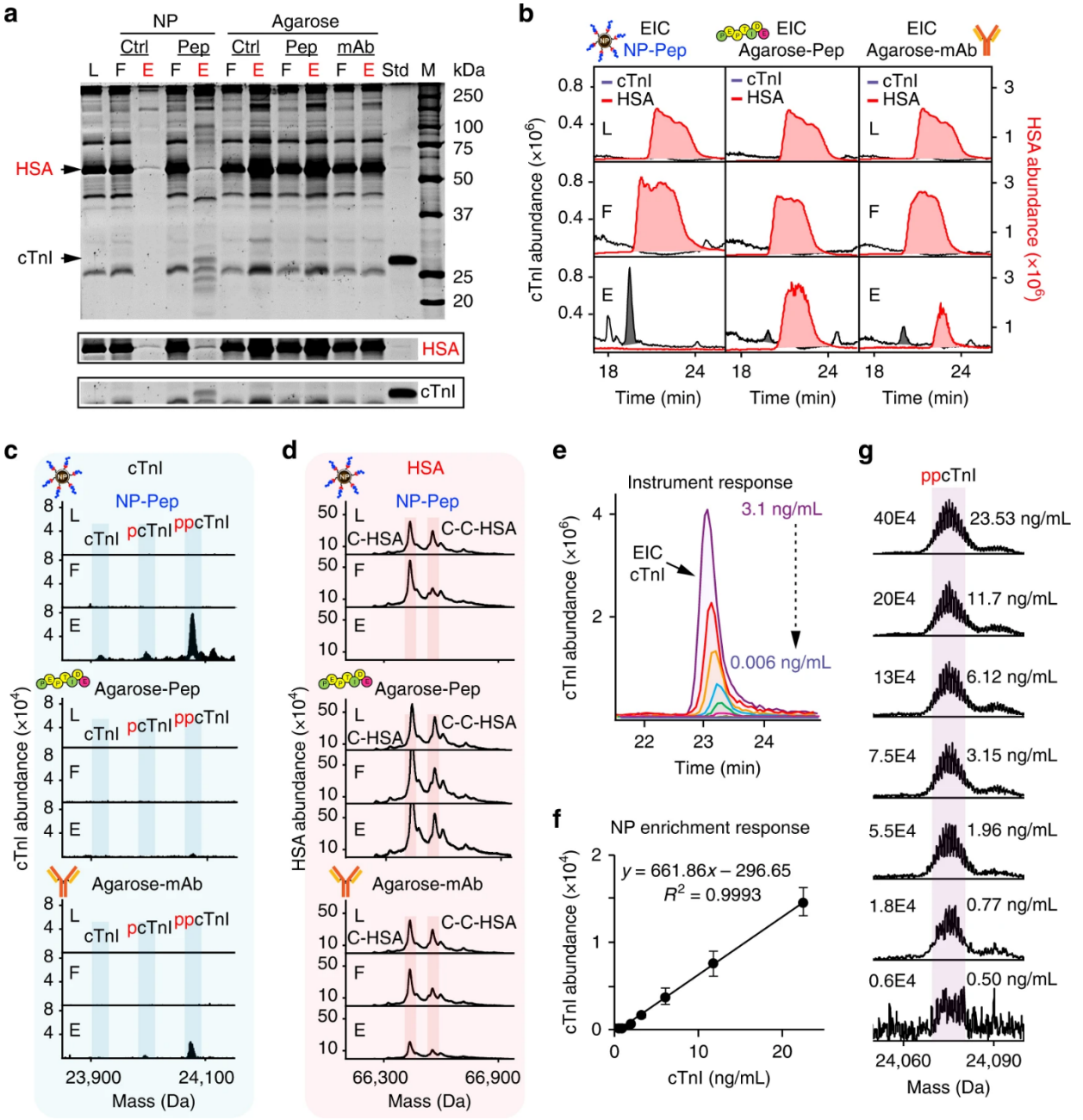


Fig. 2.3| Direct enrichment of cTnI from human serum by NP-Pep with high specificity and sensitivity. Human serum with spike-in cTnI was used for the evaluation of cTnI enrichment performance using NP-Pep and the conventional agarose-mAb. **a,** SDS-PAGE visualizing the effectiveness of cTnI enrichment from human serum using different affinity platforms. The cTnI (306 ng) was obtained from a sarcomeric protein extract from a healthy donor heart and spiked into human serum (10 mg). Loading mixture (L), flow through (F), and elution mixture after enrichment (E) were equally loaded on the gel (500 ng) using NP-Ctrl, NP-Pep, Agarose-Ctrl, Agarose-Pep, and Agarose-mAb. The NP-Pep enabled nearly complete depletion of human serum albumin (HSA), allowing for more specific enrichment of cTnI from human serum. **b**, Normalized extracted ion chromatograms (EICs) of cTnI (black) and HSA (red) corresponding to the NP-Pep, Agarose-Pep, and Agarose-mAb shown in (**a**). **c-d**, Normalized deconvoluted mass spectra corresponding to enriched cTnI (**c**) and depleted HSA (**d**), illustrating the abundance of

cTnI and HSA before and after enrichment using NP-Pep, Agarose-Pep, and Agarose-mAb corresponding to (**a**, **b**). **e,** Normalized EICs illustrating the MS response for cTnI (3.1 ng/ml to 0.006 ng/ml) obtained from a human heart tissue extract (n = 2 independent experiments) using a Bruker maXis II ETD. **f,** Evaluation of the sensitivity performance of the nanoproteomics assay. MS response against cTnI concentration is shown for various NP-Pep enrichment fractions from serum mixtures spiked with cTnI (22.53 – 0.50 ng/mL). The deconvoluted peak intensities of ppcTnI were used for the analyses. Data are representative of n = 3 independent experiments with error bars indicating standard error of the mean. **g**, Representative deconvoluted mass spectra illustrating the MS response for ppcTnI corresponding to the plot in (**f**). M., protein marker; Std., endogenous cTnI protein standard. p*.,* phosphorylation. pp*,* bisphosphorylation. HSA, human serum albumin. C-HSA, cysteinylated human serum albumin. C-C-HSA, doubly cysteinylated human serum albumin. NP-Ctrl, unfunctionalized NP; NP-Pep, high affinity peptide-functionalized NP, Agarose-Ctrl, unfunctionalized agarose; Agarose-Pep, high affinity peptide-functionalized agarose; Agarose-mAb), antibody (mAb M46) functionalized with agarose. Source data are provided as a Source Data file.

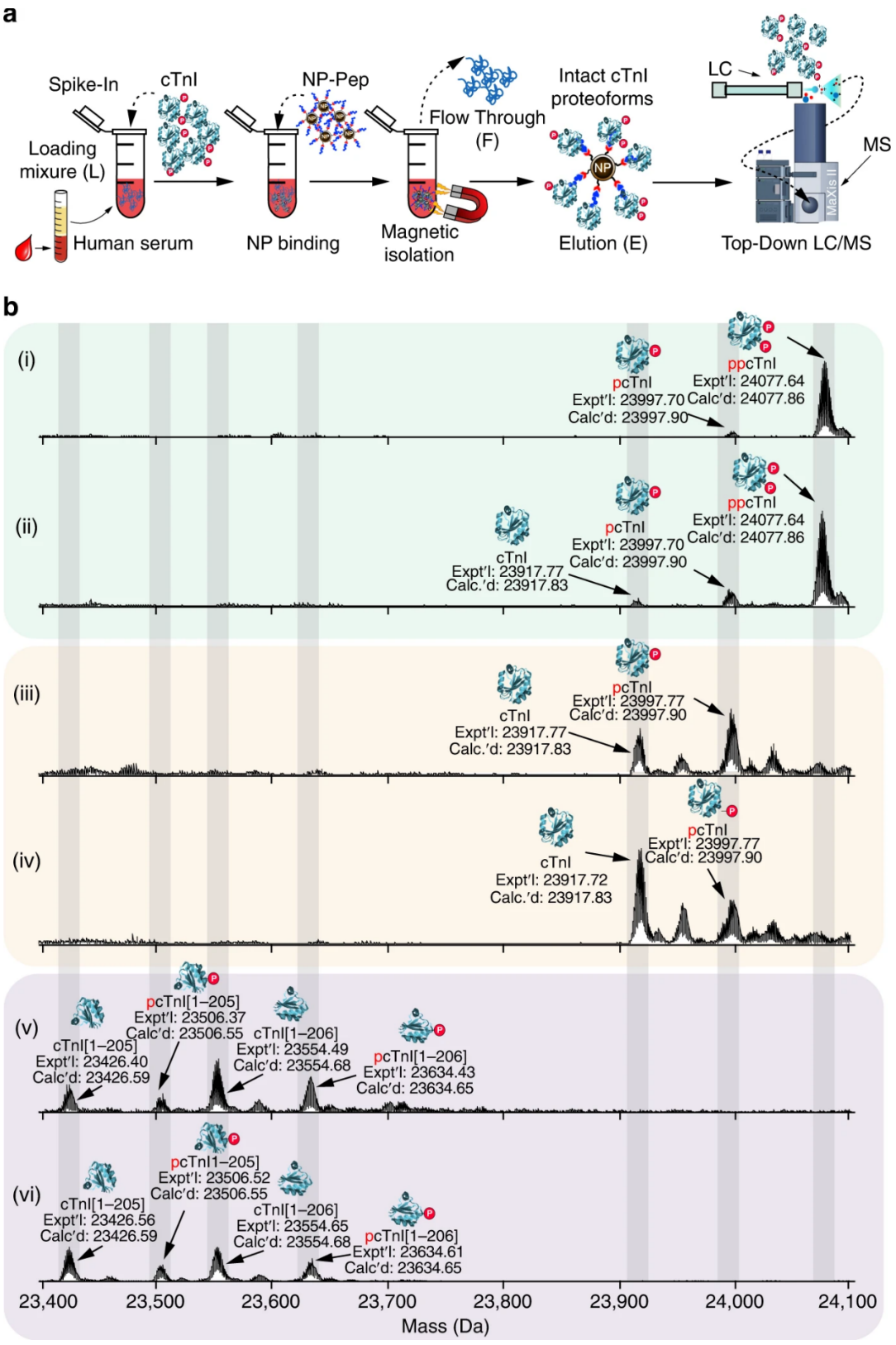
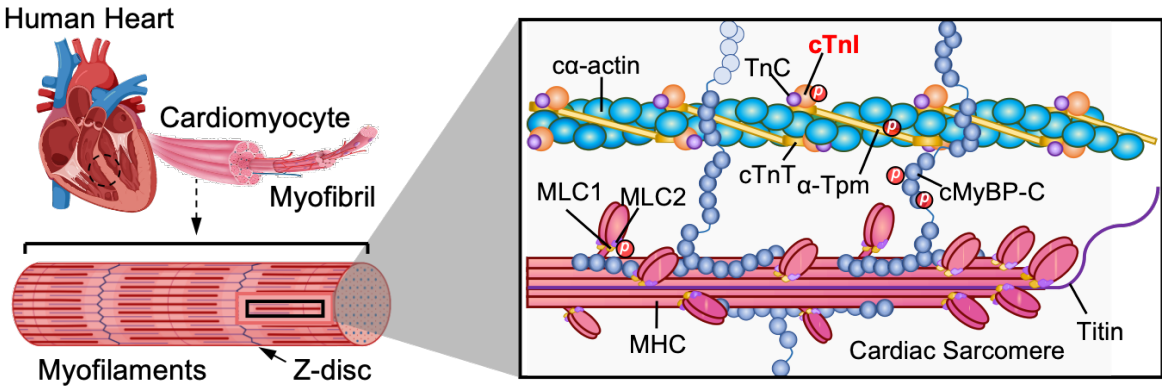
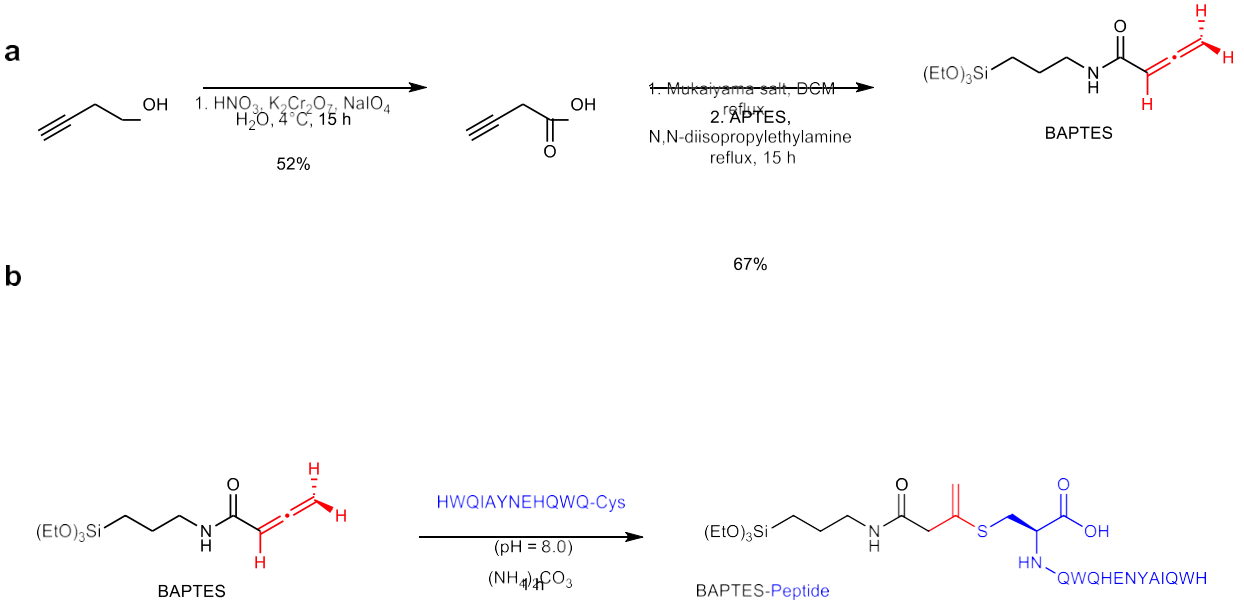

a
Spike-In
cTnI
Loading mixure (L)
Human serum
NP-Pep
NP binding
Flow Through (F)
Magnetic isolation
Intact cTnI proteoforms
NP
Elution (E)
LC
MS
MaXis II
Top-Down LC/MS
b
(i)
pcTnI
Expt'l: 23997.70
Calc'd: 23997.90
ppcTnI
Expt'l: 24077.64
Calc'd: 24077.86
(ii)
cTnI
Expt'l: 23917.77
Calc.'d: 23917.83
pcTnI
Expt'l: 23997.70
Calc'd: 23997.90
ppcTnI
Expt'l: 24077.64
Calc'd: 24077.86
(iii)
cTnI
Expt'l: 23917.77
Calc.'d: 23917.83
pcTnI
Expt'l: 23997.77
Calc'd: 23997.90
(iv)
cTnI
Expt'l: 23917.72
Calc.'d: 23917.83
pcTnI
Expt'l: 23997.77
Calc'd: 23997.90
(v)
cTnI[1–205]
Expt'l: 23426.40
Calc'd: 23426.59
pcTnI[1–205]
Expt'l: 23506.37
Calc'd: 23506.55
cTnI[1–206]
Expt'l: 23554.49
Calc'd: 23554.68
pcTnI[1–206]
Expt'l: 23634.43
Calc'd: 23634.65
(vi)
cTnI[1–205]
Expt'l: 23426.56
Calc'd: 23426.59
pcTnI1–205]
Expt'l: 23506.52
Calc'd: 23506.55
cTnI[1–206]
Expt'l: 23554.65
Calc'd: 23554.68
pcTnI[1–206]
Expt'l: 23634.61
Calc'd: 23634.65
23,400
23,500
23,600
23,700
23,800
23,900
24,000
24,100
Mass (Da)

Fig. 2.4 | Nanoproteomics enables comprehensive analysis of cTnI proteoforms from human serum. a, Nanoproteomics assay utilizing NP-Pep for specific enrichment of cTnI from serum and subsequent top-down MS analysis of cTnI proteoforms. cTnI is first spiked into human serum to prepare the loading mixture (L). The NPs are then incubated with the serum loading mixture, the cTnI-bound NPs are magnetically isolated, the unwanted and nonspecific proteins are removed as flow-through (F). The captured cTnI is then eluted and the final elution fraction after enrichment is analyzed by top-down LC/MS. **b**, Deconvoluted MS corresponding to cTnI proteoforms enriched from human serum. The cTnI (~10-20 ng/mL) spiked in the human serum (10 mg) were extracted from various human hearts: (i) and (ii), donor hearts; (iii) and (iv), diseased hearts with dilated cardiomyopathy, (v) and (vi), post-mortem hearts. cTnI proteoforms were identified using accurate intact mass measurement, using the most abundant mass based on the amino acid sequence of entry name TNNI3_human from the UniProtKB sequence database. p, phosphorylation. pp, bisphosphorylation. Data are representative of n = 3 independent experiments. A summary of the enriched proteoforms from serum and their respective mass measurements by top-down MS are listed in Supplementary Table 8**.** Source data are provided as a Source Data file.



Supplementary Figure 2.1

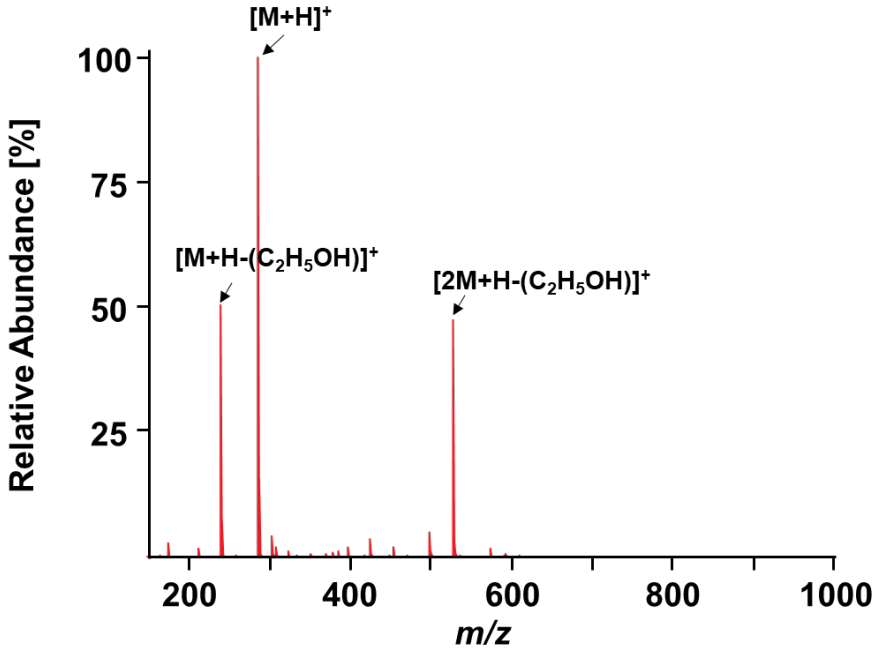
Schematic molecular composition of the cardiac sarcomere with the location of cardiac troponin I (cTnI) highlighted in Red**.** Cardiomyocytes are arranged into tubular structures called myofibrils and further sub-organized into cardiac sarcomeres, the smallest contractile units in themyofibril. Sarcomeres are composed of myofilaments (thin and thick filaments) flanked by denseprotein structures called Z-discs. Myofilaments include thin filament proteins [cTnI (cardiac troponin I), TnC (troponin C), cTnT (cardiac troponin T), c□-actin (cardiac alpha actin), and □- Tpm (alpha-tropomyosin)] and thick filament proteins [MHC (myosin heavy chain), MLC1 (essential light chain), MLC2 (regulatory light chain), cMyBP-C (cardiac myosin binding proteinC), and titin]. p denotes protein phosphorylation.



Supplementary Figure 2.2

Synthesis of N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide (BAPTES) organosilane monomer and peptide modification with BAPTES.

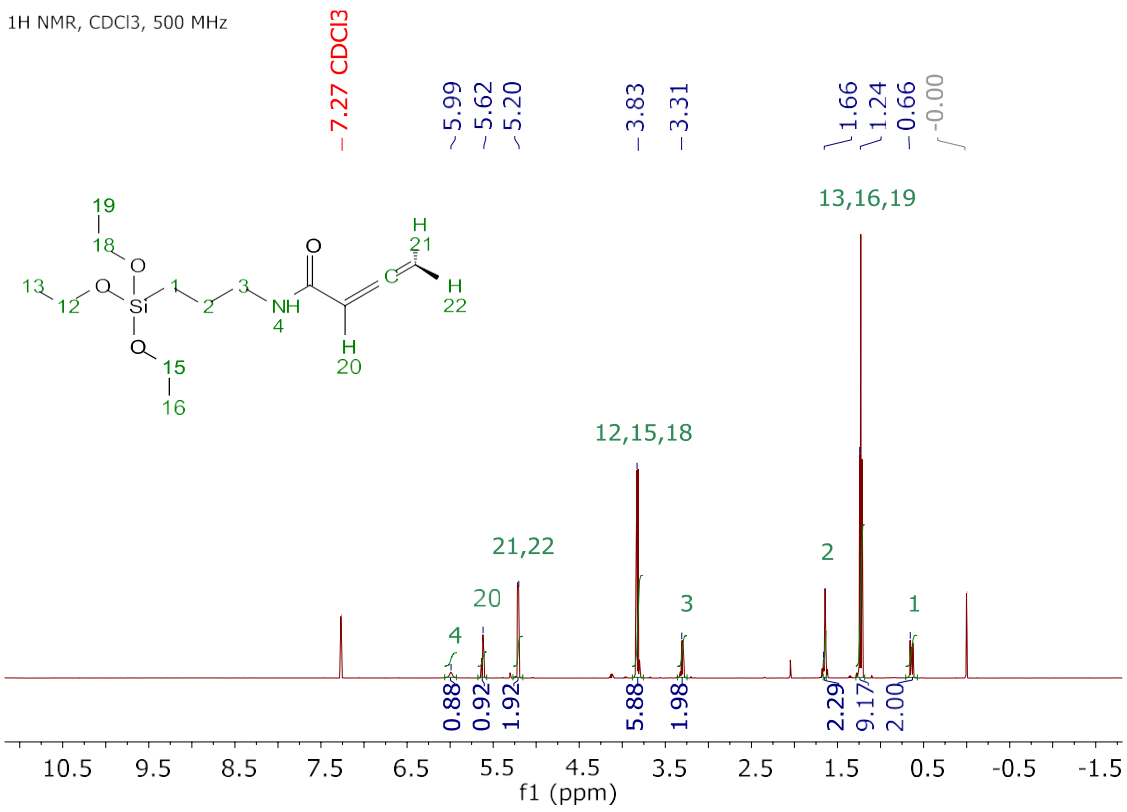
a, Scheme illustrating the facile synthesis of the BAPTES organosilane monomer used to silanize the 8 nm superparamagnetic Fe3O4 NPs. **b**, Cysteine modification of the high affinity cTnI-binding peptide with BAPTES monomer, illustrating potential functionalization of BAPTES surface-silanized NPs with cTnI-binding peptide. Yields are reported as the average of three independent



synthesis experiments.

Supplementary Figure 2.3

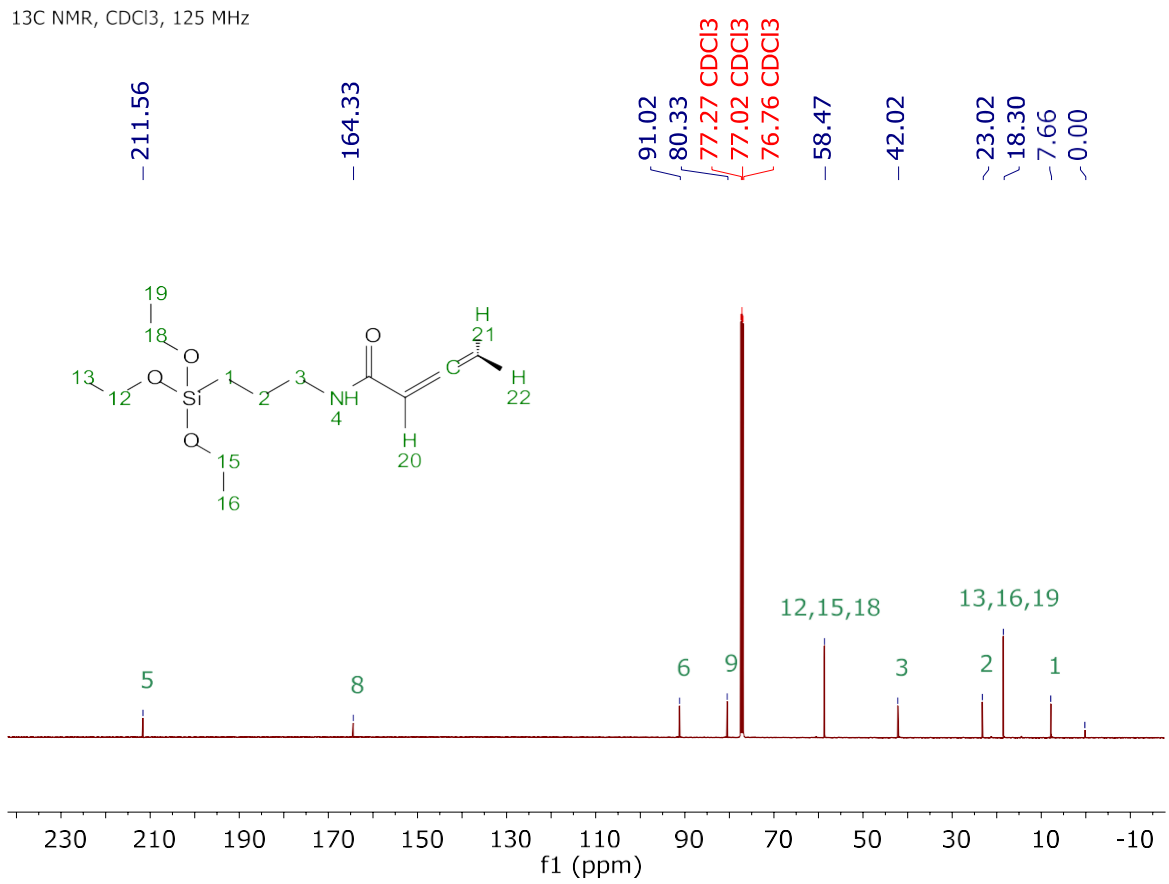
High resolution ESI-MS of BAPTES monomer. ESI-MS for $C_{13}H_{25}NO_4Si$ $[M+H]^+$ observed: 288.162 *m/z*, $[M+H]^+$ calculated: 288.162 *m/z*. Neutral mass losses are denoted on spectra: $[M+H-C_2H_5OH]^+$ observed: 242.120 *m/z,* $[2M+H-C_2H_5OH]^+$ observed: 529.276 *m/z*.



Supplementary Figure 2.4

^{1}H NMR of BAPTES monomer.

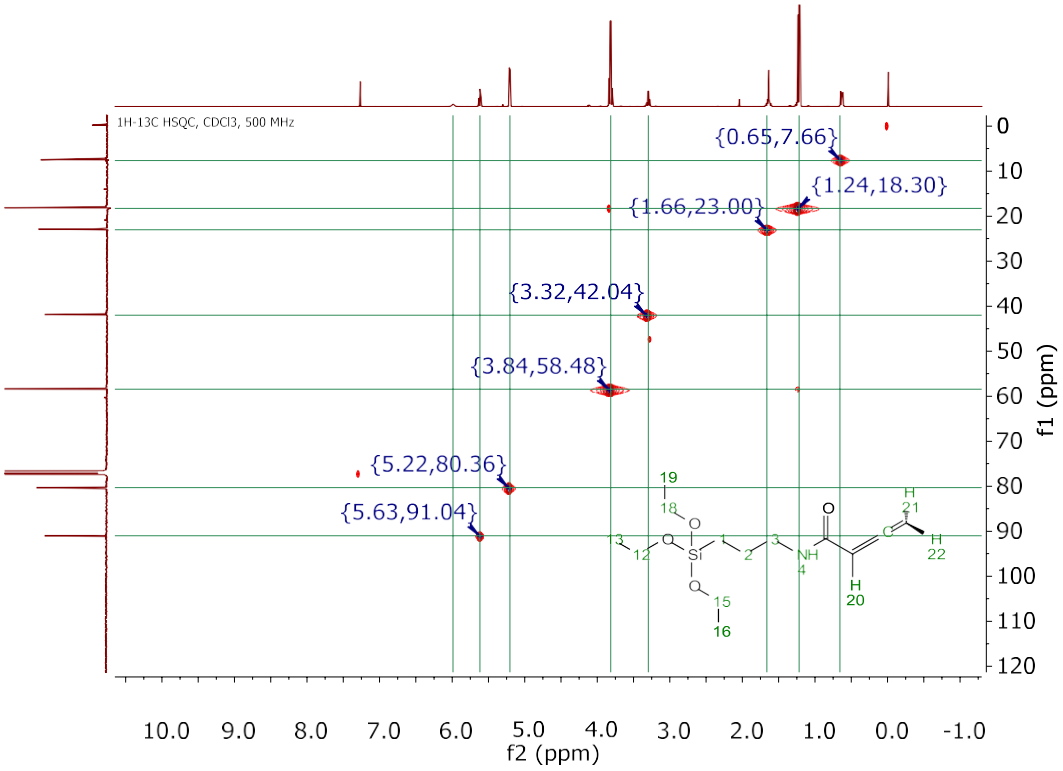
^{1}H NMR (500 MHz, Chloroform-*d*) δ (ppm) 5.99 (s, 1H), 5.62 (t, *J* = 6.6 Hz, 1H), 5.20 (d, *J* = 6.7Hz, 2H), 3.83 (q, *J* = 7.0 Hz, 6H), 3.31 (q, *J* = 6.7 Hz, 2H), 1.66 (m, 2H), 1.24 (t, *J* = 7.0 Hz, 9H),
0.66 (m, 2H). ^{1}H integrations confirm isolation and synthesis of BAPTES monomer.



Supplementary Figure 2.5

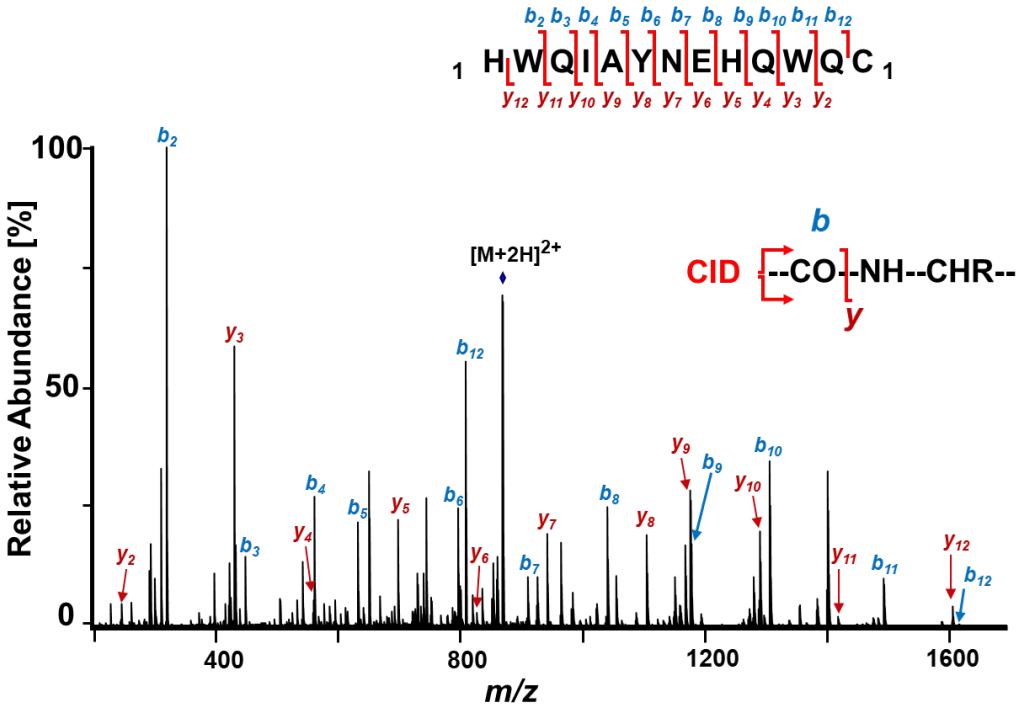
^{13}C NMR of BAPTES monomer.

^{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. ^{13}C NMR confirms the synthesis of the BAPTES monomer and the presence of 9 unique ^{13}C chemical environments.



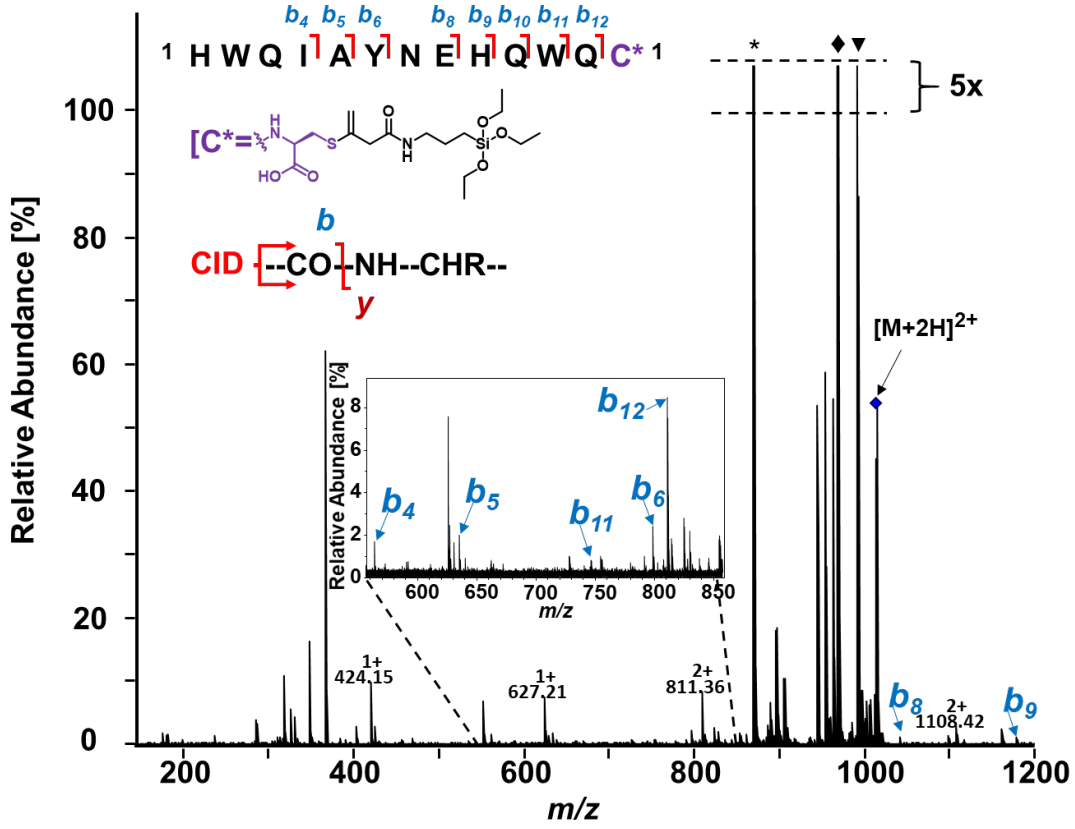
^{1}H-^{13}C Heteronuclear Single Quantum Coherence (HSQC) NMR of BAPTES monomer.

^{1}H-^{13}C HSQC (500 MHz, Chloroform-*d*) correlates the observed ^{1}H and ^{13}C signals present in the BAPTES molecule, confirming purification and synthesis of BAPTES monomer.



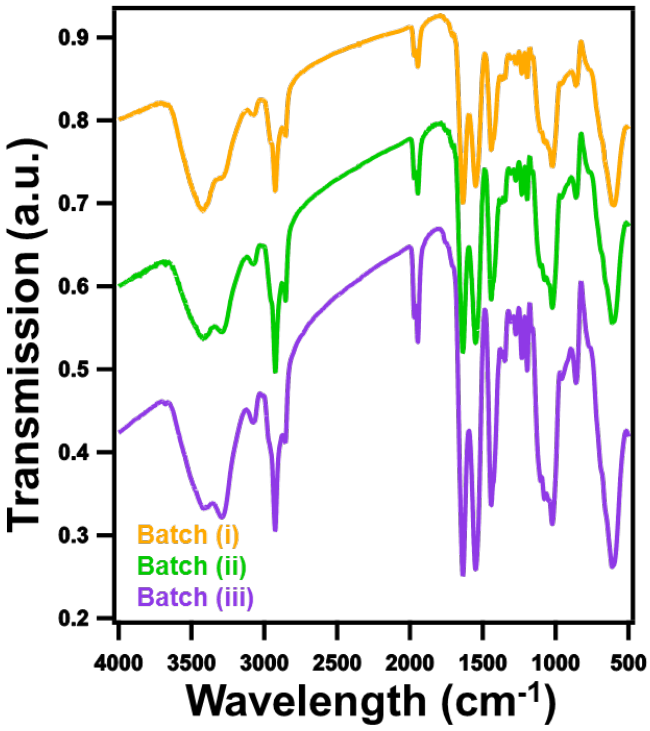
Supplementary Figure 2.7

ESI-MS of cTnI affinity peptide showing full sequence coverage.



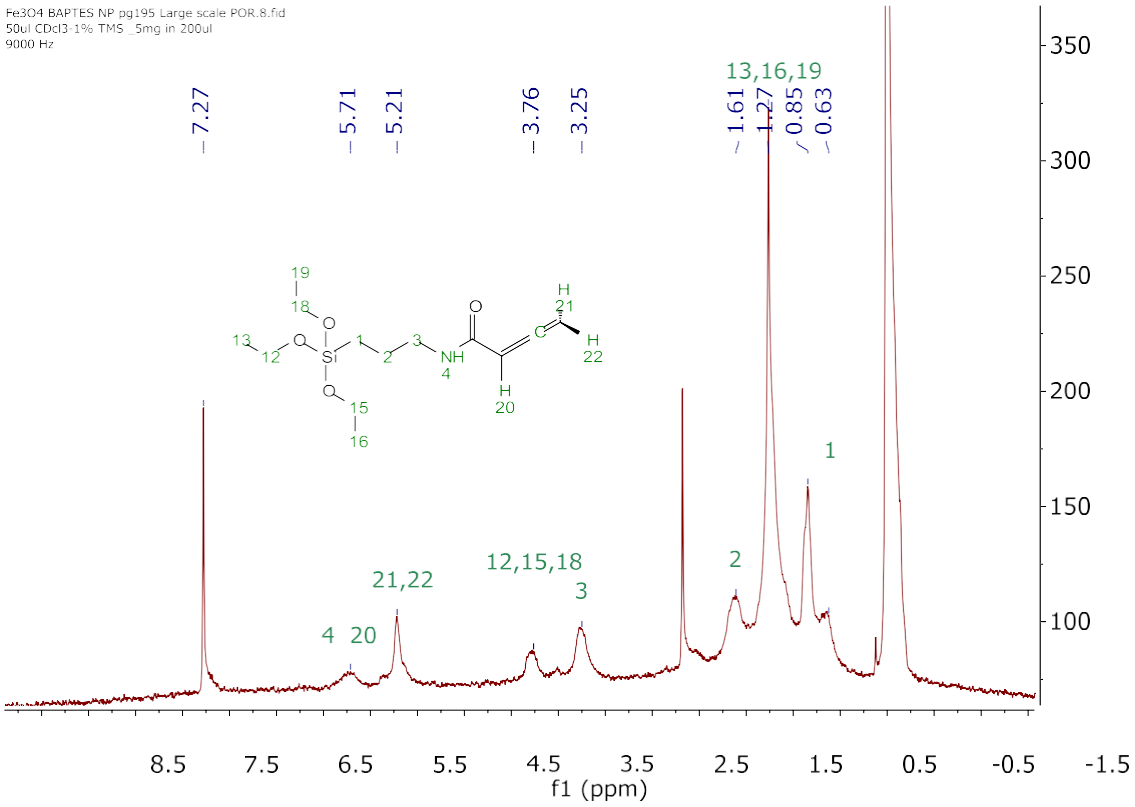
Supplementary Figure 2.8

The sequence of the BAPTES ($C_{13}H_{25}NO_4Si$)-modified cTnI affinity peptide was validated by tandem MS (MS/MS) sequence coverage using a 12 T Bruker Fourier transform ion cyclotron resonance (FTICR) mass spectrometer. Representative ions are denoted in the spectrum, with the inset showing an expansion of the 550-850 *m/z* range. A mass accuracy cut-off of 10 ppm was used for MS/MS fragment ion assignments. All ion identifications were manually validated. ESI-MS for $C_{92}H_{128}N_{24}O_{25}SiS$: $[M+2H]^{2+}$ observed: 1015.451 *m/z*, $[M+2H]^{2+}$ calculated: 1015.455 *m/z*. Neutral mass losses are denoted on spectra: Solid triangle (▼), $[M+2H-C_2H_5OH]^{2+}$; Solid diamond (♦), $[M+2H-2(C_2H_5OH)]^{2+}$; Asterisk (*), $[M+2H-C_{13}H_{25}NO_4Si]^{2+}$.



Supplementary Figure 2.9

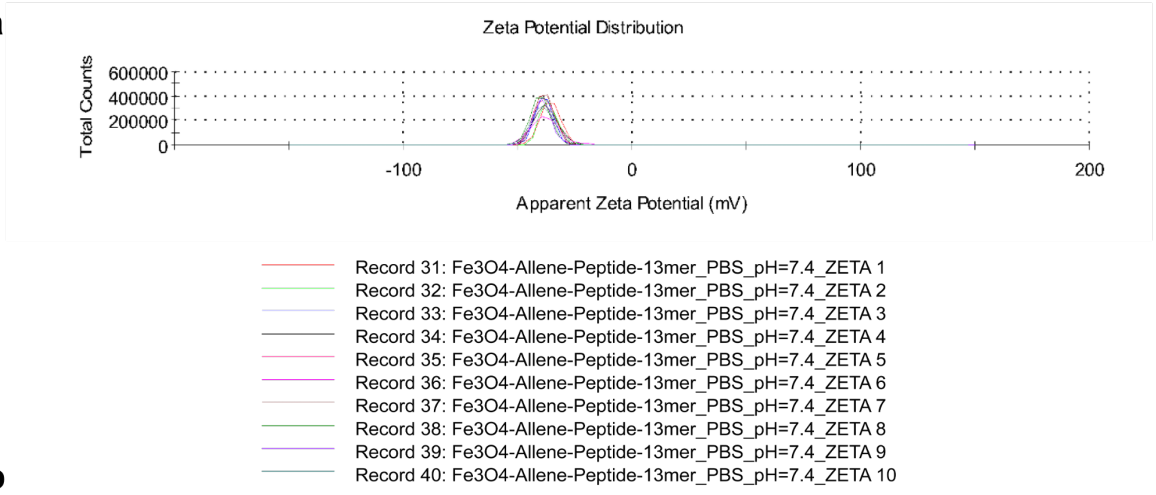
FTIR traces of three representative batches (i-iii) of Fe3O4-BAPTES NPs synthesized under the same reaction conditions. Traces are offset along the *y*-axis for clarity.



High resolution magic angle spinning nuclear magnetic spectroscopy (HRMAS) analysis of silanized NPs.

^{1}H NMR spectrum of Fe3O4-BAPTES NPs performed in Chloroform-*d* at a 9 kHz magic angle spinning (MAS) rate. δ (ppm) 5.71, 5.21, 3.76, 3.25, 1.61, 1.27, 0.85, 0.63. Characteristic allene carboxamide signatures belonging to the BAPTES monomer indicate the successful surface silanization of the starting Fe3O4-OA NPs to the Fe3O4-BAPTES NPs (i.e., unfunctionalized NPs, later referred to as NP-Ctrl).

a



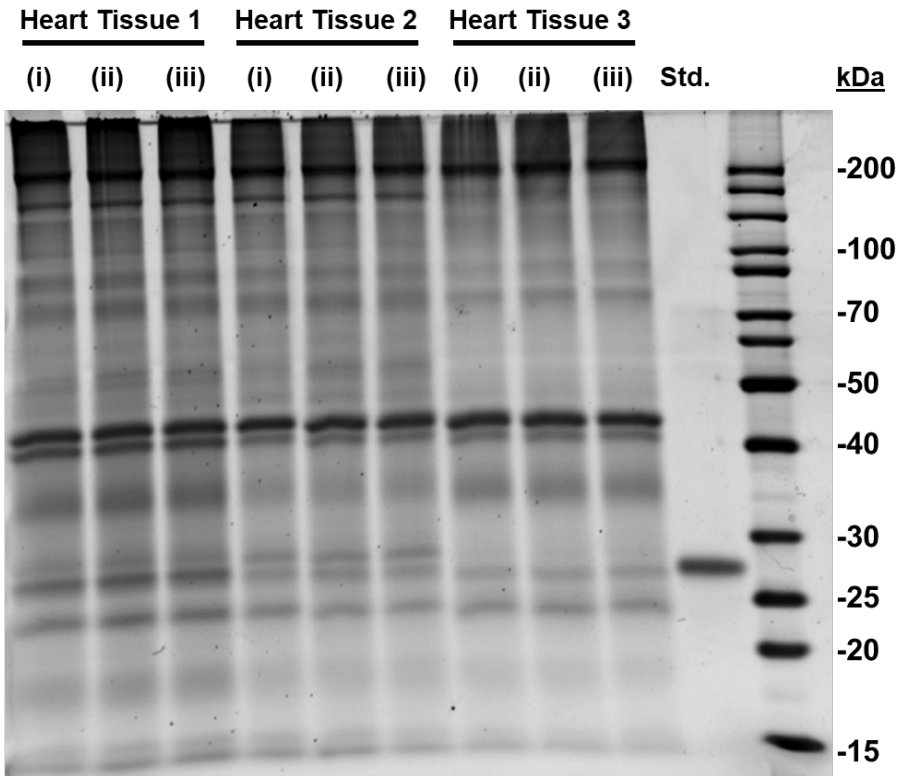
b

	Electrophoretic Mobility (μmcm/Vs)	Conductivity (mS/cm)	Zeta Potential (mV)
Average	-3.01	0.30	-38.2
Standard Deviation	0.10	0.0044	1.2
Relative Standard Deviation (%)	3.3	1.4	3.2

Supplementary Figure 2.11

Summary of zeta potential measurements on NP-Pep.

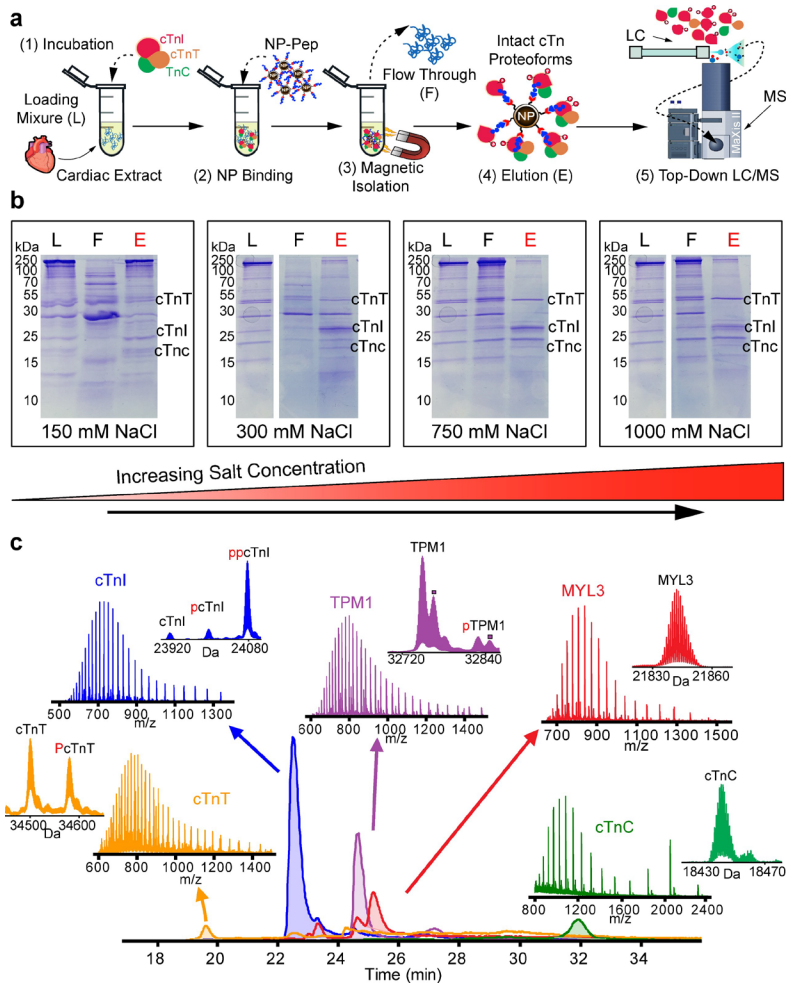
a, Plots of apparent zeta potentials for NP-Pep suspended in 0.1x PBS buffer (pH = 7.4) at a mass loading concentration of 0.25 mg/mL. Data are representative of n = 10 independent NP-Pep syntheses from the same NP-BAPTES batch. **b**, Table summary of electrophoretic mobility, conductivity, and zeta potential results as shown in (**a**).



Supplementary Figure 2.12

Reproducibility of sarcomeric protein extractions from heart tissue.

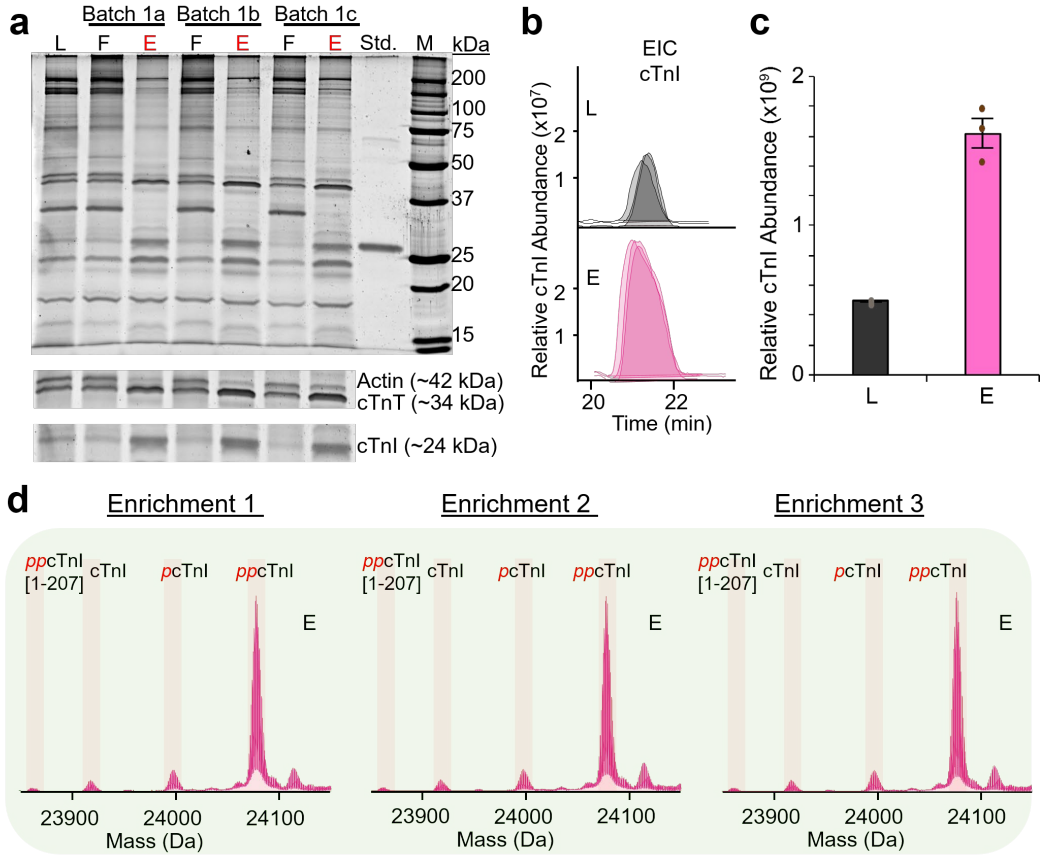
SDS-PAGE stained with SYPRO Ruby visualizing sarcomeric proteins obtained from the heart tissue protein extraction procedure. The heart tissue protein extraction procedure is listed in Supplementary Table 4. Three independent protein extractions (i-iii) were performed on three different heart tissue samples (1, donor tissue; 2, ischemic tissue; 3, post-mortem tissue). Equal protein amount (1 μg) was loaded in each lane, demonstrating the reproducibility of protein extractions within three biological samples. Std., human cTnI protein standard.



Optimization of salt concentration in the wash buffer for effective cTnI enrichment.

a, Schematic illustration of the nanoproteomics strategy for enrichment of cTnI and its associating protein-protein interactors and top-down LC/MS analysis of proteins associated with cTnI. **b**, SDS-PAGE stained with Coomassie blue, visualizing that human cardiac sarcomere protein interacting partners of cTnI can be obtained during enrichment by simply tuning the salt concentration (150-1000 mM NaCl) of the wash buffer during the enrichment workflow. Equal amount (2 µg) of the loading mixture (L), flow-through (F), and elution mixture (E) after enrichment was loaded on the gel. **c,** Normalized extracted ion chromatograms and MS analysis of sarcomeric proteins obtained from the elution fraction after NP-Pep enrichment using 300 mM NaCl washing buffer. Deconvoluted MS for each protein is presented. 300 mM NaCl was ultimately used for all sarcomeric protein extract enrichment as it provided an optimal balance between cTnI enrichment specificity and overall cTnI recovery. TMP1 (□-tropomyosin), and MYL3 (myosin light chain 3) are named from the UniProtKB sequence database. Squares indicate

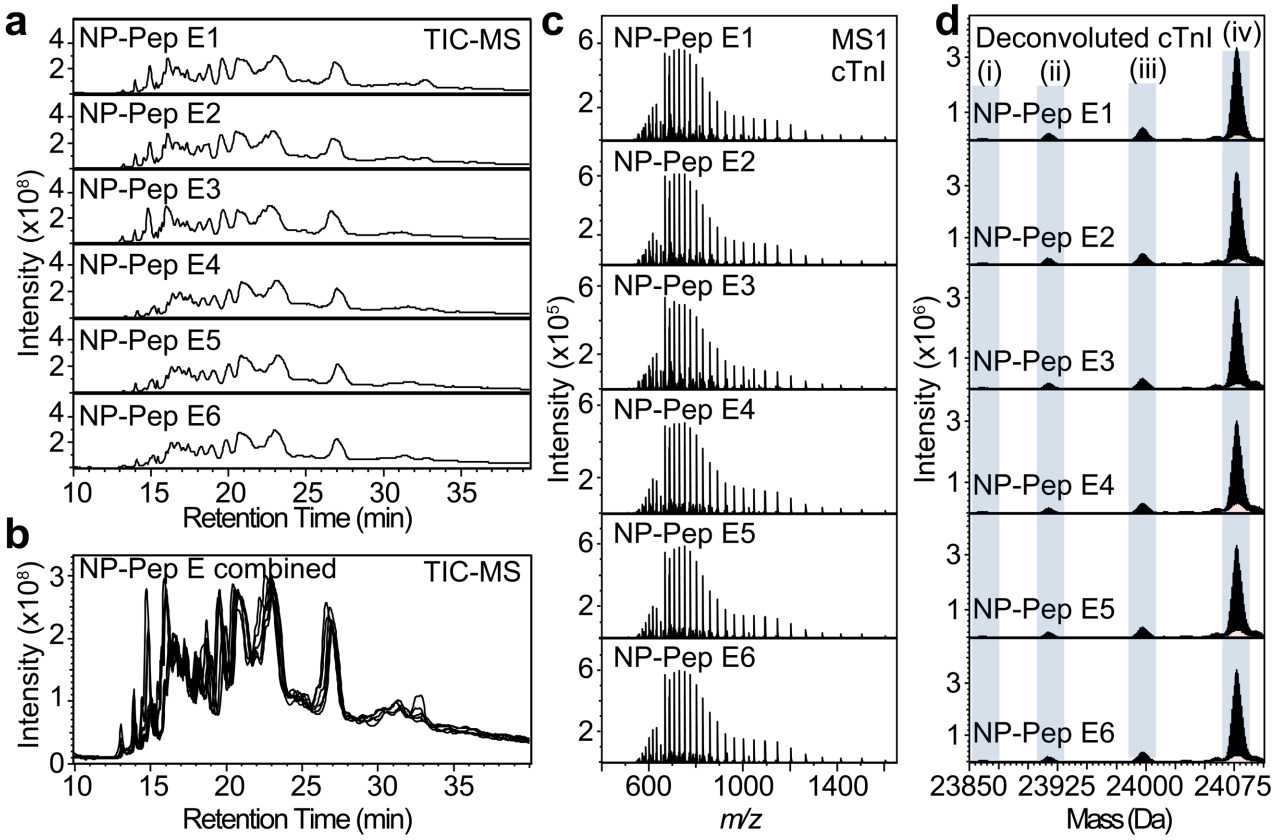
+16 Da proteoform. p, phosphorylation; pp, bisphosphorylation.



Supplementary Figure 2.14

a, SDS-PAGE stained with SYPRO Ruby visualizing the cTnI enrichment performance and demonstrating the high reproducibility obtained from the same synthesis batch of Fe3O4-BAPTES NPs (intra-batch). Equal amounts of NP-Pep (5 mg) from the same synthesis batch were used for cTnI enrichment from sarcomeric protein extracts (300 µg) containing 0.3% cTnI obtained from a human donor heart. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. Lower panels display enlarged SDS-PAGE strips highlighting NP enrichment of cTnI (~24 kDa) and cTnT (~34 kDa) from the sarcomeric protein loading mixture. **b**, Normalized extracted ion chromatograms (EIC) of cTnI from the sarcomeric protein loading mixture (L) and elution mixture (E). **c**, Relative abundance of cTnI from L and E when normalizing for 500 ng total protein loading, demonstrating significant enrichment of cTnI in E. **d**, Normalized deconvoluted mass spectra of cTnI present in the E, demonstrating consistency of cTnI enrichment across three enrichment replicates performed using NPs synthesized from a single batch (intra-batch). N-terminally acetylated cTnI proteoforms following Met exclusion were identified using accurate intact mass measurement, using the most abundant mass based on the amino acid sequence of entry name TNNI3_human from the UniProtKB sequence database. p, phosphorylation; pp, bisphosphorylation. Std., human cTnI

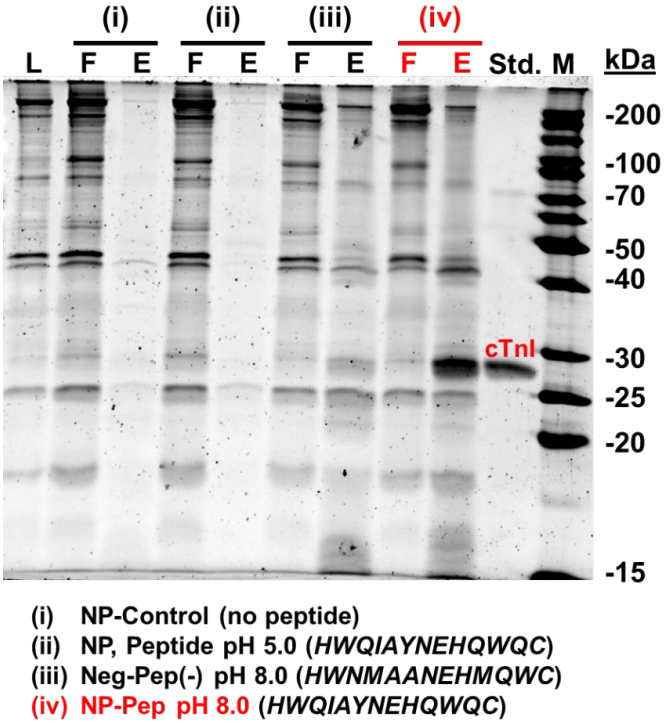
protein standard; M., protein marker. Data are representative of n = 3 independent experiments and error bars indicate standard error of the mean.



Supplementary Figure 2.15

Top-down LC/MS cTnI enrichment reproducibility by NP-Pep.

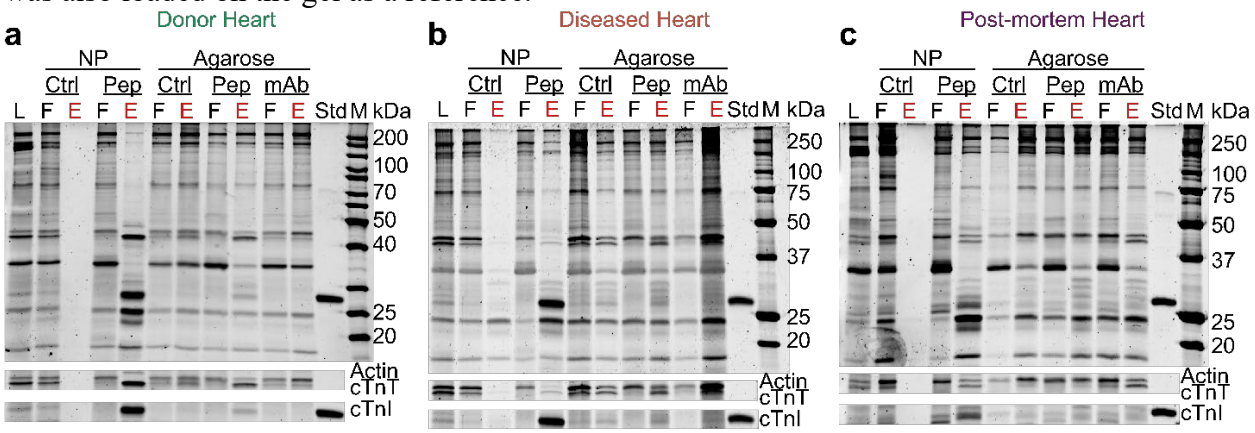
a-b, Total ion chromatogram mass spectra (TIC-MS) of six independent NP-Pep elution mixtures (E) obtained from human heart extracts. The elution mixtures are shown for each run (**a**) and overlaid in a single plot (**b**). Equal amounts of NP-Pep (5 mg) were used for cTnI enrichment containing 0.3% cTnI obtained from a human donor heart. **c**, Raw MS1 of cTnI obtained from the NP-Pep elution mixtures corresponding to (**a-b**). **d**, Deconvoluted mass spectra corresponding to enriched cTnI in (**a-c**). Roman numerals correspond to N-terminally acetylated cTnI proteoforms following Met exclusion: (i) ppcTnI[1-207]; (ii) cTnI; (iii) pcTnI; (iv) ppcTnI. cTnI proteoforms were identified using accurate intact mass measurement, using the most abundant mass based on the amino acid sequence of entry name TNNI3_human from the UniProtKB sequence database. Data correspond to the results shown in Fig. 2 and Supplementary Fig 14.



Supplementary Figure 2.16

Multiple NP controls (including unfunctionalized NPs or NPs functionalized sub-optimally) were compared against the NPs functionalized with the cTnI-binding peptide, demonstrating NPs functionalized with the correct cTnI-binding peptide (NP-Pep) are crucial for the successful enrichment. SDS-PAGE stained with SYPRO Ruby was used to compare the cTnI enrichment performance of i) unfunctionalized Fe3O4-BAPTES NPs with no peptide coupled (NP-Control); ii), Fe3O4-BAPTES NPs incubated with the high affinity cTnI peptide at pH 5.0 (NP, peptide pH 5.0); iii), Fe3O4-BAPTES NPs functionalized with a cTnI-negative control peptide (as shown in aprevious report[6]) at pH 8.0 (Neg-Pep), and iv) the Fe3O4-BAPTES NPs functionalized with the high affinity cTnI peptide at pH 8.0 (NP-Pep). In (i), the unfunctionalized NPs do not appreciably enrich protein as seen in NP-control elution lane, demonstrating the resistance of the NPs to non-specific protein adsorption. In (ii), the reaction of the cysteine-thiol with BAPTES is inhibited at acidic pH (pH 5.0), leading to poor NP-peptide conjugation and downstream enrichment of cTnI. In (iii), the reaction of the cysteine-thiol with BAPTES occurs readily and specifically at slightly alkaline pH (pH 8.0), leading to NP-peptide conjugation. However, the use of a negative-binding sequence peptide with diminished binding affinity in (iii) results in poor cTnI enrichment. In (iv), the use of the high affinity cTnI binding peptide at slightly alkaline (pH 8.0) coupling conditions results in effective cTnI enrichment. cTnI is clearly enriched only in the NP-Pep elution lane, illustrating that proper surface chemistry and functionalization of the NP platform is critical for enabling cTnI enrichment. All NP samples were subjected to the same enrichment workflow. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. 200 ng loading of endogenous human cTnI standard (Std.)

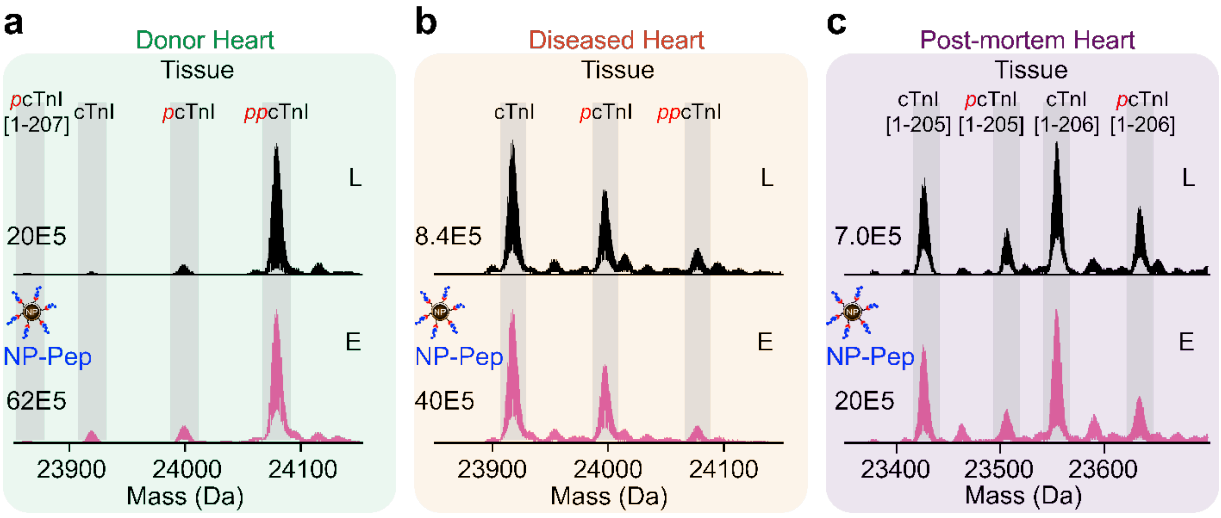
was also loaded on the gel as a reference.



Supplementary Figure 2.17

Comparison of cTnI enrichment performance between the NP and Agarose platforms using sarcomeric protein mixtures extracted from a human donor heart, a diseased dilated cardiomyopathic heart, and a post-mortem heart extract.

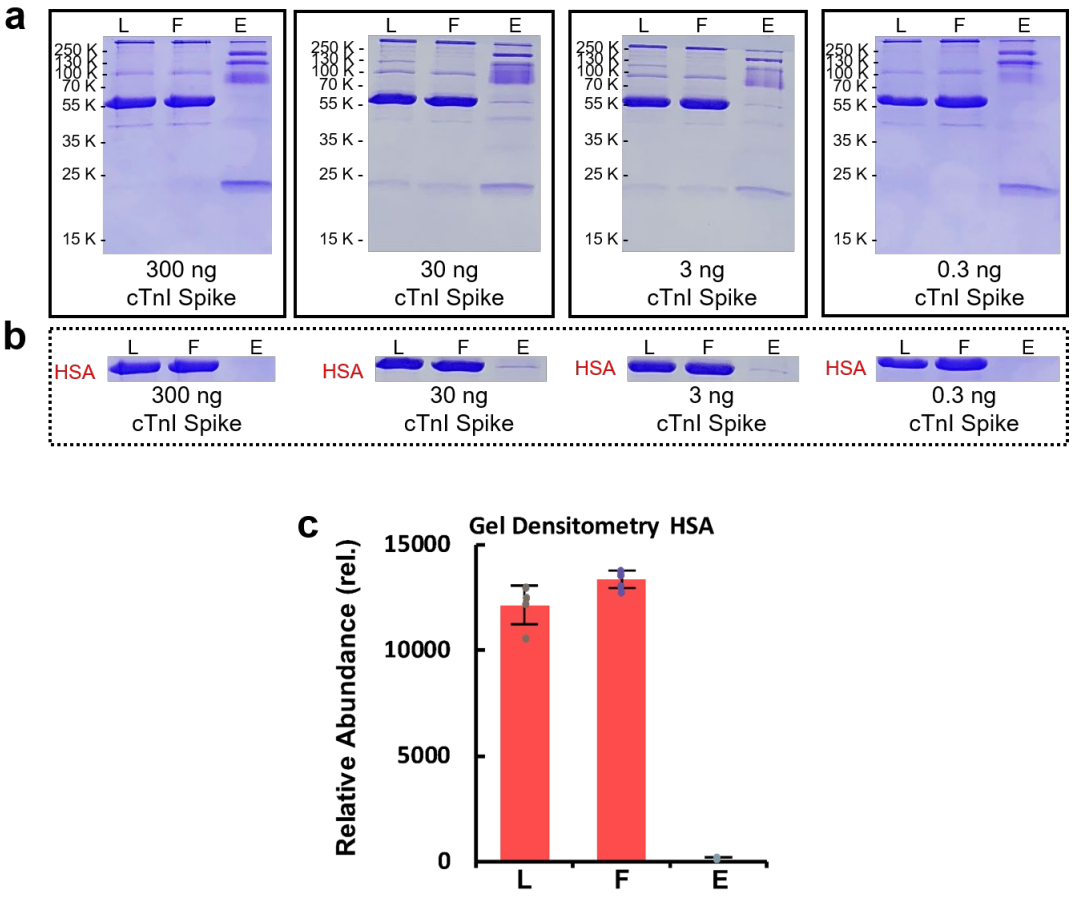
a-c, SDS-PAGE stained with SYPRO Ruby visualizing the enrichment performance of the NP control (NP-Ctrl, no peptide), high affinity peptide-functionalized NP (NP-Pep), agarose-control (Agarose-Ctrl), high affinity peptide-functionalized agarose (Agarose-Pep), and antibody (mAb M46) functionalized agarose (Agarose-mAb). Sarcomeric protein extract (100 µg) obtained from a human donor heart (**a**), a diseased dilated cardiomyopathic heart (**b**), and a post-mortem heart (**c**) were used as the loading mixture, such that cTnI was not loaded in excess relative to the targeting ligands functionalized on the NPs or agarose. cTnI concentrations are listed in Supplementary Table 6. Lower panels display enlarged SDS-PAGE strips focusing on enrichment of cTnT and cTnI from (L). Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. The NP-Pep demonstrates the most effective enrichment across the different platforms examined here.



Supplementary Figure 2.18

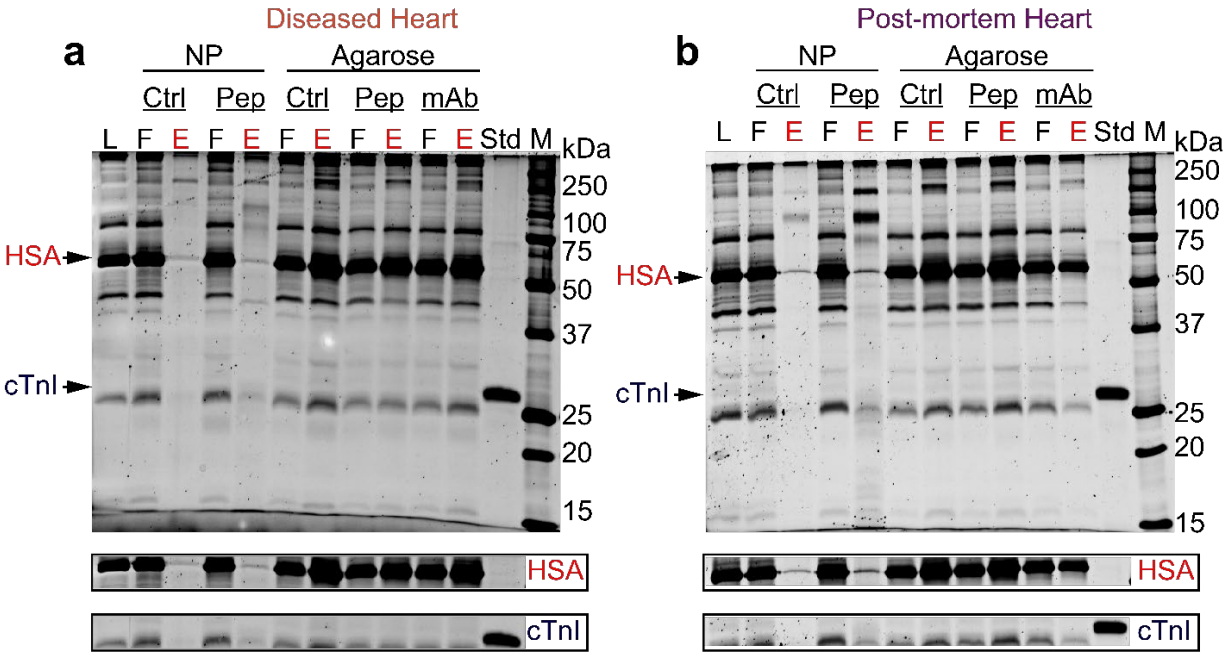
Preservation of cTnI PTM profiles by the NP-Pep platform using sarcomeric protein extracts from a human donor heart, a dilated cardiomyopathic heart (diseased heart), and a post-mortem heart.

a-c, Deconvoluted mass spectra corresponding to reference cTnI proteoforms originating from a donor heart (**a**), a diseased dilated cardiomyopathic heart (**b**), and a post-mortem heart (**c**) before and after enrichment using the high affinity peptide-functionalized NP (NP-Pep). Equal amounts (500 ng) were loaded for LC/MS analysis. Data correspond to the same cTnI proteoform enrichment data presented in Fig. 2f-h, but are displayed at the same height to facilitate comparison (signal intensity is shown for each spectrum). Concentrations of cTnI used are listed in Supplementary Table 6. p, phosphorylation; pp, bisphosphorylation.



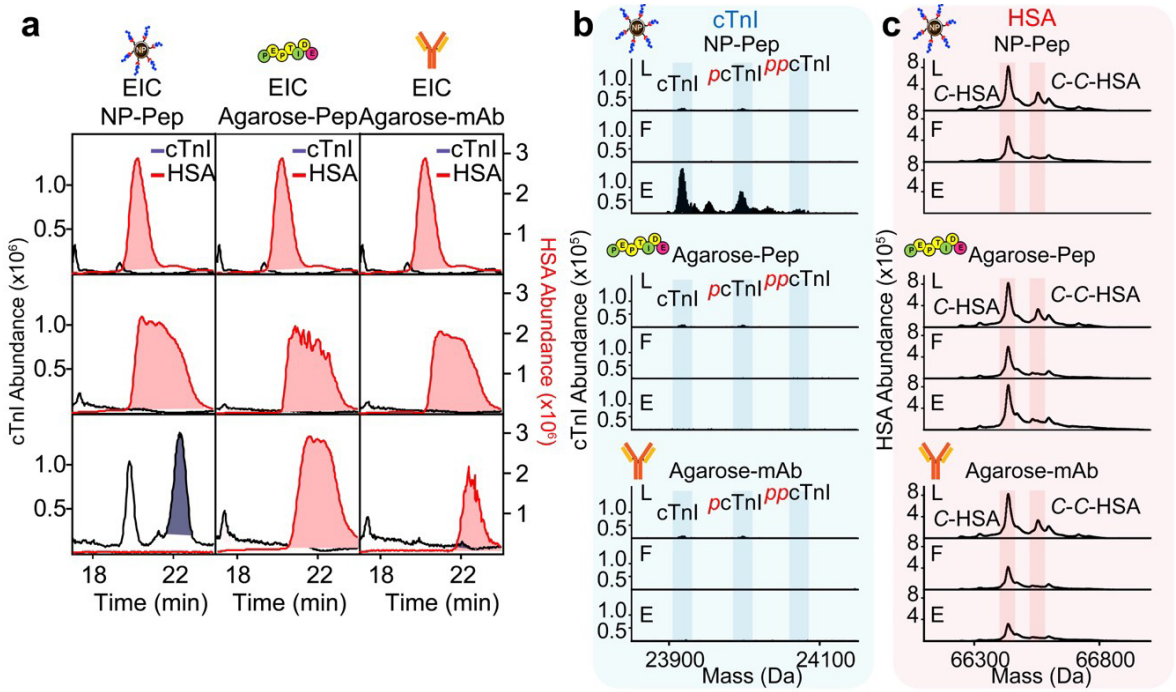
Effective simultaneous depletion of human serum albumin (HSA) from protein mixture during NP-Pep enrichment of cTnI.

a, SDS-PAGE stained with Coomassie blue demonstrating highly effective simultaneous HSA depletion during cTnI enrichment using the NP platform from human serum (20 mg) at various spike-in cTnI concentrations (0.3 to 300 ng/mL). Equal amount (5 µg) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. **b**, Magnification of the HSA lane corresponding to (**a**). **c**, Gel densitometry relative quantitation of HSA depletion corresponding to the results shown in (**a**) and (**b**). HSA was depleted ~215-fold relative to the loading mixture following NP enrichment by ImageJ analysis. Data are representative of n = 4 independent experiments and error bars indicate standard deviation from the mean.



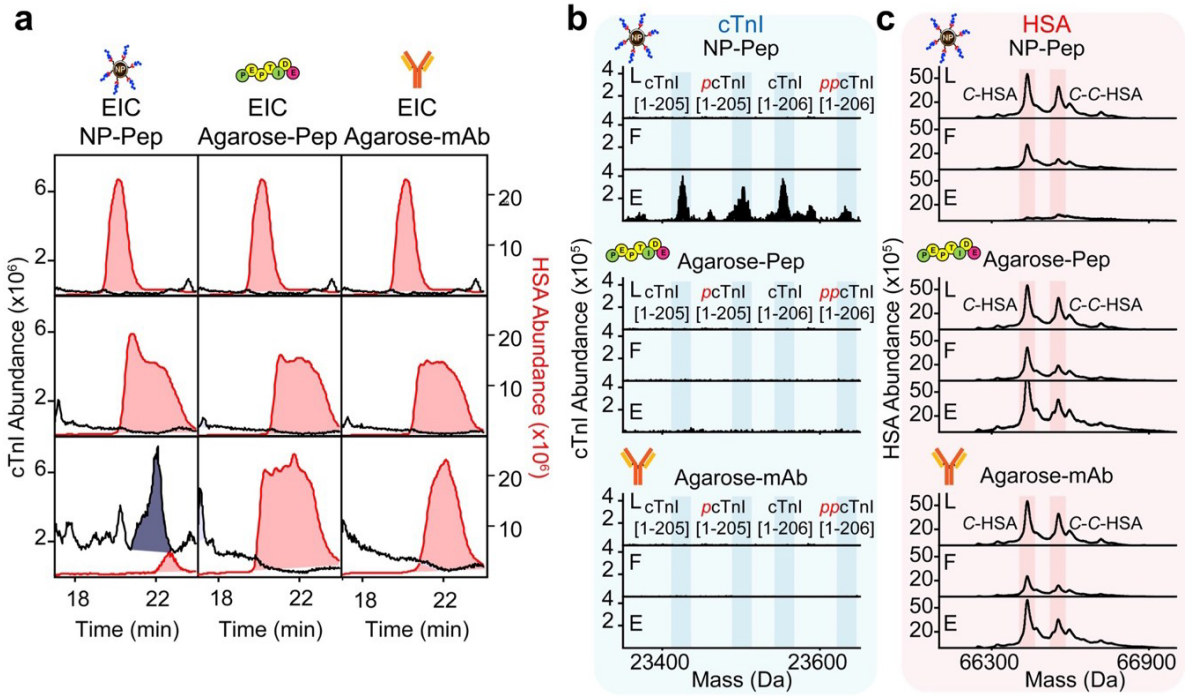
Supplementary Figure 2.20

SDS-PAGE stained with SYPRO Ruby visualizing the enrichment performance of the NP control (NP-Ctrl, no peptide), high affinity peptide-functionalized NP (NP-Pep), agarose-control (Agarose-Ctrl), high affinity peptide-functionalized agarose (Agarose-Pep), and antibody (mAb M46) functionalized agarose (Agarose-mAb). Sarcomeric protein extract (100 µg) containing reference cTnI obtained from a diseased dilated cardiomyopathic heart (**a**) and a post-mortem heart (**b**) were spiked in human serum (10 mg) and used as the loading mixture, such that cTnI was not loaded in excess relative to the targeting ligands functionalized on the NPs or agarose. cTnI concentrations are listed in Supplementary Table 6. Lower panels display enlarged SDS-PAGE strips highlighting the depletion of HSA from the NP elution and the enrichment of cTnI. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel.



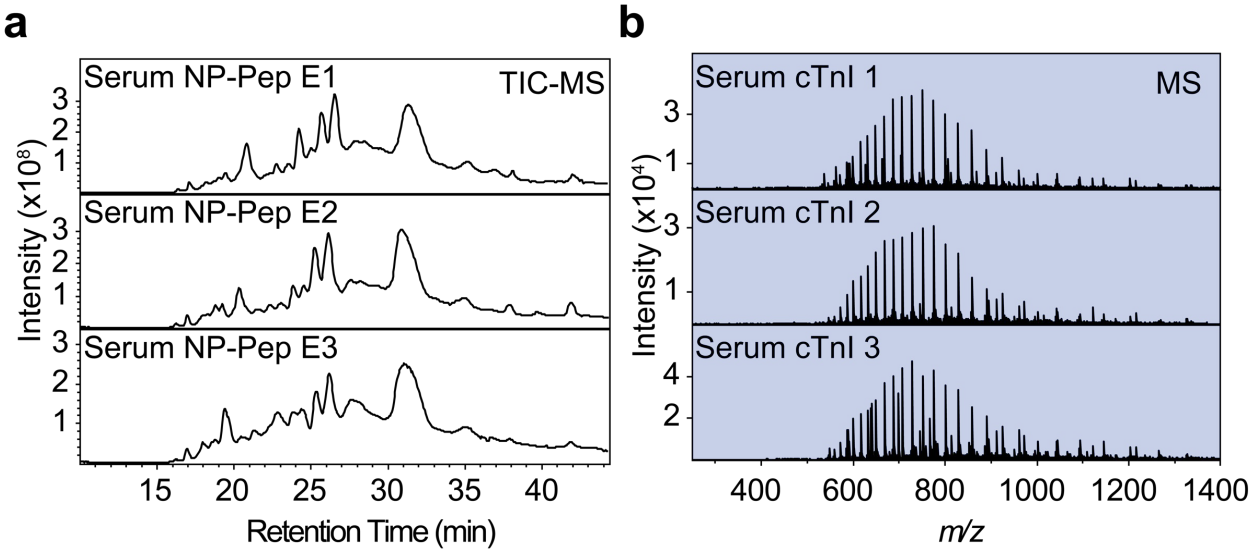
Supplementary Figure 2.21

a, Normalized extracted ion chromatograms (EICs) of cTnI (black) and HSA (red) corresponding to high affinity peptide-functionalized NP (NP-Pep), high affinity peptide-functionalized agarose (Agarose-Pep), and mAb-functionalized agarose (Agarose-mAb) enrichment from a human serum mixture with spike-in cTnI. cTnI (482 ng) from a human dilated cardiomyopathy heart extract was spiked into human serum (10 mg). **b**, Normalized deconvoluted mass spectra of cTnI obtained from the L, F, and E mixtures, corresponding to the data presented in (**a**). The high affinity peptide functionalized NPs (NP-Pep) show superior cTnI enrichment from human serum compared to the agarose functionalized with either the same high affinity peptide or a mAb targeting for the same epitope region as the peptide. **c**, Normalized deconvoluted mass spectra of HSA obtained from the L, F, and E mixtures and corresponding to the data presented in (**a**). The NP-Pep demonstrates superior depletion of the non-specific HSA to Agarose-Pep and Agarose-mAb. cTnI proteoforms (N-terminally acetylated following Met excision) were identified using accurate intact mass measurement, using the most abundant mass based on the amino acid sequence of entry name TNNI3_human from the UniProtKB sequence database. (L), loading mixture; (F), flow-through; (E), elution mixture; p, phosphorylation; pp, bisphosphorylation. *C*-HSA, cysteinylated human serum albumin; *C*-*C*-HSA, doubly-cysteinylated human serum albumin.



Supplementary Figure 2.22

LC/MS evaluation of cTnI enrichment from human serum with spike-in cTnI obtained from a post-mortem human heart extract, using the NP-Pep and Agarose-mAb platforms.



Supplementary Figure 2.23

Top-down LC/MS serum cTnI enrichment reproducibility and characterization by NP-Pep. **a**, Total ion chromatogram mass spectra (TIC-MS) of three independent NP-Pep elution mixtures

(E) obtained from a serum spike-in cTnI enrichment. Equal amounts of NP-Pep (5 mg) from separate synthesis batch were used for cTnI enrichment from human serum (10 mg) containing a minimal spike-in of cTnI (final concentration of 18.7 ng/mL), obtained from a human donor heart. **b**, Raw MS1 of cTnI obtained from the NP-Pep elution mixtures corresponding to (**a**).

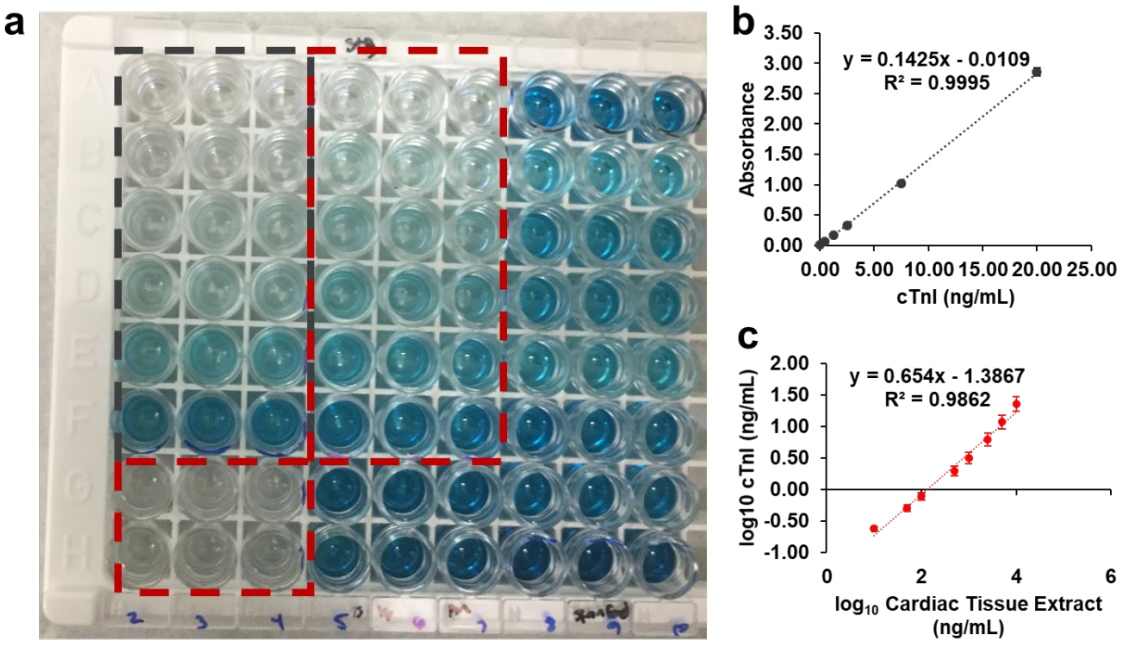
Top 12 Most Frequently Identified Proteins Enriched from Serum

Gene Code	Protein Name	Molecular Weight (Da)	E-Value
TNNI3	Troponin I	24063.6894	3.78E-10
TNNC1	Troponin C	18498.5755	1.19E-18
LDB3	Isoform 6 of LIM domain-binding protein 3	30890.7825	2.88E-15
CRYAB	Alpha-crystallin B chain	20188.4295	7.99E-17
RS10	40S ribosomal protein S10	18941.921	1.02E-14
APOA1	Apolipoprotein A-I	11299.7765	1.68E-07
MYL3	Myosin light chain 3	21846.1696	5.07E-10
MLRV	Myosin regulatory light chain 2	18726.18	1.04E-09
H2B1D	Histone H2B type 1-D	10725.7183	9.55E-09
H2A2C	Histone H2A type 2-C	11000.1674	3.78E-15
APOE	Apolipoprotein E	34263.6597	6.71E-14
TNNT2	Isoform 11 of Troponin T	34573.645	1.42E-03

Identified in all n = 5 idependent serum enrichments
Identified in n = 4 idependent serum enrichments
Identified in n = 3 idependent serum enrichments

Supplementary Figure 2.24

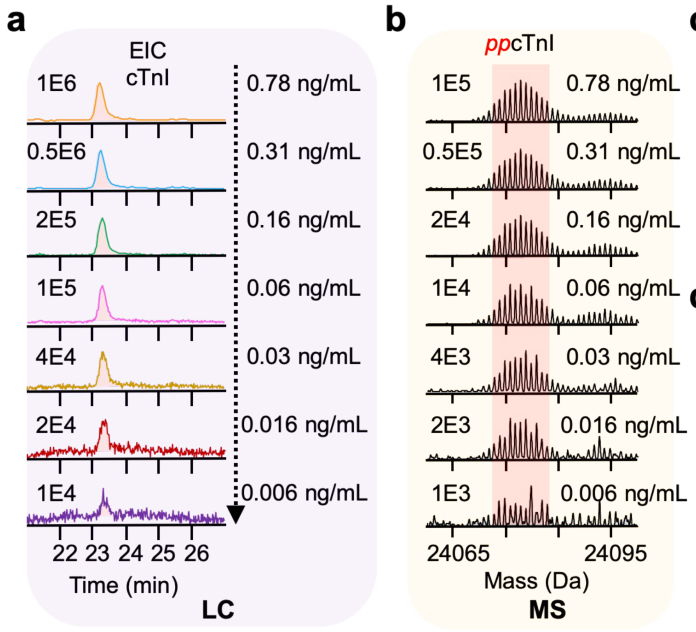
Summary and illustration of proteins identified by NP-Pep following serum enrichment. Table summarizing the top 12 most frequently identified proteins enriched by the NP-Pep from serum cTnI spike-in enrichments obtained by combining n = 5 independent enrichments. cTnI (TNNI3) is confidently identified in all enrichment trials. Protein E-Value score is reported for each identification.



Supplementary Figure 2.25

ELISA-based cTnI quantification and Bradford total protein assay-based quantification of human serum with cTnI spike-in.

a, ELISA-based colorimetric quantification of cTnI standards (gray dashed box) and serum samples incrementally spike-in with sarcomeric protein extracts containing cTnI (red dashed boxes). All samples were dispensed in triplicate and the assay was performed according to the manufacturer's instructions. The ELISA assay uses capture antibodies targeting cTnI amino acids 18-28 and 86-90, with detection antibodies targeting cTnI amino acids 41-49. **b**, ELISA-based standard curve (0 ng, 0.4 ng, 1.25 ng, 2.5 ng, 7.5 ng, 20 ng) used for quantification of cTnI spike-in human serum, and quantification of MS-based sensitivity of cTnI detection. **c**, Log-log plot of cTnI concentration (ng/ml) quantified by ELISA, as a function of total sarcomeric protein spiked into serum (ng/ml) quantified by Bradford protein assay. Approximately 1 µg of sarcomeric protein extract contains 3 ng of cTnI. Data are representative of n = 3 independent experiments and error bars indicate standard deviation from the mean.

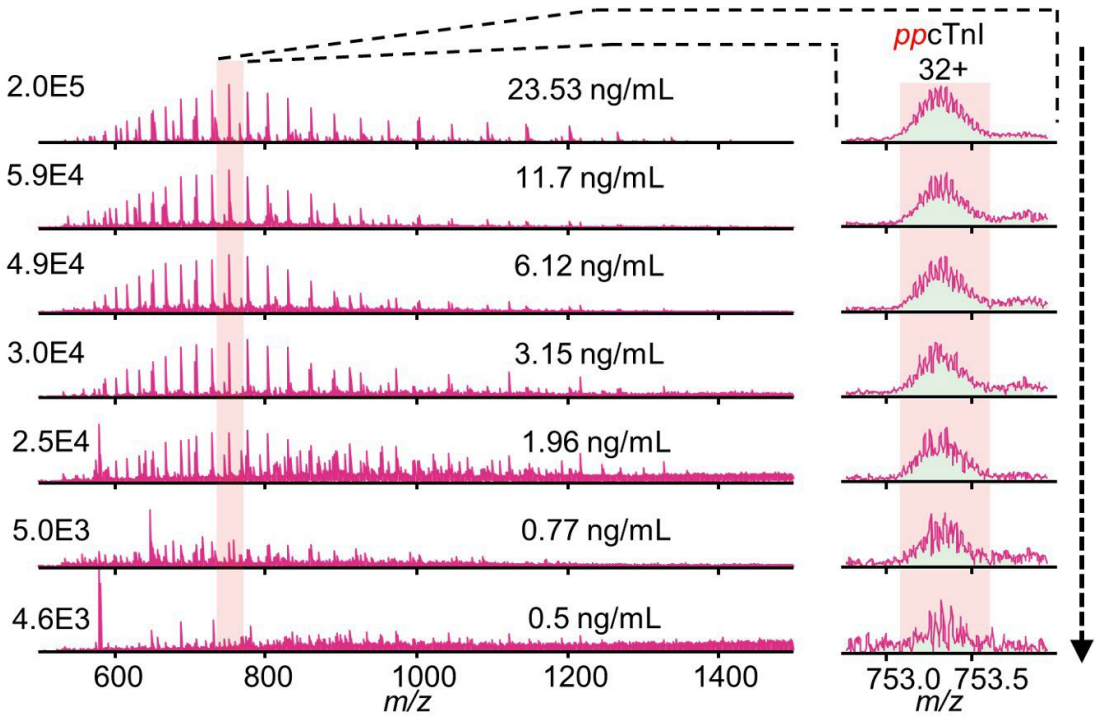


cTnI Amount (ng/mL)	Proteoform	Observed Mass (M_r, Da)	Calculated Mass (M_r, Da)	Mass Error (ppm)
0.78	*pp*cTnI	24063.66	24063.73	2.6
0.31	*pp*cTnI	24063.65	24063.73	3.3
0.155	*pp*cTnI	24063.66	24063.73	2.8
0.062	*pp*cTnI	24063.66	24063.73	2.8
0.031	*pp*cTnI	24063.66	24063.73	2.7
0.0155	*pp*cTnI	24063.64	24063.73	3.7
0.0062	*pp*cTnI	24063.67	24063.73	2.3

Supplementary Figure 2.26

Bruker MaXis II instrument limit of quantification (LOQ) and limit of detection (LOD) for cTnI.

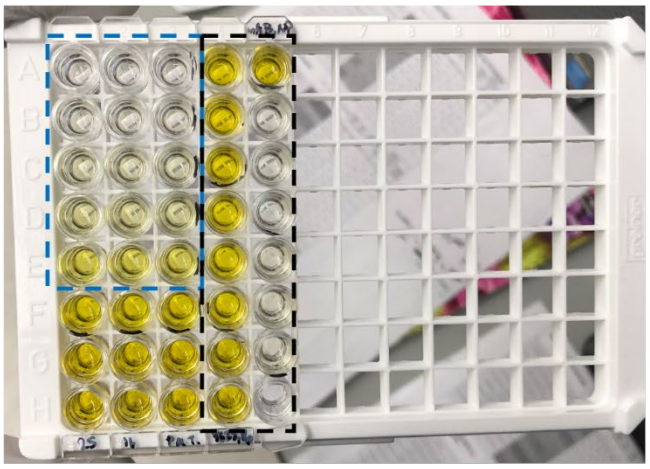
a, Extracted ion chromatograms (EICs) for cTnI at various loading amounts (3.1-0.006 ng/mL) obtained from a heart tissue extract from a healthy donor heart tissue. Data correspond to the same data presented in Fig. 3e. **b**, Deconvoluted mass spectra corresponding to the same data presented in (**a**). ppcTnI is the most abundant proteoform of cTnI and is highlighted to illustrate the linear decrease in raw signal abundance as a function of concentration 3.1-0.006 ng/ml. **c**, Plot of relative cTnI abundance obtained from (**a**) at the various loadings tested, demonstrating linearity of MS-detection and response. From this plot, the LOD (3.3 σ/s) was determined to be 0.06 ng/ml and LOQ (10 σ/s) to be 0.2 ng/ml. **d**, Table summarizing the MS analysis of the data presented in (**a-c**). The mass accuracy (monoisotopic mass) at each tested cTnI loading is shown. All data are representative of n =2 independent experiments.



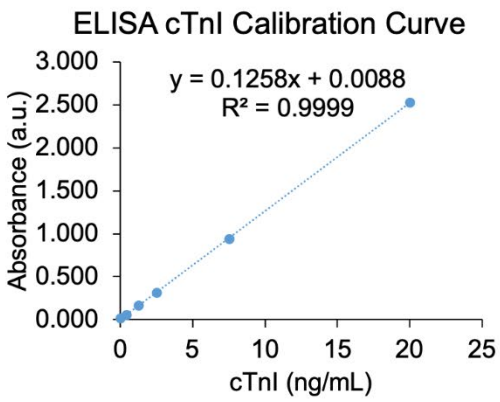
Supplementary Figure 2.27

Sensitivity of Top-Down MS-based cTnI assay using the nanoproteomics platform. Normalized raw mass spectra corresponding to the same data presented in **Fig**. **3**. The most abundant signal, ppcTnI charge state 32+, is highlighted to illustrate the raw signal abundance as a function of the concentration from 22.53 ng/ml to 0.50 ng/ml. pp, bisphosphorylation.

a



b



c

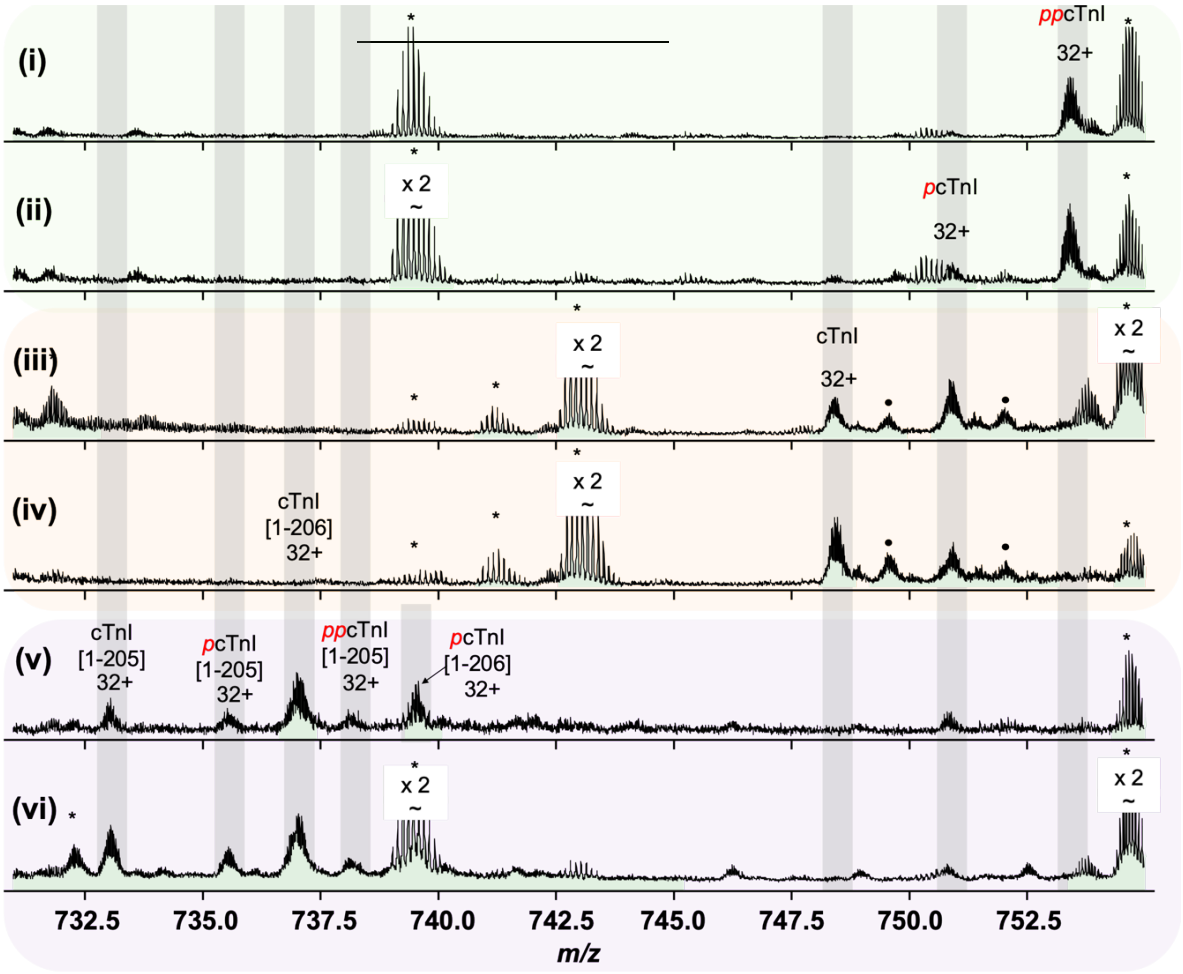
Sample	Sample Source	cTnI Enrichment Factor	Total cTnI Amount Before Enrichment (ng)	Total cTnI Amount Recovered After Enrichment (ng)	cTnI Percent Recovery (%)
NP-Pep	Tissue	1.7	302	268	89
NP-Pep	Serum	115		153	51
Agarose -mAb	Serum	39		53	17

Supplementary Figure 2.28

ELISA-based cTnI enrichment efficiency quantification of NP-Pep and Agarose-mAb.

a, ELISA-based colorimetric quantification of cTnI standards (blue dashed box) and enrichment samples (black dashed box). All samples were dispensed in triplicate and the assay was performed according to the manufacturer's instructions. The ELISA assay uses capture antibodies targeting cTnI amino acids 18-28 and 86-90, with detection antibodies targeting cTnI amino acids 41-49. **b**, ELISA-based standard curve (0 ng, 0.4 ng, 1.25 ng, 2.5 ng, 7.5 ng, 20 ng) used for quantification of cTnI amount before enrichment and after enrichment by NP-Pep or Agarose-mAb. **c**, Summary of enrichment performance results for NP-Pep (tissue/serum) and Agarose-mAb (serum).

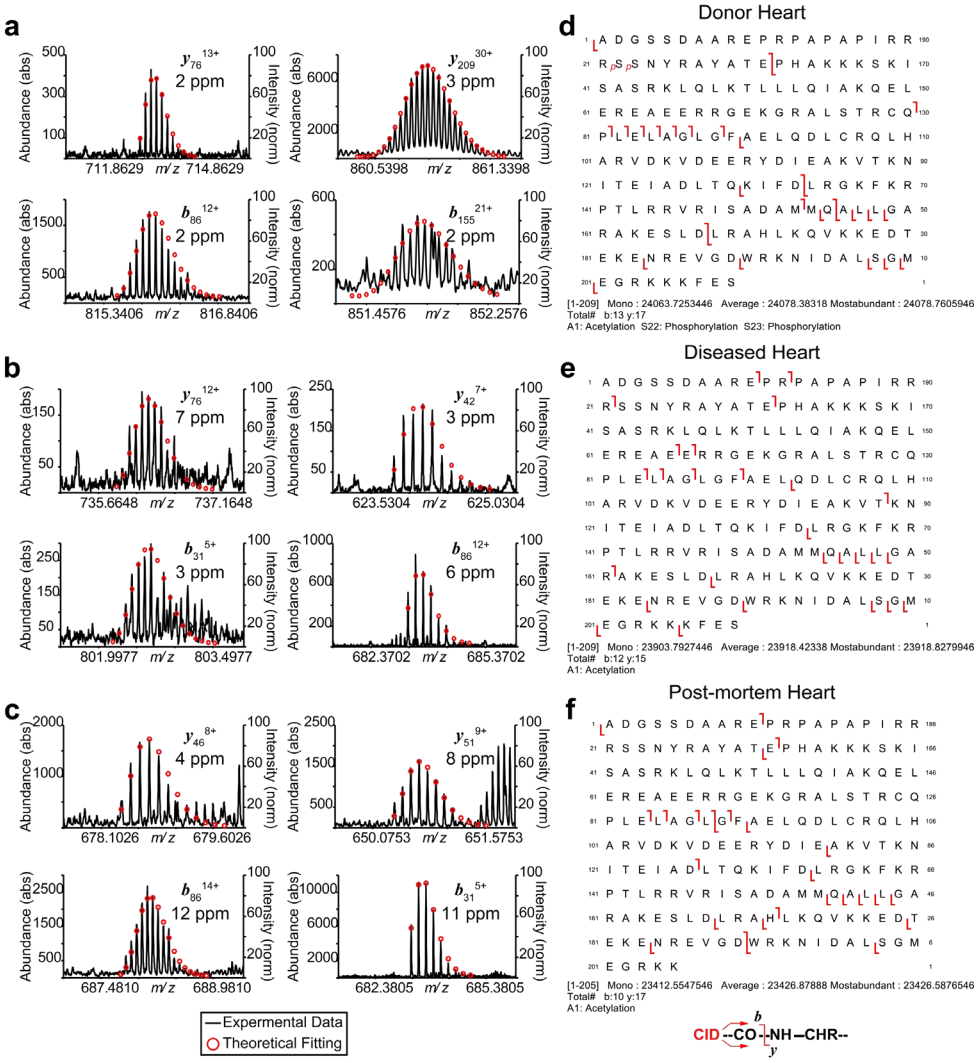
$$\text{cTnI enrichment factor} = \frac{\text{concentration cTnI After Enrichment } (\frac{\text{ng}}{\text{mL}})}{\text{concentration cTnI Before Enrichment } (\frac{\text{ng}}{\text{mL}})}.$$



Supplementary Figure 2.29

LC/MS analysis of cTnI proteoforms enriched directly from human serum by NP-Pep. Normalized raw mass spectra data corresponding to deconvoluted mass spectra shown in s The cTnI (~10-20 ng/mL) spiked in the human serum (10 mg) are extracted from various human hearts:

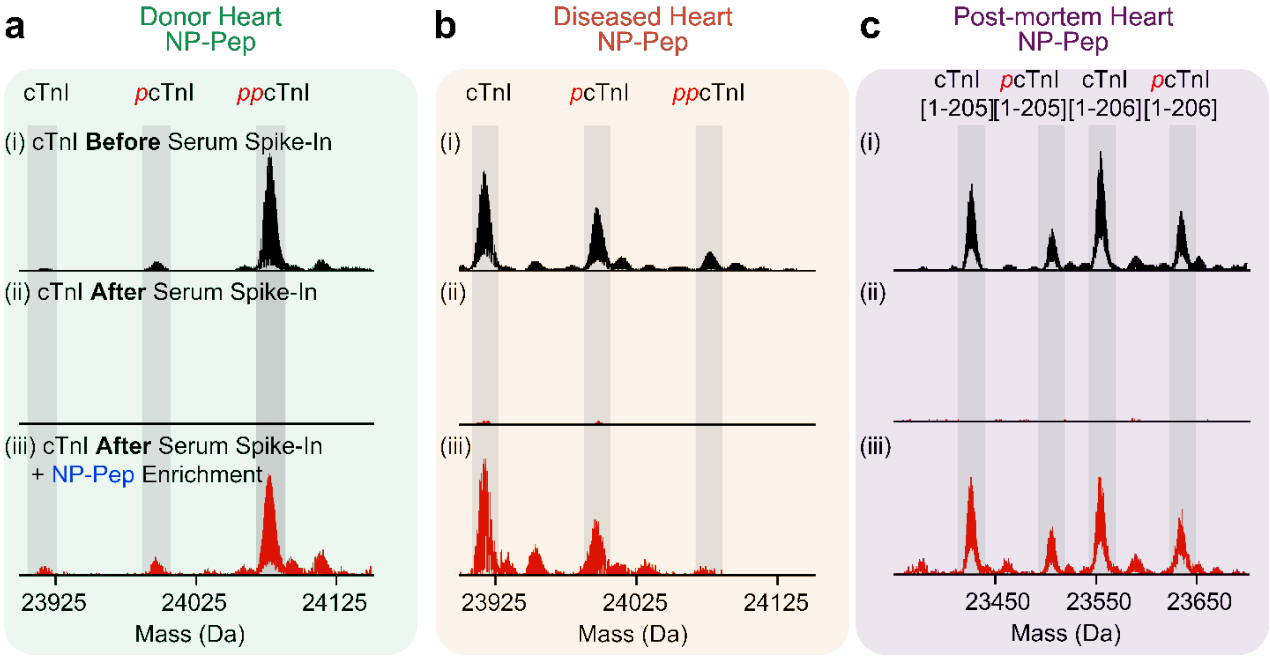
(i) and (ii), donor hearts; (iii) and (iv), diseased hearts with dilated cardiomyopathy, (v) and (vi), post-mortem hearts. p, phosphorylation. pp, bisphosphorylation. Data are representative of n = 3 independent experiments. Asterisks "*" indicate co-eluting serum proteins and circles "•" indicate other cTnI proteoforms. A summary of the enriched proteoforms from serum and their respective mass measurements by top-down MS are listed in Supplementary Table 8.



Supplementary Figure 2.30

Top-down LC-MS/MS characterization of cTnI arising from the various biological samplesafter enrichment.

a-c, Representative CID fragment ions obtained from cTnI arising from the donor heart (y_{76}^{13+}, y_{209}^{30+}, b_{86}^{12+}, b_{155}^{21+}), the disease heart (y_{76}^{12+}, y_{42}^{7+}, b_{31}^{5+}, b_{86}^{12+}), and the post-mortem heart (y_{76}^{8+}, y_{51}^{9+}, b_{86}^{14+}, b_{31}^{5+}) sources. cTnI was found to be primarily in its bisphosphorylated state state in the donor heart (ppcTnI; Ser22, and Ser23), unphosphorylated state in the diseased heart (cTnI), and in a proteolytically degraded form in the post-mortem heart (cTnI[1-206]). Theoreticalion distributions are indicated by the red dots and mass accuracy errors are listed for each fragment ion. **d-f**, Protein sequence fragmentation mapping of the specific proteoform of cTnI corresponding to the fragment ion data obtained from each cTnI source (**a-c**). CID fragmentations are shown as red cleavages. All matched sequences contained N-terminally acetylated cTnI proteoforms following Met exclusion. Amino acid sequence was based on the entry nameTNNI3_human obtained from the UniProtKB sequence database.



Supplementary Figure 2.31

Preservation of cTnI PTM profiles by NP-Pep from serum reflecting different cardiac pathophysiologies.

a-c, Representative deconvoluted mass spectra corresponding to cTnI proteoforms originating from a donor heart (**a**), a diseased dilated cardiomyopathic heart (**b**), and a post-mortem heart (**c**) before serum spike-in (i), after serum spike-in (ii), and after NP-Pep enrichment of human serumspiked with cTnI (iii). Equal amounts of NP-Pep (5 mg) were used for the cTnI enrichments and equal amounts (500 ng) were loaded for LC/MS analysis. Data correspond to the same cTnI proteoform data presented in Fig. 2 and Fig. 4 (Samples (i), (iv), and (v)). Concentrations of cTnIused are listed in Supplementary Table 6 and Supplementary Table 8. p, phosphorylation; pp, bisphosphorylation.

Supplementary Tables

Supplementary Table 2.1. Summary of various materials characterization and surface functionalization analysis of various functionalized Fe_3O_4 NPs synthesized at a large-scale (120 mg).

Particle	TEM Diameter (nm)	TGA Weight Loss (%)	Surface Ligand Coverage (ligands/nm^2)	Surface Ligand Relative Concentration (μmol ligand/mg NP)
Fe_3O_4-OA	8.03 ± 0.08	44	12	1.6
Fe_3O_4-BAPTES	8.2 ± 0.2	26	5	0.90
Fe_3O_4-Peptide	8.3 ± 0.2	32	0.2	0.034

All errors indicate standard deviations from measurements on three independent batches of particles.

Supplementary Table 2.2. Summary showing the details of using TGA analysis to estimate the surface density of BAPTES ligands on the surface of NPs.

Nanoparticle	NP-BAPTES
Average radius (obtained from TEM), *r*	4.0 nm
The mass of BAPTES ligand molecules (quantified by TGA % mass loss), *m*	26% × 4.313 mg (Sample size used for TGA) = 1.121 mg BAPTES
Total number of BAPTES ligand molecules = $\frac{m}{MW} \times N_A$ where MW, ligand molecular weight; N_A, Avogadro's number ($6.022 \times 10^{20} \frac{molecules}{mmol}$)	$\frac{1.121 \text{ mg BAPTES}}{287.12 \frac{\text{mg}}{\text{mmol}}} \times 6.022 \times 10^{20} \frac{\text{molecules}}{\text{mmol}}$ = ~ 2.35×10^{18} BAPTES ligand molecules
The total mass (*M*) of NPs in TGA sample	4.313 mg sample − 1.121 mg BAPTES = 3.192 mg NPs
The total volume (*V*) of NPs from TGA sample $\left(\rho = \frac{5.18\,g}{cm^3}\right)$	$V = M / \rho$; 6.162×10^{-4} cm^3
The volume (*v*) of a single NP	$v = \frac{4}{3}\pi (4 \times 10^{-7}\text{cm})^3 = 2.68 \times 10^{-19}\text{cm}^3$
Total # of NPs from TGA sample, N:	$\frac{V}{v} = \frac{6.95 \times 10^{-4}\ \text{cm}^3}{2.68 \times 10^{-19}\text{cm}^3} = 2.298 \times 10^{15}$ NPs
# of BAPTES ligands per NP	~1023
SA (total surface area) N × $4\pi r^2$	4.622×10^{17} nm^2
# of BAPTES ligands per nm^2	~5
µmol BAPTES ligands per mg of nanoparticle	~0.90 µmol BAPTES/mg of NP

Supplementary Table 2.3. Summary showing the details of using TGA analysis TGA analysis to estimate the surface density of cTnI peptide ligands on the surface of NPs.

Nanoparticle	NP-Pep
Average radius (obtained from TEM), *r*	4.0 nm
The mass of BAPTES and peptide ligand molecules (quantified by total TGA % mass loss)	32% × 2.434 mg (Sample size used for TGA) = 0.779 mg BAPTES + Peptide ligand
The mass of peptide ligand molecules (quantified by TGA % mass loss (i.e. 26%) compared to NP-BAPTES), *m*	6 % × 2.434 mg (Sample size used for TGA) = 0.146 mg peptide
Total number of peptide ligand molecules = $\frac{m}{MW} \times N_A$ where MW, ligand molecular weight; N_A, Avogadro's number ($6.022 \times 10^{20} \frac{\text{molecules}}{\text{mmol}}$)	$\frac{0.146 \text{ mg peptide}}{1742.89 \frac{\text{mg}}{\text{mmol}}} \times 6.022 \times 10^{20} \frac{\text{molecules}}{\text{mmol}}$ $= \sim 5.046 \times 10^{16}$ peptide molecules
The total mass (*M*) of NPs in TGA sample	2.434 mg sample − 0.779 mg BAPTES + peptide = 1.655 mg NPs
The total volume (*V*) of NPs from TGA sample ($\rho = \frac{5.18 \text{ g}}{\text{cm}^3}$)	$V = M / \rho$; 3.195×10^{-4} cm^3
The volume (v) of a single NP	$v = \frac{4}{3}\pi (4 \times 10^{-7}\text{cm})^3 = 2.68 \times 10^{-19}\text{cm}^3$
Total # of NPs from TGA sample, N:	$\frac{V}{v} = \frac{3.197 \times 10^{-4} \text{ cm}^3}{2.68 \times 10^{-19}\text{cm}^3} = 1.191 \times 10^{15}$ NPs
# of peptide ligands per NP	~42
SA (total surface area) N × $4\pi r^2$	2.396×10^{17} nm^2
# of peptide ligands per nm^2	~0.2
µmol peptide ligands per mg of nanoparticle	~0.034 µmol peptide/mg of NP

Supplementary Table 2.4. Summary of sarcomeric protein extraction buffers and reagents. All heart tissue extracts were prepared using the described extraction protocol.

Wash Buffer (pH 7.0)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
NaH_2PO_4 (500 mM)	5 mM	Maintain buffering capacity
Na_2HPO_4 (100 mM)	5 mM	Maintain buffering capacity
$MgCl_2$ (100 mM)	5 mM	Maintain ionic strength of medium
NaCl (137 mM)	100 mM	Maintain ionic strength of medium
EGTA (100 mM)	0.5 mM	Chelation of metal ions; inhibition of metalloproteases
Add remaining reagents immediately prior to performing wash protocol		
Triton™ X-100 (10% v/v)	1% (v/v)	Solubilization of poorly soluble proteins
DTT (1 M)	5 mM	Reducing agent
Halt™ protease inhibitor cocktail (100x)	1x	Inhibition of Ser/Cys/Asp proteases
PMSF (25 mM in ethanol)	1 mM	Inhibition of Ser proteases
Phosphatase inhibitor cocktail A (100x)	1x	Inhibition of Ser/Thr/Tyr phosphatases
Protein Extraction Buffer (pH 7.5)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
Tris (1 M)	25 mM	Maintain buffering capacity
$CaCl_2$ (100 mM)	0.1 mM	Maintain ionic strength of medium
LiCl (3.5 M)	700 mM	Protein extraction
EGTA (100 mM)	5 mM	Chelation of metal ions; inhibition of metalloproteases
Add remaining reagents immediately prior to performing extraction protocol		
DTT (1 M)	5 mM	Reducing agent

Halt™ protease inhibitor cocktail (100x)	1x	Inhibition of Ser/Cys/Asp proteases
PMSF (25 mM in ethanol)	1 mM	Inhibition of Ser proteases
Phosphatase inhibitor cocktail A (100x)	1x	Inhibition of Ser/Thr/Tyr phosphatases

Supplementary Table 2.5. Summary of accurate mass measurement by top-down MS analysis of NP-Peptide enriched cTnI proteoforms (i)-(iv) corresponding to the data presented in Fig. 2e. Most abundant masses are shown.

Sample	Proteoform	Observed Mass (Da)	Calculated Mass (Da)	Mass Error (ppm)
Batch 1	(i) ppcTnI[1-207]	23861.66	23861.68	0.8
	(ii) cTnI	23917.80	23917.83	1.3
	(iii) pcTnI	23997.78	23997.90	5.0
	(iv) ppcTnI	24077.73	24077.86	5.4
Batch 2	(i) ppcTnI[1-207]	23861.75	23861.68	2.9
	(ii) cTnI	23917.92	23917.83	3.8
	(iii) pcTnI	23997.89	23997.90	0.4
	(iv) ppcTnI	24077.85	24077.86	0.4
Batch 3	(i) ppcTnI[1-207]	23861.79	23861.68	4.6
	(ii) cTnI	23917.89	23917.83	2.5
	(iii) pcTnI	23997.89	23997.90	0.4
	(iv) ppcTnI	24077.85	24077.86	0.4

Supplementary Table 2.6. Summary of accurate mass measurement by top-down MS analysis of NP-Peptide enriched cTnI proteoforms (i)-(iv) corresponding to the data presented in Fig. 2f-h. Most abundant masses are shown.

Heart Sample	Proteoform	[cTnI]	Observed Mass (Da)	Calculated Mass (Da)	Mass Error (ppm)
Donor	(i) ppcTnI[1-207]	306 ng/ml	23861.73	23861.68	2.1
	(ii) cTnI		23917.87	23917.83	1.7
	(iii) pcTnI		23997.82	23997.90	3.3
	(iv) ppcTnI		24077.80	24077.86	2.5
Dilated cardiomyopathy	(ii) cTnI	482 ng/ml	23917.71	23917.83	5.0
	(iii) pcTnI		23997.77	23997.90	5.4
	(iv) ppcTnI		24077.73	24077.86	5.4
Post-mortem	(v) cTnI[1-205]	465 ng/ml	23426.67	23426.59	3.4
	(vi) pcTnI[1-205]		23506.65	23506.55	4.3
	(vii) cTnI[1-206]		23554.76	23554.68	3.4
	(viii) pcTnI[1-206]		23634.72	23634.65	3.0

Supplementary Table 2.7. Summary of the population characteristics of various non-diseased donor, diseased (dilated cardiomyopathy), and post-mortem heart samples corresponding to the data presented in Figure 2.4b and Supplementary Fig. 2.29.

Sample Code (Fig. 2.4b; Fig. S2.29)	Decoded ID	Specimen Type	Age	Sex	Ethnicity	ß-blocker?	Known medications
(i)	YG25	Donor Heart	61	M	--	--	--
(ii)	YG40	Donor Heart	25	F	--	--	Duonebs, Budesonide, Dexamethasone, Flovent, Sertraline, Mirena, Albuterol, Epipen
(iii)	YG13	Dilated Cardiomyopathy Heart	22	M	--	Yes	--
(iv)	YG16	Dilated Cardiomyopathy Heart	--	M	Caucasian	--	--
(v)	A08-089	Post-mortem Heart	58	F	Caucasian	Yes	aspirin, atenolol, budesonide, bupropion, calcium, cyancobalamin, duloxetine, ergocalciferol, ferrous sulfate, gabapentin, lamotrigine, levothyroxine, loratadine, methadone, modafinil, multivitamin, oxybutynin, pantoprazole, prednisone, piperacillin tazob
(vi)	A08-451	Post-mortem Heart	39	M	Caucasian	Yes	oxycodone, insulin, zocor, diflucan, macrobid, plavix, mycelex, aspirin, tums, colace, cardura, zetia, lasix, norco, prevacid, metoprolol, myfortic, lovaza, KCl, prednisone, flomax, valsartan, flexeril, lexapro, xanax, aranesp, ergocalciferol, nitroglycerin

Supplementary Table 2.8. Summary of accurate mass measurement by top-down MS analysis of NP-Peptide enriched cTnI proteoforms from human serum corresponding to the data presented in Fig. 4. Most abundant masses are shown.

Heart Sample	Proteoform	[cTnI]	Observed Mass (Da)	Calculated Mass (Da)	Mass Error (ppm)
(i) Donor	pcTnI	11.05 ng/ml	23997.70	23997.90	8.3
	ppcTnI		24077.64	24077.86	9.1
(ii) Donor	cTnI	10.69 ng/ml	23917.77	23917.83	2.5
	pcTnI		23997.70	23997.90	8.3
	ppcTnI		24077.64	24077.86	9.1
(iii) Dilated cardiomyopathy	cTnI	20.72 ng/ml	23917.77	23917.83	2.5
	pcTnI		23997.77	23997.90	5.4
(iv) Dilated cardiomyopathy	cTnI	22.84 ng/ml	23917.72	23917.83	4.6
	pcTnI		23997.77	23997.90	5.4
(v) Post-mortem	cTnI[1-205]	11.66 ng/ml	23426.40	23426.59	8.1
	pcTnI[1-205]		23506.37	23506.55	7.7
	cTnI[1-206]		23554.49	23554.68	8.1
	pcTnI[1-206]		23634.43	23634.65	9.3
(vi) Post-mortem	cTnI[1-205]	15.22 ng/ml	23426.56	23426.59	1.3
	pcTnI[1-205]		23506.52	23506.55	1.3
	cTnI[1-206]		23554.65	23554.68	1.3
	pcTnI[1-206]		23634.61	23634.65	1.7

Chapter 3

Native Mass Spectrometry of Human Cardiac Troponin Complexes Enabled by Functionalized Nanoparticle Enrichment and Online Size Exclusion Chromatography

Abstract

The cardiac troponin complex (cTn, ~77 kDa) consists of three subunits (troponin I, T, and C), and plays critical roles in cardiac contraction and relaxation. In the study of cardiovascular diseases, recombinantly expressed cardiac troponin subunits are frequently utilized and incorporated into reconstituted cTns or cardiac thin filaments to investigate protein-protein interactions and the structure-function relationships of myofilament proteins. However, methods to characterize intact, endogenous cTn from human heart tissues remain lacking due to the challenges in the isolation and preservation of cTn under non-denaturing conditions. Herein, we report a "native nanoproteomics" strategy incorporating functionalized nanomaterials for specific enrichment of cTn under physiological pH and native size exclusion chromatography (SEC) for rapid online buffer exchange (OBE), thereby enabling the native MS analysis of cTn from human heart samples for the first time. We demonstrate that this robust and reproducible 'native nanoproteomics' technology can be applied for the study of cardiac troponin complexes from human heart samples representative of various cardiomyopathies, as a powerful structural biology tool to interrogate protein stoichiometry, non-covalent cofactor binding, and post-translational modifications.

Introduction

The cardiac troponin complex (cTn) is composed of three subunits: troponin C (TnC), the calcium (Ca^{2+}) -binding regulatory subunit; troponin I (cTnI), the inhibitory subunit; troponin T (cTnT), the tropomyosin-binding subunit, and critically functions as the molecular switch that associates intracellular Ca^{2+} concentration with cardiac muscle contraction and relaxation[1-3]. In the presence of Ca^{2+}, cTnI binds to TnC in its Ca^{2+}-occupied state and ultimately enables actin-myosin interactions that produce muscle contractions. In the absence of Ca^{2+}, the C-terminus of cTnI is released from TnC, and interacts with actin and tropomyosin to inhibit muscle contraction[3, 4]. Independently of Ca^{2+} concentration, the speed and force of contraction are further modified by myofilament protein post-translational modifications (PTMs) which alter protein-protein interactions[3, 5]. Post-translational modification of cTnI by phosphorylation regulates cTnI function by inducing structural changes that directly result in reduced Ca^{2+} sensitivity and faster dissociation of Ca^{2+} from TnC[5, 6]. As a result, mutations in troponins which alter the intrinsic Ca^{2+} or cTnI binding to TnC often have profound and complex downstream effects on cardiac function, but the correlation between specific alterations in cTn structure and the mutations which give rise to complex phenotypes remain poorly understood[3, 7]. Thus, it is critical to understand how the structural alterations and Ca^{2+} within the cTn ultimately influence downstream cardiac function for a greater understanding of the molecular mechanisms behind cardiovascular diseases[2].

In the study of various cardiovascular diseases, such as hypertrophic cardiomyopathy[7-9] and dilated cardiomyopathy[8, 10], recombinant cardiac troponin subunits are frequently utilized and incorporated into reconstituted cTns or cardiac thin filaments to investigate protein-protein interactions and the structure-function relationships of myofilament proteins[7, 11]. However,

reconstituting and refolding recombinant expressed cardiac troponin subunits together is often challenging and is prone to misfolding or misassembly of the intact protein complexes, which can reduce protein functionality when attempting to study a protein system.[12] Moreover, the lack of post-translational machinery in prokaryotic cell expression systems[13] requires additional considerations when recapitulating eukaryotic, post-translationally modified proteins such as phosphorylated cardiac troponin subunits, using recombinant protein systems[12, 13]. Therefore, a comprehensive, protein complex-resolved analysis which can analyze cTns from human tissue samples would be powerful to directly elucidate and interrogate PTM structure-function relationships within endogenous complexes.

Native mass spectrometry (MS) has emerged as an important and powerful technique to characterize endogenous protein/protein complexes and their noncovalent interactions by transferring them into the gas phase without loss of higher-order structure[14-16]. By directly acquiring and determining the molecular mass (from the *m/z* range), native MS can provide rich information to simultaneously resolve multiple coexisting protein species, the stoichiometry of Ca^{2+} cofactor[17] or protein-ligand binding kinetics and affinities, and the cooperativity of various interactions—circumventing the need for chemical/fluorescent labeling or immobilization strategies used in common biochemical and biophysical assays[14]. Moreover, when native MS is paired with ion mobility, a technique that separates gas-phase ions based on their differential mobility arising from their size and shape[18], ion-mobility-mass spectrometry(IM-MS) can additionally provide gas-phase structural information of biomolecules and their assemblies that is largely consistent with the structures obtained by X-ray crystallography and solution-phase nuclear magnetic resonance (NMR) [16, 19, 20]. Thus, native MS is well-suited for the analysis of endogenous

cTns and their cofactors, to elucidate the structure-function relationship of cTn with respect to various cardiovascular disease states.

Previously, we reported the demonstration of an antibody-free, nanoproteomics approach integrating surface-functionalized nanoparticles (NPs) and top-down mass spectrometry-based proteomics to specifically enrich and analyze low-abundance cardiac troponin–I proteoform monomers directly from human serum[21]. Nanoparticles (NPs) are highly effective for such sensitive and specific proteoform enrichment because: (1) they are commensurate in size and diffusion kinetics with proteins allowing effective penetration through complex biological mixtures without causing denaturation; (2) they have high surface-to-volume ratios to enhance protein interaction; (3) they are versatile scaffolds to couple diverse affinity ligands for protein binding and capture[21-24]. Herein, we report a "native nanoproteomics" strategy for the affinity purification of endogenous, intact cTns from human heart tissues, with high-resolution MS coupled online with native size exclusion chromatography (SEC)[25], thereby enabling the native MS analysis of cTns from human heart samples for the first time (Figure 3.1). Native MS simultaneously revealed cTn monomers, cTn dimers (cTnI-cTnC), and cTn ternary complexes (cTnI-cTnT-TnC), and the varying stoichiometry of non-covalently bound Ca^{2+} among them. Using native MS, we qualitatively probed the differential binding affinity of Ca^{2+} ions within the different protein species, demonstrating that dimer and trimer cTn complexes decrease the dissociation rate of Ca^{2+} compared to free monomer, and that the loss of Ca^{2+} ions can destabilize cardiac troponin complex stability. Finally, we demonstrate that this robust and reproducible 'native nanoproteomics' technology can be applied for the study of cardiac troponin complexes from human heart samples representative of various cardiomyopathies, as a powerful structural

biology tool to interrogate protein stoichiometry, non-covalent cofactor binding, and post-translational modifications.

Experimental Methods:

Chemicals and Supplies

3-butynoic acid was acquired from Chem-Impex Int'l Inc. (Wood Dale, IL, USA). Buffer components, such as molecular biology grade L-arginine, L-methionine, L-glutamic acid, and imidazole were acquired from MillporeSigma, Inc. (St. Louis, MO, USA). Amicon Ultra Centrifugal Filters (regenerated cellulose, 10 kDa and 30 kDa cutoffs) and MS-SAFE Protease and Phosphatase inhibitor cocktail were acquired from MillporeSigma, Inc.. HPLC grade ammonium acetate was acquired from Fisher Scientific (Fair Lawn, NJ, USA). Solvents such as HPLC grade water, acetonitrile, isopropanol, methanol, were acquired from Fisher Scientific. Size exclusion chromatography columns (e.g., 100 x 2.1 mm, 3 μm 200 Å PolyHEA) were acquired from PolyLC, Inc. (Columbia, MD, USA). Bovine Serum Albumin (BSA) was acquired from Fisher Scientific. All other reagents and consumables were acquired from MilliporeSigma and were used without further purification.

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 (6.39 g, 1 eq., 25 mmol), and dichloromethane (250 mL) were added to an oven dried 500 mL three-neck flask fitted with a reflux condenser. The solution was heated to reflux under N_2. In a separate flask, (3-aminopropyl)triethoxysilane (5.53 g, 1 eq., 25 mmol) and N,N-diisopropylethylamine (6.46 g, 2 eq., 50 mmol) were diluted with dichloromethane (125 mL), and added to the refluxing three-neck flask by a syringe. The reaction mixture was allowed to reflux for 1 hour, and then N,N-diisopropylethylamine (6.46 mL, 2 eq., 50 mmol) was additionally added to effectively isomerize the propargylic isomer to N-(3-(triethoxysilyl)propyl)buta-2,3-dienamide. The reaction was allowed to reflux overnight, and the

product was concentrated by rotary evaporation. The crude product was redispersed in ethyl acetate, centrifuged at 5,500 x *g* for 5 min, and the supernatant was concentrated by rotary evaporation. The product was purified with flash column chromatography using a gradient of 50:50 ethyl acetate: n-hexane to 100:0 ethyl acetate: n-hexane to yield BAPTES as a clear, orange oil (4.80 g, 67% yield). R_f (Ethyl Acetate)= 0.67; ^{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*.

Synthesis of iron-oleate precursor. Iron oleate was synthesized using a previously established method[26]. 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-hour 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 by rotary evaporation. 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.

Synthesis of 8 nm magnetite (Fe_3O_4) nanocrystals. Iron oleate (10 mmol, 9.0 g) and oleic acid (5.5 mmol, 1.56 g) were added into a three-neck round bottom flask (250 mL) with a solvent mixture of 1-octadecene:1-tetradecene (40 g: 10 g). The mixture was stirred and degassed on a Schlenk line at 110 °C for 3 h. The mixture was then placed under nitrogen flow and the reaction solution was heated to 300 °C at a heating rate of 3.3 °C/min using a temperature controller. Reaction time was counted starting from when 300 °C was reached. After 1 h, the reaction solution was quickly cooled to room temperature by blowing air across the reaction flask. The resulting iron-oxide nanocrystals were precipitated using ethanol, isolated via centrifugation (5,500 xg, 20 min), and then washed with three precipitation/redispersion cycles (5,500 x *g*, 20 min) using ethanol. The resulting nanocrystals were then dried under vacuum, weighed, and redispersed in n-hexane at a concentration of 20 mg/mL for further use.

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 reaction under stirring 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 (5,500 x *g*, 10 min) to remove excess silane molecules and surfactants. The NPs were then dried under vacuum for later use.

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) 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 hour and later collected into Eppendorf tubes for washing. The peptide-functionalized NPs were washed three times with nanopure water via centrifugation (15,000 x *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.

Protein extraction from human heart tissues at physiological pH for the preparation of Cardiac protein extract

Whole human heart tissue samples were maintained in cardioplegic solution prior to regional dissection, subsequent snap-freezing in liquid nitrogen, and storage at - 80 °C. Once removed from -80 °C storage, samples were kept in a vessel containing liquid nitrogen. All procedures involving sample handling were performed in a temperature-controlled cold room (~4°C to reduce artifactual modification (e.g., protein oxidation) of samples. Cardiac tissue was first homogenized on ice using a Polytron homogenizer in 10 volumes (mL/g tissue) of 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) as reported previously[27, 28], with modification to include 125 mM L-methionine as a sacrificial reductant to prevent protein oxidation, and to exclude 5 mM $MgCl_2$ and 0.5 mM EGTA, which can alter the concentration of divalent cations endogenously associating with sarcomeric proteins, as well as Triton X-100, a non-ionic surfactant that can suppress MS-signal intensity. DTT was excluded to prevent the alteration of protein tertiary structure through the reduction of protein disulfide bonds. The homogenate was centrifuged at 17,000 rcf for 3 min at 4°C, and the supernatant was discarded. The washing and homogenization step was repeated once more, and then the homogenate was centrifuged at 17,000 rcf for 3 min at 4°C, and the supernatant was discarded. To extract proteins such as the cardiac troponin (T-I-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)[27, 28], with inclusion of 125 mM L-Met and exclusion of 5 mM DTT, 0.1 mM $CaCl_2$ and 5 mM EGTA as previously described above. The resulting pellet following washing and centrifugation was homogenized in 5 vol (mL/g tissue) of LiCl extraction buffer and incubated at 4°C for 10 minutes to extract the sarcomeric proteome. The homogenate was centrifuged at 17,000 rcf for 3 min at 4°C, and then the supernatant containing the sarcomeric proteome was transferred to new Eppendorf tubes, and further centrifuged at 21,000 rcf for 30 mins at 4°C to clarify the extracts. The supernatants were finally transferred to new Eppendorf tubes, snap-frozen in liquid nitrogen, and stored at -80°C.

Cardiac troponin complex purification using functionalized nanoparticle enrichment

NP-Pep (5 mg) originally dispersed in H_2O was redispersed in a 2 mL Eppendorf Protein Lo-Bind tube with 1.0 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 rcf for 2 min at 4 °C, isolated from the solution using the DynaMag, and the supernatant was removed. 1.0 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 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 0.3 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 rcf 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) following the same centrifugation and magnetic isolation steps to remove unbound, nonspecific proteins. To elute the bound cTn without severely dissociating the troponin monomer subunits from the protein 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 15 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 rcf 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 rcf for 20 min at 4 °C, and the concentrated protein mixture

(approximately 40 µL) was transferred to a chilled LC-MS vial for automated buffer exchange using online size exclusion chromatography columns.

Calcium binding experiments using EGTA for stability

To probe the role and influence of calcium (Ca^{2+}) co-factor association with respect to the stability of the cardiac troponin complex, a metallic ion chelator with high affinity for calcium (1.0 M EGTA, pH 7.5) was added to NP-Pep elution mixtures at various final concentrations (0, 10, 25, 50, and 100 mM EGTA) in 0.5 ml Eppendorf lo-bind tubes. The total volume of each EGTA reaction was normalized to 15 µL to minimize dilution of the 750 mM L-Arginine, 750 mM imidazole, 50 mM L-glutamic acid (pH 7.5) solutions containing the cardiac troponin complex. 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, protein solutions were centrifuged at 21,000 rcf for 5 min at 4 °C to prevent loading of any possible precipitates. The protein concentration of each solution after centrifugation was quantified by Bradford assay or NanoDrop analysis for equal protein loading for online SEC-MS analysis, and the reaction mixtures were transferred to a chilled LC-MS vial for automated buffer exchange using online size exclusion chromatography columns.

Online size exclusion chromatography and native MS analysis (QTOF)

SEC experiments were performed using a nanoACQUITY UPLC system (Waters, Milford, MA, USA) PolyHYDROXYETHYL A (PolyHEA) columns from PolyLC Inc. (Columbia, MD, USA) were used for all experiments. Columns varying in their internal diameters (i.d.; 2.1 mm; 1.0 mm, 0.15 mm) with length of 100 mm, particle size of 5 µm, and pore sizes of 200 Å were evaluated using bovine serum albumin as a protein standard loaded at 500 ng using a 1 µL sample

injection, with SEC-based separations performed by proportionally varying the flowrate with respect to the column internal diameter. For example, protein samples were separated with a flow rate of 28 µl/min for a 2.1 mm i.d column, 7 µl/min for a 1.0 mm i.d. column, and 1 µl/min for a 0.15 mm i.d. column. The isocratic SEC separation consisted of 200 mM ammonium acetate (pH 7.0) as mobile phase 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 over time, which may require cleaning of the ESI spray shield, capillary cap, and MS capillary. SEC columns were flushed with a minimum of 1 column volume of HPLC H_2O, and re-equilibrated with a minimum of 1 column volume of 200 mM ammonium acetate (pH 7.0) after every protein sample containing 750 mM L-arginine, 750 mM imidazole, 50 mM L-glutamic acid (pH 7.5). It was found that the use of at least 1 column volume of HPLC H_2O and at least 1 column volume of 200 mM ammonium acetate (pH 7.0) was optimal for effectively equilibrating SEC columns for subsequent sample analyses compared to using multiple column volumes of 200 mM ammonium acetate (pH 7.0), which were unable to reproduce a stable chromatographic baseline due to the column retention of residual non-volatile salts and buffers.

The proteins buffer exchanged into 200 mM ammonium acetate (pH 7.0) were analyzed by a MaXis II quadrupole-time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) via electrospray ionization. The maXis Q-TOF was initially calibrated for mass spectrometry signal-to-noise intensity/resolution using ESI-L Low Concentration Tuning Mix (Agilent) at a flow rate of 3 µl/min. Following intensity/resolution optimization, the maXis Q-TOF was calibrated for mass accuracy using CsI clusters (2 mg/ml in 50:50 methanol:water) at a flow rate of 3 µl/min. For ion source parameters, the end plate offset and capillary voltage were 500 and

4500 V, respectively. Additionally, the nebulizer gas (nitrogen) pressure 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. Mass spectra were collected at a scan rate of 0.25 Hz over 500–8000 *m/z* range. For ion transfer parameters, the funnel RF was set to 400 V, octopole RF was set to 800 V, quadrupole energy set to 4 V, low mass cutoff set to 500 *m/z,* and the in-source CID energy set to 150 - 200 V. Additionally, collision Vpp was set to 4000 Vpp; pre-pulse storage time set to 45 µs, and transfer time (collision cell to TOF) to 145 µs. It was found that increasing in-source CID energy provided the greatest increase in MS signal-to-noise for more effective desolvation of protein ions in ammonium acetate. Increasing the collision Vpp to the highest possible instrument setting of 4000 Vpp, as well as increasing the pre-pulse and transfer time, favored the transmission of larger m/z ions such as those expected for the cardiac troponin complex (3500 – 5000 *m/z*) resulting in a greater increase in MS signal-to-noise of the cardiac troponin complex.

Offline Size exclusion chromatography and native MS protein characterization (FTICR)

SEC experiments for offline protein fractionation were performed exactly as those described in the previous section for online SEC. After establishing consistent and reproducible SEC chromatography by monitoring the elution time and MS-signal intensity of the protein of interest across three independent sample injections, the PEEK tubing connected from the MS divert valve to the ESI source was disconnected at the end closest to the ESI source immediately prior to the elution of the protein of interest. The eluate from the PEEK tubing containing the protein of interest was directed into a chilled Eppendorf Lo-bind tube placed on ice, and the PEEK tubing was reconnected to the ESI source following fraction collection. Offline protein fractions were either immediately snap-frozen in liquid nitrogen and stored at – 80 °C for longer-term

storage use, or held on ice and analyzed within an hour to prevent degradation or precipitation of sample over time.

The protein samples collected by offline fraction collection were analyzed by nanoelectrospray ionization via direct infusion using a TriVersa Nanomate system (Advion BioSciences, Ithaca, NY, USA) coupled to a solariX XR 12-T Fourier Transform Ion Cyclotron Resonance mass spectrometer (FTICR-MS, Bruker Daltonics, Billerica, MA, USA). For the nanoelectrospray ionization source using a TriVersa Nanomate, the desolvating gas pressure was set at 0.45 PSI and the voltage was set to 1.45–1.65 kV versus the inlet of the mass spectrometer. The source dry gas flow rate was set to 2 L/min at 120 °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, and collision cell RF amplitude were optimized at 190 V, 200 V, 150 V, 100 V, 300 Vpp, 2 MHz, 490 Vpp, 2 MHz, and 2000 Vpp, respectively. Mass spectra were acquired with an acquisition size of 16M-words of data (with a resolution of 530 000 at 400 m/z) in the mass range between 200 and 6000 m/z, and a minimum of 250 scans were accumulated for each sample. Ions were accumulated in the collision cell for 3.0 s, and a time-of-flight of 2.0 ms was used prior to their transfer to the ICR cell.

Denaturing Reverse Phase Chromatography and Top-Down MS Analysis

For denatured protein sample preparation of the cardiac protein loading mixtures, flow through, and NP-Pep elution mixtures, Amicon Ultra Centrifugal filters (10 kDa, 0.5 mL) for each respective sample were pre-equilibrated with a cartridge volume of 0.2% formic acid in water (mobile phase A) and centrifuged at 14,000 rcf 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%

formic acid in water to the cartridge filter volume, and concentrated by centrifugation at 14,000 rcf 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. The protein concentration was determined by either using Bradford protein assay or NanoDrop protein analysis. 500 ng of the desalted protein sample was equally loaded for all samples for consistent analysis, and injected onto a PLRP column (1000 Å, 200 mm × 0.5 mm, 10 µm, 1000Å) and separated by RPC using a nanoACQUITY UPLC system (Waters, Milford, MA, USA). Proteins were eluted by a 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.

The denatured, eluted proteins were analyzed by a MaXis II quadrupole-time-of-flight (Q-TOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) via electrospray ionization. The maXis Q-TOF was calibrated for mass accuracy and signal-to-noise intensity/resolution using ESI-L Low Concentration Tuning Mix (Agilent) at a flow rate of 3 µl/min. For ion source parameters, the end plate offset and capillary voltage were 500 and 4500 V, respectively. Additionally, the nebulizer gas (nitrogen) pressure was set to 0.5 bar, the dry gas flow rate set to 4.0 L/min, and the dry gas temperature set to 200°C. Mass spectra were collected at a scan rate of 0.5 Hz over 200–3000 *m/z* range. For ion transfer parameters, the funnel RF was set to 400 V, octopole RF was set to 400 V, quadrupole energy set to 4 V, low mass cutoff set to 500 *m/z,*and the in-source CID energy set to 10 V. Additionally, collision Vpp was set to 2000 Vpp; pre-pulse storage time set to 15 µs, and transfer time (collision cell to TOF) to 100 µs.

Results and Discussion

Evaluation of NP-Pep for affinity purification of endogenous cardiac troponin complexes

Affinity purification strategies, which rely on the principle of antigen-affinity reagent recognition, enable specific and effective enrichment of an antigen by its respective affinity reagent. Previously, we developed surface-functionalized superparamagnetic iron-oxide (magnetite, Fe_3O_4) NPs functionalized with a 12-mer peptide (NP-Pep) with high affinity towards cTnI to specifically enrich cTnI from complex mixtures[21, 29]. Because of the impressive binding affinity (K_d, 270 pM) of the 12-mer peptide for cTnI, strong elution conditions are required to disrupt the intermolecular interactions between the binding peptide and the antigen of interest, cTnI. In our initial report, we used an acidic buffer to disrupt the intermolecular interactions between NP-Pep and bound cTnI. Acidic, low pH buffers, which are commonly used to alter affinity interactions by inducing changes in net charge of both the affinity reagent and antigen, not only cause the antigen-affinity association to dissociate and release, but potentially also disrupt the intermolecular forces which keep the affinity-captured protein monomers in intact protein assemblies and complexes[30, 31]. Upon eluting the proteins with an acidic buffer, we found that the intermolecular interactions between the troponin monomers (cTnI, cTnT, and cTnC) were also disrupted, and were unable to immediately reform as an intact trimer despite immediate neutralization to physiological pH. Alternatively, strategies such as the incorporation of affinity tags, or amino acid sequences that are genetically grafted onto proteins, can be leveraged to enable dissociation of protein antigens from their affinity substrates without harsh conditions, but are limited to recombinant protein and expressed cell systems and cannot be used for the analysis of endogenous, non-recombinant proteins from tissue samples[32, 33]. To address the challenges of

eluting non-recombinant, affinity-purified protein complexes that maintain their quaternary structures, we utilized a strategy involving the use of amino acids at physiological pH to competitively elute protein complexes[34]. Because of the high occurrence of aromatic amino acids (nearly half) in interacting residues at the typical antigen-antibody surface, an emerging strategy is to use a solution of amino acids composed of those more frequently found in the interaction area to competitively elute immunoaffinity purified protein-complex resulting in the disruption of the antigen-antibody complex. The addition of imidazole, an aromatic small molecule which functions as the side chain of histidine[35], could be used to competitively elute the cTn complex from NP-Pep without disrupting the intermolecular interactions between the troponin monomers. The addition of L-arginine and L-glutamic acid, which are known to increase protein solubility and in-solution stability, and can also improve protein elution, were included in the elution buffer[34, 36]. These amino acids not only serve to dissociate non-covalently bound proteins and protein complexes from their affinity ligands, but can additionally stabilize difficult to solubilize proteins that are prone to aggregation or dissociation, by shielding hydrophobic residues on proteins which tend to associate and cause precipitation in aqueous solutions[37]. We hypothesized that the sequence of our high affinity binding 12-mer peptide, which contained many aromatic residues, could be amenable to competitive elution using amino acids.

To evaluate the use of amino acids for competitive elution from NP-Pep, we used sarcomeric protein extracts prepared from human heart tissues, with care to maintain intact disulfide bonds while avoiding artificial oxidation during the protein extraction procedure (Figure S3.1). The cTnI enrichment experiments using the functionalized NP-Pep proceed as following: (1) incubating the NP-Pep in the protein mixture, (2) magnetically isolating the NP-Pep to wash and remove unbound nonspecific proteins, and (3) eluting captured proteins off of the NP-Pep

using amino acids. We screened a variety of conditions at increasing molarity (100 mM L-arginine, 100 mM imidazole, 50 mM L-glutamic acid, pH 7.5 to 750 mM L-arginine, 750 mM imidazole, 50 mM L-glutamic acid) to evaluate which concentration was most suitable for effective elution and used SDS-PAGE analysis to initially evaluate the enrichment performance using various elution conditions (Figure S3.2). If successful, the monomers of cTnT, cTnI, and cTnc should be approximately in a 1:1:1 ratio, as this would reflect the composition of the cTn in the sarcomere and would demonstrate that the NP-Pep can effectively capture the protein interactors of cTnI. When analyzing the elution lanes consisting of < 500 mM L-arginine, < 500 mM imidazole, and 50 mM L-glutamic acid (pH 7.5), the lanes were much fainter when normalizing by equal protein loading, indicating that higher concentration is required for effective enrichment elution, suggesting that the large excess of L-arginine and imidazole out-competes the NP-pep for protein binding, thereby releasing the bound proteins. This enrichment was reproducible across multiple NP-Pep synthesis batches and multiple replicates (Figure S3.3) indicating the robustness and reproducibility of this enrichment workflow, demonstrating effective enrichment of cTnI and its interacting partners.

We next evaluated the cTn enrichment performance of the NP-Pep across n = 4 apparently healthy donor heart samples by SDS-PAGE and by top-down reverse-phase liquid chromatography (LC)/MS analysis of the initial protein loading mixture, the resulting flow through, and the final elution mixtures after cTn enrichment (Figure 3.2). As demonstrated by SDS-PAGE, the NP-Pep was able to significantly and consistently enrich cTnT, cTnI, and cTnC from across the n = 4 samples, while avoiding non-specific binding of other proteins that were prevalent in the loading mixture (Figure 3.2a) . Additionally, top-down MS of the denatured and desalted protein samples provided a bird's eye view of all cTn monomer proteoforms present for direct quantification of the

relative abundance of individual cTn subunit proteoforms relative to its respective total subunit proteoforms. As demonstrated by the total ion chromatogram traces representative of 500 ng equal protein loading across each sample (Figure 3.2b), the elution analysis (black trace) confirms significantly enriched cTnI, cTnT and cTnC when present in all samples, demonstrating higher enrichment by absolute intensity comparison, while analysis by relative intensity demonstrates that NP-Pep preserved the endogenous cTnI, cTnT, and cTnC proteoform distributions and PTM profiles without artifactual modifications (Figure 3.2c).

Analysis of intact cTn complexes by Native-PAGE

After determining that the NP-Pep can effectively enrich for cTnI monomers as well as its interactors such as cTnT, and cTnC, we sought to determine whether the enriched and eluted cTnI was retained in an intact binary (cTnI-cTnC) or ternary (cTnT-cTnI-cTnC) protein complex. Analysis of the elution mixture by Native-PAGE[38], an electrophoresis technique which separates proteins by molecular weight without the use of denaturing surfactants, suggested that the predominant form of the cTn complex after amino acid elution was the ternary cTn complex, as seen by the strong dominant signal of the cTn at ~77 kDa (Figure 3.3a). However, upon attempting direct MS analysis of this NP-Pep elution mixture containing the ternary cTn visualized by Native-PAGE, we were unable to obtain a high-resolution MS spectrum, due to the interference of non-volatile salts and buffers which impeded direct MS analysis of the protein complex (Figure S3.4).

Online SEC optimization for rapid buffer exchange of proteins and protein complexes

After NP-pep protein enrichment and elution, removal of non-volatile buffers and salts is critical for preparing high quality protein samples for native mass spectrometry analysis[39]. However, amino acids such as L-arginine form protonated clusters which compete for protein

ionization during the electrospray ionization process, and imidazole can act as a charge-reducing agent which may alter the charge distribution and signal intensity of native protein complexes[40, 41] (Figure S3.4). This is conceivably why this elution strategy has not yet been used for the native mass spectrometry analysis of affinity-purified protein complexes. Therefore, to effectively analyze the enriched proteins by native mass spectrometry analysis, we employed size exclusion chromatography (SEC) for rapid online buffer exchange (OBE)[25], to effectively separate proteins of interest from non-volatile salts and buffers, while transferring proteins into MS-compatible solutions such as ammonium acetate. Here, we utilized an SEC column with a neutral, polar stationary phase (PolyHYDROXYLETHYL A), a 200 Å matrix pore size to elute proteins in the volume void volume, and 200 mM ammonium acetate (pH 7.0) as mobile phase, optionally equipped with a stainless steel-tee splitter for offline MS characterization by direct infusion (Figure S3.5). Upon analysis of the TIC chromatograms, we successfully observed effective size-based fractionation of endogenous cardiac troponin complexes (Figure 3.3b). Moreover, small molecules and non-volatile buffers were fully removed by SEC, eluting later and subsequently diverted away from the ESI source to avoid fouling of the mass spectrometer (Figure S3.6, S3.7).

Native MS analysis of the intact cTn complex

With native SEC for rapid OBE utilizing 200 mM ammonium acetate (pH 7.0) as mobile phase, proteins and protein complexes were effectively separated from non-volatile buffer components, thereby enabling detection on the Bruker maXis II Q-TOF. Native SEC-MS revealed subunit monomers, cTn dimers (cTnI-cTnC), and native cTn complex trimers (cTnT-cTnI-cTnC) that we further analyzed. With respect to the subunit monomers of cTnC, a calcium-binding protein with three divalent cation binding sites, the total calcium binding (Ca_{Total}) was determined to be

1.60 mol Ca^{2+} (+ 40 Da mass shift from the unmodified cTnC proteoform) per mol cTnC, with the doubly Ca^{2+}-bound cTnC proteoform as the most abundant protein species (Figure 3.4a). With respect to cTn dimers (cTnI-cTnC), we detected bisphosphorylated cTnI interacting only with the doubly or triply Ca^{2+}-bound proteoform of cTnC, a finding in agreement with existing literature on the physiological interactions of cTnI with Ca^{2+}-bound cTnC (Figure 3.4b, 3.4c). With respect to cTn trimers (cTnT-cTnI-cTnC), we detected unphosphorylated and monophosphorylated cTnT, despite the presence of multiple cTnT proteoforms detected from the same human heart sample by denaturing reverse phase liquid chromatography-MS. Monophosphorylated cTnT interacted with bisphosphorylated cTnI, together with doubly and triply Ca^{2+}-bound proteoforms of cTnC, with the triply Ca^{2+}-bound state of cTnC being the most abundant protein species of the cTn heterotrimer complex (Figure 3.4d).

The role of Ca^{2+} in the cTn complex

To probe the role of Ca^{2+} in stabilizing the ternary (cTnT-cTnI-cTnC) cTn, we introduced EGTA as a metallic ion chelator to remove bound Ca^{2+} ions from eluted cTn. Interesting, the triply Ca^{2+}-bound proteoforms of cTnC within the cTn ternary complex substantially decreased upon addition of EGTA, with only the doubly Ca^{2+}-bound forms retained after reaction with EGTA (Figure 3.5a) . Despite the troponin complex containing up to three Ca^{2+} binding sites, additional EGTA (up to 100 mM) was unable to dissociate or reduce the doubly bound Ca^{2+} proteoform in the ternary complex to the singly bound Ca^{2+} proteoform while simultaneously keeping the ternary complex intact (Figure 3.5a). This is in good agreement with prior literature stating that two Ca^{2+} ions exist within cTnC that function as essential structural ions which promote association of cTnC with cTnI and cTnT[42-44]. However, as the concentration of EGTA increased in the reaction of

EGTA with the eluted proteins, the relative intensity of the intact ternary cardiac troponin complex decreased relative to free troponin monomers (Figure 3.5b). This may imply that the third Ca^{2+} bound to cTnC which is known to have a functional role in the ternary complex, may also have a structural role that when perturbed, causes cTnC to dissociate from cTnI and cTnT[45]. Future experiments, such as using IM-MS, will be critical to elucidate whether the differences in calcium binding contribute to substantially different collisional cross-sections arising from intrinsic structural differences when analyzing cTnC with 2 Ca^{2+} or 3 Ca^{2+} bound in the ternary cTn complex. Additionally, native MS reveals that the binding affinities of Ca^{2+} is lower in free cTnC monomer compared to Ca^{2+} bound in ternary cTn complexes. At 10 mM EGTA, the Ca_{Total} of cTnC is approximately 1.02 Ca^{2+} per mol cTnC, where the doubly bound Ca^{2+} proteoform is present at 40% relative abundance to all cTnC proteoforms present. However, upon addition of up to >25 mM EGTA, the doubly bound Ca^{2+} proteoform in cTnC monomer almost completely diminishes, resulting in a Ca_{Total} of approximately 0.11 Ca^{2+} per mol cTnC, where the doubly bound Ca^{2+} proteoform is present at <10% relative abundance (Figure 3.5c). This is in good agreement with previous reports that show the Ca^{2+} dissociation rate from cTnC significantly slows upon addition of cTnI and cTnT, indicating that cTnI and cTnT influence the Ca^{2+} binding properties of the cTn in solution, as recapitulated from our data[42-44, 46].

Native Nanoproteomics enables analysis of cTn complexes from multiple disease states

Finally, to demonstrate that this technology can be used for the study of cTn interactions from a variety of human heart samples representative of cardiovascular disease states, we performed enrichment on tissue samples that are representative of ischemic cardiomyopathy (ICM) and dilated cardiomyopathy (DCM) tissues. As visualized by SDS-PAGE, NP-Pep was able to

effectively enrich the cTn monomers while avoiding pull-down of other highly abundant proteins across the three tissue types (Figure 3.6a). While the most abundant proteoforms in cTnC and cTnT are consistent among the three different disease-state tissues, there is significant difference in the phosphorylation level of cTnI across the three different disease states (Figure 3.6b). Analysis of the protein complex proteoform profiles between the donor heart, ICM, and DCM heart, illustrate 80 Da mass shifts that align well with differences in the phosphorylation level in cTnI monomers, with 40 Da shifts likely indicating doubly or triply Ca^{2+} bound proteoforms of cTnC (Figure 3.6c). Notably, cTnI phosphorylation is considered as an integral mechanism in cardiac homeostasis and has potential to reveal the functional status of the heart[47]. Altered PTM profiles of cTnI are associated with dysregulated cellular signaling during the onset and progression of diseases, thus disease-induced cTnI proteoforms, combined with protein complex stoichiometry and Ca^{2+}-cofactor binding, may provide new insights into PTM structure-function relationships intrinsic to cTn, and has the potential to serve as the next-generation cardiac biomarkers for diagnosis of specific cardiovascular syndromes.

Conclusion

Herein, we demonstrate a “native nanoproteomics” strategy incorporating functionalized nanomaterials for specific affinity-based enrichment of intact cardiac troponin complexes (~77 kDa) under physiological pH and native size exclusion chromatography (SEC) for rapid online buffer exchange (OBE), thereby enabling the native MS analysis of cardiac troponin complexes from human heart samples for the first time. Using native MS, we revealed the varying stoichiometry (Ca_{Total}) of non-covalent Ca^{2+} cofactors which regulate cardiac contraction and relaxation, in free cTn monomers, dimers, and trimers. This robust and reproducible ‘native

nanoproteomics' technology can be used as a structural biology tool for proteomics applications to interrogate protein stoichiometry, non-covalent cofactor binding interactions, and post-translational modifications and can be used for the analysis of cardiac troponin complexes from human heart samples representative of various cardiomyopathies and disease states.

Acknowledgement

The authors would like to thank National Institute of Health for financial support, R01 GM117058 (to S. J. and Y. G.) and R01 GM125085 (to Y. G.). Y. G. also would like to acknowledge the NIH grants, R01 HL096971, R01 HL109810, and S10 OD018475.

References:

1. 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.
2. 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.
3. Messer, A. E.; Marston, S. B., Investigating the role of uncoupling of troponin I phosphorylation from changes in myofibrillar Ca2+-sensitivity in the pathogenesis of cardiomyopathy. *Frontiers in physiology* **2014,** *5*, 315.
4. Farah, C. S.; Reinach, F. C., The troponin complex and regulation of muscle contraction. *The FASEB Journal* **1995,** *9* (9), 755-767.
5. Zhang, R.; Zhao, J.; Mandveno, A.; Potter, J. D., Cardiac troponin I phosphorylation increases the rate of cardiac muscle relaxation. *Circulation research* **1995,** *76* (6), 1028-1035.
6. Peng, Y.; Gregorich, Z. R.; Valeja, S. G.; Zhang, H.; Cai, W. X.; Chen, Y. C.; Guner, H.; Chen, A. J.; Schwahn, D. J.; Hacker, T. A.; Liu, X. W.; Ge, Y., Top-down Proteomics Reveals Concerted Reductions in Myofilament and Z-disc Protein Phosphorylation after Acute Myocardial Infarction. *Mol Cell Proteomics* **2014,** *13* (10), 2752-2764.
7. Cheng, Y.; Regnier, M., Cardiac troponin structure-function and the influence of hypertrophic cardiomyopathy associated mutations on modulation of contractility. *Archives of Biochemistry and Biophysics* **2016,** *601*, 11-21.
8. 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.
9. 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.
10. Robinson, P.; Griffiths, P. J.; Watkins, H.; Redwood, C. S., Dilated and Hypertrophic Cardiomyopathy Mutations in Troponin and α-Tropomyosin Have Opposing Effects on the Calcium Affinity of Cardiac Thin Filaments. *Circulation Research* **2007,** *101* (12), 1266-1273.
11. Metskas, L. A.; Rhoades, E., Order-Disorder Transitions in the Cardiac Troponin Complex. *Journal of molecular biology* **2016,** *428* (15), 2965-2977.
12. 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.
13. Sahdev, S.; Khattar, S. K.; Saini, K. S., Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry* **2008,** *307* (1), 249-264.
14. Tamara, S.; den Boer, M. A.; Heck, A. J. R., High-Resolution Native Mass Spectrometry. *Chem Rev* **2021**.
15. 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.

16. 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.
17. Crittenden, C. M.; Novelli, E. T.; Mehaffey, M. R.; Xu, G. N.; Giles, D. H.; Fies, W. A.; Dalby, K. N.; Webb, L. J.; Brodbelt, J. S., Structural Evaluation of Protein/Metal Complexes via Native Electrospray Ultraviolet Photodissociation Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2020,** *31* (5), 1140-1150.
18. Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E., The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry* **2014,** *6* (4), 281-294.
19. Larson, E. J.; Roberts, D. S.; Melby, J. A.; Buck, K. M.; Zhu, Y.; Zhou, S.; Han, L.; Zhang, Q.; Ge, Y., High-Throughput Multi-attribute Analysis of Antibody-Drug Conjugates Enabled by Trapped Ion Mobility Spectrometry and Top-Down Mass Spectrometry. *Analytical Chemistry* **2021,** *93* (29), 10013-10021.
20. 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.
21. 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.
22. Roberts, D. S.; Chen, B.; Tiambeng, T. N.; Wu, Z.; Ge, Y.; Jin, S., Reproducible large-scale synthesis of surface silanized nanoparticles as an enabling nanoproteomics platform: Enrichment of the human heart phosphoproteome. *Nano research* **2019,** *12* (6), 1473.
23. Chen, B.; Hwang, L.; Ochowicz, W.; Lin, Z.; Guardado-Alvarez, T. M.; Cai, W.; Xiu, L.; Dani, K.; Colah, C.; Jin, S.; Ge, Y., Coupling functionalized cobalt ferrite nanoparticle enrichment with online LC/MS/MS for top-down phosphoproteomics. *Chemical Science* **2017,** *8* (6), 4306-4311.
24. Hwang, L.; Ayaz-Guner, S.; Gregorich, Z. R.; Cai, W.; Valeja, S. G.; Jin, S.; Ge, Y., Specific Enrichment of Phosphoproteins Using Functionalized Multivalent Nanoparticles. *Journal of the American Chemical Society* **2015,** *137* (7), 2432-2435.
25. 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.
26. Park, J.; An, K.; Hwang, Y.; Park, J.-G.; Noh, H.-J.; Kim, J.-Y.; Park, J.-H.; Hwang, N.-M.; Hyeon, T., Ultra-large-scale syntheses of monodisperse nanocrystals. *Nature Materials* **2004,** *3* (12), 891-895.
27. 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.
28. 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. Academic Press: 2019; Vol. 626, pp 347-374.

29. Xiao, X.; Kuang, Z.; Slocik, J. M.; Tadepalli, S.; Brothers, M.; Kim, S.; Mirau, P. A.; Butkus, C.; Farmer, B. L.; Singamaneni, S.; Hall, C. K.; Naik, R. R., Advancing Peptide-Based Biorecognition Elements for Biosensors Using in-Silico Evolution. *ACS Sensors* **2018,** *3* (5), 1024-1031.
30. Firer, M., Efficient elution of functional proteins in affinity chromatography. *Journal of biochemical and biophysical methods* **2001,** *49* (1-3), 433-442.
31. Moser, A. C.; Hage, D. S., Immunoaffinity chromatography: an introduction to applications and recent developments. *Bioanalysis* **2010,** *2* (4), 769-790.
32. Hengen, P. N., Purification of His-Tag fusion proteins from Escherichia coli. *Trends in Biochemical Sciences* **1995,** *20* (7), 285-286.
33. Einhauer, A.; Jungbauer, A., The FLAG™ peptide, a versatile fusion tag for the purification of recombinant proteins. *Journal of Biochemical and Biophysical Methods* **2001,** *49* (1), 455-465.
34. 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.
35. Hansen, A. L.; Kay, L. E., Measurement of histidine pK_a values and tautomer populations in invisible protein states. *Proceedings of the National Academy of Sciences* **2014,** *111* (17), E1705-E1712.
36. 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.
37. 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.
38. Schagger, H.; Cramer, W.; Vonjagow, G., Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Analytical biochemistry* **1994,** *217* (2), 220-230.
39. 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., Best practices and benchmarks for intact protein analysis for top-down mass spectrometry. *Nature methods* **2019,** *16* (7), 587-594.
40. Zhang, D.; Wu, L.; Koch, K. J.; Cooks, R. G., Arginine clusters generated by electrospray ionization and identified by tandem mass spectrometry. *European Mass Spectrometry* **1999,** *5* (5), 353-361.
41. 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.
42. Tikunova, S. B.; Davis, J. P., Designing Calcium-sensitizing Mutations in the Regulatory Domain of Cardiac Troponin C*. *Journal of Biological Chemistry* **2004,** *279* (34), 35341-35352.
43. Davis, J. P.; Tikunova, S. B., Ca2+ exchange with troponin C and cardiac muscle dynamics. *Cardiovascular Research* **2007,** *77* (4), 619-626.
44. 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.

45. Cox, J. A.; Comte, M.; Stein, E. A., Calmodulin-free skeletal-muscle troponin C prepared in the absence of urea. *Biochemical Journal* **1981,** *195* (1), 205-211.
46. Johnson, J. D.; Collins, J. H.; Robertson, S. P.; Potter, J. D., A fluorescent probe study of Ca2+ binding to the Ca2+-specific sites of cardiac troponin and troponin C. *Journal of Biological Chemistry* **1980,** *255* (20), 9635-9640.
47. Solaro, R. J.; Henze, M.; Kobayashi, T., Integration of Troponin I Phosphorylation With Cardiac Regulatory Networks. *Circulation Research* **2013,** *112* (2), 355-366.

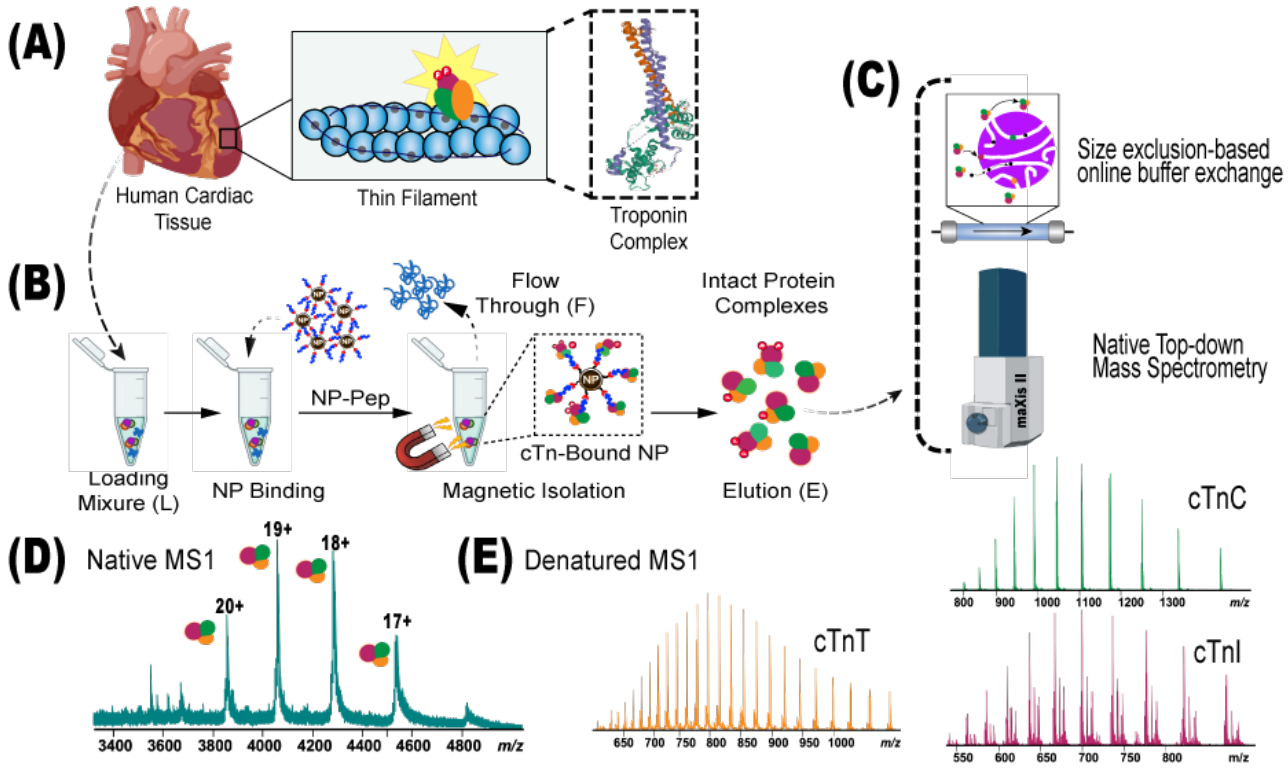


Figure 3.1. Native Nanoproteomics for the affinity purification of endogenous cTns from human heart tissues. a) Schematic cartoon of the cardiac troponin complex bound on the actin thin filament. b) Schematic illustration of the native nanoproteomics strategy for cTn complex enrichment from heart tissue extract. The loading mixture, L, is first incubated with the NP-Pep. Following magnetic isolation, the nonspecific proteins are removed as flow through (F). The NPs are washed and the NP-bound proteins of interest are eluted (E). c) Native proteins and protein complexes are separated by size exclusion chromatography for rapid buffer exchange into 200 mM ammonium acetate , coupled online to Bruker QTOF Maxis for native top-down mass spectrometry. d) Native MS1 profile of the ternary cardiac troponin (T-I-C) complex, ~77 kDa. e) Denatured MS1 profile of the cardiac troponin monomer subunits.

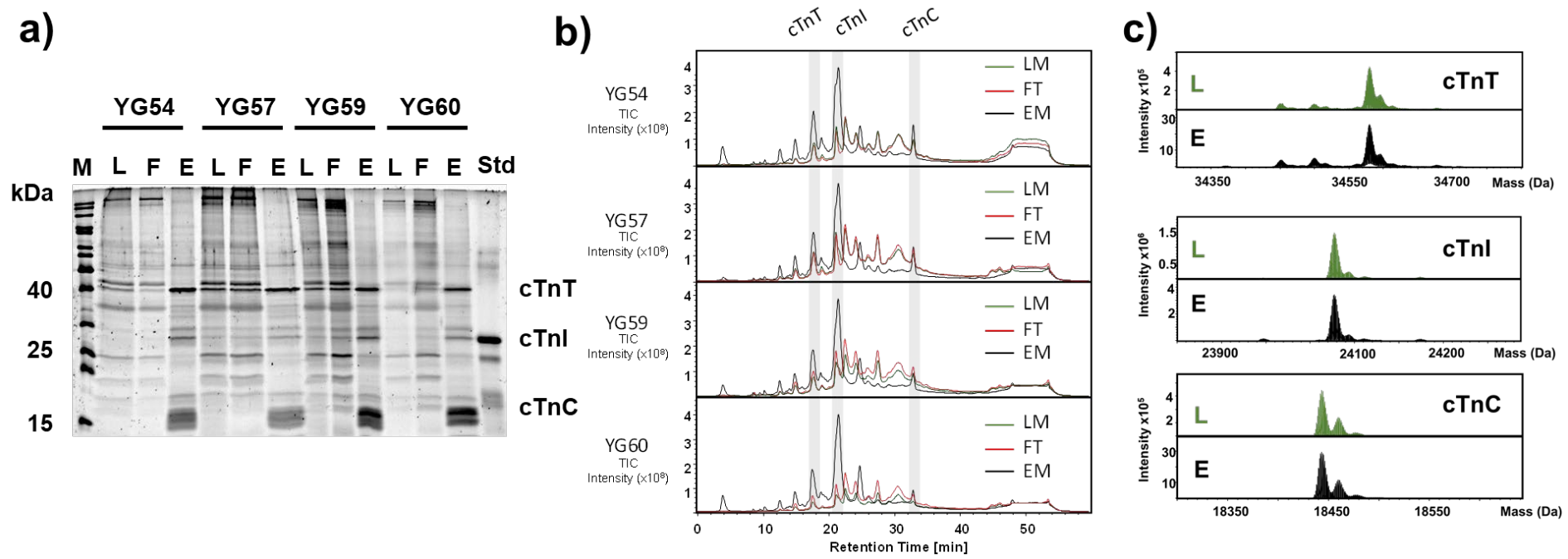


Figure 3.2. Evaluation of enrichment performance using native elution buffer (750 mM L-Arg, 750 mM imidazole, 50 mM L-Glu, pH 7.5). a) SDS-PAGE stained with SYPRO Ruby visualizing the cTnI enrichment performance from sarcomeric protein extracts obtained from n = 4 apparently health donor heart samples. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel. M, marker. b) TIC-MS analysis of LM, FT, EM, demonstrating enrichment of cTnT, cTnI, and cTnC. c) MS-based evaluation of cTnT, cTnI, and cTnC enrichment by NP-Pep, demonstrating preservation of enriched proteoforms while significantly enriching cTnT, cTnI, and cTnC compared to the initial loading mixture.

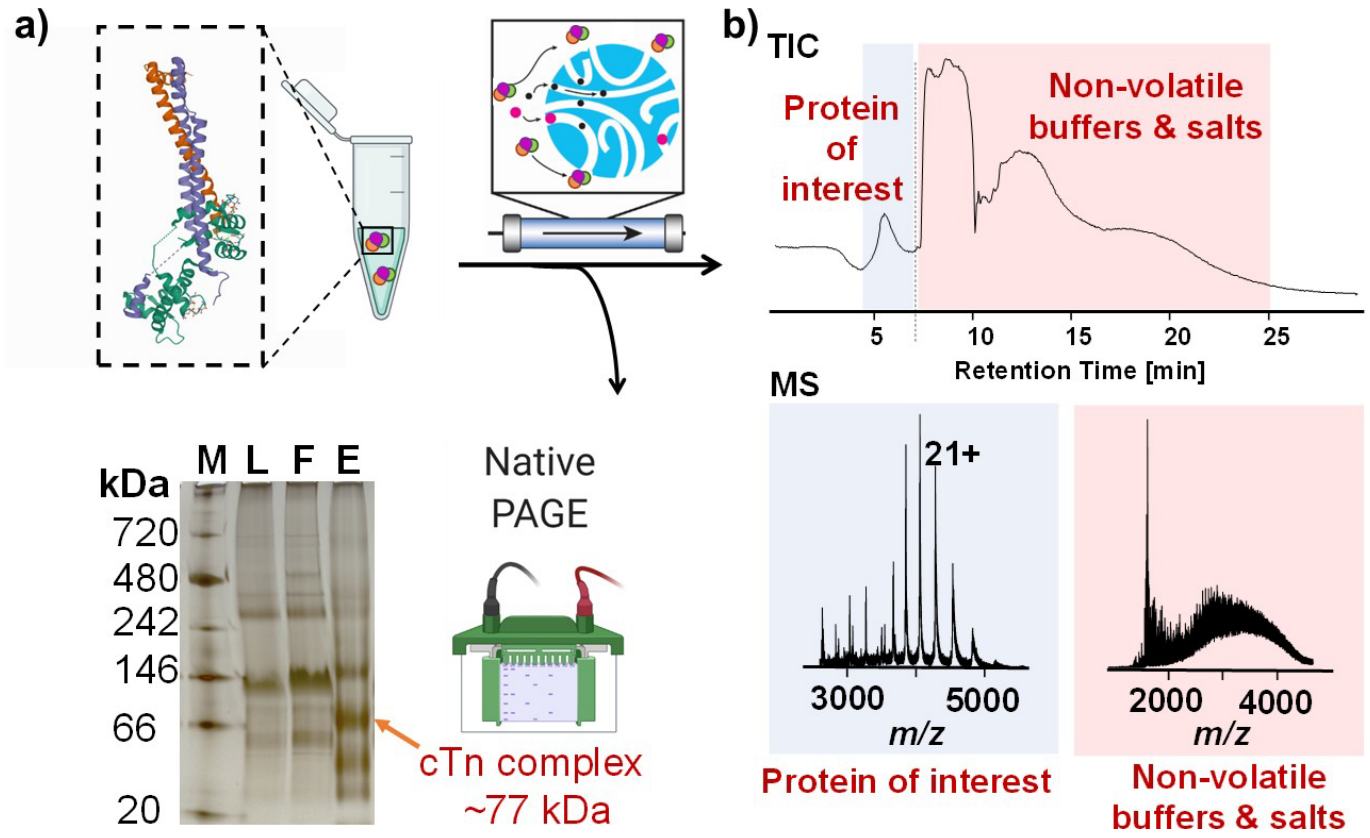
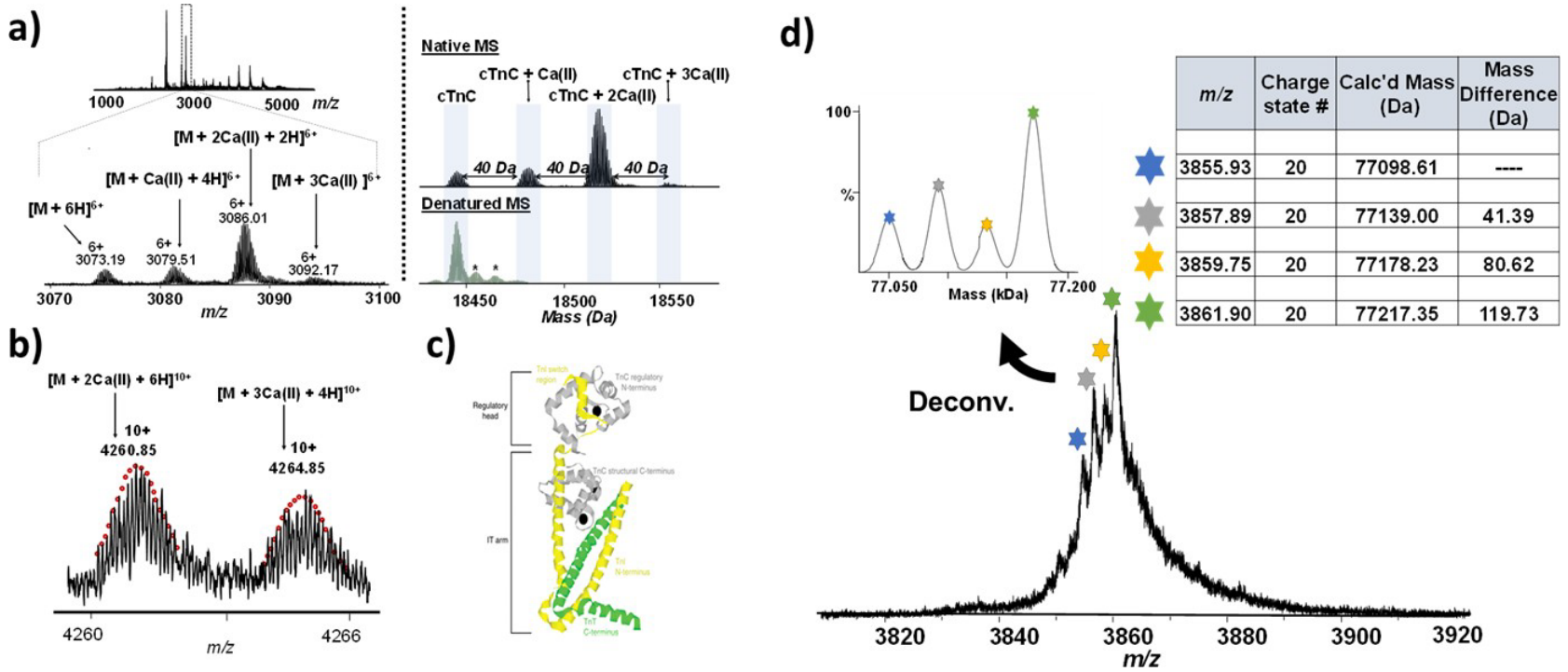


Figure 3.3. Size exclusion chromatography for online buffer exchange enables native MS analysis of enriched protein and protein complexes. a) Schematic representation of NP-Pep enriched cTn for analysis by native PAGE or online buffer exchange-MS (a) Native PAGE analysis demonstrates enrichment of cTn complex (77 kDa) from initial protein loading mixture. M, marker; L, loading mixture; F, flow-through; E, elution. b) Total ion chromatogram (TIC) of enriched cTn mixture, where proteinselute first, followed by non-volatile buffers and salts.



m/z	Charge state #	Calc'd Mass (Da)	Mass Difference (Da)
3855.93	20	77098.61	----
3857.89	20	77139.00	41.39
3859.75	20	77178.23	80.62
3861.90	20	77217.35	119.73

Figure 3.4. Native MS reveals non-covalent Ca^{2+} binding of cTnC, cTn(I-C) dimers, and cTn(T-I-C) ternary complexes. a) Enlarged view of charge state 6+ of cTnC proteoforms, revealing noncovalent Ca^{2+} cofactor binding. The total calcium binding (Ca_{Total}) was 1.60 mol Ca^{2+} per mol cTnC, with the doubly Ca^{2+}- bound cTnC proteoform as the most abundant protein species. Deconvoluted spectra of enriched cTnC by native MS (top) and denatured MS analysis, revealing Ca^{2+} only by native MS. **b)** Enlarged view of charge state 10+ of cTn(I-C) dimer proteoforms (~42 kDa) by native MS analysis. The total calcium binding of cTn(I- C) was 2.41 mol Ca^{2+} per mol cTn(I-C), with the doubly Ca^{2+}-bound cTnC proteoform as the most abundant protein species **c)** Crystal structure of Ca^{2+} (black circles) saturated cTn(I-C) complex, depicting two Ca^{2+} ions which are structural (IT arm) and one Ca^{2+} which is functional (regulatory head). **d)** Enlarged view of charge state 20+ of the cTn(I-T-C) heterotrimeric complex (~77 kDa) enriched from a donor heart sample, with the zero-charge deconvoluted spectrum (UniDec) shown in the inset.

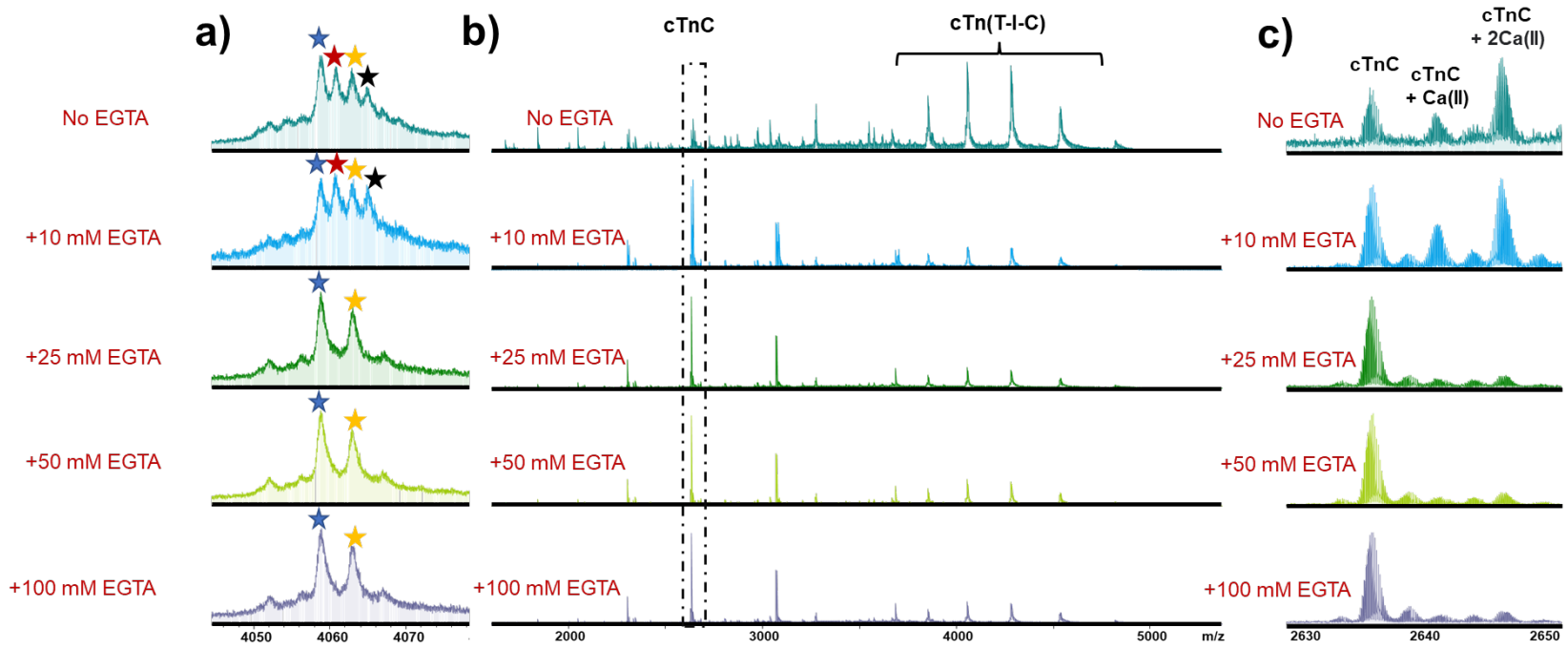


Figure 3.5. The role of Ca^{2+} in the cTn complex. a) Enlarged view of charge state 19+ of the ternary cTn(T-I-C) complex, demonstrating the change in Ca^{2+}-bound proteoforms of cTnC within the ternary cTn complex as EGTA increases. Blue star; *pp*cTnI + cTnT + cTnC + 2Ca (II); Red star; *pp*cTnI + cTnT + cTnC + 3Ca (II); Yellow star; *pp*cTnI + *p*cTnT + cTnC + 2Ca (II); Black star; *pp*cTnI + *p*cTnT + cTnC + 3Ca (II). As EGTA increases, triply Ca^{2+} bound proteoforms of cTnC decrease, but doubly Ca^{2+} bound proteoforms of cTnC are retained. **b)** Native MS spectra highlighting changes in relative intensity of the intact ternary cTn(T-I-C) complex as EGTA concentration increases. **c)** Enlarged view of charge state 7+ of cTnC monomer, demonstrating change in Ca^{2+}-bound proteoforms of cTnC within the free monomer as EGTA increases. Doubly Ca^{2+} bound proteoforms decrease, indicating different binding affinities of Ca^{2+} in the free cTnC monomer compared to cTnC in a ternary complex.

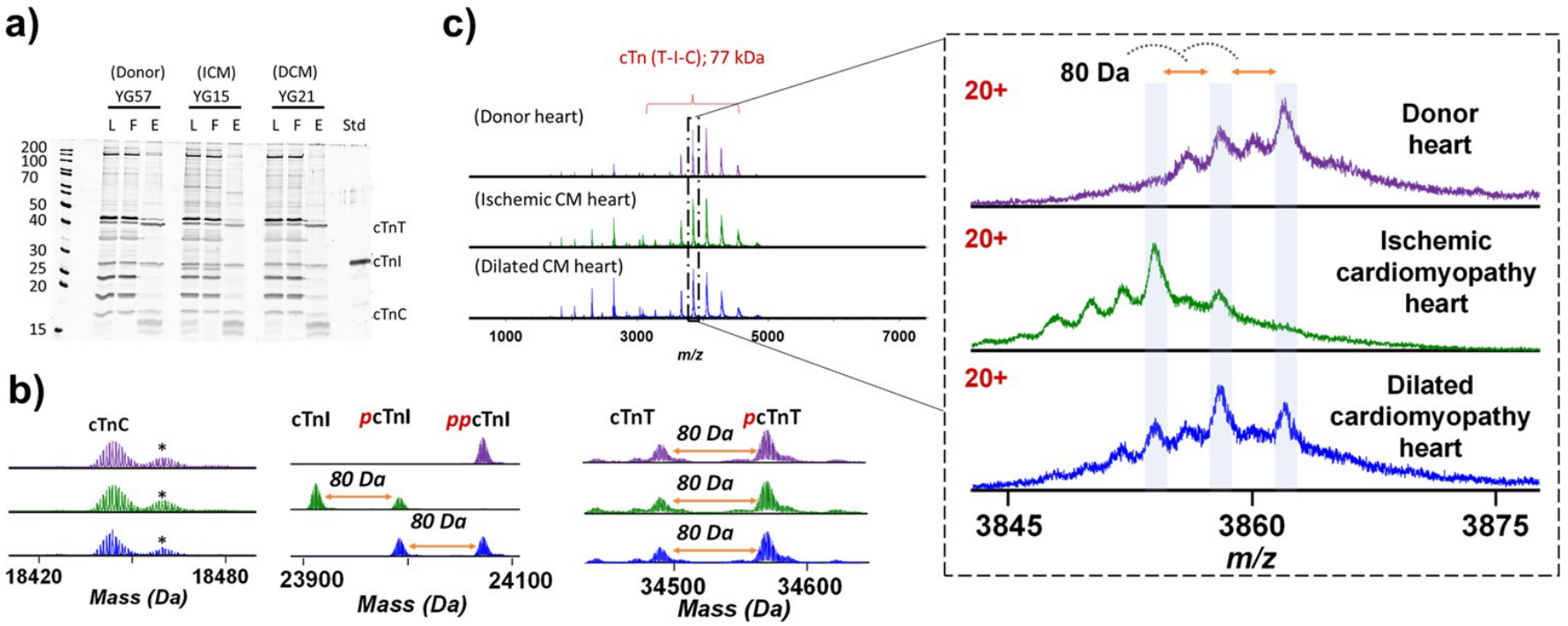


Figure 3.6. Native Nanoproteomics enables analysis of cTn complexes from multiple disease states

a) SYPRO Ruby visualizing the cTnI enrichment performance using NP-Pep across three different tissue samples representative of an apparently healthy donor heart, an ischemic cardiomyopathy heart, and a dilated cardiomyopathy heart. b) Deconvoluted MS spectra (denatured) of the cTnC, cTnI, and cTnT proteoforms present in each heart sample. c) Native MS spectra of the ternary cTn (T-I-C) complexes, featuring an enlarged view of charge state 20+. 80 Da mass shifts highlight the difference in phosphorylation profile present in cTnI among the different tissue samples.

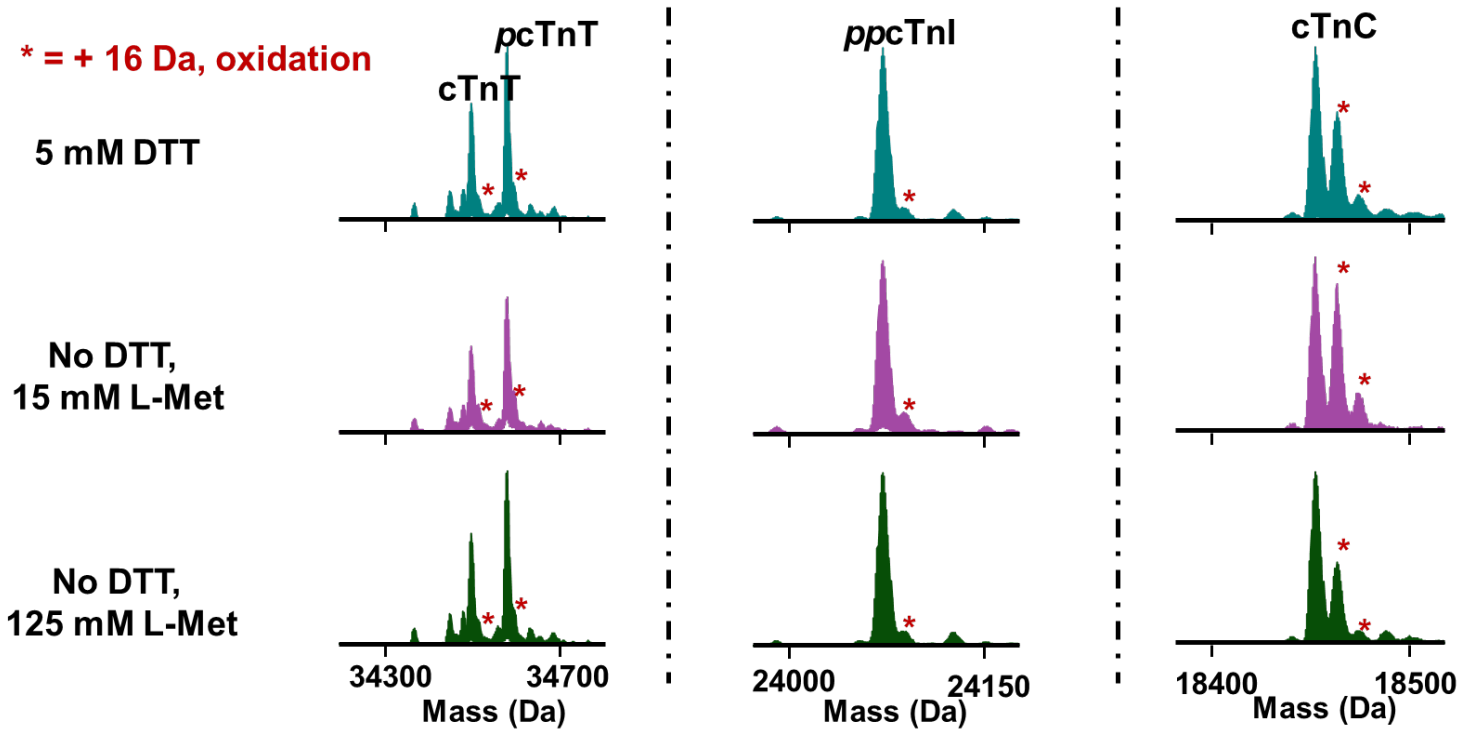


Figure S3.1. Optimization of L-methionine concentration as anti-oxidant additive on protein oxidation. Deconvoluted mass spectra of cTnT (34 kDa), cTnI (24 kDa) and cTnC (18 kDa), illustrating the differences in protein oxidation (+ 16 Da; red stars) after sarcomeric protein extraction procedure with different reducing agents. 125 mM L-methionine is the most effective additive to reduce protein oxidation during the extraction procedure, without potentially reducing disulfide bonds using stronger reducing agents such as DTT.

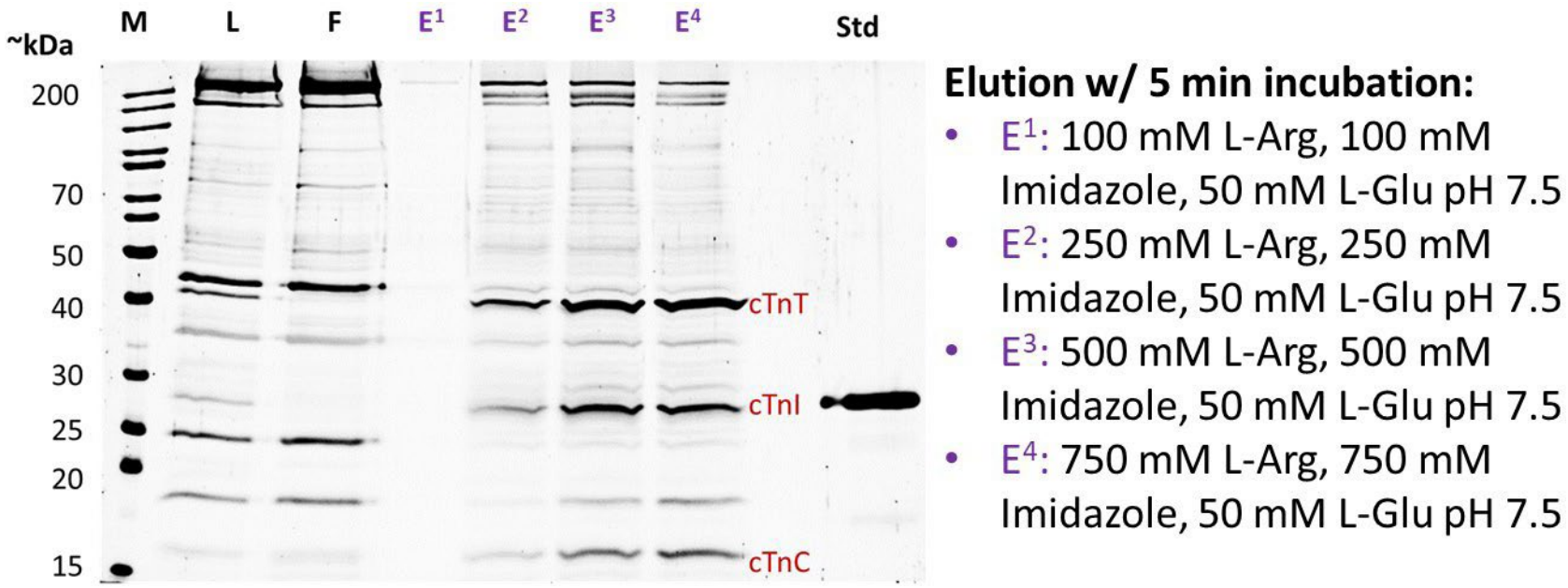


Figure S3.2. Optimization of amino acid elution buffer (L-arginine, imidazole, L-glutamic acid) from NP-Pep. SDS-PAGE stained with SYPRO Ruby visualizing sarcomeric proteins after the NP-Pep enrichment procedure from sarcomeric proteins extracted from an apparently healthy donor heart sample. Equal amount (500 ng) of the loading mixture (L), flow-through (F), and elution mixture after enrichment (E) was loaded on the gel.. Higher concentrations of imidazole and L-arginine (> 500 mM L-Arg, 500 mM imidazole, 50 mM L-Glu pH 7.5) are required for more effective elution from the NP-Pep, suggesting a mechanism of competitive elution. M, marker.

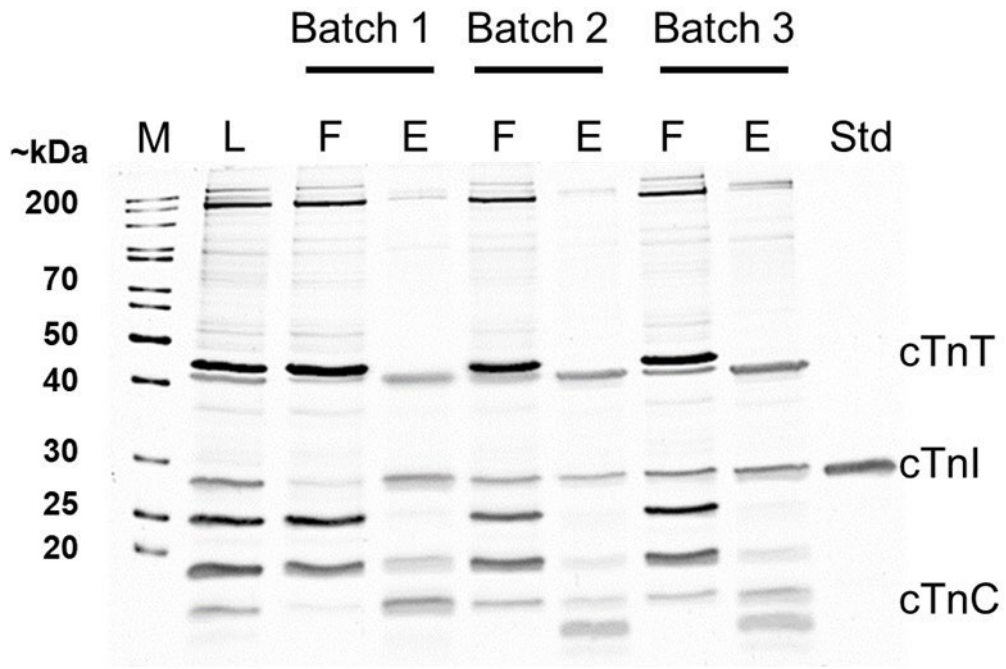


Figure S3.3. High reproducibility of inter-batch enrichment by NP-Pep. SDS-PAGE stained with SYPRO Ruby visualizing the cTnI enrichment performance and demonstrating the high reproducibility obtained from three different synthesis batches of NP-Pep, using 750 mM L-Arg, 750 mM Imidazole, 50 mM L-Glu, pH 7.5 as 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, marker.

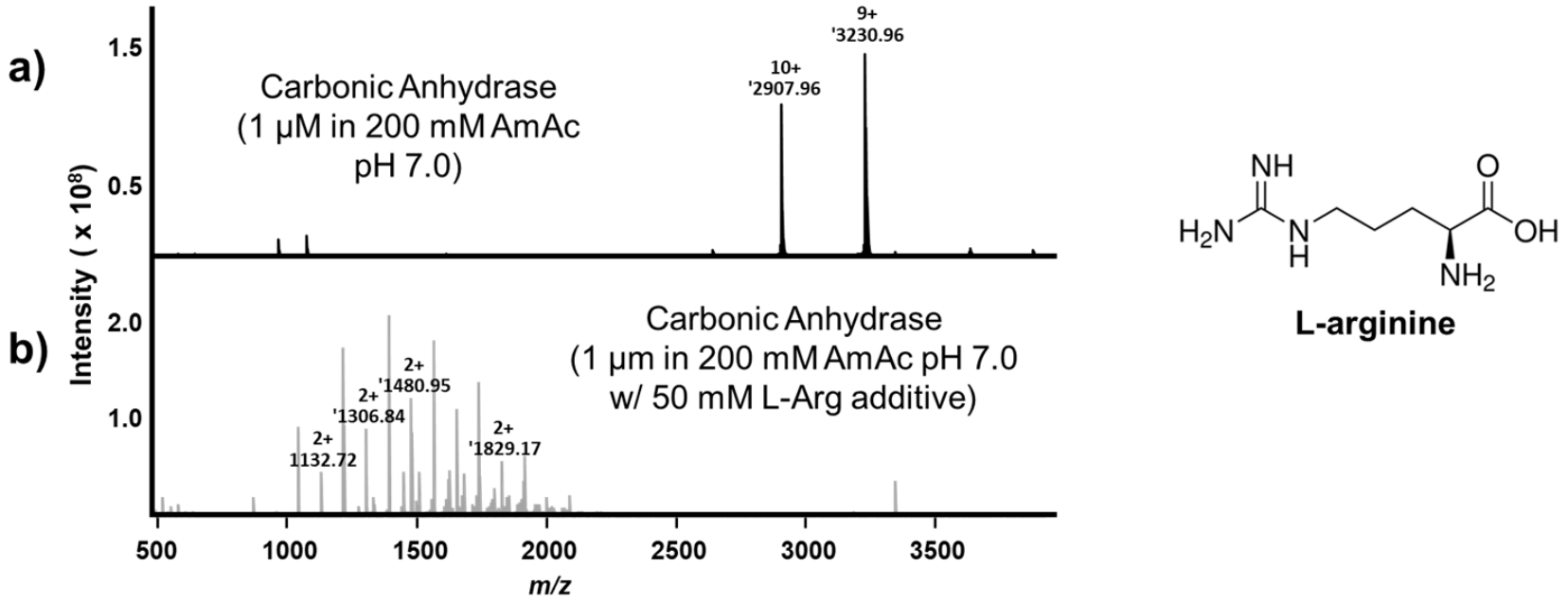


Figure S3.4. Native mass spectra of carbonic anhydrase (29 kDa), illustrating interference of L-arginine with protein mass spectrometry. a) Native MS of carbonic anhydrase reconstituted in 200 mM AmAc pH 7.0. **b)** Native MS of carbonic anhydrase reconstituted in 200 mM AmAc, 50 mML-Arg pH 7.0, demonstrating that L-arginine forms protonated clusters by electrospray ionization and suppresses native MS signal of proteins. Non- volatile salts and buffers must be removed prior to nMS analysis of enriched protein complexes.

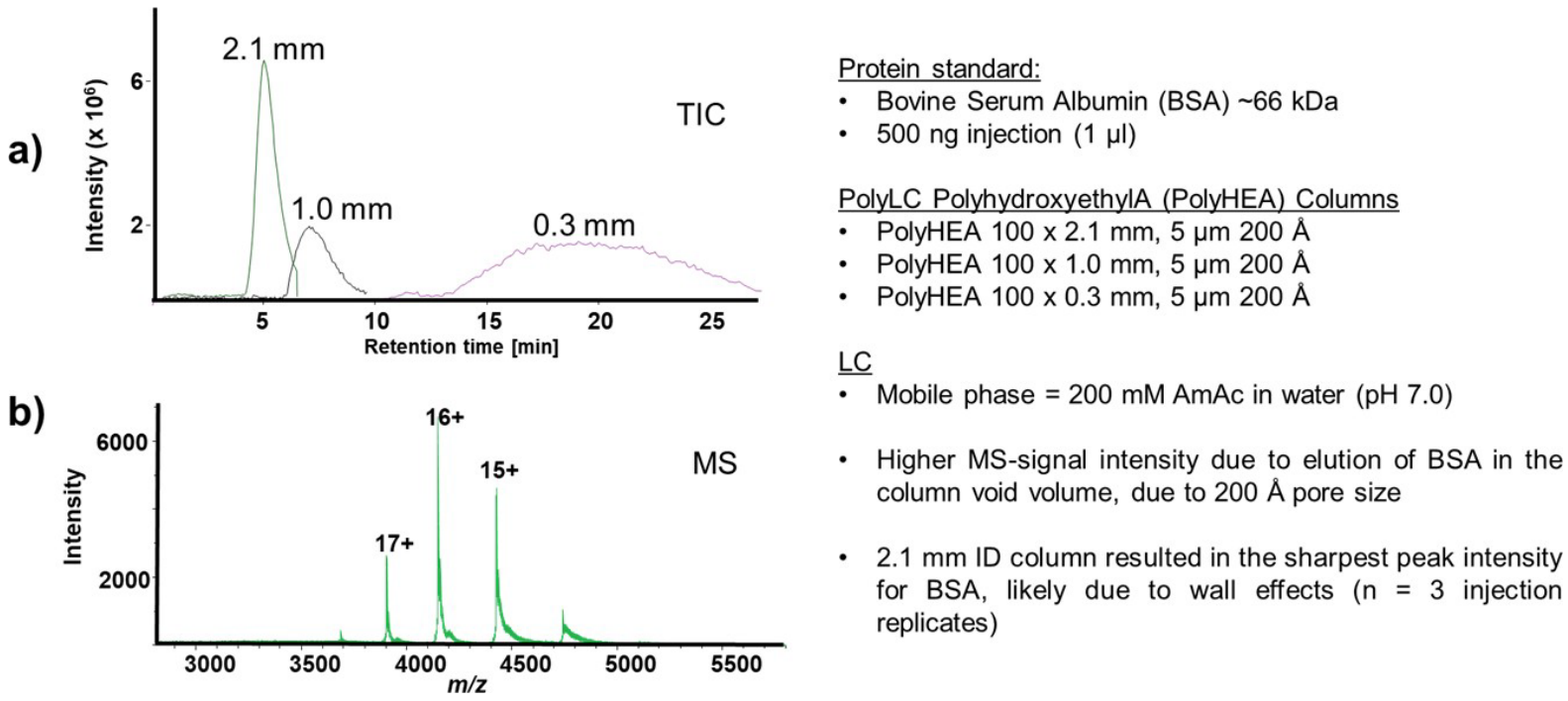


Figure S3.5. Online native size exclusion chromatography (SEC) column screening and optimization. A) Total ion chromatograms (TICs) of 500 ng bovine serum albumin (66 kDa) injected on a 2.1 mm, 1.0 mm, or 0.3 mm I.D. column. b) Representative **n**ative MS spectrum of BSA corresponding to the 2.1 mm I.D. column.

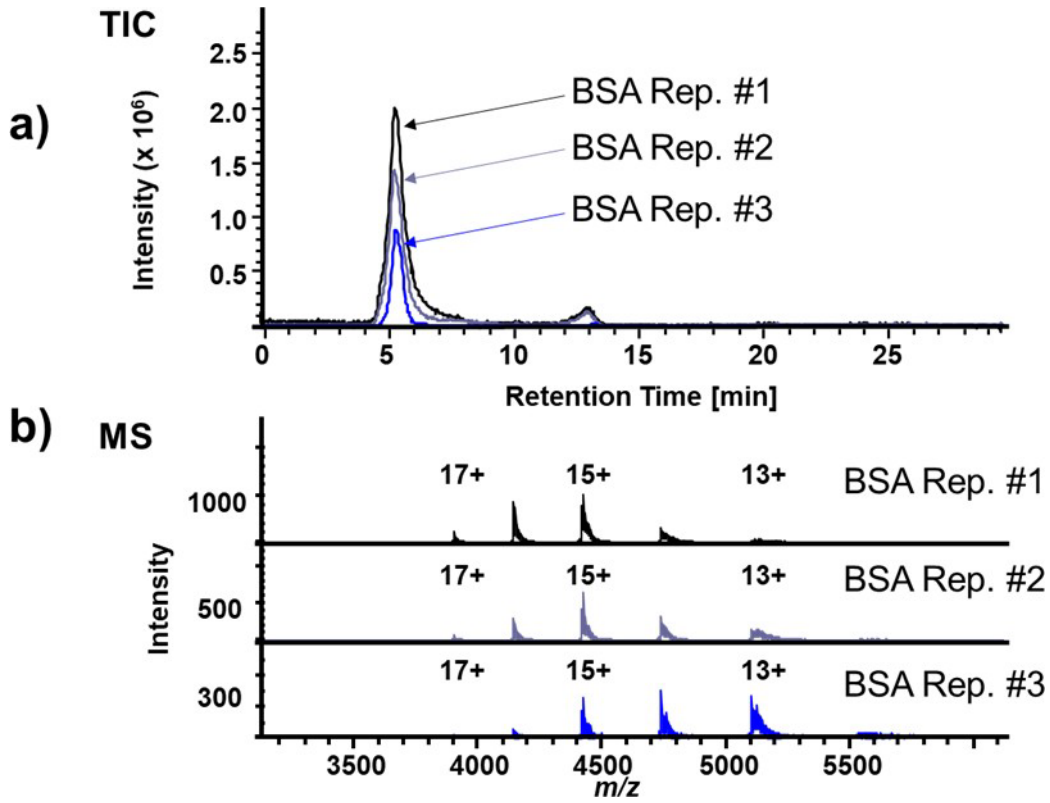


Figure S3.6. Native online SEC-MS illustrating that loading samples with high concentration of amino acid buffers can foulMS source. a) Total ion chromatogram mass spectra (TIC-MS) of three replicate injections of BSA loaded in between samples containing 750 mM L-Arg, 750 mM imidazole, 50 mM L-Glu, without column diverting and re-equilibration. b) Native MS spectra corresponding to the TIC plots in a), demonstrating that repeated online SEC-MS of protein mixtures containing high salt (750 mM L- Arg, 750 mM imidazole, 50 mM L-Glu) can foul the MS capillary resulting in loss of MS signal intensity and induce protein charge state reduction from residual salts on column (imidazole).

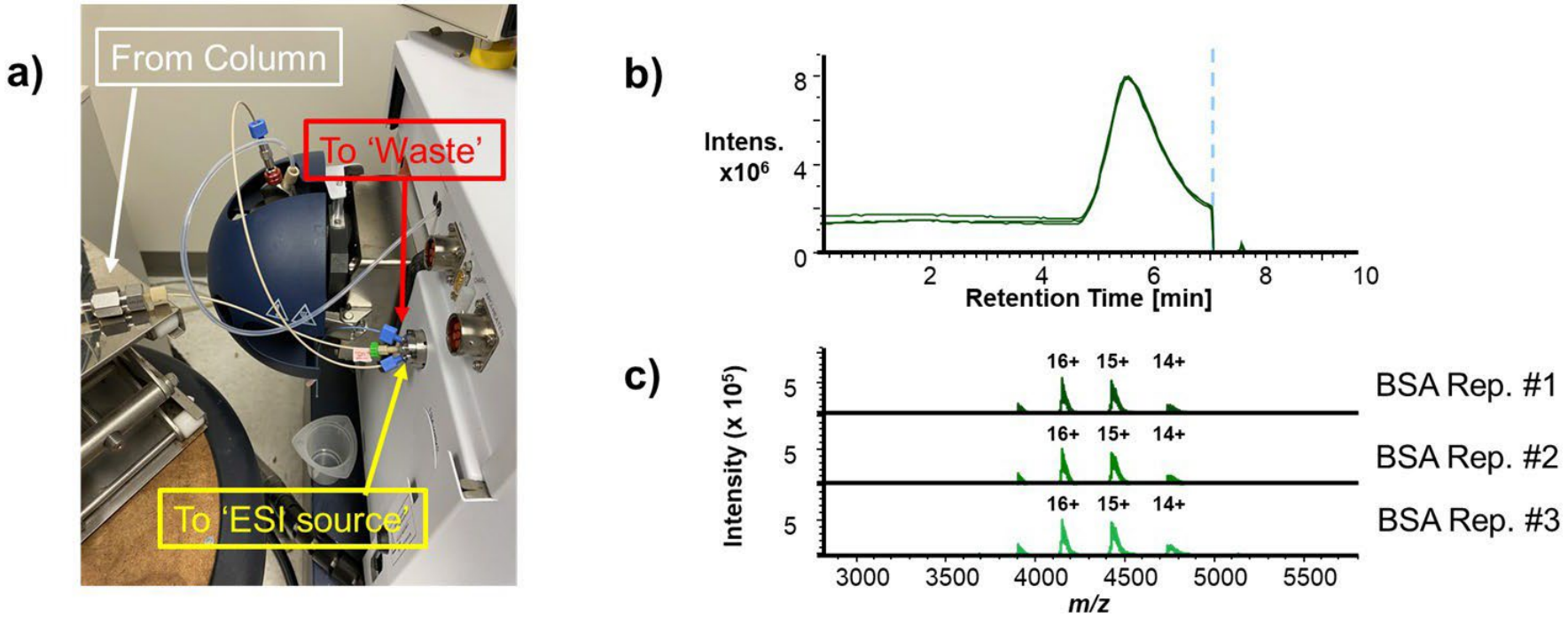


Figure S3.7. Implementation of a 'divert' valve and re-equilibrating the SEC column between samples improves reproducibility. a) Image of divert valve setup to divert SEC column eluate to waste after protein elutes and is detected by MS. b) Total ion chromatogram mass spectra (TIC-MS) of three replicate injections of BSA loaded in between samples containing 750 mM L-Arg, 750 mM imidazole, 50 mM L-Glu, with column diverting and re-equilibration. c) Native MS spectra corresponding to the TIC plots in b), demonstrating high reproducibility with no change in the protein charge state distribution

Chapter 4

Conclusion and Future Directions

Cardiac troponin I (cTnI) is a universally-accepted, clinical biomarker for the diagnosis of acute myocardial infarction (a.k.a., heart attack). cTnI is released into the bloodstream following cardiovascular injury where it circulates with low abundance (ng/ml) and in myriad proteoforms (e.g., phosphorylated, acetylated, oxidized, and truncated). Top-down mass spectrometry (MS)-based proteomics has emerged as a powerful technology for characterizing and quantifying proteoforms arising from post-translational modifications, alternative splicing, and sequence variations at the intact protein level. However, the large dynamic range of the blood proteome makes the MS-detection of low-abundance biomarker proteins extremely challenging. Thus, protein enrichment strategies are necessary to specifically capture and enrich low-abundance proteins from complex mixtures prior to MS-analysis.

Chapter 2 described the development of an integrated "nanoproteomics" method coupling peptide-functionalized superparamagnetic nanoparticles (NPs) with top-down MS for the enrichment and comprehensive analysis of cTnI proteoform monomers, directly from serum for the first time. We demonstrated that carefully designed peptide-functionalized NPs can directly capture and enrich cTnI from human serum with high specificity and reproducibility, while simultaneously depleting highly abundant blood proteins, such as human serum albumin. These NPs not only outperform conventional monoclonal antibody platforms for serum cTnI enrichment, but also faithfully and holistically preserve all endogenous cTnI proteoforms. By further applying to a large human cohort, patient blood samples can be analyzed to comprehensively detect all cTnI proteoforms and establish the relationships between cTnI proteoforms and underlying disease etiology.

Although our current nanoproteomics assay has a limit of cTnI detection (LOD) of 0.50 ng/mL, our technology has not yet met the diagnostic cutoff values comparable with contemporary ELISAs (≤ 0.04 ng/mL). To further develop our technology toward a clinical assay, it is essential to improve its specificity and its sensitivity to better detect cTnI in various concentrations representing a broad spectrum of the disease severity. Although functionalized NPs (NP-Pep) demonstrated impressive resistance to nonspecific adsorption from the highly abundant human serum albumin (~ 90% of total blood mass), we still observed non-specific binding of other serum proteins: namely apolipoprotein which is also abundant in human serum. We can further optimize the enrichment procedures and the surface chemistry of our NP-Pep to completely reduce the non-specific binding. We can use salt concentration as an optimization parameter since it plays a key role in promoting effective cTnI enrichment by screening against nonspecific electrostatic interactions from blood proteins. Additionally, we can investigate changing the wash buffer (currently Tris) to other buffer systems (PBS, HEPES, etc.) which possess different ionic strengths and potentially may promote more effective cTnI enrichment. We may also consider further functionalization of our NP-Pep with poly(ethylene glycol) (PEG), which is a known antifouling agent to further improve resistance to nonspecific binding. To improve the sensitivity of detection, we can optimize the sample preparation steps using miniature devices and automate the NP-Pep enrichment workflow to reduce sample loss during the handling and transfer steps. Moreover, we can improve our front-end separation strategies with optimized nano-liquid chromatography methods and implementing nanoESI strategies to significantly improve MS sensitivity.

Additionally, if we are to clinically translate our nanoproteomics assay, we will likely need to demonstrate enrichment of the native cTn complex from clinical blood samples. It is known that once released into the blood, cTnI is commonly found in its cTnI-cTnT-cTnC (I-T-C) or cTnI-

cTnC (I-C) complex. Importantly, ELISAs are incapable of distinguishing among modified circulating cTnI proteoforms or cTn complexes. Thus, it would be advantageous to demonstrate enrichment and native MS analysis of intact cTn complexes from human serum using our nanoproteomics platform to apply this technology toward clinical applications, since we have already demonstrated that our technology can enrich intact cardiac troponin protein complexes from human heart tissues. An ideal starting point would be the use of NP-Pep to enrich cTnI from blood samples obtained from patients who arrive to the emergency room with ST elevation observed on electrocardiogram (STEMI patients), which typically represent the most severe cases of heart attacks, as these patients can be successfully diagnosed just by the abnormalities in their electrocardiogram alone. We can use our assay to analyze the cTnI from both healthy and STEMI subjects and discover potential proteoform specific biomarkers for AMI. We can analyze the data obtained from our subject samples for statistically significant trends in cTnI proteoform variations to find correlations in patient-specific variables such as age, gender, co-morbidities, etc. These results will guide future studies into top-down MS analysis of cTnI proteoforms from clinical blood samples in a large human cohort and how proteoforms change in different cardiac pathologies with elevated cTnI, such as hypertension, myocarditis, cardiomyopathy, and renal failure.

Chapter 3 utilizes the NP platform developed in chapter 2 to purify and enable the native MS analysis of intact, non-recombinant cardiac troponin complexes (cTnT-cTnI-cTnC) from human heart tissues, with native MS coupled online with size exclusion chromatography (SEC), thereby enabling the native MS analysis of cardiac troponin complexes for the first time. This robust and reproducible ‘native nanoproteomics’ technology can be used as a structural biology tool for proteomics applications to interrogate protein stoichiometry, non-covalent cofactor

binding interactions, and post-translational modifications and can be used for the analysis of cardiac troponin complexes from human heart samples representative of various cardiomyopathies and disease states.

From a structural biology viewpoint, it would be interesting to understand how calcium-binding differences (two or three Ca^{2+} bound) in cTnC present in the ternary cTn complexes potentially lead to substantially different collisional cross-sections/structural features in the complex, by using ion-mobility mass spectrometry. As calcium sensitivity and perhaps calcium occupancy is often altered in the cardiac troponin complexes representative of various disease state tissues, it would be fascinating if the data acquired from the combination of PTMs, protein complex stoichiometry, cofactor binding, and collisional cross section can be used for distinguishing among disease states that share, for example, certain combinations like PTMs and protein-complex stoichiometry but differ in other combinations like cofactor binding. Additionally, with the development of this native nanoproteomics tool combined with ion-mobility mass spectrometry, new biochemistry insights may be gained by applying kinases or phosphatases to phosphorylate or to dephosphorylate cardiac troponin complexes, to study how phosphorylation modifies the cardiac troponin complex structure. Lastly, a powerful feature of native mass spectrometry is that it can simultaneously detect and allow for relative quantification of protein and protein complex heterogeneity, such as differences in non-covalent cofactor binding. Because of the limitations in other detection modes such as fluorescence, which potentially may average protein-complex heterogeneity altogether, the information gained from studying the cardiac troponin complex by native MS can lead to interesting new research directions and be leveraged to give insight into longstanding structural biology questions.

Appendix A

Analysis of Cardiac Troponin Proteoforms by Top-Down Mass Spectrometry

Adapted from: Timothy N. Tiambeng[1], Trisha Tucholski[1], Zhijie Wu[1], Yanlong Zhu[2,3], Stanford D. Mitchell[2,4], David S. Roberts[1], Yutong Jin[1], Ying Ge[1,2,3,4] , *Methods Enzymol*. **2019**, 626, 347.

Abstract

The cardiac troponin complex, composed of three regulatory proteins (cTnI, cTnT, TnC), functions as the critical regulator of cardiac muscle contraction and relaxation. Myofilament protein-protein interactions are regulated by post-translational modifications (PTMs) to the protein constituents of this complex. Dysregulation of troponin PTMs, particularly phosphorylation, results in altered cardiac contractility. Altered PTMs and isoforms have been increasingly recognized as the molecular mechanisms underlying heart diseases. Therefore, it is essential to comprehensively analyze cardiac troponin proteoforms that arise from PTMs, alternative splicing, and sequence variations. In this chapter, we describe two detailed protocols for the enrichment and purification of endogenous cardiac troponin proteoforms from cardiac tissue. Subsequently, mass spectrometry (MS)-based top-down proteomics utilizing online liquid chromatography (LC)/quadrupole time-of-flight (Q-TOF) MS for separation, profiling, and quantification of the troponins was demonstrated. Characterization of troponin amino acid sequence and the localization of PTMs was shown using Fourier-transform ion cyclotron resonance (FT-ICR) MS with electron capture dissociation (ECD) and collisionally activated dissociation (CAD). Furthermore, we described the use of MASH software, a comprehensive and free software package developed in our lab, for top-down proteomics data analysis. The methods we described can be applied for the analysis of troponin proteoforms in cardiac tissues, from animal models to human clinical samples, for heart disease.

Keywords

Cardiac troponin proteoforms, phosphorylation, top-down proteomics, protein extraction

Contents

1. Introduction

2. Enrichment and purification of endogenous cardiac troponin proteins from cardiac tissue

2.1 Myofilament protein extraction

2.1.1 Equipment

2.1.2 Stock buffers and reagents

2.1.3 Procedure for tissue homogenization and myofilament protein extraction

2.2 Immunoaffinity purification of the cardiac troponin complex

2.2.1 Equipment

2.2.2 Stock buffers and reagents

2.2.3 Procedure for tissue homogenization and myofilament protein extraction

2.2.4 Procedure for immunoaffinity purification of the cTn complex

3. Top-down mass spectrometry analysis of cardiac troponin proteoforms

3.1 Separation, profiling, and quantification of troponin proteoforms using online liquid chromatography (LC)/quadrupole time-of-flight (Q-TOF) MS

3.1.1 Equipment

3.1.2 Buffers and reagents

3.1.3 Procedure for data acquisition

3.2 Characterization of troponin sequence and PTMs using offline Fourier-transform ion cyclotron resonance (FT-ICR) MS

3.2.1 Equipment

3.2.2 Buffers and reagents

3.2.3 Procedure for sample preparation and data acquisition

4. Top-down mass spectrometry data analysis

4.1 Relative proteoform quantification

4.1.1 Procedure for relative proteoform quantification

4.2 Proteoform identification and PTM characterization

4.2.1 Procedure for MS/MS data analysis

5. Conclusions

Abbreviations

CAD, collisionally activated dissociation; cTn, cardiac troponin; cTnI, cardiac troponin I; cTnT, cardiac troponin T; DTT, dithiothreitol; ECD, electron capture dissociation; EDTA, Ethylenediaminetetraacetic acid; EGTA, (ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid); FT-ICR, Fourier-transform ion cyclotron resonance; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); LC, liquid chromatography; MOPS, (3-(N-morpholino)propanesulfonic acid); MS, mass spectrometry; MS/MS, tandem mass spectrometry; MWCO, molecular weight cut off; PMSF, phenylmethylsulfonyl fluoride; PTMs, post-translational modifications; Q-TOF, quadrupole time-of-flight; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride; TFA, trifluoroacetic acid; TnC, troponin C

1.Introduction

Post-translational modifications (PTMs) of the cardiac troponin (cTn) complex subunits play an important role in regulating cardiac contractility (Layland, Solaro, & Shah, 2005; Streng, de Boer, van der Velden, van Dieijen-Visser, & Wodzig, 2013). Cardiac troponin I (cTnI) and cardiac troponin T (cTnT), which are the cardiac-specific isoforms present in the cTn complex, are clinically recognized as sensitive and specific "gold-standard" biomarkers for cardiac injury (Apple, Sandoval, Jaffe, & Ordonez-Llanos, 2017; Thygesen et al., 2018). Modifications of the amino acid residues in both cTnI and cTnT have substantial consequences for modulating cardiac functions and are increasingly recognized as 'proteoform biomarkers' underlying heart diseases (Peng, Ayaz-Guner, Yu, & Ge, 2014; Soetkamp et al., 2017; Zhang et al., 2011). For instance, phosphorylation of cTnI plays an important role in the regulation of contractility and affects cardiac regulatory networks (Kooij et al., 2010; Solaro, Henze, & Kobayashi, 2013). C-terminal cleavage and proteolysis of cTnI is present in irreversibly-injured myocardium and may play a role in contractile dysfunction (Marston & Redwood, 2003). Furthermore, Van Eyk and coworkers have reported extensive cTnI and cTnT modifications in the serum of patients with acute myocardial infarction (Labugger, Organ, Collier, Atar, & Van Eyk, 2000). While the ramifications of cTnT phosphorylation in disease are still under intense debate (Katrukha & Gusev, 2013), amino acid mutations of cTnT have been associated with familial hypertrophic cardiomyopathy and dilated cardiomyopathy (Chang & Potter, 2005; Tardiff, 2005) . Therefore, robust methods that enable comprehensive analysis of the intact cardiac troponins and their proteoforms are important for understanding molecular mechanisms in cardiac diseases.

Mass spectrometry (MS)-based top-down proteomics is arguably the most powerful method for characterizing proteoforms—forms of a single gene product which arise from

sequence variations due to mutations/polymorphisms, alternative splicing of the mRNA transcript, and PTMs (Aebersold et al., 2018; Smith et al., 2013). Compared to the bottom-up proteomics approach, which analyzes digested peptides, top-down MS provides a "bird's eye" view of all existing proteoforms by directly analyzing intact proteins after fractionation and separation of the proteome (Chen, Brown, Lin, & Ge, 2018; Toby, Fornelli, & Kelleher, 2016). Acquisition of intact mass spectra prior to tandem MS (MS/MS) allows proteoforms to be quantified relative to each other, providing information such as total protein phosphorylation (Dong et al., 2012; Gao et al., 2018; Peng, Gregorich, et al., 2014). Proteoforms of interest can then be isolated in the gas phase and subsequently subjected to MS/MS. On a liquid chromatography (LC) timescale, energy-based dissociation, such as collisionally activated dissociation (CAD), is typically implemented to obtain preliminary information about proteoform primary structure and proteoform identifications. Electron-based options, such as electron capture dissociation (ECD) and electron transfer dissociation (ETD), are more commonly used in the offline mode and provide in-depth sequence coverage. Additionally, ECD and ETD preserve labile PTMs, such as phosphorylation, allowing site localization (Ayaz-Guner, Zhang, Li, Walker, & Ge, 2009; Chen et al., 2017; Ge, Rybakova, Xu, & Moss, 2009; Zhang, Dong, Hacker, & Ge, 2010).

In this chapter, we described two detailed protocols for the enrichment and purification of cardiac troponin proteoforms, myofilament-enriched extraction, and immunoaffinity purification from cardiac tissue samples (Ayaz-Guner et al., 2009; Dong et al., 2012; Guy et al., 2013; Peng, Gregorich, et al., 2014; Zhang et al., 2011). Subsequently, top-down online LC/MS was utilized for troponin proteoform separation, profiling, and quantification (Gregorich et al., 2015; Peng, Gregorich, et al., 2014). Sequence and PTM site characterization were demonstrated

with offline Fourier-transform ion cyclotron resonance (FT-ICR) MS (Gregorich et al., 2015; Lin et al., 2018; Xu et al., 2011; Zabrouskov, Ge, Schwartz, & Walker, 2008; Zhang et al., 2010). Finally, we demonstrated the use of MASH software, a comprehensive and free software package developed in our lab, for top-down proteomics data analysis (Figure 1) (Cai, Guner, et al., 2016; Guner et al., 2014). These methods are generally applicable for the analysis of cardiac troponin proteoforms in tissues from humans and other species (Lin et al., 2018; Xu et al., 2011; Zhang et al., 2011).

2. Enrichment and purification of endogenous cardiac troponin proteins from cardiac tissue

2.1 Myofilament protein extraction

Myofilament proteins, which constitute cardiac sarcomeres along with Z-disc proteins in cardiac myocytes, can be easily enriched from small amounts of cardiac tissue (5-10 mg) for top-down analysis using a two-step extraction procedure. Our lab implements an extraction procedure originally developed in the Van Eyk lab (Neverova & Van Eyk, 2002), but modified to favor downstream top-down proteomics (Cai et al., 2018; Peng, Gregorich, et al., 2014) . First, cardiac tissue is homogenized in a neutral-pH buffer to deplete cytoplasmic proteins, leaving behind a sarcomere protein-enriched pellet, including all cardiac troponin proteoforms. This sarcomere-enriched pellet is re-homogenized in a low-pH buffer to solubilize and extract the myofilament proteins and Z-disc proteins, resulting in a membrane protein-enriched pellet. The fraction containing sarcomeric proteins can be used directly for online LC/MS for separation, profiling, and relative quantification of cardiac troponin proteoforms.

2.1.1 Equipment

- Teflon pestle (1.5 mL tube rounded tip; Scienceware, Pequannock, NJ, USA)
- Polytron electric homogenizer (Model PRO200; PRO Scientific, Oxford, CT, USA)
- Side-angle centrifuge (Centrifuge 5415R, Eppendorf, Hamburg, Germany)

2.1.2 Stock buffers and reagents

All reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA), unless otherwise specified. All solutions were prepared in Milli-Q water (Millipore, Billerica, MA, USA).

HEPES Extraction Buffer (pH 7.4)		
Reagent (stock concentration)	**Final concentration**	**Purpose**
HEPES (500 mM)	25 mM	Maintain buffering capacity
NaF (500 mM)	50 mM	Inhibition of Ser/Thr phosphatases
EDTA (100 mM)	2.5 mM	Inhibition of metalloproteases; chelates metal ions
Na_3VO_4 (200 mM)	1 mM	Inhibition of Tyr phosphatases
Add remaining reagents immediately prior to performing extraction protocol		
DTT (1 M)	5 mM	Reducing agent
Halt ™ protease inhibitor cocktail (100x)	1x	Inhibition of Cys/Ser/Asp proteases
PMSF (25 mM in ethanol)	1 mM	Inhibition of Ser proteases

TFA Extraction Buffer		
Reagent (stock concentration)	**Final concentration**	**Purpose**

Trifluoroacetic acid (TFA)	1% (v/v)	Protein extractant
Add remaining reagents immediately prior to performing extraction protocol		
TCEP (100 mM)	2 mM	Reducing agent

- Halt™ protease inhibitor cocktail (Thermo Fisher Scientific, Rockford, IL, USA)
- Pierce™ 660 nm Protein Assay Reagent (Thermo Fisher Scientific, Rockford, IL, USA)

2.1.3 Procedure for tissue homogenization and myofilament protein extraction

Small-scale extraction (5-50 mg tissue)

1. Prepare the HEPES extraction buffer from **section 2.1.2** with final concentrations of 25 mM HEPES (pH 7.4), 50 mM NaF, 2.5 mM EDTA, and 1 mM Na_3VO_4. Reducing agents and protease inhibitors are added immediately prior to the protocol to maximize their efficacy, with final concentrations of 5 mM DTT, 1 mM PMSF, and 1x Halt™ protease inhibitor cocktail. We typically prepare a total volume of HEPES extraction buffer proportional to the amount of tissue homogenized (20 volumes; 20 µL HEPES extraction buffer/ 1 mg tissue). **Note:** Buffers should be chilled on ice prior to use. The entire extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming (Cai et al., 2018).
2. Prepare the TFA extraction buffer from **section 2.1.2** with final concentrations of 1% TFA (v/v) and 2 mM TCEP. We typically prepare a total volume of TFA extraction buffer proportional to the amount of tissue homogenized (10 volumes; 10 µL TFA extraction buffer/ 1 mg tissue). **Note:** Buffers should be chilled on ice prior to use. The

entire extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming.

3. Remove the cardiac tissue from -80 °C and place it in a container filled with liquid nitrogen for transportation to a cold room. Use long forceps to retrieve samples.
4. Cut cardiac tissue into small pieces on a metal plate cooled over dry ice, being careful to cut the striated, or "striped", pieces of muscle tissue and avoiding pieces of fat (yellow) which could interfere with downstream proteomics.
5. Weigh the desired amount of tissue (typically between 5-50 mg) into a 1.5 mL microcentrifuge tube.
6. Add 10 volumes of chilled HEPES extraction buffer to the microcentrifuge tube containing tissue (e.g., 50 µL for 5 mg of tissue) immediately after cutting/weighing tissue. At all times, hold the microcentrifuge tube at the base of the opening and place the sample on ice to avoid warming of the tissue sample.
7. Using the Teflon homogenizer, move the tissue to the side of the tube and carefully apply pressure and shear the tissue against the wall of the tube approximately 6-8 times. Lift the homogenizer to readjust tissue on the tube wall, and then shear for an additional 2-4 times.
8. Visually inspect the homogenized tissue, holding the microcentrifuge tube at the base of the opening to avoid warming of the tissue sample. No visible pieces of tissue should remain.
9. Centrifuge the homogenate at 17,000 g for 15 min at 4 °C.
10. Remove and discard the supernatant, and then add 10 volumes of chilled HEPES extraction buffer. Re-suspend the pellet in buffer using a clean pipette tip and centrifuge

the homogenate again at 17,000 for 15 minutes at 4 °C. Remove and discard the resulting supernatant.

11. Add 10 volumes of chilled TFA extraction buffer to the tissue pellet (e.g., 50 µL for 5 mg of tissue).
12. Using the Teflon homogenizer, move the tissue to the side of the tube and carefully apply pressure and shear the tissue against the wall of the tube approximately 6-8 times. Lift the homogenizer to readjust tissue on the tube wall, and then shear for an additional 2-4 times.
13. Centrifuge the homogenate at 17,000 g for 15 min at 4 °C.
14. Remove and store the supernatant containing myofilament proteins in a separate microcentrifuge tube. Measure total protein concentration using Pierce™ Protein 660 nm Assay. The sample can be stored at -80 °C until analysis. *

    ***Special note:** For best results, perform top-down proteomics on fresh tissue extracts. Freeze/thaw cycles compromise sample stability and often result in sample loss due to protein precipitation.

Large-scale extraction (50 mg - 2 g tissue)

1. Prepare the HEPES extraction buffer from **section 2.1.2** with final concentrations of 25 mM HEPES (pH 7.4), 50 mM NaF, 2.5 mM EDTA and 1 mM Na_3VO_4. Reducing agents and protease inhibitors are added immediately prior to the protocol to maximize their efficacy, with final concentrations of 5 mM DTT, 1x Halt™ protease inhibitor cocktail and 1 mM PMSF. We typically prepare a total volume of HEPES extraction buffer proportional to the amount of tissue homogenized (20 volumes; 20 mL HEPES extraction buffer/ 1 g tissue). **Note:** Buffers should be chilled on ice prior to use. The

entire extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming (Cai et al., 2018).

2. Prepare the TFA extraction buffer from **section 2.1.2** with final concentrations of 1% TFA (v/v) and 2 mM TCEP. We typically prepare a total volume of TFA extraction buffer proportional to the amount of tissue homogenized (10 volumes; 10 mL TFA extraction buffer/ 1 g tissue). **Note:** Buffers should be chilled on ice prior to use. The entire extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming.
3. Remove the cardiac tissue from -80 °C and place it in a container filled with liquid nitrogen for transportation to a cold room. Use long forceps to retrieve samples.
4. Cut cardiac tissue into small pieces on a metal plate cooled over dry ice, being careful to cut the striated, or "striped", pieces of muscle tissue and avoiding pieces of fat (yellow) which could interfere with downstream proteomics.
5. Weigh the desired amount of tissue for extraction (typically between 50 mg - 2 g) into a 50 mL conical centrifuge tube.
6. Add chilled HEPES extraction buffer to the conical tube (e.g., 2.0 mL for 200 mg of tissue) immediately after cutting/weighing tissue. At all times, hold the conical tube at the base of the opening and place the sample on ice to avoid warming of the tissue sample.
7. With the tube placed on ice, use the electric homogenizer to homogenize the tissue in HEPES extraction buffer using swirling and pulsing motions in brief, 1 s intervals (there should be no visible pieces of tissue in the resulting homogenate). **Note:** Brief pulses of the electric homogenizer are preferred to reduce overheating.

8. Transfer the homogenized tissue mixture to 1.5 mL microcentrifuge tubes in equal amounts of volume (e.g., 1.0 mL in each tube).
9. Centrifuge the homogenate at 17,000 g for 30 min at 4 °C.
10. Remove and discard the supernatant, and then add 10 volumes of chilled HEPES extraction buffer. Re-suspend the pellets in the 1.5 mL tubes and centrifuge again at 17,000 g for 15 minutes at 4 °C. Remove and discard the resulting supernatant.
11. Transfer the sarcomere-enriched pellets back into a clean 50 mL conical tube.
12. Add 10 volumes of chilled TFA extraction buffer to the tube (e.g., 2.0 mL for 200 mg of tissue).
13. With the tube placed on ice, use the electric homogenizer to homogenize the pellet in chilled TFA extraction buffer using swirling and pulsing motions in brief, 1 s intervals. **Note:** Brief pulses of the electric homogenizer are preferred to reduce overheating.
14. Transfer the homogenate to 1.5 mL microcentrifuge tubes in equal amounts of volume (e.g., 1.0 mL in each tube).
15. Centrifuge the homogenate at 17,000 g for 30 min at 4 °C.
16. Remove and store the supernatant containing myofilament proteins in separate microcentrifuge tubes. Measure total protein concentration using Pierce™ Protein 660 nm Assay. The sample can be stored at -80 °C until analysis*.

    ***Special note:** For best results, perform top-down proteomics on fresh tissue extracts.

2.2 Immunoaffinity purification of the cardiac troponin complex

Immunoaffinity purification of the cardiac troponin (cTn) complex greatly reduces sample complexity prior to MS analysis, and enables the detection of low abundance cTn proteoforms. First, we perform a modified extraction procedure, originally developed by the Marston lab, which

enriches sarcomere-associated proteins using a neutral-pH, high-salt buffer (Messer, Jacques, & Marston, 2007). The subsequent immunoaffinity purification procedure is performed using monoclonal mouse anti- cTnI antibodies conjugated onto amine reactive Sepharose® beads, such as cyanogen bromide (CNBr) activated Sepharose®. Due to the strong interactions of cTnI with cTnT and TnC, the entire cTn complex is isolated when the proteins are maintained at physiological pH and in non-denaturing conditions for antibody binding. With careful consideration of the anti-cTnI antibody epitope (Guy et al., 2013), we have also successfully purified cTnI from sarcomeric extracts from other species such as swine and rat, due to the high sequence homology of cTnI among different species. After MS-incompatible salts are removed (usually using a MWCO filter), offline high-resolution top-down MS analysis can be performed.

2.2.1 Equipment

- Polytron electric homogenizer (Model PRO200; PRO Scientific, Oxford, CT, USA)
- Side-angle centrifuge (Centrifuge 5415R, Eppendorf, Hamburg, Germany)

2.2.2 Stock Buffers and Reagents

All reagents were purchased from Sigma Chemical Co. (St. Louis, MO, USA) unless otherwise specified. All solutions were prepared in Milli-Q water (Millipore, Billerica, MA, USA).

Wash Buffer (pH 7.0)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
NaH_2PO_4 (500 mM)	5 mM	Maintain buffering capacity
Na_2HPO_4 (100 mM)	5 mM	Maintain buffering capacity
$MgCl_2$ (100 mM)	5 mM	Maintain ionic strength of medium

NaCl (137 mM)	100 mM	Maintain ionic strength of medium
EGTA (100 mM)	0.5 mM	Chelation of metal ions; inhibition of metalloproteases
Add remaining reagents immediately prior to performing wash protocol		
Triton™ X-100 (10% v/v)	1% (v/v)	Solubilization of poorly soluble proteins
DTT (1 M)	5 mM	Reducing agent
Halt™ protease inhibitor cocktail (100x)	1x	Inhibition of Ser/Cys/Asp proteases
PMSF (25 mM in ethanol)	1 mM	Inhibition of Ser proteases
Phosphatase inhibitor cocktail A (100x)	1x	Inhibition of Ser/Thr/Tyr phosphatases

Protein Extraction Buffer (pH 7.5)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
Tris (1 M)	25 mM	Maintain buffering capacity
$CaCl_2$ (100 mM)	0.1 mM	Maintain ionic strength of medium
LiCl (3.5 M)	700 mM	Protein extractant
EGTA (100 mM)	5 mM	Chelation of metal ions; inhibition of metalloproteases
Add remaining reagents immediately prior to performing extraction protocol		
DTT (1 M)	5 mM	Reducing agent
Halt™ protease inhibitor cocktail (100x)	1x	Inhibition of Ser/Cys/Asp proteases
PMSF (25 mM in ethanol)	1 mM	Inhibition of Ser proteases

Phosphatase inhibitor cocktail A (100x)	1x	Inhibition of Ser/Thr/Tyr phosphatases

Elution Buffer (pH 2.0)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
Glycine hydrochloride (1 M)	100 mM	Weakens antibody-antigen non-covalent interactions

Neutralization Buffer (pH 9.0)		
Reagent (Stock concentration)	**Final concentration**	**Purpose**
MOPS (1 M)	1 M	Neutralization of acidic buffer

- Halt™ Protease inhibitor cocktail (Thermo Fisher Scientific, Rockford, IL, USA)
- Phosphatase inhibitor cocktail A (Santa Cruz Biotechnology, Inc. Dallas, TX, USA)
- CNBr-activated Sepharose® CL-4B (Sigma-Aldrich, St. Louis, MO, USA)
- Monoclonal mouse anti-cTnI antibodies (MF4; HyTest, Turku, Finland)
- Pierce™ 660 nm Protein Assay Reagent (Thermo Fisher Scientific, Rockford, IL, USA)

2.2.3 Procedure for tissue homogenization and myofilament protein extraction

1. Prepare the wash buffer from **section 2.2.2** with final concentrations of 5 mM NaH_2PO_4, 5 mM Na_2HPO_4 (pH 7.0), 5 mM $MgCl_2$, 100 mM NaCl and 0.5 mM EGTA. Triton™ X-100, reducing agents, protease inhibitors, and phosphatase inhibitors are

added immediately prior to the protocol to maximize their efficacy, with final concentrations of 1% (v/v) Triton™ X-100, 5 mM DTT, 1x Halt™ protease inhibitor cocktail, 1 mM PMSF, and 1x phosphatase inhibitor cocktail A. We typically prepare wash buffer proportional to the amount of tissue homogenized (20 volumes; 20 mL wash buffer/ 1 g tissue). **Note:** Buffers should be chilled on ice prior to use. Extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming (Cai et al., 2018).

2. Prepare the protein extraction buffer from **section 2.2.2** with final concentration of 25 mM Tris (pH 7.5), 0.1 mM $CaCl_2$, 700 mM LiCl, and 5 mM EGTA. Reducing agents, protease inhibitors, and phosphatase inhibitors are added immediately prior to the protocol to maximize their efficacy, with final concentrations of 5 mM DTT, 1% (v/v) Halt™ protease inhibitor cocktail, 1 mM PMSF, and 1% (v/v) phosphatase inhibitor cocktail A. We typically prepare protein extraction buffer proportional to the amount of tissue homogenized (10 volumes; 10 mL protein extraction buffer/ 1 g tissue). **Note:** Buffers should be chilled on ice prior to use. Extraction procedure should be performed in a cold room (4 °C) to avoid tissue warming.
3. Remove the cardiac tissue from -80 °C and place it in a container filled with liquid nitrogen for transportation to a cold room. Use long forceps to retrieve samples.
4. Cut cardiac tissue into small pieces on a metal plate cooled over dry ice, being careful to cut the striated, or “striped”, pieces of muscle tissue and avoiding pieces of fat (yellow) which could interfere with downstream proteomics.
5. Weigh the desired amount of tissue for extraction (typically between 50 mg - 2 g) into a 50 mL conical centrifuge tube.

6. Add chilled wash buffer to the conical tube (e.g., 2 mL for 200 mg of tissue) immediately after cutting/weighing tissue. At all times, hold the conical tube at the base of the opening and place the sample on ice to avoid warming of the tissue sample.
7. With the tube placed on ice, use the electric homogenizer to homogenize the tissue in wash buffer using swirling and pulsing motions in brief, 1 s intervals (there should be no visible pieces of tissue in the resulting homogenate). **Note:** Brief pulses of the electric homogenizer are preferred to reduce overheating.
8. Transfer the homogenized tissue mixture to 1.5 mL microcentrifuge tubes in equal amounts of volume (e.g., 1 mL in each tube).
9. Centrifuge the homogenate at 17,000 g for 15 min at 4 °C.
10. Remove and discard the supernatant, and then add 10 volumes of chilled wash buffer. Re-suspend the pellets placed in the 1.5 mL tubes and centrifuge again at 17,000 g for 15 min at 4 °C. Remove and discard the resulting supernatant.
11. Transfer the myofilament-enriched pellets back into a clean 50 mL conical tube.
12. Add 10 volumes of chilled protein extraction buffer to the tube (e.g., 2 mL for 200 mg of tissue).
13. With the tube placed on ice, use the electric homogenizer to homogenize the pellet in chilled protein extraction buffer using swirling and pulsing motions in brief, 1 s intervals. **Note:** Brief pulses of the electric homogenizer are preferred to reduce overheating.
14. Transfer the homogenate to 1.5 mL microcentrifuge tubes in equal amounts of volume (e.g., 1 mL in each tube).

15. Allow the suspension to agitate on a rotating shaker at 4 °C for 40 min, and then centrifuge the suspension at 17,000 g for 60 min at 4 °C to remove all tissue debris prior to immunoaffinity purification of the troponin complex.
16. Remove and store the supernatant containing myofilament proteins in clean microcentrifuge tubes. Measure the total protein concentration using Pierce™ Protein 660 nm Assay. The sample can be stored at -80 °C until immunoaffinity purification.

2.2.4 Procedure for immunoaffinity purification of the cTn complex

1. For immunoaffinity purification, transfer 300 µL of CNBr-activated Sepharose® CL-4B conjugated to monoclonal anti-cTnI antibodies into an affinity column (Dong et al., 2012; Guy et al., 2013; Zhang et al., 2011). While we purchased these antibody-conjugated beads as a customized order (HyTest), we recommend the following reference for CNBr-activated Sepharose® conjugation (Kavran & Leahy, 2014).
2. Prepare the column by rinsing the beads with 1 mL of protein extraction buffer.
3. Transfer the supernatant prepared from **section 2.2.3** (750 µL) to the affinity column loaded with antibody-beads and allow the mixture to incubate for 40 min on a rotating shaker at 4 °C.
4. After incubating and allowing the unbound proteins to flow through the column, wash the column twice with 800 µL of protein extraction buffer to elute the unbound proteins.
5. Elute the troponin complex from the affinity column using five, 400 µL volumes of 100 mM glycine hydrochloride buffer (pH 2.0).
6. Collect the elution in five equal fractions, and immediately neutralize the glycine hydrochloride buffer with 40 µL of 1 M MOPS (pH 9.0). The fractions may be stored at -80 °C if desired. Note: We see most of the troponin complex elute in fractions 2 and

3 (Figure 2). Measure the protein concentration of each fraction using a Pierce™ 660 nm Protein Assay.

7. Wash the column by filling it with 2 mL of protein extraction buffer, and letting it flow through. Leave approximately 1 mL on the column, ensuring that the wash is neutral-pH to avoid denaturation of the antibody-bead conjugates.
8. Place a cap and stopper on the column, and store it at 4 °C.

3. Top-down mass spectrometry analysis of cardiac troponin proteoforms

3.1 Separation, profiling, and quantification of troponin proteoforms using online liquid chromatography (LC)/quadrupole time-of-flight (Q-TOF) MS

Online LC/MS profiling using a Q-TOF mass spectrometer is powerful due to its fast scan rate and ability to generate a quality mass spectrum in 1 s or less, which is valuable in applications where liquid chromatography strategies are coupled to a mass spectrometer (Beck et al., 2015; Cai, Tucholski, Gregorich, & Ge, 2016; Liang et al., 2019). Liquid chromatography-based strategies and other front-end separation strategies have burgeoned as effective methods to reduce the complexity of the proteome by separating proteins based on their physicochemical properties, enabling MS-detection and analysis of a diverse array of proteins and their proteoforms. Using online reverse-phase LC/MS, proteins can be separated by their hydrophobicity, observed intact, and their proteoforms quantified by their relative abundances. This workflow is effective for quantifying changes in the total amount of phosphorylation of a single protein (Dong et al., 2012; Gao et al., 2018; Peng, Ayaz-Guner, et al., 2014). Furthermore, energy-based dissociation such as collisionally activated dissociation (CAD) can be implemented on an LC/MS timescale to obtain sequence information and enable proteoform identification (Figure 3). This workflow is

especially useful in profiling abundant myofilament proteoforms, including the cardiac troponin complex, using limited tissue samples (5-50 mg).

3.1.1 Equipment

- Quadrupole time-of-flight mass spectrometer, such as the Impact II (Bruker Scientific LLC Billerica, MA, USA)
- nanoAcquity UPLC System (Waters, Milford, MA, USA)
- Reverse-phase chromatography column (we recommend using a PLRP-S column (200 mm bed length, 500 µm or 250 µm i.d., 10 µm particle size, 1000 Å pore size, Agilent Technologies, Santa Clara, CA, USA)

3.1.2 Buffers and reagents

All reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise specified.

- LC/MS grade 0.1% formic acid in water
- LC/MS grade 0.1% formic acid in 50:50 acetonitrile: ethanol

3.1.3 Procedure for data acquisition

1. Dilute the myofilament protein-enriched extracts (**section 2.1.3**) to a total protein concentration of 100 ng/µL total protein with 0.1% formic acid in water containing 2 mM TCEP.
2. Load 1 µg of total protein (10 uL injection) for a 500 µm i.d. column, or 500 ng of total protein for a 250 µm i.d. column. Separate the myofilament proteins

chromatographically using a nanoAcquity UPLC system equipped with a PLRP-S column.

3. Elute the cTn subunits using a mobile phase gradient going from 5% B to 95% B over 45 min at a flow rate of 4 µL/min (mobile phase A: 0.1% formic acid in water, mobile phase B: 0.1% formic acid in 50:50 acetonitrile: ethanol). cTnT proteoforms elute first at approximately 50% mobile phase B; cTnI proteoforms elute at approximately 60% mobile phase B (Figure 4).

Source Settings

4. For the electrospray ionization source, set the 'End Plate Offset' at 500 V, the 'Capillary' at 4500 V, the 'Nebulizer' at 0.5 bar, the 'Dry Gas' at 4.0 L/min, and the 'Dry Temp' at 220 °C.

Tune Settings

5. For the Impact II Q-TOF mass spectrometer, collect the intact mass spectra over a 500 to 2000 *m/z* range and a scan rate of 0.5 Hz.
6. Set the 'Transfer' parameters with the 'Funnel 1 RF' at 400 Vpp, the 'isCID Energy' at 10 eV, and the 'Multipole RF' at 400 Vpp.
7. Set the 'Quadrupole' parameters with the 'Ion Energy' at 4 eV, and the 'Low Mass' at 500 *m/z*.
8. Set the 'Collision Cell' parameters with the 'Collision Energy' at 4 eV, the 'Collision RF' at 2000 Vpp, the 'Transfer Time' at 110 µs, and the 'Pre Pulse Storage' at 10 µs.

MS/MS Settings

9. For targeted CAD-based MS/MS experiments such as 'MRM', input the precursor ion(s) *m/z* of interest, set the 'Width' to 5 *m/z*, the 'Mode' to CID, the 'isCID' to 10 eV, and the 'Collision' to 20 eV.
10. Right-click the chromatographic spectrum to select the retention window for performing targeting MS/MS experiments on the precursor ion(s) *m/z* of interest.
11. Acquire the data.
12. Identify proteins using MASH software and quantify relative proteoforms and total phosphorylation. (**Section 4**).

3.2 Characterization of troponin sequence and PTMs using offline Fourier-transform ion cyclotron resonance (FT-ICR) MS

Offline, deep sequencing of proteins using a Fourier-transform-based instrument is powerful due to its high mass resolving power and high mass accuracy (Marshall, Hendrickson, & Jackson, 1998). An instrument with high mass resolving power is essential for accurate mass measurement, while high mass accuracy of detected fragments from MS/MS spectra enables higher confidence of proteoform identifications (Scigelova, Hornshaw, Giannakopulos, & Makarov, 2011). Complimentary to energy-based dissociation, such as CAD, electron-based options such as electron capture dissociation (ECD) and electron transfer dissociation (ETD) can provide in-depth sequence coverage while preserving labile PTMs, allowing site localization, and are often implemented in the offline mode because an increased number of MS/MS spectra can be acquired, improving fragment ion S/N. Ultimately, offline deep sequencing is useful for comprehensive characterization of alternative splicing isoforms and PTMs present in various cTnI and cTnT proteoforms. However, due to the relatively slower scan speeds required for high resolution on an FT-ICR instrument, this method is more effective when paired with the

immunoaffinity purified cardiac troponin complex rather than the complex myofilament-protein extracts which require LC separation.

3.2.1 Equipment

- Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometer, such as the 12T solariX XR™ (Bruker Scientific LLC, Billerica, MA, USA)
- Nano-electrospray ionization source (e.g. Triversa NanoMate®, Ithaca, NY, USA)

3.2.2 Buffers and reagents

All reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise specified.

- LC/MS grade 0.1% formic acid in water
- LC/MS grade 0.1% formic acid in *MS-grade acetonitrile. Note: *HPLC-grade acetonitrile can contain trace impurities which affect ionization efficiency during direct infusion.

3.2.3 Procedure for sample preparation and data acquisition

1. Remove MS-incompatible salts from the immunoaffinity-purified cTn samples using Pierce™ protein concentrators PES (10K MWCO, 0.5 mL, Rockford, IL, USA) and perform a buffer exchange using 0.1% formic acid in water. Typically five, 400 µL volumes of 0.1% formic acid in water are added for optimal buffer exchange and removal of MS-incompatible salts.
2. After desalting, mix the sample with an equal volume of 0.1% FA in acetonitrile for direct infusion on a FT-ICR mass spectrometer. An ideal protein concentration for analysis is ~ 0.1 µg/µl.

3. For the nano-electrospray ionization source (Triversa Nanomate®), set the voltage to 1.2-1.6 kV versus the inlet of the mass spectrometer, and the de-solvating gas pressure at 0.5 PSI.
4. For intact mass acquisition on the SolariX XR™ FTICR-MS, collect the mass spectra over a 200 to 2500 *m/z* range with 2 mega-words (resolution 270,000 at 400 *m/z*; transient length 1.1534 s), with an ion accumulation time of approximately 0.100 s*. For acquisition of intact mass spectra, typically 50-100 scans should be accumulated to ensure the collection of high-resolution mass spectra for data analysis.

   ***Special note**: Because the Bruker SolariX XR™ instrument is not equipped with automatic gain control, the ion accumulation time will need to be optimized and varied based on the protein concentration, to avoid space-charge effects that compromise mass accuracy.

API Source Settings

5. Change the 'API Source' to Nano ESI offline. Set the 'Source Gas Tune' parameters with the 'Dry Gas' at 4.0 L/min, and the 'Dry Temp' at 180 °C.

Ion Transfer Settings

6. Set the 'Source Optics' parameters with the 'Capillary Exit' at 240 V, 'Deflector Plate' at 220 V, 'Funnel 1' at 150 V, 'Skimmer 1' at 20V, and the 'Funnel RF Amplitude' at 150 Vpp. **Note:** The 'Skimmer 1' voltage can be varied to modulate analyte fragmentation and the presence of salt adducts prior to its introduction into the collision cell, but the sum of the 'Skimmer 1' and 'In Source CID' voltage should not exceed the 'Funnel 1' voltage (e.g., 150 V).

7. Set the ‘Octopole’ parameters with the ‘Frequency’ at 5 MHz, and the ‘RF Amplitude’ at 350.0 Vpp.
8. Set the ‘Collision Cell’ parameters with the ‘Collision Voltage’ at -5 V, and the ‘DC Extract Bias’ at 0.5 V. The ‘RF Frequency’ should be set at 2 MHz, and the ‘Collision RF Amplitude’ at 2000 Vpp.
9. Set the ‘Transfer Optics’ parameters with the ‘Time of Flight’ at 1.200 ms, the ‘Frequency’ at 4 MHz, and the ‘RF Amplitude’ at 350 Vpp.
10. Set the ‘Gas Control’ parameters with the ‘Flow’ at 20%, with ‘Enabled’ selected.

Excitation Control Settings

11. Set the ‘Standard Excitation’ parameters with the ‘Sweep Excitation Power’ at 20%, the ‘Sweep Step Time’ at 15 µs, and the ‘Sweep Freq. Density (XBB)’ at 100.0.

Source MS/MS Settings

12. Set the ‘Quadrupole’ parameters with the ‘Q1 Mass’ set at 50 *m/z* less than the lowest observable ion of interest for intact mass acquisition. **Note:** For ion isolation acquisition and MS/MS acquisition, the ‘Q1 Mass’ should be set at the m/z of interest.
13. Set the ‘Collision Cell’ parameters with the ‘Collision Voltage’ at -5 V, and the ‘DC Extract Bias’ at 0.5 V. The ‘RF Frequency’ should be set at 2 MHz, and the ‘Collision RF Amplitude’ at 2000 Vpp.
14. Prior to MS/MS experiments, acquire isolation spectra of the precursor ion of interest. Set the ‘Quadrupole MS/MS’ parameters with ‘Isolate’ selected, ‘CID’ deselected, the ‘Q1 Mass’ set at the *m/z* of interest, and the ‘Isolation Window’ at 3 *m/z*. For isolation scanning acquisition, typically 50-100 scans should be taken to acquire high-quality

spectra of the isolated precursor ion. **Note:** The 'Isolation Window' should be varied to ensure clean isolation of the *m/z* of interest.

15. For CAD-based MS/MS experiments, set the 'Quadrupole MS/MS' parameters with 'Isolate' selected, 'CID' selected, the 'Collision Energy' at 10 V, the 'Q1 Mass' at the m/z of interest, the 'Isolation Window' at 3 *m/z*, the 'Collision RF Amplitude' at 2000 Vpp, and the 'RF Frequency' at 2 MHz. Collect approximately 200 scans, with an ion accumulation time of approximately 1.0 s. **Note:** This accumulation time will need to be optimized and varied based on the protein concentration, to avoid space charge effects that compromise mass accuracy. The 'Collision Energy' voltage can be increased to improve fragmentation of the selected precursor ion, but the precursor ion should still be visible for optimal MS/MS data analysis.

<u>In-Cell MSn Settings (for ECD-based experiments)</u>

16. For ECD-based MS/MS experiments, select 'ECD Heater' and allow the ECD heater to reach its operating current (10-15 min). Set the 'Internal MS/MS' parameters with 'ECD' selected, and set the 'ECD' parameters, starting with the 'ECD Pulse Length' at 20-50 ms, the 'ECD Bias' at 0.8-1.0 V, and the 'ECD Lens' at 10 V (Chen et al., 2017; Lin et al., 2018). Collect approximately 500-1000 scans, with an ion accumulation time of approximately 0.5-1.0 s*. ***Special Note:** The ion accumulation time will need to be optimized and varied based on the protein concentration, to avoid space charge effects that compromise mass accuracy. The 'ECD Bias' and 'ECD pulse length' can be modified to improve fragmentation of the selected precursor ion, but the precursor ion should still be visible

for optimal MS/MS data analysis. Optimal ECD parameters will be different based on the size and properties of the proteoform targeted for MS/MS analysis.

17. Process the offline MS and MS/MS data (Figure 5) using MASH software (**Section 4**).

4. Top-down MS data analysis

4.1 Relative proteoform quantification

Top-down MS is attractive for performing relative quantification of protein modifications because PTMs have less influence on the physicochemical properties of proteins compared with that of peptides (Dong et al., 2012; Pesavento, Mizzen, & Kelleher, 2006; Zabrouskov et al., 2008; Zhang et al., 2011). This phenomenon suggests the ionization and detection efficiency of an intact protein and its respective proteoforms do not vary significantly, and relative quantification of proteoforms can be achieved by top-down MS analysis (Ansong et al., 2013; Wu et al., 2018). For quantification of protein expression levels across multiple samples by top-down proteomics, our group has recently demonstrated a robust LC/MS-based method for simultaneously quantification of both protein expression and PTM changes in the sarcomeric subproteome (Lin et al., 2019). In this section, the following procedure describes a general process for relative proteoform quantification using online LC/MS data, with quantification of a single protein's total amount of phosphorylation as an example (Dong et al., 2012; Gao et al., 2018; Peng, Ayaz-Guner, et al., 2014).

4.1.1 Procedure for relative proteoform quantification

1. Average or sum all of the intact MS scans across the chromatographic region of interest, corresponding to where specific cardiac troponins elute chromatographically.
2. Perform charge-state deconvolution of the averaged intact MS scans to produce a spectrum that correlates intact proteoform masses with their respective ion intensity.

3. Using the deconvoluted intact mass spectrum, record the intensity values of each proteoform of interest, and determine the total protein intensity by summing all proteoform intensities that correspond to the same protein. Minor oxidation products and other non-covalent adducts belonging to a proteoform should be taken into consideration for summing the total protein population intensity, especially if they account for >5% of the total ion intensity (Figure 6).
4. Calculate the relative abundance percentage of each individual proteoform by using equation 1:

$$(1)\ Rel.\ proteoform\ abundance\ (\%) = \left(\frac{individual\ proteoform\ intensity}{total\ protein\ intensity}\right) * 100$$

5. Calculate the total amount of phosphorylation P_{total}, defined as (mol phosphate, P_i / mol protein) by summing the relative abundance percentages of each phosphorylated species, weighted by an integer factor corresponding to their respective phosphorylation state. Divide this sum by the total protein intensity (represented as 100%) by using equation 2 (Peng, Ayaz-Guner, et al., 2014:

$$(2)\ P_{total} = \left(\frac{(1 * \%P_{mono}) + (2 * \%P_{bis}) + (3 * \%P_{tri}) + (n * \%P_{n\ \ldots})}{100\%}\right)$$

4.2 Proteoform identification and PTM characterization

MASH software is a free, user-friendly, and comprehensive data analysis platform developed by our lab for top-down proteomics (Cai, Guner, et al., 2016; Guner et al., 2014). MASH allows for data importing from a variety of vendor data formats (Thermo .raw, Bruker .baf and .ascii, Waters .raw directory and universal file formats such as .mgf and .mzXML), enabling MS, MS/MS, LC/MS, and LC/MS/MS spectra processing. MASH supports a variety of deconvolution algorithms (MS-Deconv, TopFD, eTHRASH, pParseTD, ProMex) which broadly function to

deduce the monoisotopoic mass from a protein's isotopomer envelopes. One of MASH's strongest features is its ability to allow manual user-validation of fragment ions from computational algorithm outputs, enabling corrections for both charge states and monoisotopic masses. MASH software is compatible with most of the common fragment ion types (*a, x, b, y, c, z•* ions). This section will discuss the use of MASH's 'Targeted Mode' for cardiac troponin proteoform identification and PTM mapping from MS/MS datasets.

4.2.1 Procedure for MS/MS data analysis

Targeted Mode

1. Use the 'Import Data' function to import the appropriate MS/MS file, (e.g., Bruker .baf or .ascii file).
2. Use the 'Deconvolution' function to populate MASH's mass list with MS/MS fragment ions that fit the selected deconvolution algorithm's original parameters, such as maximum charge, maximum mass, the *m/z* error tolerance, and S/N ratio. These parameters can be changed in the 'Advanced Deconvolution' settings.
3. Click on individual ion entries in the mass list to visualize the MS/MS spectrum against its theoretical isotopomer envelope distribution. Manually validate the curated ions, and apply a mass shift correction, a charge state correction, or delete false ion assignments, if needed.
4. Use the 'Database Search' function to apply the manually validated MS/MS ion list against the MSAlign+ database search. FASTA files from different genomes can be loaded and applied for the database search. Top-down identification workflows normally will require searching the dataset against a database; If the target sequence is known, it is not necessary to search the deconvoluted mass list against the database.

5. Use the ‘Characterization’ function to import the database search results, or to load in your desired protein sequence in the ‘Sequence Table’. Select the fragment ion type used in the experiment (e.g., *b, y* ions), and click on ‘Show Assignments’ to view the fragmentation coverage for the selected protein sequence (Figure 7).
6. Under ‘Data Reporting’, use the ‘Save Files’ function to save the annotated results, the mass list, and the protein search results list as an XML file.
7. Additional step-by-step manuals and video tutorials for using MASH Explorer for top-down proteomics applications are available at http://ge.crb.wisc.edu/software.html.

5. Conclusions

This chapter describes sample preparation, data acquisition, and data analysis for the intact protein analysis of cardiac troponin proteoforms using top-down MS. Top-down MS has emerged as the most powerful method for unambiguously characterizing proteoforms. With a quadrupole time-of-flight mass spectrometer equipped with front-end LC separation, cardiac troponin proteoforms can be rapidly profiled and relatively quantified to measure changes in phosphorylation of single proteins. With a FT-ICR mass spectrometer, cardiac troponin proteoforms can be comprehensively characterized to understand their sequences and the location of their PTMs. As PTMs and amino acid variants of cardiac troponin proteins alter protein-protein interactions within the sarcomere and modulate cardiac contractility and function, the methods described in this chapter enable reproducible and comprehensive analysis of intact cardiac troponin proteoforms from biological samples.

References:

Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., … Zhang, B. (2018). How many human proteoforms are there? *Nature Chemical Biology*, *14*, 206–214.

Ansong, C., Wu, S., Meng, D., Liu, X., Brewer, H. M., Deatherage Kaiser, B. L., … Pasa-Tolic, L. (2013). Top-down proteomics reveals a unique protein S-thiolation switch in Salmonella Typhimurium in response to infection-like conditions. *Proceedings of the National Academy of Sciences*, *110*, 10153–10158.

Apple, F. S., Sandoval, Y., Jaffe, A. S., & Ordonez-Llanos, J. (2017). Cardiac Troponin Assays: Guide to Understanding Analytical Characteristics and Their Impact on Clinical Care. *Clinical Chemistry*, *63*, 73–81.

Ayaz-Guner, S., Zhang, J., Li, L., Walker, J. W., & Ge, Y. (2009). 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*, *48*, 8161–8170.

Beck, S., Michalski, A., Raether, O., Lubeck, M., Kaspar, S., Goedecke, N., … Mann, M. (2015). The Impact II, a Very High-Resolution Quadrupole Time-of-Flight Instrument (QTOF) for Deep Shotgun Proteomics. *Molecular & Cellular Proteomics*, *14*, 2014–2029.

Cai, W., Guner, H., Gregorich, Z. R., Chen, A. J., Ayaz-Guner, S., Peng, Y., … Ge, Y. (2016). MASH Suite Pro: A Comprehensive Software Tool for Top-Down Proteomics. *Molecular & Cellular Proteomics*, *15*, 703–714.

Cai, W., Hite, Z. L., Lyu, B., Wu, Z., Lin, Z., Gregorich, Z. R., … Ge, Y. (2018). Temperature-sensitive sarcomeric protein post-translational modifications revealed by top-down proteomics. *Journal of Molecular and Cellular Cardiology*, *122*, 11–22.

Cai, W., Tucholski, T. M., Gregorich, Z. R., & Ge, Y. (2016). Top-down Proteomics: Technology Advancements and Applications to Heart Diseases. *Expert Review of Proteomics*, *13*, 717–730.

Chang, A. N., & Potter, J. D. (2005). Sarcomeric protein mutations in dilated cardiomyopathy. *Heart Failure Reviews*, Vol. 10, pp. 225–235.

Chen, B., Brown, K. A., Lin, Z., & Ge, Y. (2018). Top-Down Proteomics: Ready for Prime Time? *Analytical Chemistry*, *90*, 110–127.

Chen, B., Guo, X., Tucholski, T., Lin, Z., McIlwain, S., & Ge, Y. (2017). The Impact of Phosphorylation on Electron Capture Dissociation of Proteins: A Top-down Perspective. *Journal of the American Society for Mass Spectrometry*, *28*, 1805.

Dong, X., Sumandea, C. A., Chen, Y. C., Garcia-Cazarin, M. L., Zhang, J., Balke, C. W., … Ge, Y. (2012). Augmented phosphorylation of cardiac troponin I in hypertensive heart failure. *Journal of Biological Chemistry*, *287*, 848–857.

Gao, L., Gregorich, Z. R., Zhu, W., Mattapally, S., Oduk, Y., Lou, X., … Zhang, J. (2018).

Large cardiac muscle patches engineered from human induced-pluripotent stem cell-derived cardiac cells improve recovery from myocardial infarction in swine. *Circulation*, *137*, 1712–1730.

Ge, Y., Rybakova, I. N., Xu, Q., & Moss, R. L. (2009). 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 of the United States of America*, *106*, 12658–12663.

Gregorich, Z. R., Peng, Y., Lane, N. M., Wolff, J. J., Wang, S., Guo, W., … Ge, Y. (2015). Comprehensive assessment of chamber-specific and transmural heterogeneity in myofilament protein phosphorylation by top-down mass spectrometry. *Journal of Molecular and Cellular Cardiology*, *87*, 102–112.

Guner, H., Close, P. L., Cai, W., Zhang, H., Peng, Y., Gregorich, Z. R., & Ge, Y. (2014). MASH Suite: A User-Friendly and Versatile Software Interface for High-Resolution Mass Spectrometry Data Interpretation and Visualization. *Journal of The American Society for Mass Spectrometry*, *25*, 464–470.

Guy, M. J., Chen, Y.-C. C., Clinton, L., Zhang, H., Zhang, J., Dong, X., … Ge, Y. (2013). The impact of antibody selection on the detection of cardiac troponin I. *Clinica Chimica Acta*, *420*, 82–88.

Katrukha, I. A., & Gusev, N. B. (2013). Enigmas of cardiac troponin T phosphorylation. *Journal of Molecular and Cellular Cardiology*, *65*, 156–158.

Kavran, J. M., & Leahy, D. J. (2014). Coupling Antibody to Cyanogen Bromide-Activated Sepharose. *Methods in Enzymology*, *541*, 27.

Kooij, V., Saes, M., Jaquet, K., Zaremba, R., Foster, D. B., Murphy, A. M., … Stienen, G. J. M. (2010). Effect of troponin I Ser23/24 phosphorylation on Ca2+-sensitivity in human myocardium depends on the phosphorylation background. *Journal of Molecular and Cellular Cardiology*, *48*, 954–963.

Labugger, R., Organ, L., Collier, C., Atar, D., & Van Eyk, J. E. (2000). Extensive troponin I and T modification detected in serum from patients with acute myocardial infarction. *Circulation*, *102*, 1221–1226.

Layland, J., Solaro, R., & Shah, A. (2005). Regulation of cardiac contractile function by troponin I phosphorylation. *Cardiovascular Research*, *66*, 12–21.

Liang, Y., Jin, Y., Wu, Z., Tucholski, T., Brown, K. A., Zhang, L., … Ge, Y. (2019). Bridged Hybrid Monolithic Column Coupled to High-Resolution Mass Spectrometry for Top-Down Proteomics. *Analytical Chemistry*, *91*, 1743–1747.

Lin, Z., Guo, F., Gregorich, Z. R., Sun, R., Zhang, H., Hu, Y., … Ge, Y. (2018). Comprehensive Characterization of Swine Cardiac Troponin T Proteoforms by Top-Down Mass Spectrometry. *Journal of The American Society for Mass Spectrometry*, *29*, 1284–1294.

Lin, Z., Wei, L., Cai, W., Zhu, Y., Tucholski, T., Mitchell, S. D., … Ge, Y. (2019). Simultaneous

Quantification of Protein Expression and Modifications by Top-down Targeted Proteomics: A Case of the Sarcomeric Subproteome. *Molecular & Cellular Proteomics*, *18*, 594–605.

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

Marston, S. B., & Redwood, C. S. (2003). Modulation of Thin Filament Activation by Breakdown or Isoform Switching of Thin Filament Proteins: Physiological and Pathological Implications. *Circulation Research*, Vol. 93, pp. 1170–1178.

Messer, A. E., Jacques, A. M., & Marston, S. B. (2007). Troponin phosphorylation and regulatory function in human heart muscle: dephosphorylation of Ser23/24 on troponin I could account for the contractile defect in end-stage heart failure. *Journal of Molecular and Cellular Cardiology*, *42*, 247–259.

Neverova, I., & Van Eyk, J. E. (2002). Application of reversed phase high performance liquid chromatography for subproteomic analysis of cardiac muscle. *Proteomics*, *2*, 22–31. John Wiley & Sons, Ltd.

Peng, Y., Ayaz-Guner, S., Yu, D., & Ge, Y. (2014). Top-down mass spectrometry of cardiac myofilament proteins in health and disease. *Proteomics - Clinical Applications*, *8*, 554–568.

Peng, Y., Gregorich, Z. R., Valeja, S. G., Zhang, H., Cai, W., Chen, Y.-C., … Ge, Y. (2014). Top-down proteomics reveals concerted reductions in myofilament and Z-disc protein phosphorylation after acute myocardial infarction. *Molecular & Cellular Proteomics : MCP*, *13*, 2752–2764.

Pesavento, J. J., Mizzen, C. A., & Kelleher, N. L. (2006). Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: Human histone H4. *Analytical Chemistry*, *78,* 4271-4280.

Scigelova, M., Hornshaw, M., Giannakopulos, A., & Makarov, A. (2011). Fourier transform mass spectrometry. *Molecular & Cellular Proteomics : MCP*, *10*, M111.009431.

Smith, L. M., Kelleher, N. L., Proteomics, T. C. for T. D., Linial, M., Goodlett, D., Langridge-Smith, P., … Marshall, A. (2013). Proteoform: A single term describing protein complexity. *Nature Methods*, *10*, 186–187.

Soetkamp, D., Raedschelders, K., Mastali, M., Sobhani, K., Merz, C. N. B., & Van Eyk, J. (2017). The continuing evolution of cardiac troponin I biomarker analysis: From protein to proteoform. *Expert Review of Proteomics*, Vol. 14, pp. 973–986.

Solaro, R. J., Henze, M., & Kobayashi, T. (2013). Integration of Troponin I Phosphorylation With Cardiac Regulatory Networks. *Circulation Research*, *112*, 355–366.

Streng, A. S., de Boer, D., van der Velden, J., van Dieijen-Visser, M. P., & Wodzig, W. K. W. H. (2013). Posttranslational modifications of cardiac troponin T: An overview. *Journal of Molecular and Cellular Cardiology*, *63*, 47–56.

Tardiff, J. C. (2005, September). Sarcomeric proteins and Familial Hypertrophic Cardiomyopathy: Linking mutations in structural proteins to complex cardiovascular

phenotypes. *Heart Failure Reviews*, Vol. 10, pp. 237–248.

Thygesen, K., Alpert, J. S., Jaffe, A. S., Chaitman, B. R., Bax, J. J., Morrow, D. A., … Infarction, E. G. on behalf of the J. E. S. of C. (ESC)/American C. of C. (ACC)/American H. A. (AHA)/World H. F. (WHF) T. F. for the U. D. of M. (2018). Fourth Universal Definition of Myocardial Infarction (2018). *Journal of the American College of Cardiology*, *72*, 2231–2264.

Toby, T. K., Fornelli, L., & Kelleher, N. L. (2016). Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annual Review of Analytical Chemistry*, *9*, 499–519.

Wu, Z., Tiambeng, T. N., Cai, W., Chen, B., Lin, Z., Gregorich, Z. R., & Ge, Y. (2018). Impact of Phosphorylation on the Mass Spectrometry Quantification of Intact Phosphoproteins. *Analytical Chemistry*, *90*, 4935–4939.

Xu, F., Xu, Q., Dong, X., Guy, M., Guner, H., Hacker, T. A., & Ge, Y. (2011). 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*, *305*, 95–102.

Zabrouskov, V., Ge, Y., Schwartz, J., & Walker, J. W. (2008). Unraveling Molecular Complexity of Phosphorylated Human Cardiac Troponin I by Top Down Electron Capture Dissociation/Electron Transfer Dissociation Mass Spectrometry. *Molecular & Cellular Proteomics*, *7*, 1838–1849.

Zhang, J., Dong, X., Hacker, T. A., & Ge, Y. (2010). Deciphering modifications in swine cardiac troponin I by top-down high-resolution tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry*, *21*, 940–948.

Zhang, J., Guy, M. J., Norman, H. S., Chen, Y.-C. C., Xu, Q., Dong, X., … Ge, Y. (2011). Top-down quantitative proteomics identified phosphorylation of cardiac troponin i as a candidate biomarker for chronic heart failure. *Journal of Proteome Research*, *10*, 4054–4065.

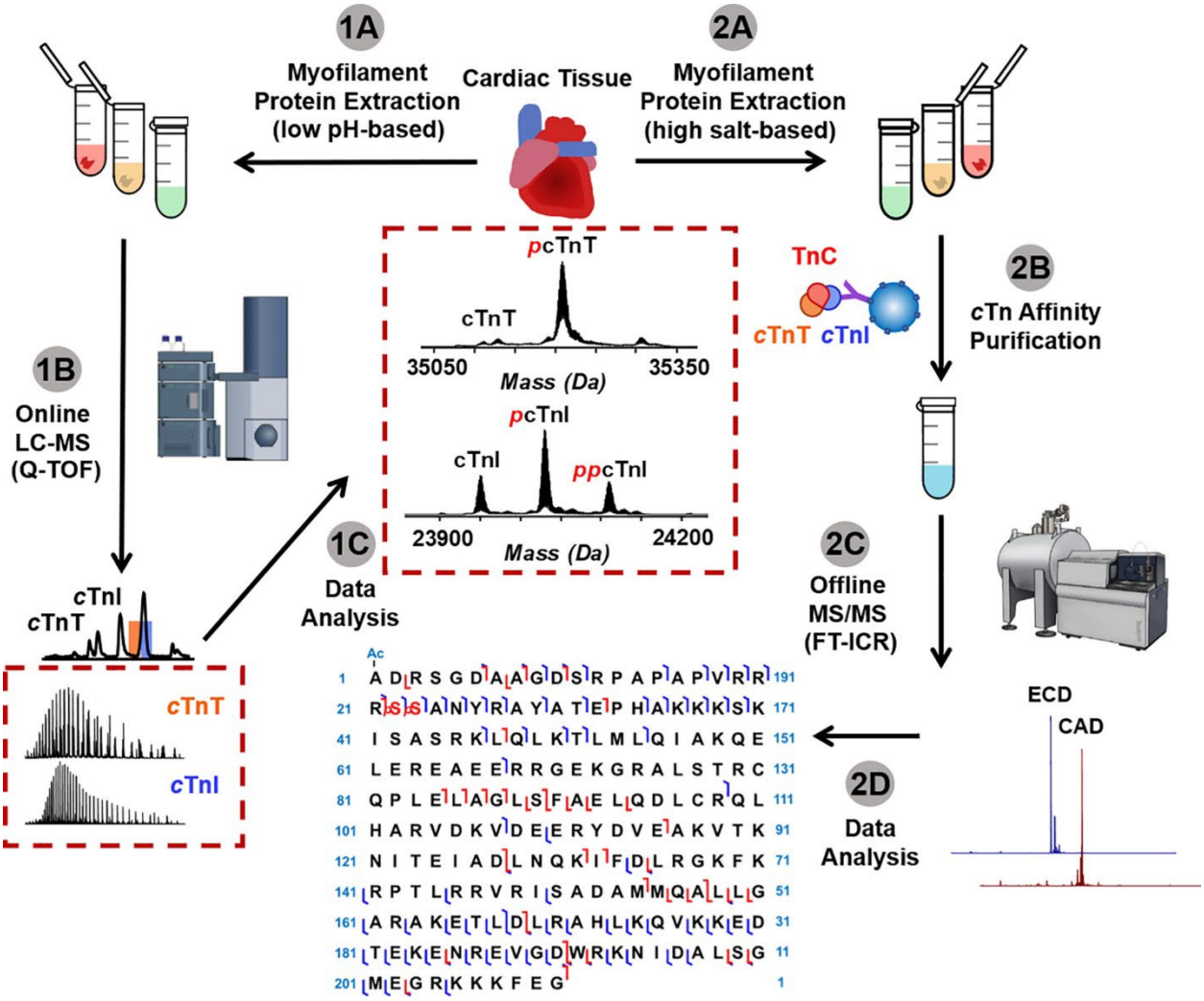


Figure 1: Characterization of intact *c*Tn proteoforms by top-down mass spectrometry.

Schematic illustration of two workflows for intact cardiac troponin analysis: (1a) Myofilament subproteome extracted from cardiac tissue (pH-based); (1b) Online LC/MS proteoform profiling; (1c) Data analysis and relative proteoform quantification. (2a) Myofilament subproteome extracted from cardiac tissue (salt-based); (2b) Immunoaffinity purification of cTn complex; (2c) Offline MS/MS using CAD/ECD; (2d) Data analysis for characterization of cardiac troponin proteoforms.

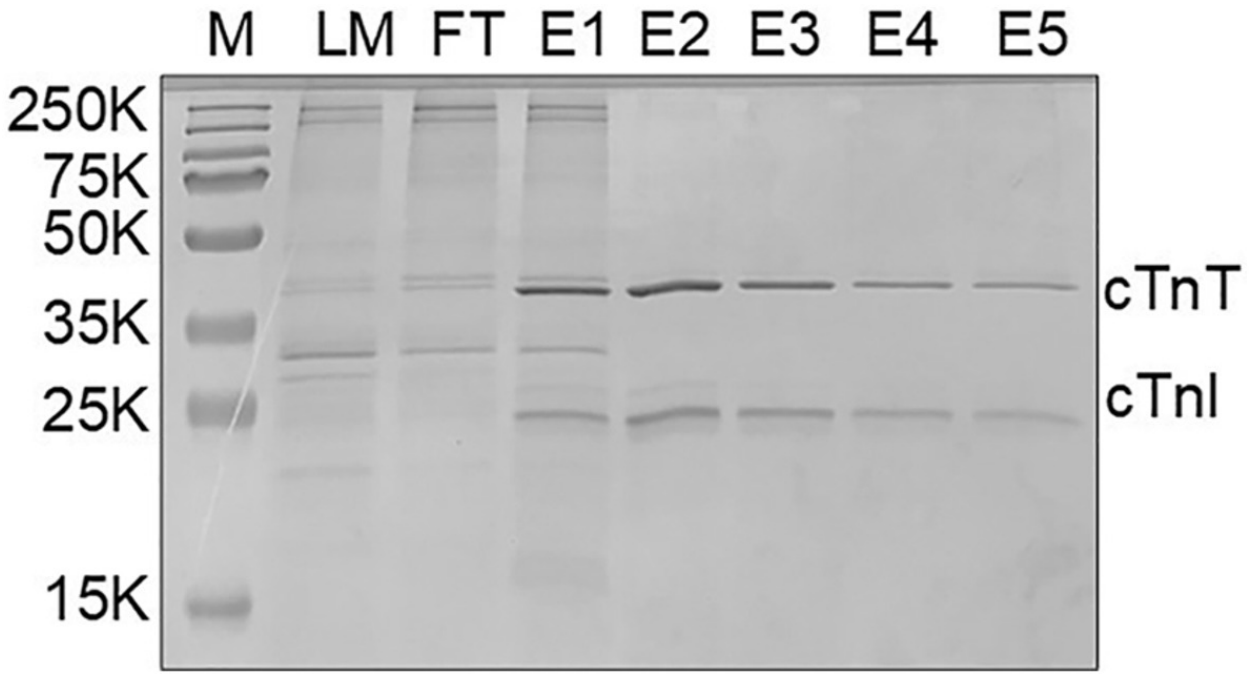


Figure 2: SDS-PAGE visualization of immunoaffinity purified cTn complex.

SDS-PAGE of immunoaffinity purified swine cTn complex from a cardiac tissue sample. M, molecular weight marker; LM, myofilament loading mixture; FT, flow through; E, elution fraction (1-5).

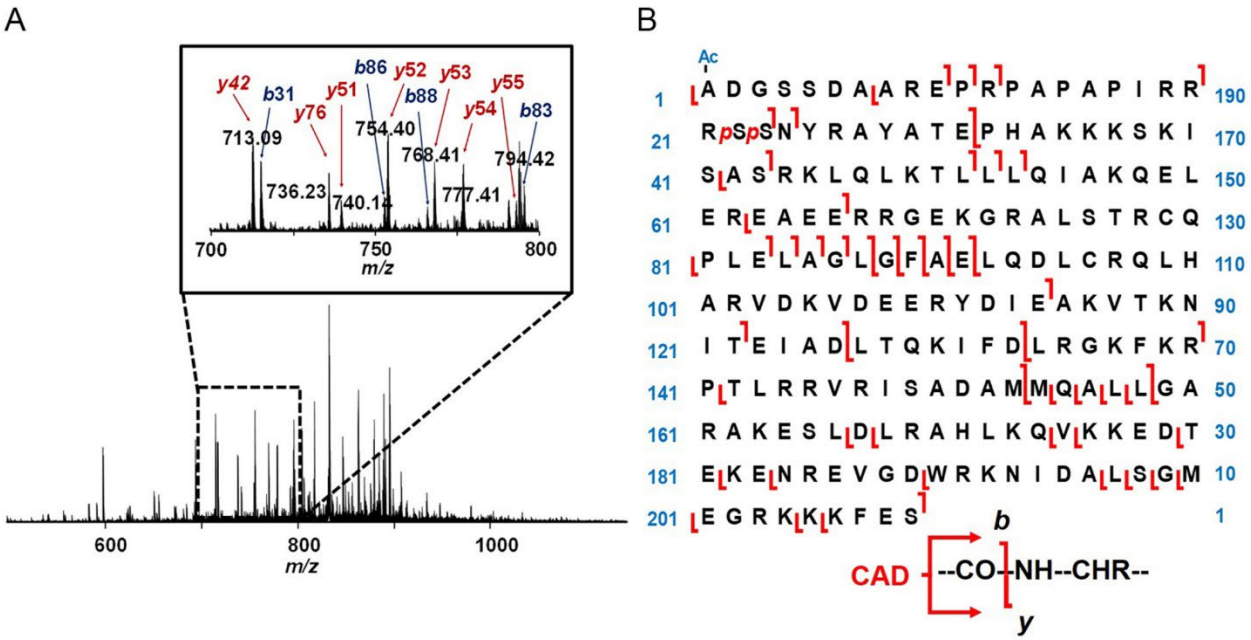


Figure 3: Online LC/MS/MS of cardiac troponin I from human myofilament extracts using pH-based extraction.

A cTnI sequence map using online targeted CAD-based MS/MS for proteoform identification. From gel densitometry experiments (not depicted), approximately 100 ng of cTnI was loaded for online LC/MS/MS analysis using a Q-TOF instrument such as the Impact II.

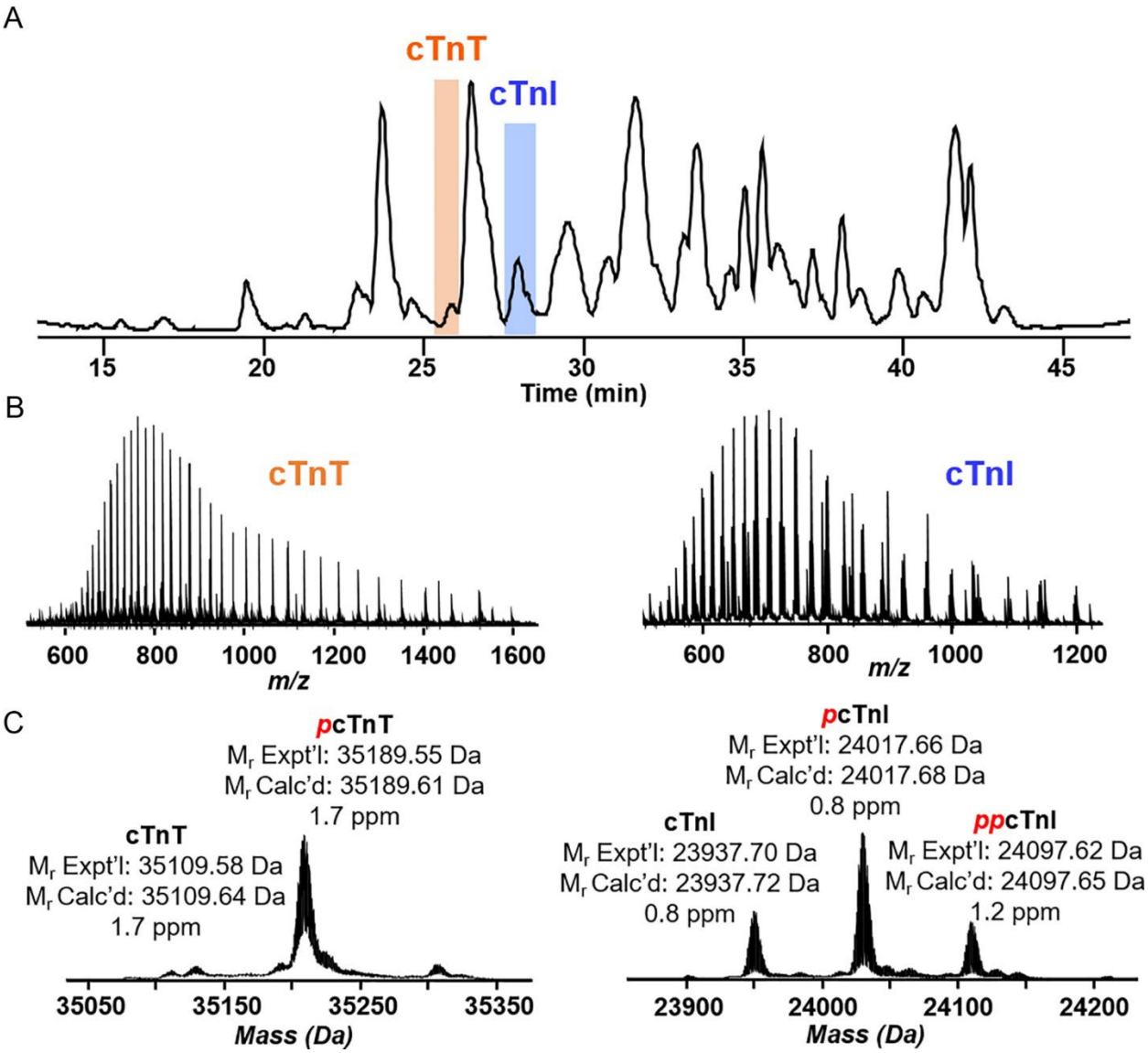


Figure 4: Online LC/MS of cardiac troponins from swine myofilament extracts using pH-based extraction.

(a) BPC, base peak chromatogram of swine myofilament-enriched extracts containing the cardiac troponins, separated by reverse-phase LC. (b) Online LC/MS spectra of cTnT and cTnI using an Impact II Q-TOF. The charge state envelope of cTnT and cTnI are depicted above. (c) Deconvoluted mass spectra of cTnT and cTnI, illustrating online LC/MS-based proteoform profiling. The red "p" represents phosphorylation.

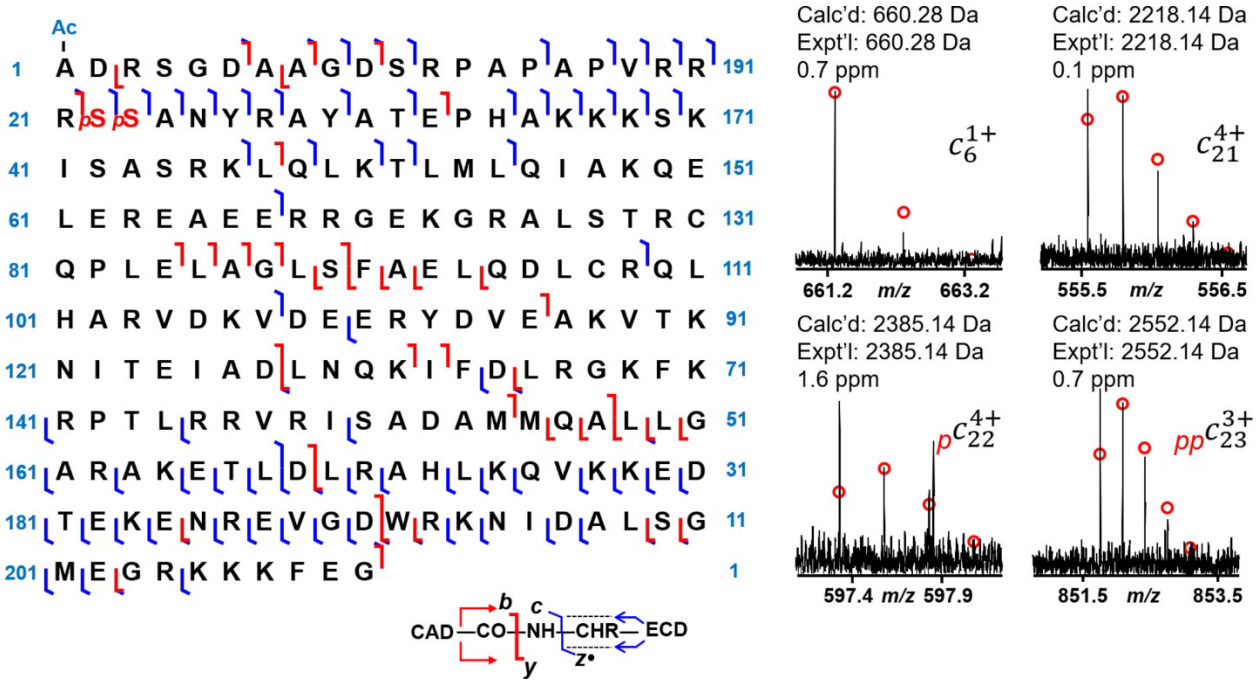


Figure 5: Offline MS/MS analysis of immunoaffinity purified cTnI using CAD and ECD.

ECD and CAD fragment ions and sequence map of a bis-phosphorylated swine cTnI proteoform with Met excision and N-terminal acetylation. A series of *c* ions including c_{21} with no phosphorylation, c_{22} with one phosphorylation, and c_{23} with two phosphorylations, confirmed phosphorylation sites at Ser22 and Ser 23. The presence of acetylation was localized within the first six residues by a c_6 ion. The red "p" represents phosphorylation; Ac, acetylation.

A

Proteoform Name	Proteoform Intensity (ion intensity)	Relative Proteoform Abundance (%)
cTnI	202114	24.7
*p*cTnI	443832	54.1
*pp*cTnI	174164	21.2

B

Proteoform Name	Relative Proteoform Abundance (%)	P_{total}
cTnT	8.9	0.91
*p*cTnT	91.1	

Figure 6: Quantification of relative proteoform abundance (cTnI) and total phosphorylation level (cTnT) using online LC-MS data.

(a) Relative proteoform abundance calculation using the proteoform intensities of cTnI, *p*cTnI, and *pp*cTnI obtained from a deconvoluted mass spectra of cTnI (Figure 4c). (b) Total phosphorylation level calculation using the relative proteoform abundance percentage of *p*cTnT obtained from a deconvoluted mass spectra of cTnT (Figure 4c).

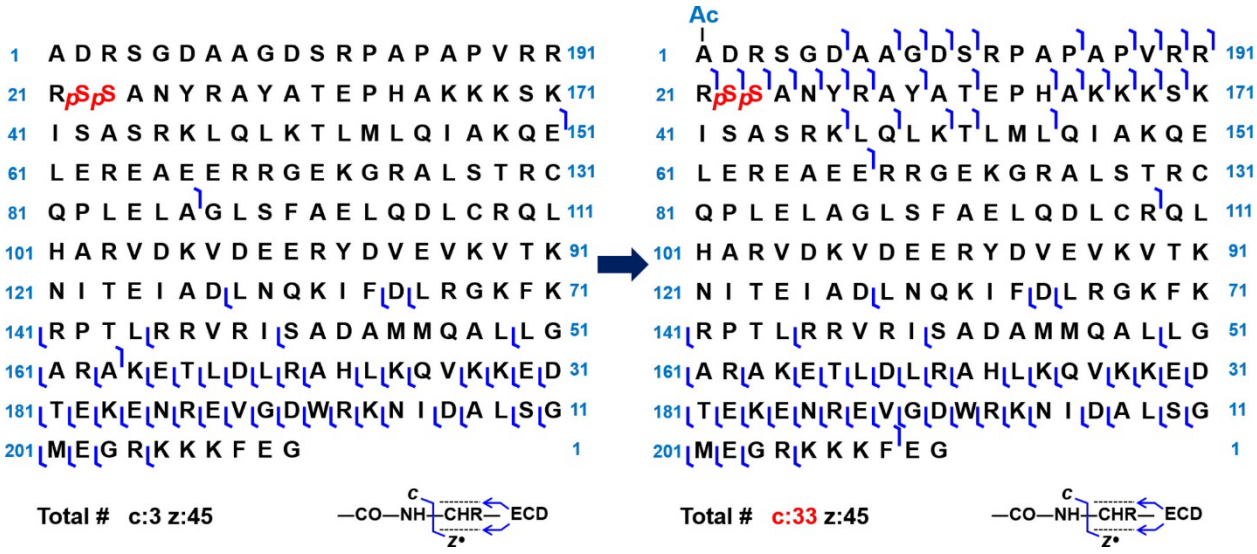


Figure 7: Offline MS/MS PTM characterization of cTnI using MASH software. MASH software facilitates the characterization of PTMs and sequence variations by allowing users to add modifications to individual amino acids in a protein sequence. Without considering the presence of N-terminal acetylation, only 3 *c* ions and 45 *z*• ions could be identified from an offline ECD-based experiment of an immunoaffinity purified swine cTnI proteoform. With the addition of N-terminal acetylation, an additional 30 *c* ions were assigned.

Appendix B

Size Exclusion Chromatography Strategies and MASH Explorer for Targeted Proteoform Characterization

Timothy N. Tiambeng,[1] Zhijie Wu,[1] Jake A. Melby,[1] Ying Ge[1,2,3*]

Abstract:

Top-down mass spectrometry (MS)-based analysis of larger proteoforms (> 50 kDa) is typically challenging due to an exponential decay in the signal-to-noise ratio with increasing protein molecular weight (MW) and co-elution with low-MW proteoforms. Size exclusion chromatography (SEC) fractionates proteins based on their size, separating larger proteoforms from those of smaller size in the proteome. In this protocol, we initially describe the use of SEC to fractionate high MW proteoforms from low MW proteoforms. Subsequently, the SEC fractions containing the proteoforms of interest are subjected to reverse phase liquid chromatography (RPLC) coupled online with high-resolution MS. Finally, proteoforms are characterized using MASH Explorer, a user-friendly software environment for in-depth proteoform characterization.

Key Words: Size exclusion chromatography, proteoforms, data analysis, Top-down proteomics

Introduction:

Top-down mass spectrometry (MS)-based proteomics is a powerful technology due to its capability to characterize proteoforms arising from sequence variations, alternative splicing, and post-translational modifications (PTMs) [1-8]. Despite the capability of top-down proteomics, high molecular weight (MW) proteoforms are often under-represented in the MS analysis of the proteome [9-11]. This challenge arises from both the high dynamic range of the proteome, in which protein expression can vary in orders of magnitude, and the exponential decay in the MS signal-to-noise ratios of large proteoforms, resulting mainly from the increased number of charge states observed with electrospray ionization of proteoforms, the greater contribution of heavy isotopes at higher precursor mass, and the detrimental influence of adducting and interfering species [12]. MS-identification and characterization of large proteoforms are particularly challenging if they are co-eluted with smaller proteoforms [9,10]. Moreover, user-friendly software tools remain under-developed for proteoform identification and comprehensive characterization of large proteoforms.

Size exclusion chromatography (SEC) addresses a critical challenge in top-down MS-based analysis of large, high-MW proteoforms[9,10,13,14]. This chromatographic technique fractionates the proteome by separating proteoforms based on their size or hydrodynamic volumes, while reducing sample loss due to minimal interactions of proteins with the SEC stationary phase [13,14]. SEC is usually performed using hydrophilic stationary phases with well-defined pore diameters [15]. Typically, smaller proteoforms more readily diffuse into the pores, while relatively larger proteoforms are less likely to enter the pores. Thus, proteoforms separate based on their size as they pass through the column and are eluted in order of decreasing molecular weight [16]. SEC can be used as an effective and versatile separation method that is highly compatible with top-

down MS-based proteomics and is orthogonal to other chromatographic methods. SEC experiments can be performed in both denatured and native modes, using mobile phases such as formic acid in water [10] and ammonium acetate solution [17], respectively. The resulting SEC fractions containing proteoforms eluting at specific chromatographic elution times can then be subjected to high-resolution MS with simple sample processing. As a notable example, Cai and Tucholski *et. al.,* developed serial size exclusion chromatography (sSEC) by connecting together SEC columns in series with different stationary phase pore diameters, to achieve efficient size-based separation over a broad MW range [9]. When sSEC fractions were further separated by RPLC coupled online with high resolution MS, it enabled a 15-fold improvement in the observation of proteoforms greater than 60 kDa compared to one-dimensional RPLC-MS alone. Additionally, Tucholski *et. al.,* utilized sSEC fractionation directly with Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) analysis to enable sequence characterization of large proteoforms without RPLC separation or extensive protein purification [10].

For characterizing the primary sequence and localizing the PTMs of proteoforms, MASH Explorer is a universal, user-friendly, and freely available software environment for top-down proteomics [18]. MASH Explorer is built upon the previous successes of MASH Suite [19] and MASH Suite Pro [20], to bolster the continued growth of the top-down proteomics community. In comparison to MASH Suite Pro, MASH Explorer can process MS, tandem MS (MS/MS), and liquid chromatography tandem MS (LC-MS/MS) across multiple vendor-specific-formats, with automated database searching for protein identification, and tools for proteoform characterization and data validation. MASH Explorer incorporates various deconvolution and database searching algorithms to assist fragment ion identification and large proteoform characterization. MASH Explorer provides an effective and comprehensive solution to process many types of MS data, and

can be used in combination with other freely available software tools, such as Proteoform Suite [21,22], to assist identification of large proteoforms [11] that are not monoisotopically resolved.

Here in this protocol, we demonstrate that SEC can be generally used to fractionate high-MW proteoforms from low-MW proteoforms. SEC fractions are then subjected to online LC-MS and LC-MS/MS analysis using a quadrupole time-of-flight (QTOF) instrument. Subsequently, MASH Explorer [18] is used to process and analyze MS data to characterize the sequence and PTMs of proteoforms.

2. Materials

All reagents were purchased from MilliporeSigma Inc. (St. Louis, MO, USA) unless otherwise noted. HPLC grade solvents, such as water, acetonitrile, ethanol, isopropanol, and formic acid were purchased from Fisher Scientific (Fair Lawn, NJ, USA) (*see* **Note 1**).

2.1. Size Exclusion Chromatography

1) Waters ACQUITY H-Class UPLC system equipped with a UV detector and an automatic fraction collector (Waters Corporation, Milford, MA, USA).
2) PolyHYDROXYETHYL A (PolyHEA) columns (PolyLC Inc., Columbia, MD, USA) (*see* **Note 2**).
3) Mobile phase, such as 1% formic acid in water (v/v) (*see* **Note 3**).
4) Ultra-centrifugal 10 kDa molecular weight cutoff (MWCO) filters (0.5 mL) (Fisher Scientific, Fair Lawn, NJ, USA).

5) 1/16” O.D. (outer diameter) PEEK tubing with 100 µm I.D. (inner diameter), 2 cm (VICI Valco instruments, Houston, TX, USA)
6) Supelco® 1/16” O.D. PEEK ferrules and fittings (MilliporeSigma Inc. (St. Louis, MO, USA)

2.2 . Online Reverse Phase Liquid Chromatography Mass Spectrometry

1) maXis II quadrupole time-of-fight (QTOF) mass spectrometer (Bruker Daltonics, Bremen, Germany).
2) Waters nanoACQUITY UPLC system (Waters Corporation, Milford, MA, USA).
3) Home-packed PLRP column, (200 x 0.5 mm I.D with PLRP-S bulk media (10 µm, 1000 Å, Agilent Technologies, Santa Clara, CA, USA) (*See* **Note 4**)
4) 1/16” O.D. PEEK tubing with 100 µm I.D., 20 cm (VICI Valco instruments, Houston, TX, USA)
5) Supelco® 1/16” O.D. PEEK ferrules and fittings (MilliporeSigma Inc. (St. Louis, MO, USA)
6) Mobile Phase A: 0.1% formic acid in water (v/v).
7) Mobile Phase B: 0.1% formic acid in 50:50 acetonitrile/ethanol (v/v).
8) Absorbance microplate reader, such as the BioTek ELx808 Absorbance Reader (Winooski, VT, USA).
9) Bradford protein assay reagents, such as the Quick Start™ Bradford Protein Assay Kit (Hercules, CA, USA).

2.3. MASH Explorer Data Analysis

1) Personal computer meeting the minimum hardware requirements (*See* **Note 5**).
2) MASH Explorer software, version 2.1.1. (https://labs.wisc.edu/gelab/software.html).

3) MASH Explorer User Installation Guide: (https://labs.wisc.edu/gelab/MASH_Explorer/doc/User%20Installation%20Guide.pdf) (*See* **Note 6**)

3. Methods

3.1 . Size Exclusion Chromatography

1) For new and previously unused PolyHEA columns, flush the column with 15 column volumes of water to remove methanol.
2) Condition the column(s) by flushing with the intended mobile phase overnight at room temperature to ensure reproducible retention times.
3) For serial SEC, connect the desired number of SEC columns together using 2 cm segments of 1/16" O.D. PEEK tubing (100 μm I.D.), and 1/16" O.D. PEEK ferrules and fittings.
4) Ensure that an appropriately sized sample loop is installed on the UPLC system for SEC or serial SEC applications (*See* **Note 7**)
5) Set the flow rate to 0.5 mL/min and perform three sample injections using water at the same flow rate and isocratic gradient as the intended sample to ensure stable baseline and column backpressure. Ensure that the sample manager is kept at 4°C to reduce artifactual temperature-induced protein modification, such as oxidation. (*See* **Note 8)**
6) As standards for reproducibility, inject 5 μg of protein solution that is buffer exchanged with the starting SEC mobile phase. (*See* **Note 9**)
7) Monitor protein elution from the SEC column by recording the change in UV-absorbance at 280 nm as a function of retention time. (*See* **Note 10**)
8) For fraction collection, begin collecting fractions when the absorbance intensity increases about 10 times the signal-to-noise ratio, or at the approximate elution time of the protein of

interest based on the elution times of the standard proteins (*See* **Note 11**). Ensure that the fraction manager is kept at 4 °C to reduce artifactual temperature-induced protein modification.

9) Concentrate the desired protein-containing SEC fractions using 10 kDa MWCO (*See* **Note 12**) for SDS-PAGE (*See* **Note 13)** and LC-MS/MS analysis (Fig. 1).

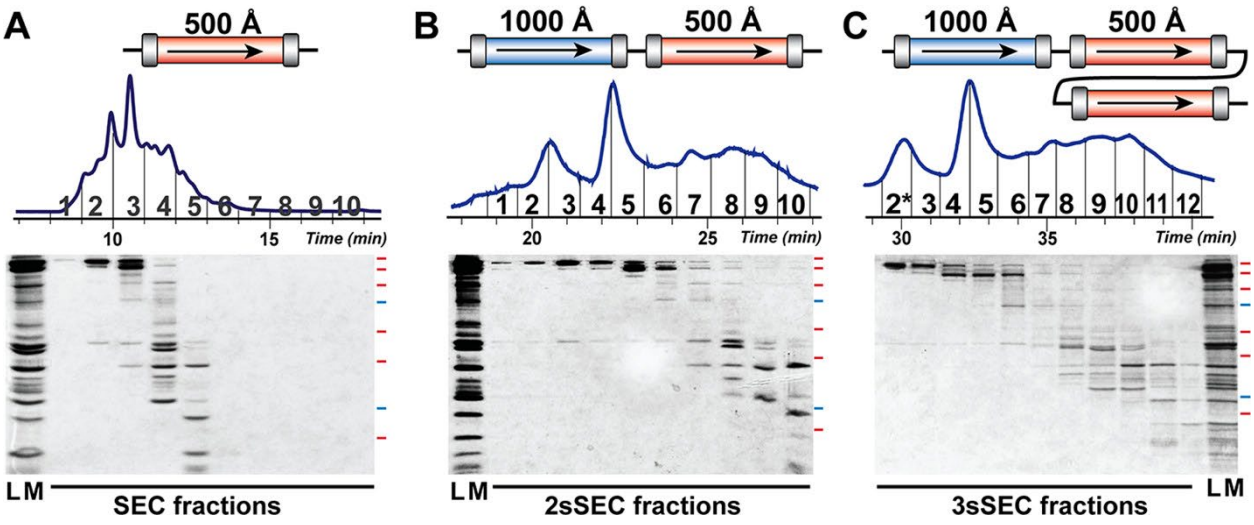


Fig. 1 Example of an SEC separation of a complex protein mixture. Comparison of (**A**) single column SEC (500 Å), (**B**) SEC with two columns connected serially (1000 Å–500 Å), and (**C**) SEC with three columns connected serially (1000 Å–500 Å–500 Å) for the fractionation of the same protein loading mixture (LM). The top panel illustrates the UV chromatogram for the corresponding SEC experiment, annotated with numbers corresponding to the collected SEC fractions. The bottom panel illustrates the SDS-PAGE analysis corresponding to the collected and annotated SEC fractions (Adapted and reprinted with permission from [9] Copyright 2017 American Chemical Society.)

3.2. Online Reverse Phase Chromatography and Top-down MS Analysis:

1) Prepare LC-MS grade 0.1% formic acid in water (v/v) and LC-MS grade 0.1% formic acid in 50:50 acetonitrile:ethanol (v/v).

2) Dilute the protein fractions to a total protein concentration of 100 ng/μL total protein with 0.1% formic acid in water containing 2 mM TCEP. (*See* **Note 14**)

3) Load 500 ng of total protein for a 0.5 mm I.D. PLRP-S column. Separate the proteins chromatographically using a nanoAcquity UPLC system equipped with a PLRP-S column. Connect the PLRP-S column to the electrospray ionization source using 1/16" O.D. PEEK tubing with 100 μm I.D., and 1/16" O.D. PEEK ferrules and fittings.

4) Elute proteins using a mobile phase gradient going from 5% B to 95% B over 45 min at a flow rate of 8 μL/min (mobile phase A: 0.1% formic acid in water, mobile phase B: 0.1% formic acid in 50:50 acetonitrile:ethanol). (*See* **Note 15**)

5) For the electrospray ionization source on the QTOF mass spectrometer, set the "End Plate Offset" at 500 V, the "Capillary" at 4500 V, the "Nebulizer" at 0.5 bar, the "Dry Gas" at 4.0 L/min, and the "Dry Temp" at 220 °C. Collect mass spectra at a scan rate of 0.5 Hz over 500-2000 *m/z* range.

6) Complete an initial LC-MS run on a SEC protein fraction of interest and perform data analysis to determine the charge state ions, molecular weight, and retention time of each target protein.

7) For targeted protein analysis, complete a second LC-MS run using online collisionally activated dissociation (CAD) by inputting the precursor ion(s), ion isolation range (*m/z*), CAD energy, and the chromatographic elution interval. The precursor ion(s) will be isolated and fragmented in the targeted MS/MS experiment for protein identification (Fig. 2) (*See* **Note 16**).

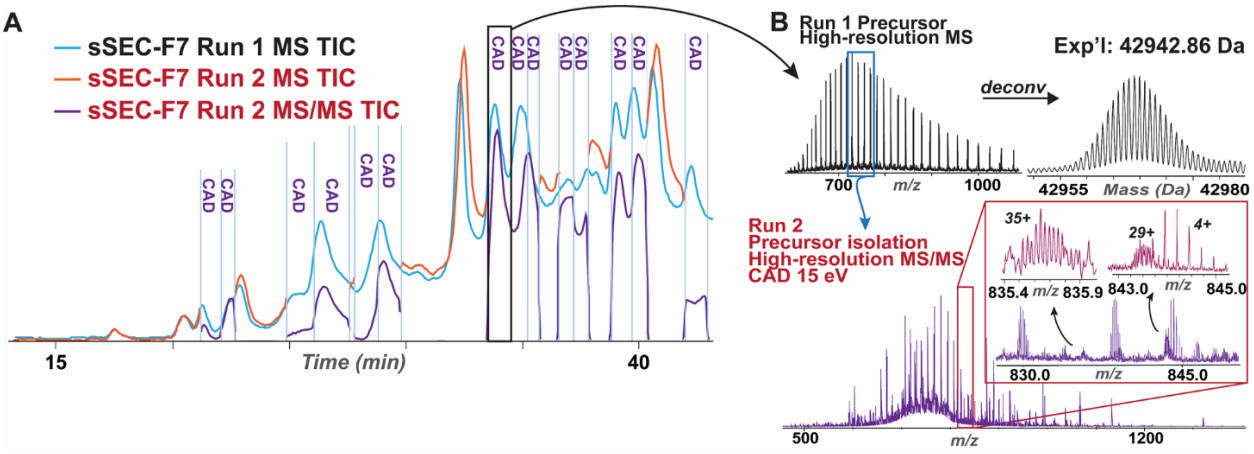


Fig. 2 Top-down targeted MS/MS for protein identification in a selected sSEC fraction. (**A**) CAD-based MS/MS experiments are performed on selected LC retention time windows containing proteoforms of interest. (**B**) High-resolution MS demonstrating representative precursor ion isolation and effective CAD-based fragmentation of a 42.9 kDa protein, yielding high-resolution tandem mass spectra. (Adapted and reprinted with permission from [9] Copyright 2017 American Chemical Society.)

3.3 Data Analysis using MASH Explorer's Targeted Mode:

1) Import raw MS/MS data files into the MASH Explorer software under Targeted Mode (Fig. 3), (*See* **Note 17)**.

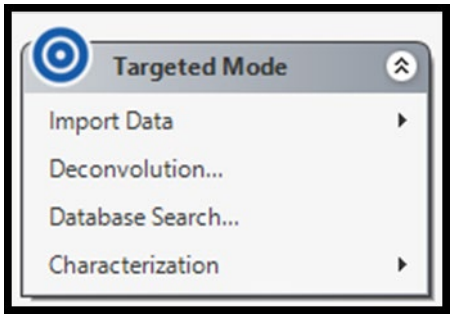
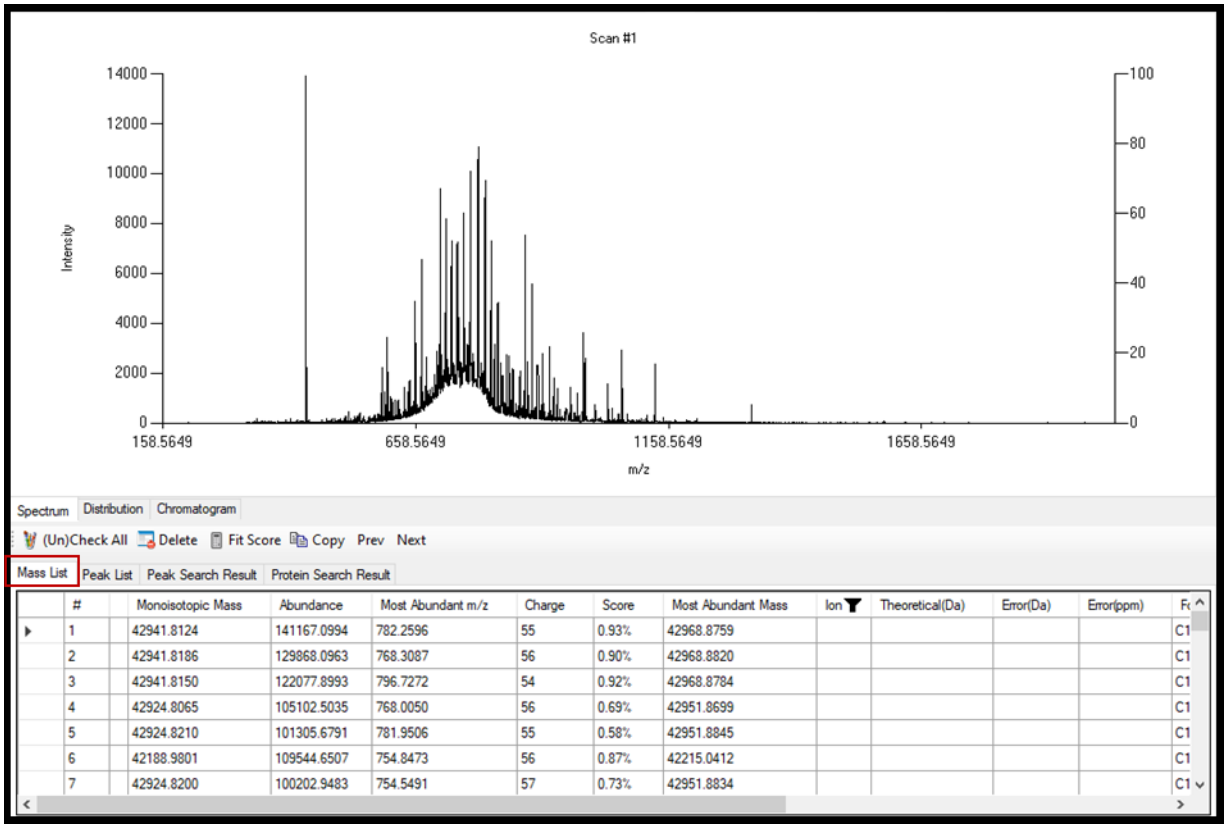


Fig. 3 MASH Explorer's Targeted Mode interface for MS/MS data import.

2) Perform spectral deconvolution using algorithms such as THRASH [23] or TopFD [24] (*See* **Note 18**). After selecting a deconvolution method, proceed to the Advanced tab to change settings such as the precursor ion *m/z* and charge in the General subheading, as well as specific algorithm parameters such as Max Charge and Max Mass in the Deconvolution subheading. A list of deconvoluted fragment ions will be generated by the selected algorithm and automatically imported back to the MASH Explorer software interface. (Fig. 4)



#	Monoisotopic Mass	Abundance	Most Abundant m/z	Charge	Score	Most Abundant Mass	Ion	Theoretical(Da)	Error(Da)	Error(ppm)	Fc
1	42941.8124	141167.0994	782.2596	55	0.93%	42968.8759					C1
2	42941.8186	129868.0963	768.3087	56	0.90%	42968.8820					C1
3	42941.8150	122077.8993	796.7272	54	0.92%	42968.8784					C1
4	42924.8065	105102.5035	768.0050	56	0.69%	42951.8699					C1
5	42924.8210	101305.6791	781.9506	55	0.58%	42951.8845					C1
6	42188.9801	109544.6507	754.8473	56	0.87%	42215.0412					C1
7	42924.8200	100202.9483	754.5491	57	0.73%	42951.8834					C1

Fig. 4 MS/MS spectral deconvolution and the calculated mass list imported by MASH Explorer software from the selected deconvolution algorithm, displayed in MASH Explorer's Mass List panel (highlighted in red).

3) Verify the deconvoluted fragment ions in the Mass List by adjusting their charge states and monoisotopic mass. A more precise list of fragment ions will allow for shorter processing time and more accurate proteoform identifications in the database search.
4) Using the verified fragment ion list, perform database search using algorithms such as MS-Align+ [25] (*See* **Note 19**). The precursor ion information obtained during data analysis in section 3.2, such as the precursor ion *m/z* and charge will assist the database search algorithm to identify targeted proteoforms more accurately.
5) Import the identified proteoform identification after completing the database search. Under the Characterization heading, select Import Database Search Results, and select the database algorithm used for database search (e.g., TopPIC Search Results). Select the folder labeled "MASH_Workflows" and select the appropriate MassList folder labeled with the database search algorithm (e.g., '…MassList_TopPIC_1_1) (Fig. 5). The folder contains a .CSV file with the Protein Search results. After opening the .CSV file, the protein sequence will be uploaded to MASH Explorer's Sequence Table.

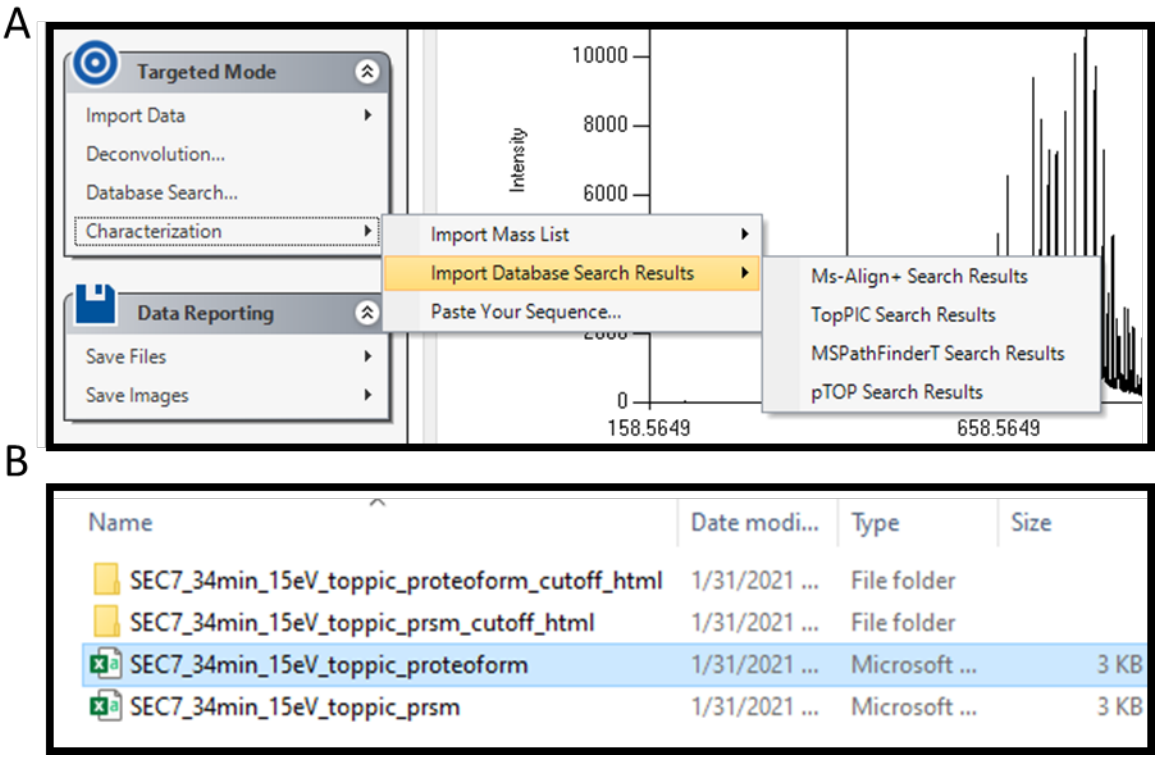


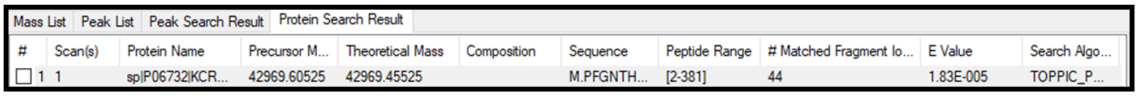
Fig. 5 MASH Explorer enables import of database search results. A) MASH Explorer supports multiple database search algorithms including TopPIC [24], MS-Align+ [25], pTop [26], and Informed-Proteomics [27] workflows. B) Highlighted .CSV file containing database search results from the "MASH_Workflows" folder.

6) Using MASH Explorer software, visualize the verified fragment ion list against the imported proteoform sequence in the Sequence Table for proteoform characterization (Fig. 6). A warning message will appear from MASH Explorer reminding the user to correct the protein sequence to obtain the correct fragmentation list.

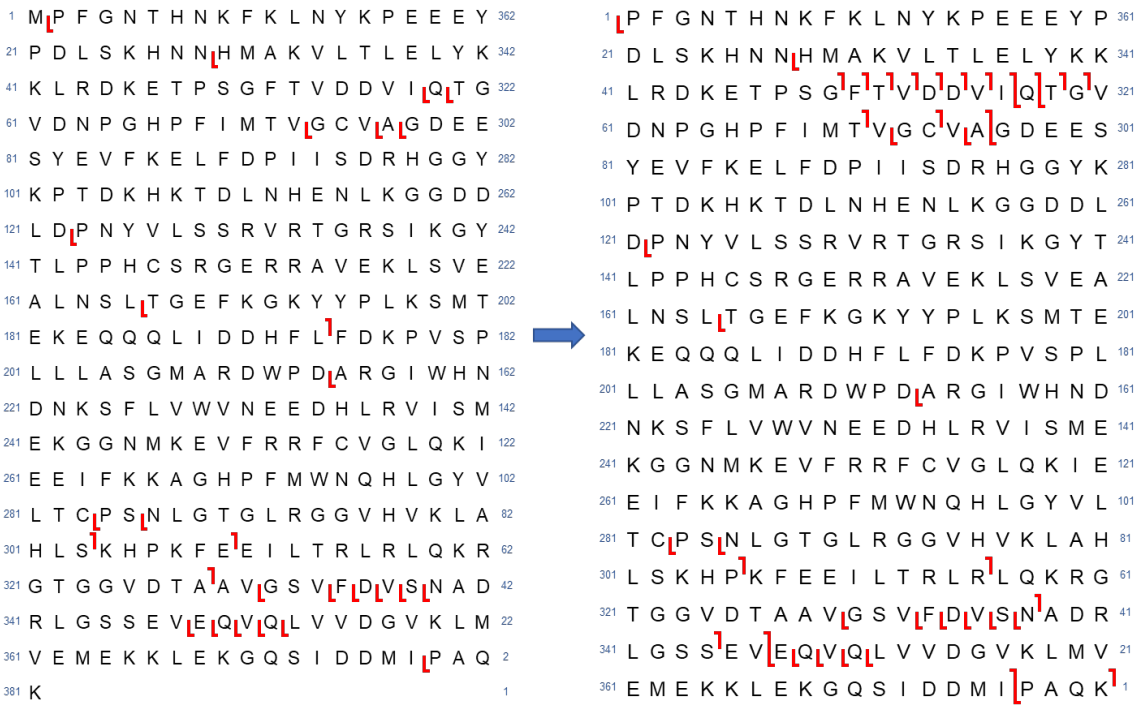
7) Characterize the protein post-translational modifications. Users can obtain this information by intact mass analysis during section 3.2, or by database search algorithm results (*See* **Note 20**).

8) Use database search algorithms in MASH Explorer to identify proteoforms (*See* **Note 21**).

A


Mass List
Peak List
Peak Search Result
Protein Search Result
| Scan(s) | Protein Name | Precursor M... | Theoretical Mass | Composition | Sequence | Peptide Range | # Matched Fragment Io... | E Value | Search Algo...
1 | 1 | sp|P06732|KCR... | 42969.60525 | 42969.45525 | | M.PFGNTH... | [2-381] | 44 | 1.83E-005 | TOPPIC_P...

B


1 M P F G N T H N K F K L N Y K P E E E Y 362
21 P D L S K H N N H M A K V L T L E L Y K 342
41 K L R D K E T P S G F T V D D V I Q T G 322
61 V D N P G H P F I M T V G C V A G D E E 302
81 S Y E V F K E L F D P I I S D R H G G Y 282
101 K P T D K H K T D L N H E N L K G G D D 262
121 L D P N Y V L S S R V R T G R S I K G Y 242
141 T L P P H C S R G E R R A V E K L S V E 222
161 A L N S L T G E F K G K Y Y P L K S M T 202
181 E K E Q Q Q L I D D H F L F D K P V S P 182
201 L L L A S G M A R D W P D A R G I W H N 162
221 D N K S F L V W V N E E D H L R V I S M 142
241 E K G G N M K E V F R R F C V G L Q K I 122
261 E E I F K K A G H P F M W N Q H L G Y V 102
281 L T C P S N L G T G L R G G V H V K L A 82
301 H L S K H P K F E E I L T R L R L Q K R 62
321 G T G G V D T A A V G S V F D V S N A D 42
341 R L G S S E V E Q V Q L V V D G V K L M 22
361 V E M E K K L E K G Q S I D D M I P A Q 2
381 K 1

1 P F G N T H N K F K L N Y K P E E E Y P 361
21 D L S K H N N H M A K V L T L E L Y K K 341
41 L R D K E T P S G F T V D D V I Q T G V 321
61 D N P G H P F I M T V G C V A G D E E S 301
81 Y E V F K E L F D P I I S D R H G G Y K 281
101 P T D K H K T D L N H E N L K G G D D L 261
121 D P N Y V L S S R V R T G R S I K G Y T 241
141 L P P H C S R G E R R A V E K L S V E A 221
161 L N S L T G E F K G K Y Y P L K S M T E 201
181 K E Q Q Q L I D D H F L F D K P V S P L 181
201 L L A S G M A R D W P D A R G I W H N D 161
221 N K S F L V W V N E E D H L R V I S M E 141
241 K G G N M K E V F R R F C V G L Q K I E 121
261 E I F K K A G H P F M W N Q H L G Y V L 101
281 T C P S N L G T G L R G G V H V K L A H 81
301 L S K H P K F E E I L T R L R L Q K R G 61
321 T G G V D T A A V G S V F D V S N A D R 41
341 L G S S E V E Q V Q L V V D G V K L M V 21
361 E M E K K L E K G Q S I D D M I P A Q K 1

C

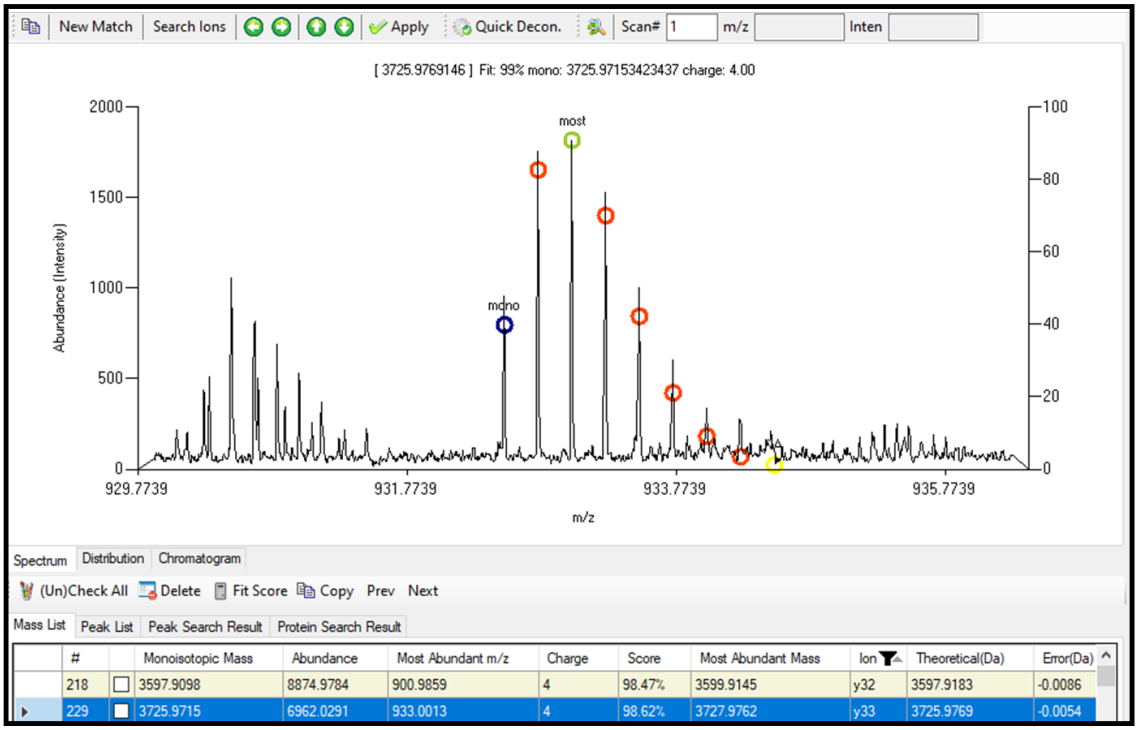

New Match
Search Ions
Apply
Quick Decon.
Scan# 1
m/z
Inten
[3725.9769146] Fit: 99% mono: 3725.97153423437 charge: 4.00
most
mono
Abundance (Intensity)
m/z
2000
1500
1000
500
0
100
80
60
40
20
0
929.7739
931.7739
933.7739
935.7739
Spectrum
Distribution
Chromatogram
(Un)Check All
Delete
Fit Score
Copy
Prev
Next
Mass List
Peak List
Peak Search Result
Protein Search Result
| Monoisotopic Mass | Abundance | Most Abundant m/z | Charge | Score | Most Abundant Mass | Ion | Theoretical(Da) | Error(Da)
218 | 3597.9098 | 8874.9784 | 900.9859 | 4 | 98.47% | 3599.9145 | y32 | 3597.9183 | -0.0086
229 | 3725.9715 | 6962.0291 | 933.0013 | 4 | 98.62% | 3727.9762 | y33 | 3725.9769 | -0.0054

Fig. 6 Visualizing and verifying fragment ions against an imported protein sequence in MASH Explorer. A) Imported database search results can be found under the Protein Search Result tab. B) MASH Explorer's 'Protein Sequence Table' allows users to visualize fragment ion mapping for different MS/MS techniques to characterize the protein sequence and PTMs. The protein amino acid sequence corresponding to the identified protein is directly imported from the database search. Correcting the protein sequence according to the database search (N-terminal methionine excision), results in significantly improved ion fragmentation lists. C) MASH Explorer enables manual verification of MS/MS fragment ion data and adjustment of charge state and monoisotopic mass.

4. Notes:

1) Ensure that all mobile phases are prepared in thoroughly cleaned glassware intended for HPLC solvents. Always clean glassware with the same mobile phase/solvent used for the chromatography; never wash glassware with detergent as this can introduce contaminants during LC-MS analysis.
2) A variety of column dimensions can be used for SEC and sSEC experiments. Typical column dimensions include 200 mm x 9.4 mm, 3 μm particles with pore sizes of 500 Å, and 1000 Å. Ideally for SEC, the use of small particles, wider I.D. columns, and shorter column lengths can result in columns where wall effects are negligible and protein band spreading decreases [28]. Strategies such as minimizing sample injection volume and increasing initial protein loading can assist in reducing significant dilution and loss of UV signal during SEC.

3) A variety of MS-compatible mobile phase compositions can be used for the fractionation of different protein-classes. Previously, we have demonstrated that 1% formic acid is effective for acid-soluble and water-soluble proteins [9,10]. For membrane protein fractionation, our lab has shown that addition of 40% isopropanol (1% formic acid, 40% isopropanol, 59% water v/v) can facilitate membrane protein solubility in the absence of surfactants [29]. For proteins that strongly interact with each other, even under denaturing conditions, the addition of hexafluoroisopropanol can be helpful to disrupt protein-protein interactions and can be added to a final concentration of 100 mM. Samples with hexafluoroisopropanol can be centrifuged in MWCO filters at 15,000 x *g* for 30 min at 4 °C prior to SEC.
4) The maximum backpressure of this PLRP-S material is approximately 3000 psi; exceeding this pressure will reduce the performance and lifetime of the column.
5) MASH Explorer requires Microsoft Windows operating system with installation with .NET framework. It is recommended the personal computer to have at least central processing units with two cores and two threads, and four gigabytes of random-access memory for operation. For deconvolution and database search algorithms, it is suggested that a personal computer with a minimal computing power of a central processing unit with four cores and four threads, and eight gigabytes of random-access memory is used.
6) The MASH Explorer installation package does not contain any deconvolution or database search algorithms except for the THRASH deconvolution algorithm. To access other supported deconvolution and database search algorithms, they must be downloaded from the individual research groups which developed them. MASH Explorer allows for convenient import of these deconvolution and database search algorithms by accessing the Tools menu and selecting the Configuration dialog.

7) For single column SEC (200 x 4.6 mm, 5 µm, 500 Å PolyHEA), the protein loading capacity is approximately 500 µg. A sample loop volume of 50 - 250 µl can be used for sample injection. Using smaller sample loop volumes when possible helps to reduce protein band broadening. For single column SEC, we typically use a 50 µl sample loop and 100 µg protein loading at 2 µg/µl sample concentration; for two SEC columns connected serially, we use a 250 µl sample loop and 200 µg protein loading; for three SEC columns connected serially, we use a 250 µl sample loop and 300 µg protein loading. Serial SEC can improve protein separation and fractionation range but requires a larger sample injection and greater protein loading to avoid diluting protein signal.
8) Using a single SEC column (200 x 9.6 mm, 3 µm, 500 Å PolyHEA), the backpressure is approximately 1000 psi using a flow rate of 0.5 mL/min with 1% formic acid in water (v/v) as mobile phase. For two SEC columns connected serially, the pressure is approximately 1500 psi; for three SEC columns connected serially, the pressure is approximately 2000 psi. Use a KimWipe™ to detect solvent leaks at the junction of the fittings and the columns.
9) A panel of standard proteins are often used in our lab to evaluate the performance of SEC columns. Depending on native or denaturing SEC applications, these standard proteins include thyroglobulin (669 kDa), apoferritin (443 kDa), β-amylase (200 kDa), alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa), myoglobin (17 kDa). For denaturing SEC, β-amylase, bovine serum albumin, and myoglobin are typical standards used to evaluate the performance of SEC columns.
10) Proteins usually show UV-absorption maxima between 275 and 280 nm, which are caused by absorbance of the two aromatic amino acids tryptophan (Trp) and tyrosine (Tyr), and to a small extent, by the absorbance of cystine (i.e., disulfide bonds) [30].

11) We normally collect SEC fractions using 1-min intervals (~ 0.5 mL elution fractions), but longer time intervals (e.g., 1.5-, or 2-min intervals) can be collected and pooled together.

12) MWCO filters are pre-rinsed by centrifuging with the mobile phase used in the SEC separation to reduce protein loss resulting from concentration of the SEC fractions. 2 mM TCEP is added to reduce protein oxidation and is active under acidic conditions such as 1% formic acid in water (v/v).

13) For typical SDS-PAGE experiments, commercially available products such as 8-16% Mini-PROTEAN® 12.5% Tris-Glycine eXtended (TGX) gels or Novex™ 8 to 16 % Tris-Glycine Plus Midi Gels can be used according to the manufacturer's instructions (e.g., 125 V until electrophoresis is complete). For protein loading between 100 – 1000 ng per lane, it is recommended to use a highly sensitive staining method such as SYPRO Ruby or Silver Staining. For protein loading greater than 1 µg per lane we normally perform Coomassie Brilliant Blue staining.

14) To quantify and normalize total protein loading for equal protein loading in SDS-PAGE and LC-MS experiments, we typically use the Bradford microplate assay for protein samples that are prepared without detergents and measure the absorbance of standards and samples at 595 nm using an absorbance microplate reader. For samples which contain detergents, it is recommended to use the bicinchoninic acid (BCA) microplate assay, which is compatible with most ionic and non-ionic detergents, and measure the absorbance of standards and samples at 562 nm using an absorbance microplate reader.

15) While LC mobile phase gradients will require optimization and vary depending on the protein mixture, a typical LC gradient used for the separation of cardiac proteins commonly performed

in our lab is 0-5 min 20% B, 5-17 min 20-40% B, 17-25 min 40-50% B, 25-32 min 50-65%, 32-43 min 65-95 % B, 43-48 min 95% B, 48-52 min 95-5% B, 52-60 min 5% B.

16) To select the appropriate fragmentation energy for targeted MS/MS experiments using collisionally activated dissociation (CAD), the precursor ion in the resulting MS/MS spectrum should still be present at 50-70% base peak intensity to avoid over fragmentation, which may result in internal fragments and lower identified fragments. For larger MW proteins, they may require lower fragmentation energy than that required for smaller MW proteins. A good starting range for CAD energy is 10-20 eV for large MW proteins.

17) To support data import, MASH Explorer uses both ProteoWizard [31] and vendor-specific software (Thermo .RAW, Bruker .BAF, Bruker .ascii, Waters .RAW, .MzXML, .MGF, .mzML) .

18) MASH Explorer supports several deconvolution algorithms including THRASH, TopFD [24], and pParseTD [26], MS-Deconv [32],. The results from THRASH, MS-Deconv, and TopFD are most often used for fragment ion verification.

19) MASH Explorer supports multiple database search algorithms including TopPIC [24], MS-Align+ [25], pTop [26], and Informed-Proteomics [27] workflows.

20) PTMs are supported directly in MASH Explorer software. Modifications such as acetylation, trimethylation, phosphorylation, etc., can be directly added on the target amino acid in the protein sequence. Other modifications can be added using custom modification. Database search algorithms are useful tools in identifying PTMs such as N-terminal acetylation and methylation. However, for phosphorylation, manual proteoform characterization is often needed using the verified fragment ion list [33-40].

21) Identification of larger MW proteoforms may not be confidently and accurately identified by database search algorithms. Using the verified fragment ion list, a series of fragment ions in a protein sequence may be used to derive a three or four consecutive amino acid sequence, called sequence tag. This sequence tag can be as an alternative method to identify target proteoform matching the intact protein mass. The amino acid sequence for the sequence tag can be derived by mass differences of several fragment ions. The mass differences can be matched to one or the sum of multiple amino acid residue masses (https://proteomicsresource.washington.edu/mascot/help/aa_help.html). The obtained sequence tag can be searched against the database using the guide provided by UniProt (https://web.expasy.org/tagident/).

Acknowledgements:

We would like to thank Trisha Tucholski for helpful discussions. This work was supported by NIH R01 GM117058 and NIH R01 GM125085 (to Y.G.). Y.G. also would like to acknowledge R01 HL096971 and S10 OD018475. T.N.T. would like to acknowledge support from the NIH Chemistry-Biology Interface Training Program NIH T32GM008505. J.A.M. would like to acknowledge support from the Training Program in Translational Cardiovascular Science, T32 HL007936-20.

Reference Citations:

1. Smith LM, Kelleher NL (2018) Proteoforms as the next proteomics currency. Science 359 (6380):1106

2. Smith LM, Thomas PM, Shortreed MR, Schaffer LV, Fellers RT, LeDuc RD, Tucholski T, Ge Y, Agar JN, Anderson LC, Chamot-Rooke J, Gault J, Loo JA, Pasa-Tolic L, Robinson CV, Schluter H, Tsybin YO, Vilaseca M, Vizcaino JA, Danis PO, Kelleher NL (2019) A five-level classification system for proteoform identifications. Nat Methods 16:939

3. Aebersold R, Agar JN, Amster IJ, Baker MS, Bertozzi CR, Boja ES, Costello CE, Cravatt BF, Fenselau C, Garcia BA, Ge Y, Gunawardena J, Hendrickson RC, Hergenrother PJ, Huber CG, Ivanov AR, Jensen ON, Jewett MC, Kelleher NL, Kiessling LL, Krogan NJ, Larsen MR, Loo JA, Loo RRO, Lundberg E, MacCoss MJ, Mallick P, Mootha VK, Mrksich M, Muir TW, Patrie SM, Pesavento JJ, Pitteri SJ, Rodriguez H, Saghatelian A, Sandoval W, Schluter H, Sechi S, Slavoff SA, Smith LM, Snyder MP, Thomas PM, Uhlen M, Van Eyk JE, Vidal M, Walt DR, White FM, Williams ER, Wohlschlager T, Wysocki VH, Yates NA, Young NL, Zhang B (2018) How many human proteoforms are there? Nat Chem Biol 14 (3):206

4. Brown KA, Chen BF, Guardado-Alvarez TM, Lin ZQ, Hwang L, Ayaz-Guner S, Jin S, Ge Y (2019) A photocleavable surfactant for top-down proteomics. Nat Methods 16 (5):417

5. Tiambeng TN, Roberts DS, Brown KA, Zhu Y, Chen B, Wu Z, Mitchell SD, Guardado-Alvarez TM, Jin S, Ge Y (2020) Nanoproteomics enables proteoform-resolved analysis of low-abundance proteins in human serum. Nature Communications 11 (1):3903. doi:10.1038/s41467-020-17643-1

6. Chen B, Brown KA, Lin Z, Ge Y (2018) Top-Down Proteomics: Ready for Prime Time? Analytical Chemistry 90 (1):110-127. doi:10.1021/acs.analchem.7b04747

7. Toby TK, Fornelli L, Kelleher NL (2016) Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annual Review of Analytical Chemistry 9 (1):499-519. doi:10.1146/annurev-anchem-071015-041550

8. Brown KA, Melby JA, Roberts DS, Ge Y (2020) Top-down proteomics: challenges, innovations, and applications in basic and clinical research. Expert Review of Proteomics 17 (10):719-733. doi:10.1080/14789450.2020.1855982

9. Cai W, Tucholski T, Chen B, Alpert AJ, McIlwain S, Kohmoto T, Jin S, Ge Y (2017) Top-Down Proteomics of Large Proteins up to 223 kDa Enabled by Serial Size Exclusion Chromatography Strategy. Analytical Chemistry. doi:10.1021/acs.analchem.7b00380

10. Tucholski T, Knott SJ, Chen B, Pistono P, Lin Z, Ge Y (2019) A Top-Down Proteomics Platform Coupling Serial Size Exclusion Chromatography and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. Analytical Chemistry 91 (6):3835-3844. doi:10.1021/acs.analchem.8b04082

11. Schaffer LV, Tucholski T, Shortreed MR, Ge Y, Smith LM (2019) Intact-Mass Analysis Facilitating the Identification of Large Human Heart Proteoforms. Analytical Chemistry 91 (17):10937-10942. doi:10.1021/acs.analchem.9b02343

12. Compton PD, Zamdborg L, Thomas PM, Kelleher NL (2011) On the Scalability and Requirements of Whole Protein Mass Spectrometry. Analytical Chemistry 83 (17):6868-6874. doi:10.1021/ac2010795

13. Doucette AA, Tran JC, Wall MJ, Fitzsimmons S (2011) Intact proteome fractionation strategies compatible with mass spectrometry. Expert Review of Proteomics 8 (6):787-800. doi:10.1586/epr.11.67

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

15. Alpert AJ (2016) Protein Fractionation and Enrichment Prior to Proteomics Sample Preparation. In: Mirzaei H, Carrasco M (eds) Modern Proteomics - Sample Preparation, Analysis and Practical Applications, vol 919. Advances in Experimental Medicine and Biology. pp 23-41. doi:10.1007/978-3-319-41448-5_2

16. Hong P, Koza S, Bouvier ESP (2012) Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates. J Liq Chromatogr Relat Technol 35 (20):2923-2950. doi:10.1080/10826076.2012.743724

17. Hengel SM, Sanderson R, Valliere-Douglass J, Nicholas N, Leiske C, Alley SC (2014) Measurement of in Vivo Drug Load Distribution of Cysteine-Linked Antibody–Drug Conjugates Using Microscale Liquid Chromatography Mass Spectrometry. Analytical Chemistry 86 (7):3420-3425. doi:10.1021/ac403860c

18. Wu Z, Roberts DS, Melby JA, Wenger K, Wetzel M, Gu Y, Ramanathan SG, Bayne EF, Liu X, Sun R, Ong IM, McIlwain SJ, Ge Y (2020) MASH Explorer: A Universal Software Environment for Top-Down Proteomics. Journal of Proteome Research 19 (9):3867-3876. doi:10.1021/acs.jproteome.0c00469

19. Guner H, Close PL, Cai W, Zhang H, Peng Y, Gregorich ZR, Ge Y (2014) MASH Suite: a user-friendly and versatile software interface for high-resolution mass spectrometry data interpretation and visualization. J Am Soc Mass Spectrom 25 (3):464

20. Cai W, Guner H, Gregorich ZR, Chen AJ, Ayaz-Guner S, Peng Y, Valeja SG, Liu X, Ge Y (2016) MASH Suite Pro: A Comprehensive Software Tool for Top-Down Proteomics. Molecular & Cellular Proteomics 15 (2):703-714. doi:10.1074/mcp.o115.054387

21. Cesnik AJ, Shortreed MR, Schaffer LV, Knoener RA, Frey BL, Scalf M, Solntsev SK, Dai Y, Gasch AP, Smith LM (2018) Proteoform Suite: Software for Constructing, Quantifying, and Visualizing Proteoform Families. Journal of Proteome Research 17 (1):568-578. doi:10.1021/acs.jproteome.7b00685

22. Schaffer LV, Shortreed MR, Cesnik AJ, Frey BL, Solntsev SK, Scalf M, Smith LM (2018) Expanding Proteoform Identifications in Top-Down Proteomic Analyses by Constructing Proteoform Families. Analytical Chemistry 90 (2):1325-1333. doi:10.1021/acs.analchem.7b04221

23. Horn DM, Zubarev RA, McLafferty FW (2000) Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J Am Soc Mass Spectrom 11 (4):320

24. Kou Q, Xun L, Liu X (2016) TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. Bioinformatics 32 (22):3495

25. Liu X, Sirotkin Y, Shen Y, Anderson G, Tsai YS, Ting YS, Goodlett DR, Smith RD, Bafna V, Pevzner PA (2012) Protein Identification Using Top-Down Spectra. Mol Cell Proteomics 11 (6):M111.008524

26. Sun RX, Luo L, Wu L, Wang RM, Zeng WF, Chi H, Liu C, He SM (2016) pTop 1.0: A High-Accuracy and High-Efficiency Search Engine for Intact Protein Identification. Anal Chem 88 (6):3082

27. Park J, Piehowski PD, Wilkins C, Zhou M, Mendoza J, Fujimoto GM, Gibbons BC, Shaw JB, Shen Y, Shukla AK, Moore RJ, Liu T, Petyuk VA, Tolic N, Pasa-Tolic L, Smith RD, Payne SH,

Kim S (2017) Informed-Proteomics: open-source software package for top-down proteomics. Nat Methods 14 (9):909

28. Regnier FE, Gooding KM (1980) High-performance liquid chromatography of proteins. Analytical Biochemistry 103 (1):1-25. doi:https://doi.org/10.1016/0003-2697(80)90229-8

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

30. Schmid FX (2001) Biological macromolecules: UV-visible spectrophotometry. e LS

31. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak M-Y, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. Nature Biotechnology 30 (10):918-920. doi:10.1038/nbt.2377

32. Liu X, Inbar Y, Dorrestein PC, Wynne C, Edwards N, Souda P, Whitelegge JP, Bafna V, Pevzner PA (2010) Deconvolution and database search of complex tandem mass spectra of intact proteins: a combinatorial approach. Mol Cell Proteomics 9 (12):2772

33. Jin Y, Wei L, Cai W, Lin Z, Wu Z, Peng Y, Kohmoto T, Moss RL, Ge Y (2017) Complete Characterization of Cardiac Myosin Heavy Chain (223 kDa) Enabled by Size-Exclusion Chromatography and Middle-Down Mass Spectrometry. Analytical Chemistry 89 (9):4922-4930. doi:10.1021/acs.analchem.7b00113

34. Wu Z, Jin Y, Chen B, Gugger MK, Wilkinson-Johnson CL, Tiambeng TN, Jin S, Ge Y (2019) Comprehensive Characterization of the Recombinant Catalytic Subunit of cAMP-Dependent Protein Kinase by Top-Down Mass Spectrometry. Journal of the American Society for Mass Spectrometry 30 (12):2561-2570. doi:10.1021/jasms.8b06294

35. Jin Y, Lin Z, Xu Q, Fu C, Zhang Z, Zhang Q, Pritts WA, Ge Y (2019) Comprehensive characterization of monoclonal antibody by Fourier transform ion cyclotron resonance mass spectrometry. mAbs 11 (1):106-115. doi:10.1080/19420862.2018.1525253

36. Lin Z, Guo F, Gregorich ZR, Sun R, Zhang H, Hu Y, Shanmuganayagam D, Ge Y (2018) Comprehensive Characterization of Swine Cardiac Troponin T Proteoforms by Top-Down Mass Spectrometry. Journal of the American Society for Mass Spectrometry 29 (6):1284-1294. doi:10.1021/jasms.8b05834

37. Tucholski T, Cai W, Gregorich ZR, Bayne EF, Mitchell SD, McIlwain SJ, de Lange WJ, Wrobbel M, Karp H, Hite Z, Vikhorev PG, Marston SB, Lal S, Li A, dos Remedios C, Kohmoto T, Hermsen J, Ralphe JC, Kamp TJ, Moss RL, Ge Y (2020) Distinct hypertrophic cardiomyopathy genotypes result in convergent sarcomeric proteoform profiles revealed by top-down proteomics. Proceedings of the National Academy of Sciences 117 (40):24691. doi:10.1073/pnas.2006764117

38. Chen B, Guo X, Tucholski T, Lin Z, McIlwain S, Ge Y (2017) The Impact of Phosphorylation on Electron Capture Dissociation of Proteins: A Top-Down Perspective. Journal of the American Society for Mass Spectrometry 28 (9):1805-1814. doi:10.1007/s13361-017-1710-3

39. Jin Y, Diffee GM, Colman RJ, Anderson RM, Ge Y (2019) Top-down Mass Spectrometry of Sarcomeric Protein Post-translational Modifications from Non-human Primate Skeletal Muscle. J Am Soc Mass Spectrom 30 (12):2460

40. Peng Y, Gregorich ZR, Valeja SG, Zhang H, Cai W, Chen YC, Guner H, Chen AJ, Schwahn DJ, Hacker TA, Liu X, Ge Y (2014) Top-down proteomics reveals concerted reductions in myofilament and Z-disc protein phosphorylation after acute myocardial infarction. Mol Cell Proteomics 13 (10):2752-2764. doi:10.1074/mcp.M114.040675