A Novel, In-Solution Separation of Endogenous Cardiac Sarcomeric Proteins and Identification of Distinct Charged Variants of Regulatory Light Chain

Sarah B. Scruggs[1,2], Rick Reisdorph[3], Mike L. Armstrong[3], Chad M. Warren[1], Nichole Reisdorph[3], R. John Solaro[1§], Peter M. Buttrick[4]

From [1]University of Illinois at Chicago, Department of Physiology and Biophysics and Center for Cardiovascular Research, Chicago, Illinois, 60612, [2]University of California Los Angeles, Departments of Physiology and Medicine, Division of Cardiology, Los Angeles, California, 90095, [3]National Jewish Health, Department of Immunology, Denver, Colorado, 80206, [4]University of Colorado at Denver, Division of Cardiology, Aurora, CO, 80045, USA.

[§]To Whom Correspondence Should be Addressed: R. John Solaro, Ph.D., University of Illinois at Chicago, Department of Physiology and Biophysics, 835 South Wolcott Avenue, M/C 901, Chicago, Illinois 60612, Phone: 312-996-7620, Fax: 312-996-1414, Email: solarorj@uic.edu

Running Title: In-solution separation and cardiac light chain phosphorylation

Abbreviations: RP, Reverse Phase, ELC, Essential Light Chain; FHC, Familial Hypertrophic Cardiomyopathy; LMM, Light Meromyosin; MHC, Myosin Heavy Chain; MLCK, Myosin Light Chain Kinase; MyBP-C, Myosin Binding Protein-C; OGE, Offgel Electrophoresis; PTM, Post-Translational Modification; RLC, Regulatory Light Chain; SPI, Scored Peak Intensity

Summary.

The molecular conformation of the cardiac myosin motor is modulated by inter-molecular interactions among the heavy chain, the light chains, myosin binding protein-C (MyBP-C) and titin, and is governed by post-translational modifications (PTMs). In-gel digestion followed by liquid chromatography mass spectrometry (LC/MS/MS) has classically been applied to identify cardiac sarcomeric PTMs; however, this approach is limited by protein size, pI, and difficulties in peptide extraction. We report a solution-based workflow for global separation of endogenous cardiac sarcomeric proteins with a focus on the regulatory light chain (RLC) in which specific sites of phosphorylation have been unclear. Sub-cellular fractionation followed by OFFGEL electrophoresis (OGE) resulted in isolation of endogenous charge variants of sarcomeric proteins, including regulatory and essential light chains, myosin heavy chain (MHC), and MyBP-C of the thick filament. Further purification of RLC using reverse phase (RP) -HPLC separation and UV detection enriched for RLC PTMs at the intact protein level, and provided a stoichiometric and quantitative assessment of endogenous RLC charge variants. Digestion and subsequent LC/MS/MS unequivocally identified that the endogenous charge variants of cardiac RLC focused in unique OGE fractions were un-phosphorylated (78.8%), singly- (18.1%) and doubly-phosphorylated (3.1%) RLC. The novel aspects of this study are: 1) *milligram* amounts of endogenous cardiac sarcomeric sub-proteome were focused with resolution comparable to 2DE, 2) separation and quantification of post-translationally modified variants was achieved at the intact protein level, 3) separation of intact high molecular weight thick filament proteins was achieved in-solution, 4) endogenous charge variants of RLC were separated; a novel doubly-phosphorylated form was identified in mouse, and singly-phosphorylated, singly-deamidated, and deamidated/phosphorylated forms were identified and quantified in human non-failing and failing heart samples, thus demonstrating the clinical utility of the method.

Introduction. Modulation of sarcomeric protein function ensures that the heart ejects blood against a systemic resistance to supply peripheral tissues with oxygen and nutrients and removes carbon dioxide and wastes. Muscle contraction occurs by myosin motors of the thick filament reacting with actin and propelling thin filaments towards the center of the sarcomere. In cardiac muscle, contraction is activated by a release of sarcoplasmic reticular Ca^{2+} that binds to troponin C, which is positioned on the actin thin filament. It was long assumed that regulation of Ca^{2+} levels and cycling were the sole avenue by which contraction was modulated. It is now evident, however, that post-translational modifications (PTMs) are important in regulating the function of ejecting ventricles, as mechanisms downstream of Ca^{2+} fluxes at the level of the sarcomere appear to dominate ejection and sustain ventricular elastance during systole (1). Intra- and inter-molecular interactions of sarcomeric thick and thin filament proteins are modifiable by PTMs, and it has been demonstrated that the intensity and dynamics of contraction and relaxation can be finely tuned via PTMs, particularly those of thin filament proteins (2,3). Mechanisms by which PTMs tune molecular interactions of the thick filament are less well understood. Extrapolating from molecular mechanisms to the cardiac disorder requires an understanding of the endogenous charge state of proteins which can be governed by phosphorylation, acetylation, oxidation and other post-translational modifications.

The thick filament is composed mainly of myosin, a motor protein that interacts with neighboring proteins, including the essential (ELC) and regulatory light chains (RLC), and myosin binding protein-C (MyBP-C) (Figure 1). RLC binds the S1-S2 lever arm of the myosin motor and is optimally positioned at the fulcrum to modulate interactions between the globular MHC head, the coiled-coil LMM thick filament backbone and the neighboring MyBP-C (Figure 1). Cardiac RLC is phosphorylated in large part by myosin light chain kinase (MLCK) (4), which induces an increase in tension at sub-maximally activating levels of Ca^{2+} (5-7) in skinned cardiac fibers bathed in exogenous MLCK. Furthermore, mice expressing a transgenic cardiac RLC housing

alanine residues in place of N-terminal serines have significantly impaired systolic kinetics *in vivo* (8). Separation of endogenous RLC by two-dimensional electrophoresis (2DE) clearly illustrates multiple phosphorylated spots of RLC (8). The most abundant phosphorylated spot of RLC was found to represent phosphorylation at Ser-15 (9). The second, more acidic phosphorylated spot is 3-5 times lower in abundance, and has been difficult to capture using in-gel digestion and mass spectrometry, likely due to protein loss during in-gel digestion combined with poor ionization potential of the phospho-peptide. A possible approach to overcome this limitation is the use of a gel-free (in-solution) technique for separating charged species of RLC, designed to minimize loss and increase sample load.

Biochemically, proteins comprising the thick filament present unique purification challenges, in part, due to their large size (MHC, 223 kDa; MyBP-C, 150 kDa) producing difficulties in the isolation of intact proteins (10-12). In contrast, non-covalently bound ELC and RLC (ie. the light chains of myosin) are more amenable to separation using standard biochemical methods due, in part, to their moderate sizes (22.4 and 18.9 kDa, respectively), however because purification methods for either MHC or ELC/RLC are not time- and cost- effective, novel strategies for preparing/enriching these proteins in a single step are warranted.

When the research objective is to identify all proteins in a sample, a general, unbiased method is appropriate as two peptide ions with high-quality, comprehensive MS/MS are sufficient to reliably identify a protein. However in targeted proteomic approaches, where the objective is to distinguish functionally relevant charge variants—isoforms or post-translationally modified proteins—in a smaller subset of proteins, cleaving proteins into constituent peptides results in a loss of a significant amount of information and produces protein inference problems when assigning identities to modified versions of the same protein (13). A reliable and advantageous strategy to combat the issue of protein inference would be to discriminate charged variants at

the intact protein level. Reverse-phase high performance liquid chromatography (RP-HPLC) and OFFGEL electrophoresis (OGE) are two solution-based separation methods that exploit hydrophobicity and isoelectric point, respectively. Both techniques have been optimized for discriminatory separation at the peptide level (14,15), but in the past have been under-utilized in proteomic workflows for separation of intact proteins.

We report here a rapid solution-based method for purifying microgram quantities of endogenous sarcomeric proteins, allowing for the enrichment and identification of the low-abundance phospho-species of cardiac RLC. The approach employs a tandem OGE/ HPLC workflow that discriminately separates RLC at the intact protein level in a quantifiable manner. The novelty and strengths of this method are: 1) milligram quantities of sarcomeric sub-proteome are focused with resolution qualitatively similar to that of 2DE employing 100-fold less material, 2) quantifiable separation of post-translationally modified variants is achieved at the whole protein level, 3) high sequence coverage of the protein under study, RLC, is achieved due to the substantial enrichment of proteins, and 4) separation of high molecular weight cardiac thick filament proteins can be achieved while maintaining proteins in solution in their intact forms. This workflow was used to identify a novel doubly-charged phospho-species of mouse RLC phosphorylated at adjacent Ser-14 and Ser-15. Additionally, we successfully employed our method for identifying and quantifying post-translational modifications of RLC occurring in human heart failure, thus demonstrating clinical utility for future studies.

Experimental Procedures.

Animal Subjects. All experimental procedures performed using mice and mouse tissue procurement were approved by the Animal Care and Use Committee at the University of Illinois at Chicago (ACC # 05-240).

Human Subjects. All procedures and informed consent involved in the procurement of human tissue were approved by the Institutional Review Board (COMIRB # 00-242) at the University of Colorado and all clinical materials were de-identified according to HIPPA policy. Human tissue samples used were either from explanted hearts with dilated, non-ischemic cardiomyopathy (n=2) or control hearts (n=2) considered for cardiac transplantation. Samples were from the University of Colorado cardiology biobank, which includes a large collection of explanted diseased hearts as well as control heart specimens considered for transplantation. Tissue was collected not as part of a formal study design but rather to develop the biobank, which could be made available to investigators with an interest in the biology of human heart failure. Table I provides clinical details of the four samples used in the current study. A similar procurement strategy was employed for all specimens: at the time of explant, and following cardioplegia, a small portion of the left ventricle (<0.5 gm) was excised and promptly placed in liquid N_2 in the operating room and then subsequently stored at -70^0 for less than six months. Samples were shipped on dry ice from Denver to Chicago, and placed immediately at -70º C until the time of analysis.

Myofilament Enrichment and Sample Preparation. Sub-cellular fractionation proceeded according to methods previously described (16) with modifications. Hearts from 4 month-FVB male and female mice, or from human biopsies, were excised and immediately placed on ice in (mM): 75 KCl, 10 Imidazole (pH 7.2), 2 $MgCl_2$, 2 EDTA, 1 NaN_3, with 1% Triton X-100, phosphatase inhibitor cocktail I (Calbiochem, San Diego, CA), and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO). Tissue was homogenized, centrifuged (18,000 x g, 10 min, 4ºC), and the supernatant fraction removed. The pellets following two extractions were washed in the absence of Triton X-100 and then extracted in IEF buffer (8 M urea, 2 M thiourea, 4% CHAPS) and clarified by centrifuging (18,000 x g for 5 min).

OFFGEL Electrophoresis. A myofilament-enriched fraction was diluted in OFFGEL electrophoresis buffer (7 M urea, 2 M thiourea, 6% glycerol, 65 mM DTT, 1% ampholytes 4-7), and the sample was partitioned among 24-1 cm wells superimposed on a pH 4-7, 24 cm OFFGEL strip (Agilent Technologies, Santa Clara, CA). Proteins were separated using the following conditions: 8,000 V maximum voltage, 50 µA maximum current, 200 mW maximum power for 64,000 V*h. Individual fractions were aspirated and stored at -80ºC for no longer than 1 week.

Reverse-Phase HPLC Separation and Digestion. OFFGEL fractions of interest were acidified with TFA to pH 2, filtered with 0.45 µm polyvinylidene fluoride (PVDF) spin filters (Millipore, Billerica, MA), and separated using C_4-reverse phase chromatography (4.6 x 150 mm, 300 Å pore size, Grace Vydac) at a flow rate of 1.0 ml/min, 37ºC on a U3000 analytical HPLC instrument (Dionex, Bannockburn, IL). Proteins were eluted from 100% buffer A (95% H_2O/5% ACN, 0.1% TFA) to 90% buffer B (95% ACN/5% H_2O, 0.1% TFA) with the following protocol [time post-injection /buffer B]: 0 min/0%, 5min/0%, 20min/20%, 80min/65%, 85min/90%, 90min/90%, 91min/0%, 101min/0%. Fractions were collected every 1 min (1 ml), speed vacuumed to reduce volume, re-solubilized in 100 mM NH_4HCO_3 (pH 8.4)/ 1 mM DTT and digested with Lys-C (1:50 enzyme:protein, Sigma Aldrich) for 16 h at 37ºC. Digests were acidified, lyophilized and re-suspended in 3% ACN/0.1% formic acid with vigorous vortexing.

LC/MS/MS of Lys-C Peptides for Identification of Endogenous PTMs. LC/MS/MS of peptides was performed using a Q-TOF 6510 (Agilent Technologies) equipped with a nanoLC-chip cube. The HPLC consisted of a nanoflow analytical and a capillary loading pump (Agilent 1200 series). HPLC grade water and ACN used for HPLC mobile phases were obtained from Burdick and Jackson (Morristown, New Jersey). Formic acid was obtained from Sigma Aldrich Chemicals (St. Louis, Missouri). Parameters for the analytical (nano) pump were as follows:

buffer A was comprised of HPLC grade water and 0.1% formic acid, buffer B was 90% ACN, 10% HPLC water, and 0.1% formic acid. Flow rate for the nano pump was 550 nl/min and the elution gradient was as follows: 3% buffer B at 0 min, 3% buffer B at 1 min, 40% buffer B at 12 min, 80% buffer B at 13 min, 80% buffer B at 16 min, 3% buffer B at 16.1 min, and 3 min post-time. Parameters for loading the capillary pump were as follows: 3% ACN, 97% HPLC grade water and 0.1% formic acid. The loading pump was operated in isocratic mode. The QTOF instrument was externally calibrated. Peptides were enriched and separated via nano-LC (0.075 x 43 mm, packed with Zorbax 300SB-C_{18}, 5 µm material, 300 Å pore size) integrated in the HPLC Chip (G4240-62001). Full scan was acquired over a range of m/z 290-3000 at 8 spectra per sec, and MS/MS over m/z 100-2400 at 3 spectra per sec with a maximum of 5 precursors per cycle. Fragmentation energy was applied at a slope of 3.0 V/100 Da with a 2.0 offset. Spectra were deconvoluted and analyzed using Spectrum Mill, version A.03.03.080 SR1 (Agilent Technologies). Peak picking was performed with the following parameters: signal-to-noise was set at 25:1, a maximum charge state of 4 was allowed ($z \leq 4$), and the program was directed to attempt to find a precursor charge state. Spectra were searched with the following search criteria: [Database: International Protein Index (IPI) mouse version 3.32 (49,427 total proteins) or IPI human version 3.68 (87,061 total proteins) (17); Digest: Lys-C; Maximum # missed cleavages: 2; Protein pI: from 3.0 to 10.0; Modifications Variable: Oxidation of M, Phosphorylation of STY, Deamidation of NQ; Instrument: Agilent ESI Q-TOF; Precursor mass tolerance: 20 ppm; Product mass tolerance: 20 ppm; Maximum ambiguous precursor charge: 5]. All other criteria were set to software default settings. MS/MS spectra were first auto-validated with Spectrum Mill software, followed by manual inspection for confirmation of PTM identification.

Targeted Analysis and Quantification of RLC Variants in DCM vs. Control by LC/MS/MS. LC/MS/MS of peptides was performed using a QQQ 6410 (Agilent Technologies) equipped with

an electrospray source. The HPLC consisted of a binary analytical and a quaternary loading pump (Agilent 1200 series). Flow rate for the binary pump was 200 uL/min and the gradient was as follows: 20% Buffer B from 0-1 minutes, 20-40% Buffer B from 1-12 minutes. The gradient and flow rates for the quaternary pump were as follows: 3% Buffer B at 0.2mls/min for 1 min, 100% buffer B at 1ml/min from 1.01min to 12 minutes. The column switching valve was switched from enrichment to analysis at 1.0 minutes, and then back to enrichment at 9 minutes. Peptides were trapped on a Zorbax 17 x 1.0mm 5 µm SB-C_{18} enrichment column and then separated using a Zorbax 150 x 1.0mm 3.5µm 300SB-C_{18} 300 Å pore size analytical column. The Electrospray ion source parameters were 325^{o}C for the Gas temp, 10L/min for the gas flow, 30 psi for the nebulizer, 4000V for the capillary voltage and 175V for the fragmentor voltage. The collision energies for each transition were obtained from the QTOF data. The following transitions, which correspond to the resulting b_9 product ion, were monitored for the peptide (K)RAGGA**NS**NVFSMFEQTQIQEFK(E) and used to distinguish isoforms: Unmodified ($^{14/16}$NS^{15}) 830.4 → 827.4, Deamidated ($^{14/16}$nS^{15}) 830.70→ 828.4, Phosphorylated ($^{14/16}$Ns^{15}) 857.1 → 907.4, and Deamidated and Phosphorylated ($^{14/16}$ns^{15}) 857.4 → 908.4. Mass Hunter software was used to integrate peaks and determine area. Although isotopic overlap resulted in apparent chromatographic peak overlap between deamidated and non-deamidated isoforms, adequate chromatographic separation of the isoforms allowed for unadulterated integration of individual transition peaks. Individual fractions were analyzed in triplicate and the area was determined for each transition. Limited sample resulted in low signal-to-noise and poor reproducibility for 4 transitions from a total of 18 experiments (total of 72 transitions) and these values were discarded. The total area for each transition was combined between fractions for simplicity. Three experiments were conducted and the average and standard deviation were calculated. T-tests were used to determine if significant differences existed between isoforms from control and heart disease samples.

2DE. Mouse myofilament sub-proteome was prepared according to sub-fractionation methods described above. Proteins were solubilized in IEF buffer, clarified with centrifugation (18,000 x g, 5 min, 25ºC), and labeled with Cy3. Proteins were separated over a non-linear, broad pI range 3-11 on 18 cm IPG strips according to methods previously described (8). Strips were overlaid on 12% SDS-PAGE second dimension gels (18) for separation by molecular weight. Gels were rinsed in water, and imaged using a Typhoon 9410 (GE Healthcare).

Results.

Preparation of Sarcomeric Sub-Proteome. Figure 2 shows the myofilament-enrichment using the Triton X-100 extraction method, originally described by Solaro et al. (16), resulting in myofibrils that are essentially free of sarcolemmal, sarcoplasmic reticular and mitochondrial contaminants. This method was adapted to expedite fractionation by omitting the initial series of washes in 0.3 M sucrose/10 mM Imidazole (pH 7.0) and 60 mM KCl/30 mM Imidazole (pH 7.0) (16), and homogenizing heart tissue directly in physiologic salt solution containing 1% Triton X-100 with high concentrations of protease and phosphatase inhibitor cocktails using glass-on-glass homogenizers on ice. This rapid protocol resulted in a clean and reproducible enriched sarcomeric sub-proteome which was solubilized in 8 M urea, 2 M thiourea, 4% CHAPS to minimize further alterations in PTMs or degradation.

In-Solution Separation of Intact Cardiac Sarcomeric Proteins with OGE. The resolving capability of OGE at the intact protein level is illustrated in Figure 3. Separation of 3.5 mg sarcomeric sub-proteome was achieved over a narrow pH range of 4-7 and 24 cm focusing distance. Shown in 3A (left panel) is a 1D-SDS-PAGE gel with 2 µl from each OGE fraction loaded illustrating composite proteins of each fraction. OGE wells contained a volume of 100-200 µl buffer at cessation of focusing, depending on the protein concentration per well. Thus, Coomassie-stained bands shown in Figure 3 represent 1/50-1/100 of the actual protein content

harvested per OGE well. To compare the focusing pattern of OGE with traditional 2DE, 40 µg sarcomeric sub-proteome was separated using pH 3-11NL, 18 cm IPG strip, and the area containing pH 4-7 was cropped and run on a 12% second dimension SDS-PAGE gel (Figure 3A, right panel). Comparison of the OGE fractions with 2DE demonstrated similar protein profiles, despite the variation in focusing distance. Importantly, approximately 100-fold more protein was resolved using OGE with qualitatively similar resolution as 2DE. Using this method intact charge variants of thick filament proteins MHC, MyBP-C, ELC and RLC were successfully separated in-solution. To our knowledge this is the first in-solution separation of endogenous charged variants demonstrated for full-length intact MHC and My-BPC (Figure 3B). The reproducibility of OGE separation of sarcomeric proteins is illustrated in the data supplement (Figure S1). Total myofibrillar extract from 1 mouse heart was divided into two samples (approximately 2 mg each) and separated in successive OGE runs to obtain technical replicates. SDS-PAGE profiles of the fractions illustrate significant reproducibility of OGE separation (Figure S1), demonstrating the applicability of this technique to studies undergoing comparative analysis.

RP-HPLC Separation Purifies Co-Focusing Sarcomeric Proteins in OGE. To purify RLC charged species in each OGE fraction from co-migrating proteins, OGE fractions of interest were sequentially separated using RP-HPLC. Figure 4 shows overlaid chromatograms of OGE fractions 5-10. This method separated sarcomeric proteins with high resolution. Because the goal was to determine site-specific phosphorylation of RLC, to further purify RLC from contaminating proteins in OGE fractions, 1 min (1 ml) fractions from OGE 5-10 samples were collected and subjected to subsequent analyses. It was evident from Western blotting of each collected fraction for RLC (data not shown) and corresponding peak overlays that there were three isoelectric variants of RLC, denoted RLC-5, -6, and -7 (Figure 4), at the intact protein level in OGE fractions 5-7, respectively, consistent with previous 2DE. From the integration of peaks

detected via UV (214 nm), we quantified RLC species as RLC-5 [Area=75.2 mAU*min], RLC-6 [Area=446.3 mAU*min] and RLC-7 [Area=1943.5 mAU*min], thus RLC exists endogenously in cardiac mouse myofibrils as three populations (3.1% RLC-5, 18.1% RLC-6 and 78.8% RLC-7). A significant strength of this approach was in the use of UV as a quantitative measure of endogenous charged species of RLC, as UV at 214 nm offers unbiased detection of peptide bonds and is linear over several orders of magnitude.

LC/MS/MS of RLC Charged Species Reveals Two Novel Findings: Phosphorylation at Serine-14 and the Presence of a Doubly-Phosphorylated Species. To identify site-specific differences in RLC charged variants separated by OFFGEL and purified by RP-HPLC, proteins were digested with Lys-C and subjected to nano-LC/MS/MS analysis on a QTOF mass spectrometer. Peptides were eluted over a 16 min linear gradient of ACN, and scanned in a data-dependent manner with the five most abundant peaks of each scan being subjected to fragmentation. Spectra were extracted and searched against the mouse IPI database (17) using Spectrum Mill (Agilent Technologies). Confident protein hits with two or more distinct peptides detected, spectral intensities and amino acid sequence coverage from offline LC fractions RLC-5, -6 and -7 were aligned for comparison and are listed in Table II. RLC showed a striking 95% sequence coverage. The region of RLC not covered was the N-terminal MAPKKAK sequence due to the high lysine content. This high coverage demonstrates the power of using pre-fractionation and/or enrichment techniques in cases where detecting the majority of peptides is critical, as during PTM or isoform identification.

Figure 5 shows extracted ion chromatograms performed in MassHunter Qualitative Analysis software (Agilent Technologies) of the triply-charged N-terminal 9-30 peptide RIEGGSSNVFSMFEQTQIQEFK, showing its abundance as unphosphorylated, singly- and doubly-phosphorylated, corresponding to fractions RLC-7, -6 and -5. Figure 5A, left panel,

demonstrates that the majority of unphosphorylated m/z 854.7530 was present in RLC-7, with a small amount (5%) detected in RLC-6, suggesting minor amounts of overlap between fractions 6 and 7 during OGE focusing. The corresponding MS/MS spectra to demonstrate the identity of the unphosphorylated peptide as confirmed by MH+ 2562.2396 matching within 6 ppm is shown in Figure 5B. Figure 5A, middle panel, demonstrates that essentially all of the singly-phosphorylated peptide m/z 881.4070 was in RLC-6, and the corresponding MS/MS spectra in Figure 5C identified a singly-phosphorylated peptide MH+ 2642.2115 (mass error of 4.2 ppm) with the phosphorylation residing on Ser-15 or Ser-14. Our previous analysis of this singly-phosphorylated species with LTQ-FT-ICR showed that when a successful Ser-Ser break was achieved with either CID or ECD, Ser-14 and Ser-15 were alternately phosphorylated, roughly 20% and 80% of the time, respectively (unpublished observations, data not shown). This singly-phosphorylated charged species of RLC migrating in fraction 6 of OGE, thus likely represents a mixture of RLC-Ser-14 and RLC-Ser-15 phosphorylated molecules. The majority of doubly-phosphorylated RLC m/z 908.0616 shown in Figure 5A, right panel, was in RLC-5, with some overlap evident in RLC-6. MS/MS spectra of MH+ 2722.1702 (mass error of 3.0 ppm) in Figure 5D unequivocally demonstrated that these two phosphorylations were on residues Ser-14 and Ser-15. This novel finding is the first report demonstrating a doubly-phosphorylated endogenous species of cardiac RLC. We have unequivocally identified three distinct charge species of RLC with high mass accuracy of both precursor and fragment ions inherent to the Q-TOF (see low mass errors in Figure 5). Importantly, RP-HPLC chromatograms of intact RLC protein detected with UV shown in Figure 4 demonstrate doubly-phosphorylated RLC as 3.1%, the singly-phosphorylated as 18.1%, and the un-phosphorylated as 78.8% of total sarcomeric RLC. It is also worthy to note that MS/MS spectra of un-, singly-, and doubly-phosphorylated peptides shown in Figure 5 have scored peak intensities (SPI %) between 92-96% and Spectrum Mill peptide scores between 19-21. The SPI % represents the percentage of the MS/MS peak-detected spectral ion current explained by the search interpretation, demonstrating the

remarkably high confidence and fidelity of our MS/MS spectra. The peptide score reflects the information content in the MS/MS spectrum; a score of 13-25 is considered of excellent quality representing extensive peptide fragmentation, and when combined with SPI % of 70 or greater, the result is of high confidence and validity.

In-Solution Workflow Applied to Clinical Samples Reveals Significant Changes in RLC PTMs in Human Heart Failure*.* To examine whether differential modifications are present in human heart failure, non-failing and failing dilated cardiomyopathy (DCM) samples were analyzed using the OGE-HPLC-MS/MS workflow. Human samples used were either "control" left ventricular tissue from donor hearts that were considered for heart transplantation or hearts from patients with dilated, non-ischemic cardiomyopathy explanted at the time of heart transplantation. Interestingly, human RLC was identified in OGE fractions 7, 8 and 9, whereas mouse RLC was identified in OGE fractions 5, 6 and 7. *In silico* alignment of human (Swiss Prot # P10916) versus mouse (Swiss Prot # P51667) RLC demonstrated slight variations in the primary sequence of RLC, with an overall isoelectric point difference (excluding all variance due to PTMs) of 4.92 (human) to 4.86 (mouse). RLC in each fraction was subjected to MS/MS analysis on the Q-TOF (search results shown in Table III), and three arrangements of modifications on the 9-30, N-terminal peptide of RLC were detected, including deamidation, phosphorylation, and simultaneous deamidation/phosphorylation (Figure 6). Interestingly, the identified deamidation was the conversion of asparagine to aspartate at position 14 or 16, immediately adjacent of the solely identified phosphorylation site, serine-15. Note the sequence difference between the mouse (Figure 5) 9-30 and human (Figure 6) 9-30 peptide. The position 14/15 serine/serine double phosphorylation in mouse is a position 14/15 or 15/16 deamidation/phosphorylation or phosphorylation/deamidation, respectively, in human, thus the net charge on the N-terminal region of RLC in each species remains similar, albeit through different chemistries. No doubly-phosphorylated species was detected on other serine/threonine

residues in the 9-30 peptide of human RLC, as in mouse. This is the first site-specific identification of post-translational modifications in human RLC. To quantify differences in RLC PTMs in non-failing versus failing ventricular tissue, targeted multiple reaction monitoring (MRM) was employed. Figure 7 illustrates the abundance of each RLC 9-30 species, as quantified from the b9 ion transitions, and normalized to total RLC protein amount determined via UV in the offline HPLC separation of intact RLC. Failing, DCM samples demonstrated a significant decrease in the singly-phosphorylated species (Control 11411.2 +/- 2230.3 vs. DCM 5051.0 +/- 1703.5) and the combined deamidated/phosphorylated species (Control 1456.8 +/- 265.7 vs. 869.4 +/- 270.5) relative to non-failing controls, suggesting that the negative charge in the N-terminal region of RLC is reduced in human heart failure.

Discussion.

Effectiveness of a Rapid, In-Solution Method to Purify Endogenous Cardiac Proteins. While 2DE is a powerful method for protein separation and resolution, when in preparation for analysis by mass spectrometry, in-gel approaches can be less than ideal. An inherent attribute of in-gel approaches is that proteins are encased inside a matrix and must be efficiently eluted for further analysis. Although protein digestion into smaller peptides should in theory cleave proteins into easily diffusible pieces, the reality is that losses incurred by this approach due to unavoidable incomplete digestion are still greater than if kept in solution. Additional limitations of 2DE include difficulties encountered in analysis of membrane proteins; extremely basic proteins which are resistant to focusing; and high molecular weight proteins that will not migrate into the first dimension IPG strip or from the IPG strip into the second dimension gel. We have taken advantage of a relatively new methodology, OGE (19), and a classic biochemical method, RP-HPLC, and have applied these to cardiac proteins of the sarcomeric sub-proteome. With the era of “shotgun” proteomics steering the development of separation technologies, RP-HPLC of intact proteins has been widely underutilized despite the fact that analyzing proteins in their

intact form seems an attractive strategy. Nevertheless, there have been a few cases where RP-HPLC has been utilized at the protein level as the second dimension in tandem for 2D protein separation methods, following ion-exchange chromatography, chromatofocusing, electroelution, or size-exclusion chromatography (20-24), though these tandem technologies have not been routinely applied to cardiac muscle proteins. However separate from 2D-LC applications, RP-HPLC has been used alone to separate sarcomeric proteins from failing versus non-failing swine hearts by Neverova and van Eyk (25). Following buffer optimization, they were able to enhance separation of myofilaments proteins and identify alterations in MHC in heart failure. Our workflow therefore expanded on this work by incorporating OGE as a reproducible method of sample simplification prior to RP-HPLC, thus creating a 2-D, in-solution separation of sarcomeric sub-proteome which partitioned enriched pools of post-translationally modified protein. For our molecule of interest, cardiac RLC, this approach separated RLC charge variants into OFFGEL fractions with minimal overlap, thus providing an effective preparatory step for downstream PTM identification with mass spectrometry. Prior to using this approach, we had used 2DE to separate RLC charge variants followed by in-gel digestion and mass spec. While this approach was successful in identifying the highly abundant, singly-phosphorylated RLC species in rodent cardiac muscle (9), a doubly-phosphorylated peptide was undetectable with the mass spec. The current approach has proven to be more effective and will be of value for PTM identification of other sarcomeric proteins in future studies.

Advantages of OGE as a Preparatory Purification Step for High Molecular Weight Sarcomeric Proteins. Myosin and MyBP-C are high molecular weight proteins of the thick filament in which the identification of isoforms and post-translational modifications remains ambiguous. Muscle MHC exists as a heterogenous population (12) expressed as two unique isoforms, α and β (30), with post-translational modifications (31-35). MHC generates 150 tryptic peptides when unmodified and fully-cleaved, therefore the endogenous number of MHC charge

variants induced by isoforms and PTMs may be greater than 10^4. Thus, identifying and quantifying endogenous charge variants of MHC within a non-simplified cell lysate becomes difficult due to ion suppression effects and incomplete sequence coverage common to complex sample analysis, combined with ambiguities in protein identifications couched in the problem of protein inference (13). Previous in-solution methods at separating full-length MHC have been explored using hydroxyapatite HPLC (12) and hydrophobic interaction HPLC (36), however these separation chemistries have demonstrated limited and incomplete resolution of individual charged variants of MHC, and have been successful only with purified myosin sample following a lengthy, multiple-step biochemical purification. Furthermore, MHC is not amenable to separation via RP-HPLC due to irreversible binding to free silanol groups or strong hydrophobic interactions between the column packing material and myosin protein (37). MyBP-C, a second large molecular weight protein, is digested into 86 fully cleaved/unmodified tryptic peptides. Previous 2D in-gel analysis demonstrated MyBP-C to be highly modified *in vivo*, with 10 charged variants being evident in mouse (8), suggesting that MyBP-C, like MHC, possesses numerous endogenous charge variants. At least three sites of MyBP-C phosphorylation have been identified (38-40), however these cannot account for all MyBP-C charge variants observed, therefore further analysis is required. To our knowledge, no one has demonstrated an in-solution separation of full-length cardiac MyBP-C into unique charge variants. The techniques described in our study should be applicable to the separation and analysis of full-length MyBP-C and MHC.

Importance of Identifying PTMs in RLC. The importance of identifying precise sites of RLC modifications, as has been accomplished in this study, leads to a more complete understanding of intra- and inter-molecular interactions that occur within the myosin motor and contribute to disease mechanisms. It has been noted from familial hypertrophic cardiomyopathic (FHC) mutations occurring in all thick filament proteins—MHC, MyBP-C, ELC and RLC—that a change

in one amino acid side chain can have an enormous effect on cardiac morphology and function. The exchange of a Ser- or Thr-hydroxyl with a bulky, charged 80 dalton phospho-ester is chemically more substantial than the physical or chemical changes induced by a number of FHC mutations. In the case of RLC, several FHC mutations have been documented which cause mid-left ventricular obstruction and papillary muscle hypertrophy, and not coincidentally, these residues tend to cluster around the Ca^{2+} binding domain and the N-terminal phosphorylatable region (26) (Figure 8). The importance of identifying sites of endogenous phosphorylation is underscored by the observation that RLC Ca^{2+} binding, phosphorylation potential and secondary structure were significantly altered in FHC mutants A13T, F18L, E22K, R58Q, P95A, and N47Q (26). Interestingly, phosphorylation of wild-type RLC by MLCK significantly decreased the helical content of the molecule, suggesting that RLC undergoes a substantial conformational change induced by phosphorylation (26). In addition, neighboring thick filament proteins including MyBP-C and titin are phosphorylated (27,28), and there is evidence to suggest that the phosphorylation of MyBP-C governs its interaction with the S2 hinge region of MHC (29), the precise location of RLC binding. In this context, our laboratory has observed de-phosphorylation of MyBP-C in TG-RLC(P-), suggesting that the phosphorylation of MyBP-C and RLC may be functionally linked (8). Thus, it is clear that the definition of sites of endogenous modifications in RLC (as well as other cardiac thick filament proteins) will lead to a molecular description of cardiac muscle function.

RLC Modifications in Human Heart Failure. Previous studies demonstrated a decrease in RLC phosphorylation in human end-stage heart failure (41-43). Our data show a similar trend in phosphorylation, with the novel addition of alterations in deamidation in the N-terminal region of RLC in human heart failure. White et al. previously identified deamidation and phosphorylation of N-terminal RLC in rabbit cardiac muscle (44), however our study is the first to map these modifications in human RLC, and to quantify the differences in human heart failure. The

functional significance of the observed deamidation remains obscure, as this can be either a biologically relevant signal or a chemical artifact. Deamidation is a relatively common modification in human proteins, and has been generally thought of as a negative functional change via signaling proteins for degradation or replacement (45). It is surprising, therefore, that in our study the non-failing tissue shows significantly greater deamidation than the DCM samples. It is possible that cleavage of RLC at the residue 14/15 junction is occurring in DCM samples, as has been observed in human DCM (46) and rabbit ischemia-reperfusion (44), thus lowering the effective concentration of 9-30 peptide present in these samples. How the differences in each variant of RLC may contribute to the contractile dysfunction observed in heart failure is uncertain. It is unlikely that the observed decreases in RLC phosphorylation and deamidation are *causative* for heart failure, but rather the important finding reported here is that RLC is *modified* (along with other proteins) in a diseased state suggesting that this change may functionally contribute within the context of the pathogenesis of heart failure.

RLC in Therapeutic Interventions for Cardiac Failure. The myosin motor is a prime target for therapeutic intervention in clinical heart failure. Small molecules which specifically target myosin and produce a positive functional effect, termed “sarcomere activators,” are now in clinical trials (47-49). A sarcomere activator that increases the intensity of cardiac contraction per unit activating Ca^{2+} bound to troponin C, without altering the Ca^{2+} transient would be energetically favorable for the heart. The myosin light chains (ELC and RLC) bound between the active site and the lever arm fulcrum of myosin heavy chain (MHC) are ideally located to regulate enzymatic activity of the myosin head and are a plausible site of action for sarcomere activators. Endogenous MHC-ELC-RLC conformations are governed by local charge changes induced by phosphorylation (50), which subsequently affect function. Indeed, it has been demonstrated that in isolated skinned cardiac fibers treated with myosin light chain kinase (MLCK), there was an increase in MgATPase activity (51) and tension per unit Ca^{2+} (5-7). Furthermore, our laboratory

demonstrated that in mice expressing RLC with serines 14/15/19 mutated to alanine (TG-RLC(P-)), myofibrillar ATPase was significantly decreased (8). Okafor et al. examined the effect of MCI-154, a myofilament Ca^{2+}-sensitizer with unknown mechanism, on skinned fibers from non-failing, idiopathic or ischemic cardiomyopathic patients, and reported significant increases in myosin-specific ATPase activity, independent of the thin filament regulatory proteins troponin and tropomyosin (52). Additionally, our laboratory demonstrated that TG-RLC(P-) mice had altered *in vivo* systolic kinetics, measured by an increase in ejection time (8), and coincidentally the sarcomere activator CK-1827452, now in clinical trials, increases left ventricular ejection time and stroke volume in healthy patients and those with heart failure (47-49). Though the detailed myosin-specific mechanism of action of both MCI-154 and CK-1827452 are unknown, it is highly likely that alterations in charge near the binding region of RLC, which is modulated by phosphorylation, plays a role.

Concluding Remarks. We have demonstrated a novel application of current state-of-the-art in-solution separation technologies to cardiac sarcomeric proteins, allowing the novel identification and quantification of charge variants of cardiac RLC: a singly- and doubly-phosphorylated species in mouse, and a singly-phosphorylated, singly-deamidated, and deamidated-phosphorylated species in human. While it is understood that the phosphorylation of RLC is critical for cardiac function (8), future studies must address the significance of the multiple additional charge variants present in human cardiac RLC.

Acknowledgements. The authors thank Dr. Lori A. Walker from the University of Colorado at Denver for her helpful discussions, and Dr. Peipei Ping from the University of California at Los Angeles for her generous support of the project. The authors acknowledge the Proteomics and Mass Spectrometry Facility at National Jewish Health and the UIC Chicago Biomedical Consortium and Research Resources Center for their assistance. The CBC-RRC was

established by a grant from The Searle Funds at the Chicago Community Trust to the Chicago Biomedical Consortium.

References.

1. Hinken, A. C., and Solaro, R. J. (2007) *Physiology (Bethesda)* **22**, 73-80
2. Solaro, R. J. (2001) *Heart Physiology and Pathophysiology*, 4th Ed., Academic Press.
3. Solaro, R. J., and Rarick, H. M. (1998) *Circ Res* **83**, 471-480
4. Chan, J. Y., Takeda, M., Briggs, L. E., Graham, M. L., Lu, J. T., Horikoshi, N., Weinberg, E. O., Aoki, H., Sato, N., Chien, K. R., and Kasahara, H. (2008) *Circ Res* **102**(5), 571-580
5. Sweeney, H. L., and Stull, J. T. (1986) *Am J Physiol Cell Physiol* **250**, C657-C660
6. Olsson, M. C., Patel, J. R., Fitzsimons, D. P., Walker, J. W., and Moss, R. L. (2004) *Am J Physiol Heart Circ Physiol* **287**(6), H2712-H2718
7. Morano, I., Hofmann, F., Zimmer, M., and Ruegg, J. C. (1985) *FEBS Lett* **189**(2), 221-224
8. Scruggs, S. B., Hinken, A. C., Thawornkaiwong, A., Robbins, J., Walker, L. A., de Tombe, P. P., Geenen, D. L., Buttrick, P. M., and Solaro, R. J. (2009) *J Biol Chem* **284**(8), 5097-5106
9. Yuan, C., Sheng, Q., Tang, H., Li, Y., Zeng, R., and Solaro, R. J. (2008) *Am J Physiol Heart Circ Physiol* **295**(2), H647-656
10. Warren, C. M., Krzesinski, P. R., and Greaser, M. L. (2003) *Electrophoresis* **24**(11), 1695-1702
11. Libera, L. D. (1988) *Cell Biol Int Rep* **12**(12), 1089-1098
12. Kawasaki, T., Takahashi, S., and Ikeda, K. (1985) *Eur J Biochem* **152**(2), 361-371
13. Nesvizhskii, A. I., and Aebersold, R. (2005) *Mol Cell Proteomics* **4**(10), 1419-1440

14. Horth, P., Miller, C. A., Preckel, T., and Wenz, C. (2006) *Mol Cell Proteomics* **5**(10), 1968-1974

15. Ishihama, Y. (2005) *J Chromatogr A* **1067**(1-2), 73-83

16. Solaro, R. J., Pang, D. C., and Briggs, F. N. (1971) *Biochim Biophys Acta* **245**(1), 259-262

17. Kersey, P. J., Duarte, J., Williams, A., Karavidopoulou, Y., Birney, E., and Apweiler, R. (2004) *Proteomics* **4**(7), 1985-1988

18. Fritz, J. D., Swartz, D. R., and Greaser, M. L. (1989) *Anal Biochem* **180**(2), 205-210

19. Ros, A., Faupel, M., Mees, H., Oostrum, J., Ferrigno, R., Reymond, F., Michel, P., Rossier, J. S., and Girault, H. H. (2002) *Proteomics* **2**(2), 151-156

20. Wagner, K., Racaityte, K., Unger, K. K., Miliotis, T., Edholm, L. E., Bischoff, R., and Marko-Varga, G. (2000) *J Chromatogr A* **893**(2), 293-305

21. Isobe, T., Uchida, K., Taoka, M., Shinkai, F., Manabe, T., and Okuyama, T. (1991) *J Chromatogr* **588**(1-2), 115-123

22. Chong, B. E., Yan, F., Lubman, D. M., and Miller, F. R. (2001) *Rapid Commun Mass Spectrom* **15**(4), 291-296

23. Rose, D. J., and Opiteck, G. J. (1994) *Anal Chem* **66**(15), 2529-2536

24. Opiteck, G. J., Ramirez, S. M., Jorgenson, J. W., and Moseley, M. A., 3rd. (1998) *Anal Biochem* **258**(2), 349-361

25. Neverova, I., and Van Eyk, J. E. (2002) *Proteomics* **2**(1), 22-31

26. Szczesna, D., Ghosh, D., Li, Q., Gomes, A. V., Guzman, G., Arana, C., Zhi, G., Stull, J. T., and Potter, J. D. (2001) *J Biol Chem* **276**(10), 7086-7092

27. Yamasaki, R., Wu, Y., McNabb, M., Greaser, M., Labeit, S., and Granzier, H. (2002) *Circ Res* **90**(11), 1181-1188

28. Hartzell, H. C., and Glass, D. B. (1984) *J Biol Chem* **259**(24), 15587-15596

29. Gruen, M., Prinz, H., and Gautel, M. (1999) *FEBS Lett* **453**(3), 254-259

30. Palmer, B. M. (2005) *Heart Fail Rev* **10**(3), 187-197

31. Schilder, R. J., and Marden, J. H. (2007) *J Exp Biol* **210**(Pt 24), 4298-4306

32. Hedou, J., Cieniewski-Bernard, C., Leroy, Y., Michalski, J. C., Mounier, Y., and Bastide, B. (2007) *J Biol Chem* **282**(14), 10360-10369

33. Ramamurthy, B., Hook, P., and Larsson, L. (1999) *Acta Physiol Scand* **167**(4), 327-329

34. Cieniewski-Bernard, C., Bastide, B., Lefebvre, T., Lemoine, J., Mounier, Y., and Michalski, J. C. (2004) *Mol Cell Proteomics* **3**(6), 577-585

35. Ramamurthy, B., Hook, P., Jones, A. D., and Larsson, L. (2001) *Faseb J* **15**(13), 2415-2422

36. Malmqvist, U. P., Aronshtam, A., and Lowey, S. (2004) *Biochemistry* **43**(47), 15058-15065

37. Libera, L. D. (2001) *Basic Appl Myol* **11**(3), 115-118

38. Yuan, C., Guo, Y., Ravi, R., Przyklenk, K., Shilkofski, N., Diez, R., Cole, R. N., and Murphy, A. M. (2006) *Proteomics* **6**(14), 4176-4186

39. Mohamed, A. S., Dignam, J. D., and Schlender, K. K. (1998) *Arch Biochem Biophys* **358**(2), 313-319

40. Schlender, K. K., and Bean, L. J. (1991) *J Biol Chem* **266**(5), 2811-2817

41. van der Velden, J., Papp, Z., Boontje, N. M., Zaremba, R., de Jong, J. W., Janssen, P. M., Hasenfuss, G., and Stienen, G. J. (2003) *Cardiovasc Res* **57**(2), 505-514

42. van Der Velden, J., Klein, L. J., Zaremba, R., Boontje, N. M., Huybregts, M. A., Stooker, W., Eijsman, L., de Jong, J. W., Visser, C. A., Visser, F. C., and Stienen, G. J. (2001) *Circulation* **104**(10), 1140-1146

43. Morano, I. (1992) *Basic Res Cardiol* **87**, 129-141

44. White, M. Y., Cordwell, S. J., McCarron, H. C., Tchen, A. S., Hambly, B. D., and Jeremy, R. W. (2003) *J Mol Cell Cardiol* **35**(7), 833-840

45. Robinson, N. E., and Robinson, A. B. (2001) *Proc Natl Acad Sci U S A* **98**(22), 12409-12413

46. Holt, J. C., Caulfield, J. B., Norton, P., Chantler, P. D., Slayter, H. S., and Margossian, S. S. (1995) *Mol Cell Biochem* **145**(1), 89-96

47. Cleland, J. G., Coletta, A. P., and Clark, A. L. (2006) *Eur J Heart Fail* **8**(7), 764-766

48. Coletta, A. P., Cleland, J. G., Cullington, D., and Clark, A. L. (2008) *Eur J Heart Fail* **10**(9), 917-920

49. Solaro, R. J. (2009) *IDrugs* **12**(4), In Press

50. Levine, R. J. C., Kensler, R. W., Yang, Z., Stull, J. T., and Sweeney, H. L. (1996) *Biophys J* **71**, 898-907

51. Clement, O., Puceat, M., Walsh, M. P., and Vassort, G. (1992) *Biochem J* **285**, 311-317

52. Okafor, C., Liao, R., Perreault-Micale, C., Li, X., Ito, T., Stepanek, A., Doye, A., de Tombe, P., and Gwathmey, J. K. (2003) *Mol Cell Biochem* **245**(1-2), 77-89

Footnotes. This work was supported by the National Institute of Health PO1 HL62426 (RJS, PMB), RO1 HL22231 (RJS), T32 HL07962 (SBS), P01 HL80111 (PPing), and the American Heart Association 0710031Z (SBS).

Figure Legends.

Figure 1: Schematic illustrating cardiac sarcomeric proteins. The myofibrillar lattice is comprised mainly of actin thin filaments and myosin thick filaments, each bound to regulatory proteins. Activation of cardiac contraction during systole proceeds by calcium binding to troponin C (TnC) which induces conformational changes and altered interactions among troponin I (TnI), troponin T (TnT) and tropomyosin (Tm), resulting in the removal of steric inhibition over the myosin binding site on actin. An activated thin filament allows the binding of myosin heads, which then propel the actin filaments toward the center of the sarcomere. The thick filament is composed mainly of myosin heavy chain (MHC) which binds two light chains, essential and regulatory (ELC and RLC, respectively) and associates with myosin binding protein C (MyBP-C) in the hinge and light meromyosin (LMM) regions.

Figure 2: **Protocol for preparation of sarcomeric sub-proteome using Triton X-100**, originally described by from Solaro et al. (16). This method uses the non-ionic detergent, Triton X-100, to solubilize the sarcolemmal, sarcomplasmic reticular (SR), and mitochondrial membranes, allowing the release of soluble proteins into solution and pelleting of the insoluble myofibrillar lattice in a rapid and efficient manner. All fractionation is performed in the presence of phosphatase and protease inhibitors to preserve the endogenous phospho-state and integrity of sarcomeric proteins.

Figure 3: **Comparison of OGE and 2DE focusing patterns.** Cardiac sarcomeric sub-proteome, load 3.5 mg, was separated in solution by OGE over a narrow pI range 4-7, 24 cm distance. **(A)** (Left Panel) Approximately 1/75 the volume (2 µl) of each OGE fraction was run on a 12% SDS-PAGE gel to illustrate constituents of OGE fractions (labeled #1-24, x-axis) and the un-fractionated homogenate (UH). Detection is via Coomassie G250. (Right Panel) 2DE of same sarcomeric sub-proteome with a 40 µg protein load demonstrates comparable resolution

of sarcomeric proteins as OGE, method of detection is Cy3 fluorescence. Importantly, note that 3.5 mg protein separates in solution using OGE with qualitatively similar resolution as 40 µg in-gel. **(B)** OGE fractions probed with anti-MHC and anti-MyBPC antibodies illustrates the first in-solution separation of full-length endogenous high molecular weight thick filament proteins.

Figure 4: **Purification of sarcomeric proteins in OGE fractions using RP-HPLC.** Shown are OGE fractions 5-10 each run separately on a 4.6 x 150 mm C4 reverse-phase column (TP5415, Grace Vydac), detected using UV 214 nm, and the chromatograms overlaid for comparison. Separation of endogenous charge variants of essential light chain (ELC), tropomyosin (Tm), regulatory light chain (RLC) and actin are demonstrated in the retention time range 44-66 min. Quantification of the three unique RLC species demonstrated a 3.1%/18.1%/78.8% distribution of endogenous RLC in fractions RLC-5, -6 and -7, respectively, which were collected for further analysis by LC/MS/MS.

Figure 5. **MS/MS spectra showing identification of phospho-species in murine RLC. (A)** Left Panel—Extracted ion chromatogram (EIC) of unphosphorylated RLC peptide 9-30 (m/z 854.7530) showing the majority in fraction RLC-7 with minor carry-over in RLC-6; (Middle Panel) EIC of singly-phosphorylated RLC peptide 9-30 (m/z 881.4070) showing the presence of this peptide exclusively in RLC-6; (Right Panel) EIC of doubly-phosphorylated RLC peptide 9-30 (m/z 908.0616) of RLC showing the majority of this peptide in fraction RLC-5 with minor carry-over in RLC-6. **(B)** MS/MS spectra (SPI of 96%; peptide score = 20.8) of un-phosphorylated RLC peptide m/z 854.7530 (precursor mass error of 5.8 ppm), with peptide bond cleavage type denoted using the following: blue=b-ion, red=y-ion, and pink=both b- and y-ion formation. **(C)** MS/MS spectra (SPI of 91%; peptide score = 20.8) of singly-phosphorylated RLC peptide m/z 881.4070 (precursor mass error of 4.2 ppm); s = phosphorylated serine. **(D)** MS/MS spectra (SPI of 96%; peptide score = 19.2) of doubly-phosphorylated RLC peptide m/z 908.0616

(precursor mass error 3.0 ppm). Note the mass shifts and mass accuracy of b9 fragments going from B-D, illustrating a 79.966 shift in each as an additional HPO_3 is added to the peptide.

Figure 6. **MS/MS spectra showing identification of phospho- and deamidated-species in human RLC. (A)** MS/MS spectra (SPI of 100%; peptide score = 21) of un-phosphorylated, non-deamidated RLC peptide m/z 830.4007 (precursor mass error of -0.1 ppm), with peptide bond cleavage type denoted using the following: blue=b-ion, red=y-ion, and pink=both b- and y-ion formation. **(B)** MS/MS spectra (SPI of 93%; peptide score = 20.2) of singly-deamidated RLC peptide m/z 830.7321 (precursor mass error of 4.0 ppm); n = deamidated asparagine; parentheses (n) denote ambiguity in the assignment of the deamidation to the asparagine at position 14 or position 16. **(C)** MS/MS spectra (SPI of 81%; peptide score = 18.4) of singly-phosphorylated RLC peptide m/z 857.0548 (precursor mass error -1.7 ppm). s = phosphorylated serine. **(D)** MS/MS spectra (SPI of 89%; peptide score = 19.1) of singly-deamidated and singly-phosphorylated RLC peptide m/z 857.3834 (precursor mass error -1.0 ppm). Note the mass shifts and mass accuracy of b9 fragments going from A-D.

Figure 7. **Quantification of human RLC species in DCM versus control by multiple reaction monitoring reveals differences in phosphorylation and deamidation.** Heart tissue from non-failing and DCM patients was extracted and fractionated using OGE. Fractions (OGE-7, OGE-8, OGE-9) from 2 biological replicates were pooled and RLC species were separated via RP-HPLC as described above. Fractions were digested using Lys-C, and subjected to liquid chromatography tandem mass spectrometry using a triple quadrupole mass spectrometer. The following transitions were monitored for the peptide (K)RAGGA**NS**NVFSMFEQTQIQEFK(E) and used to distinguish isoforms: Unmodified ($^{(14/16)}NS^{15}$) 830.4 → 827.4, singly-deamidated ($^{(14/16)}nS^{15}$) 830.70→ 828.4, singly-phosphorylated ($^{(14/16)}Ns^{15}$) 857.1 → 907.4, and singly-deamidated/singly-phosphorylated ($^{(14/16)}ns^{15}$) 857.4 → 908.4. Each transition corresponds to the

resulting b_9 product ion. Individual fractions were analyzed, the area determined for each transition, and the total area for each transition was combined between fractions for simplicity. Three experiments were conducted and the average and standard deviation were calculated and are shown. Transitions corresponding to the singly-phosphorylated and the phosphorylated/deamidated forms are significantly different between heart disease and control while singly-deamidated and unphosphorylated showed no difference.

Figure 8. **FHC mutations identified in RLC.** Shown is a three-dimensional illustration of RLC structure (turquoise) adapted from Protein Data Bank file *1wdc*. Identified FHC point mutations are shown in red and the Ca^{2+} binding loop in yellow.

Tables.

Patient	Gender	Age	Ethnicity	Type	Subtype	Stage	EFx	Comments
NF 1	F	57	Caucasian	NF	n/a	n/a	65%	Intracranial hemorrhage
NF 2	M	37	Hispanic	NF	n/a	n/a	60%	Intraventricular hemmorrhage
DCM 3	M	20	Caucasian	DCM	non-ischemic	NYHA IV	10%	None
DCM 4	M	47	Hispanic	DCM	non-ischemic	NYHA IV	<20%	None

Table I: Definition of Diseased State of Patients. The human specimens used in this study included cardiac muscle from two non-failing (NF) and two end-stage dilated cardiomyopathic (DCM) patients. The detailed descriptions of each are given. DCM, dilated cardiomyopathy; NF, non-failing; NYHA, New York Heart Association Classification; EFx, ejection fraction; n/a, not applicable.

RLC-5	RLC-6	RLC-7	Database Acc#	%AA Coverage	Distinct Peptides (#)	Distinct Summed MS/MS Score	Protein Name
# spectra	# spectra	# spectra					
mean intensity	mean intensity	mean intensity					
40	69	78					Myl2 Myosin regulatory light chain 2, ventricular/cardiac muscle
2.11E+06	4.09E+06	5.46E+06	IPI00555015	95	24	467.94	
56	30	29					
7.20E+05	3.47E+05	3.99E+05	IPI00123316	71	33	578.58	Tpm1 33 kDa protein
9	13	7					Lamc1 Laminin subunit gamma-1 precursor
8.89E+04	2.10E+05	1.35E+05	IPI00400016	7	10	179.67	
1	2	0					
5.92E+04	1.14E+05	0.00E+00	IPI00227299	8	3	43.46	Vim Vimentin
0	2	3					
0.00E+00	4.88E+04	4.52E+04	IPI00130102	7	3	38.97	Des Desmin
1	2	2					
2.87E+04	8.35E+04	2.26E+05	IPI00346834	2	2	31.94	Similar to Keratin 2p
0	0	2					Ovgp1 Oviduct-specific glycoprotein precursor
0.00E+00	0.00E+00	2.91E+05	IPI00120900	4	2	19.35	
1	0	1					
3.29E+05	0.00E+00	9.70E+05	IPI00120455	4	2	18.53	Snx13 Sorting nexin-13

Table II: Protein-Protein Comparison Analysis of Mouse LC/MS/MS Data. Shown are constituents of HPLC fractions RLC-5, -6, and -7. Proteins detected by 2 or more distinct peptides are shown. RLC exhibited 95% sequence coverage. Database Acc# = IPI database accession number.

NF RLC-7	NF RLC-8	NF RLC-9	DCM RLC-7	DCM RLC-8	DCM RLC-9	Database Acc#	%AA Cov	Dist Pep (#)	Distinct Summed MS/MS score	Protein Name
# spectra	# spectra	# spectra	# spectra	# spectra	# spectra					
mean intensity	Mean intensity	mean intensity	mean intensity	mean intensity	mean intensity					
49	42	44	39	50	44					MYL2 Myosin regulatory light chain 2, ventricular/cardiac muscle form
1.87E+08	2.79E+08	5.21E+08	1.20E+08	6.87E+08	4.28E+08	IPI00216798	89	14	306.85	
30	0	0	42	1	0					TPM1 Isoform 1 of Tropomyosin alpha-1 chain
4.63E+07	0.00E+00	0.00E+00	4.61E+07	4.26E+05	0.00E+00	IPI00014581	54	15	287.95	
9	0	3	5	1	1					MYH7 Myosin-7
4.12E+05	0.00E+00	1.17E+06	1.66E+05	6.15E+04	3.63E+05	IPI00025880	8	11	176.6	
5	5	10	6	6	6					VIM Vimentin
2.71E+06	4.37E+06	2.81E+07	2.90E+06	8.95E+06	6.97E+06	IPI00418471	20	7	131.66	
2	1	2	2	3	5					DES Desmin
4.69E+05	7.02E+05	1.47E+06	5.12E+05	3.60E+06	5.17E+06	IPI00465084	16	5	85.47	
0	0	1	0	3	3					OGN Osteoglycin
0.00E+00	0.00E+00	1.01E+05	0.00E+00	3.33E+05	5.98E+05	IPI00025465	9	3	48.14	
2	0	0	0	0	0					MYL3 Myosin light chain 3
2.01E+05	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	IPI00243742	16	2	37.39	
1	3	2	0	3	0					APOA2 Apolipoprotein A-II
2.98E+04	1.00E+06	3.42E+05	0.00E+00	4.68E+05	0.00E+00	IPI00021854	23	2	36.96	
0	2	0	0	0	0	IPI00783987	2	2	33.56	C3 Complement C3 (Fragment)
0.00E+00	3.38E+05	0.00E+00	0.00E+00	0.00E+00	0.00E+00					
0	0	0	2	0	0	IPI00019359	4	2	30.11	KRT9 Keratin, type I cytoskeletal 9
0.00E+00	0.00E+00	0.00E+00	7.03E+04	0.00E+00	0.00E+00					

Table III: Protein-Protein Comparison of Human LC/MS/MS Data. Shown are characteristics of RLC-7, -8, and -9 in non-failing (NF) and dilated cardiomyopathy (DCM) samples. Proteins detected by 2 or more distinct peptides are shown. Database Acc# = IPI database accession number; %AA Cov = percent amino acid sequence coverage; Dist Pep = distinct peptides.

Species	Peptide Sequence	Precursor m/z	Precursor MH+	Precursor Charge	Delta ppm	SPI %	Spec Mill Score	# Detected Mods
Mouse	RIEGGSSNVFSMFEQTQIQEFK	854.75	2562.2396	3	5.8	96	20.8	0
Mouse	RIEGG**(sS)**NVFSMFEQTQIQEFK	881.40	2642.2115	3	4.2	91	20.8	1p
Mouse	RIEGG**ss**NVFSMFEQTQIQEFK	908.06	2722.1702	3	3.0	96	19.2	2p
Human	RAGGANSNVFSMFEQTQIQEFK	830.40	2489.1875	3	-0.1	100	21.0	0
Human	RAGGA**n**SNVFSMFEQTQIQEFK	830.73	2490.1817	3	4.0	93	20.2	1d
Human	RAGGAN**s**NVFSMFEQTQIQEFK	857.05	2569.1498	3	-1.7	81	18.4	1p
Human	RAGGA**ns**NVFSMFEQTQIQEFK	857.38	2570.1356	3	-1.0	89	19.1	1d, 1p

Table IV: Details of RLC N-terminal Peptide 9-30 Mouse and Human PTMs Measured via LC-MS/MS. p=phosphorylation, d=deamidation.

Figure 1

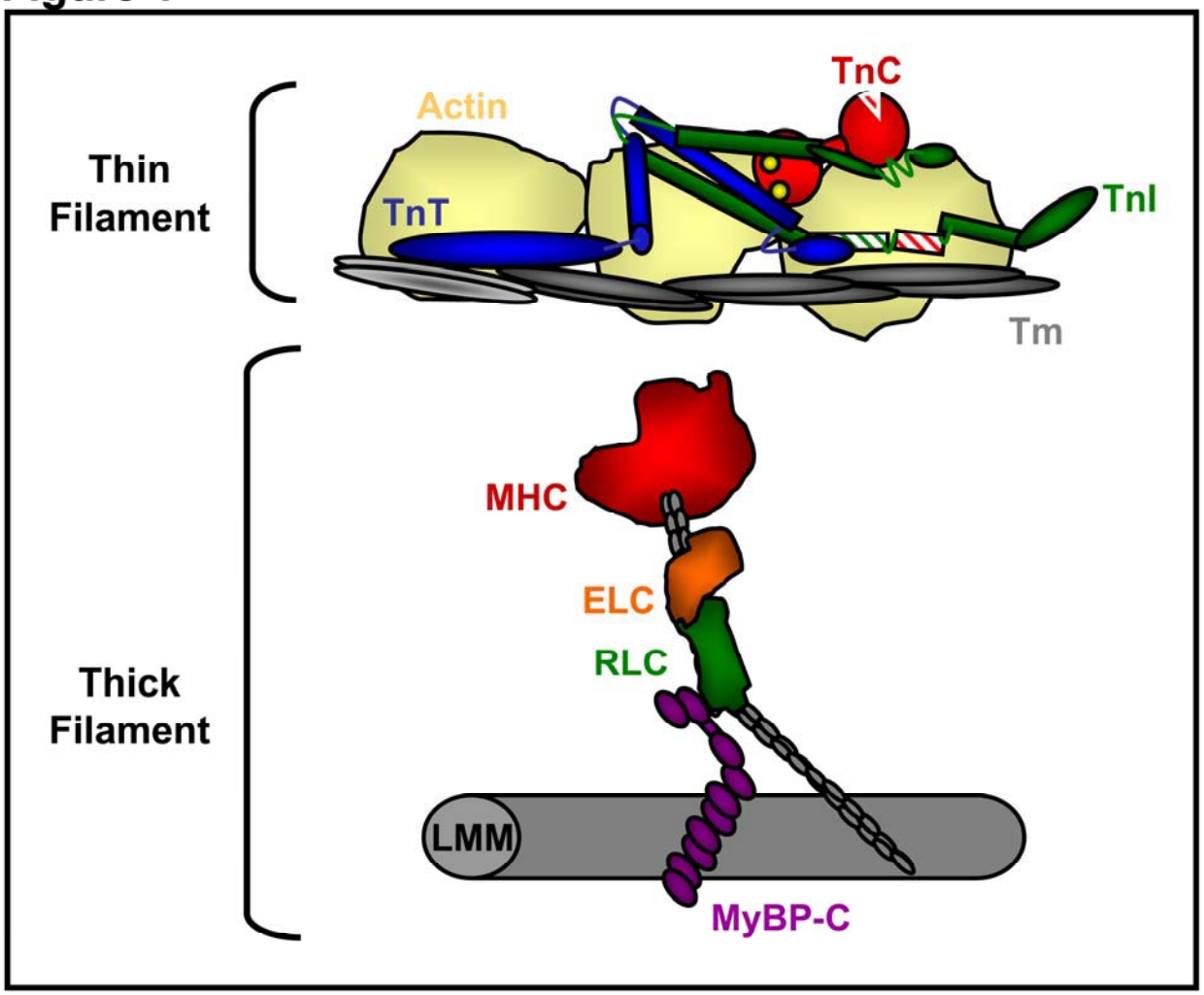

Thin Filament
Actin
TnC
TnT
TnI
Tm
Thick Filament
MHC
ELC
RLC
LMM
MyBP-C

Figure 2

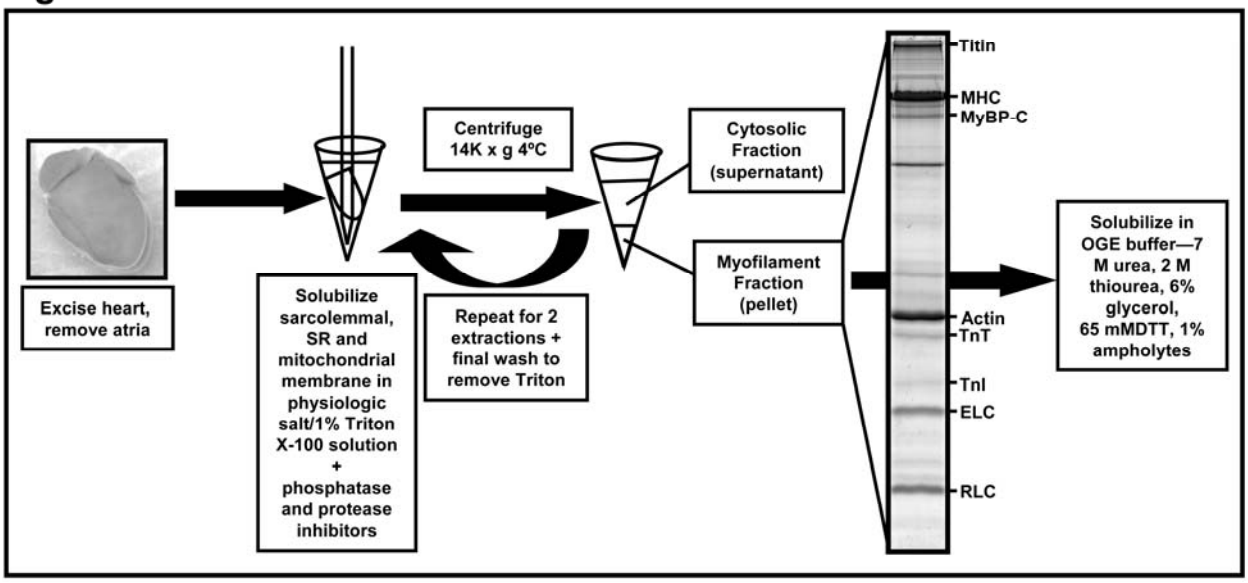
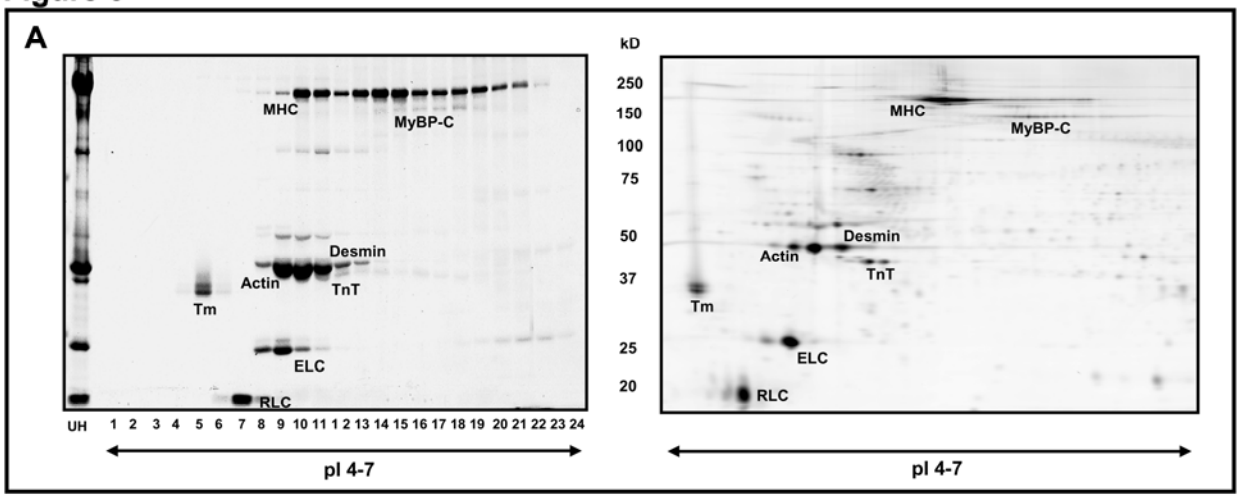

Excise heart, remove atria
Solubilize sarcolemmal, SR and mitochondrial membrane in physiologic salt/1% Triton X-100 solution + phosphatase and protease inhibitors
Centrifuge 14K x g 4°C
Repeat for 2 extractions + final wash to remove Triton
Cytosolic Fraction (supernatant)
Myofilament Fraction (pellet)
Titin
MHC
MyBP-C
Actin
TnT
TnI
ELC
RLC
Solubilize in OGE buffer—7 M urea, 2 M thiourea, 6% glycerol, 65 mMDTT, 1% ampholytes

Figure 3


A
MHC
MyBP-C
Desmin
Actin
TnT
Tm
ELC
RLC
UH 1 2 3 4 5 6 7 8 9 10 11 1 2 13 14 15 16 17 18 19 20 21 22 23 24
pI 4-7
kD
250
150
100
75
50
37
25
20
MHC
MyBP-C
Desmin
Actin
TnT
Tm
ELC
RLC
pI 4-7

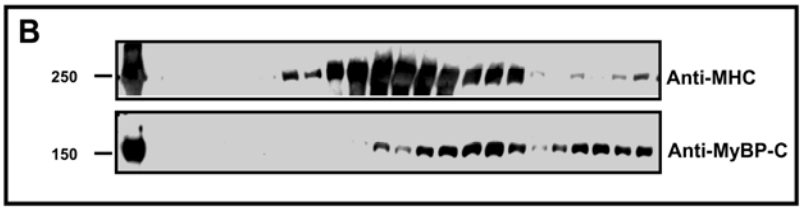

B
250
150
Anti-MHC
Anti-MyBP-C

Figure 4

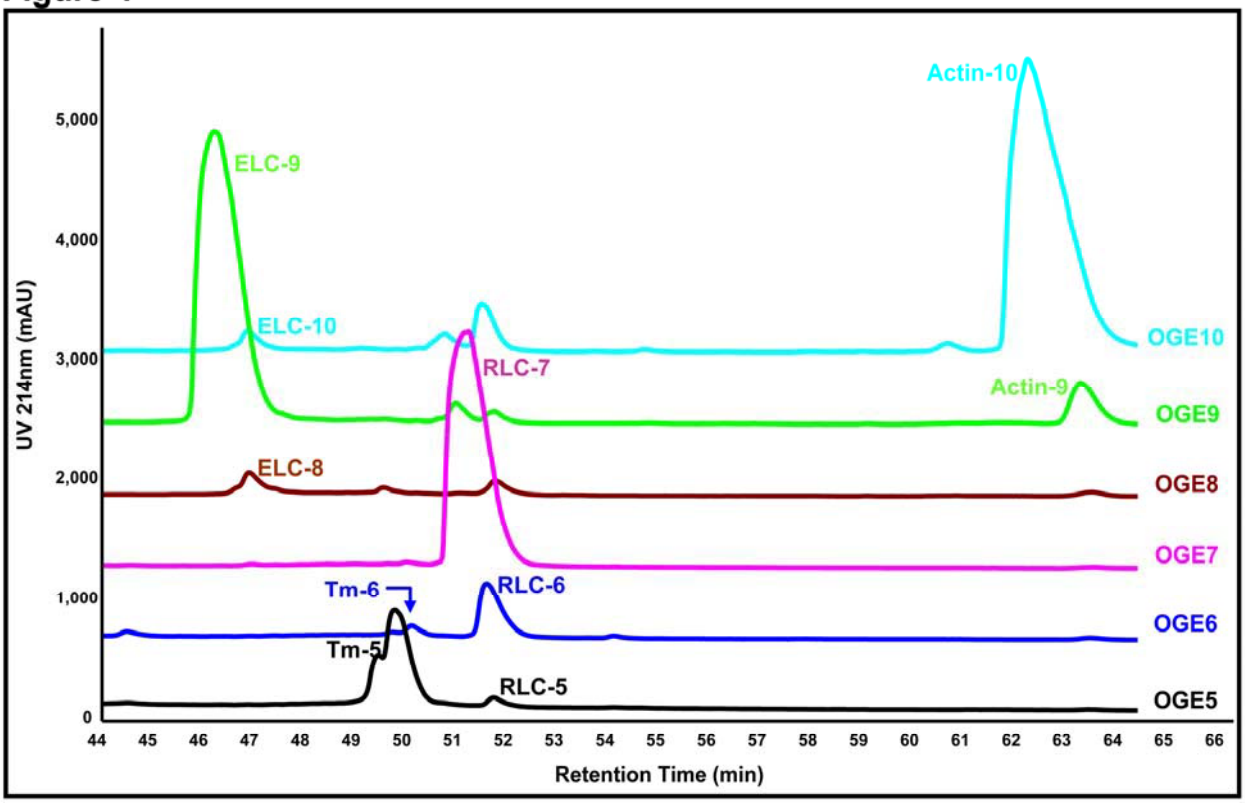

Actin-10
ELC-9
ELC-10
RLC-7
Actin-9
ELC-8
Tm-6
RLC-6
Tm-5
RLC-5
OGE10
OGE9
OGE8
OGE7
OGE6
OGE5
UV 214nm (mAU)
5,000
4,000
3,000
2,000
1,000
0
44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66
Retention Time (min)

Figure 5

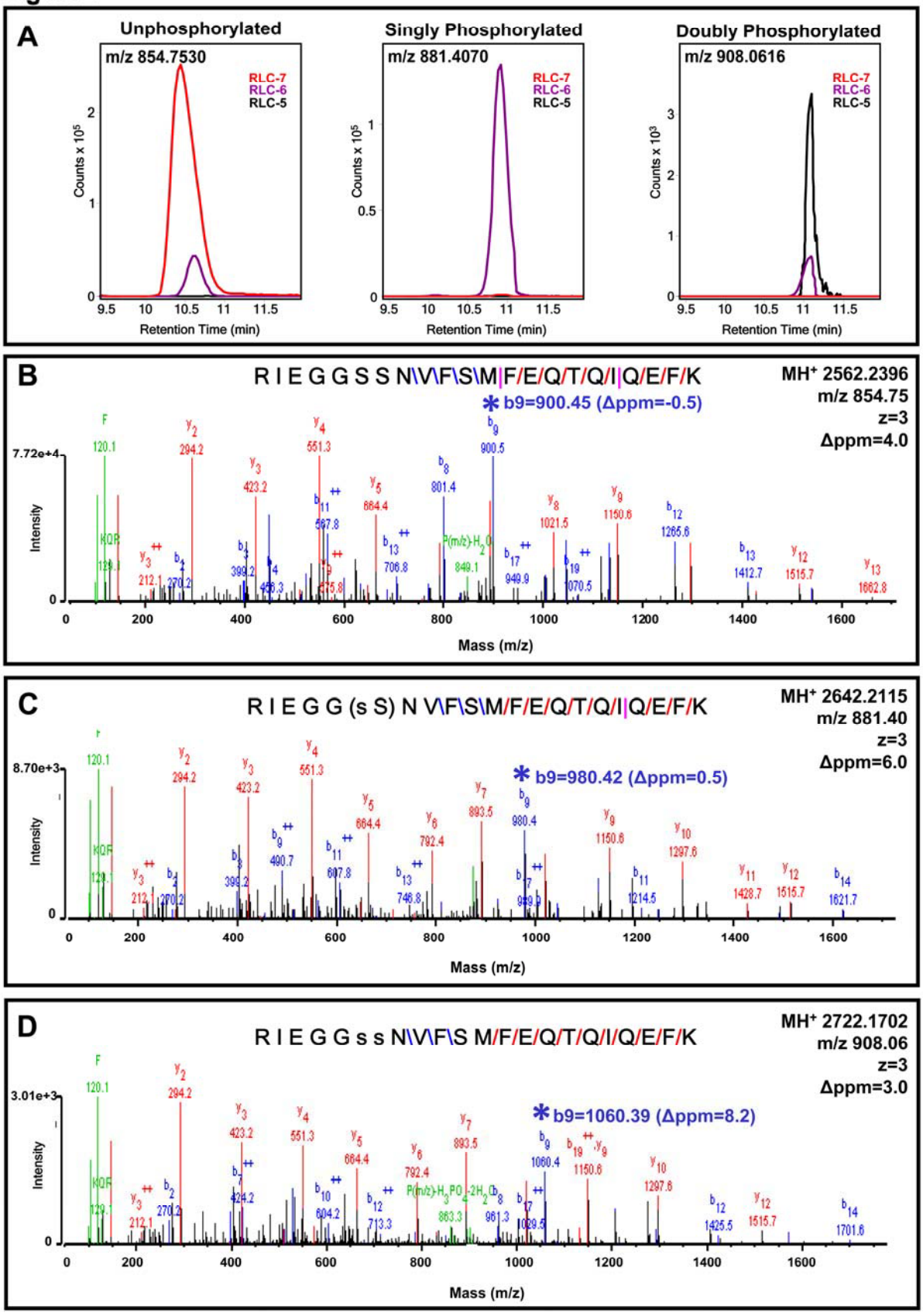

A
Unphosphorylated
m/z 854.7530
Singly Phosphorylated
m/z 881.4070
Doubly Phosphorylated
m/z 908.0616
RLC-7
RLC-6
RLC-5
Counts x 10^5
Counts x 10^3
Retention Time (min)
B
R I E G G S S N\V\F\S\M|F/E/Q/T/Q/I|Q/E/F/K
MH+ 2562.2396
m/z 854.75
z=3
Δppm=4.0
b9=900.45 (Δppm=-0.5)
7.72e+4
Intensity
Mass (m/z)
C
R I E G G (s S) N V\F\S\M/F/E/Q/T/Q/I|Q/E/F/K
MH+ 2642.2115
m/z 881.40
z=3
Δppm=6.0
b9=980.42 (Δppm=0.5)
8.70e+3
D
R I E G G s s N\V\F\S M/F/E/Q/T/Q/I/Q/E/F/K
MH+ 2722.1702
m/z 908.06
z=3
Δppm=3.0
b9=1060.39 (Δppm=8.2)
3.01e+3

Figure 6

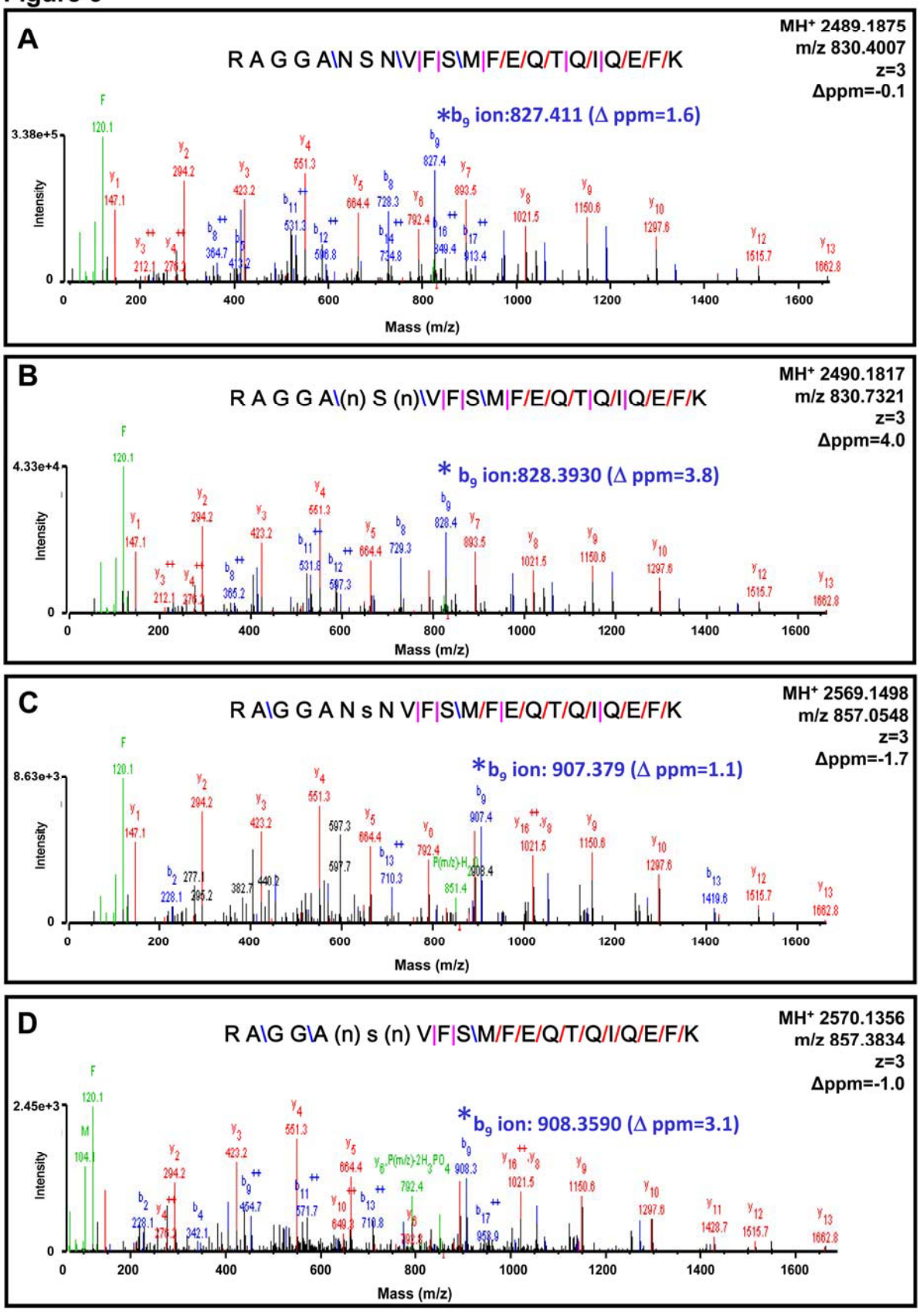

A
R A G G A\N S N\V|F|S\M|F/E/Q/T|Q/I|Q/E/F/K
MH+ 2489.1875
m/z 830.4007
z=3
Δppm=-0.1
*b9 ion:827.411 (Δ ppm=1.6)
B
R A G G A\(n) S (n)\V|F|S\M|F/E/Q/T|Q/I|Q/E/F/K
MH+ 2490.1817
m/z 830.7321
z=3
Δppm=4.0
* b9 ion:828.3930 (Δ ppm=3.8)
C
R A\G G A N s N V|F|S\M/F|E/Q/T/Q/I|Q/E/F/K
MH+ 2569.1498
m/z 857.0548
z=3
Δppm=-1.7
*b9 ion: 907.379 (Δ ppm=1.1)
D
R A\G G\A (n) s (n) V|F|S\M/F/E/Q/T/Q/I/Q/E/F/K
MH+ 2570.1356
m/z 857.3834
z=3
Δppm=-1.0
*b9 ion: 908.3590 (Δ ppm=3.1)
Intensity
Mass (m/z)

Figure 7

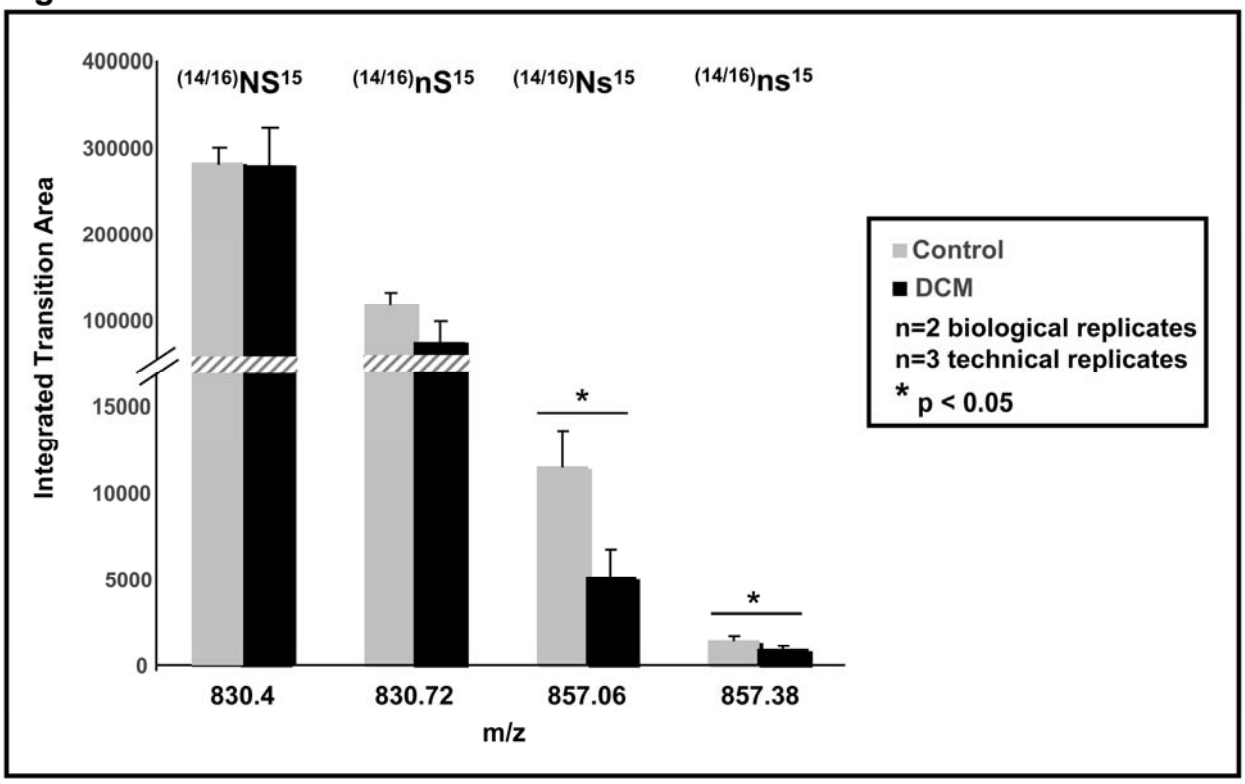
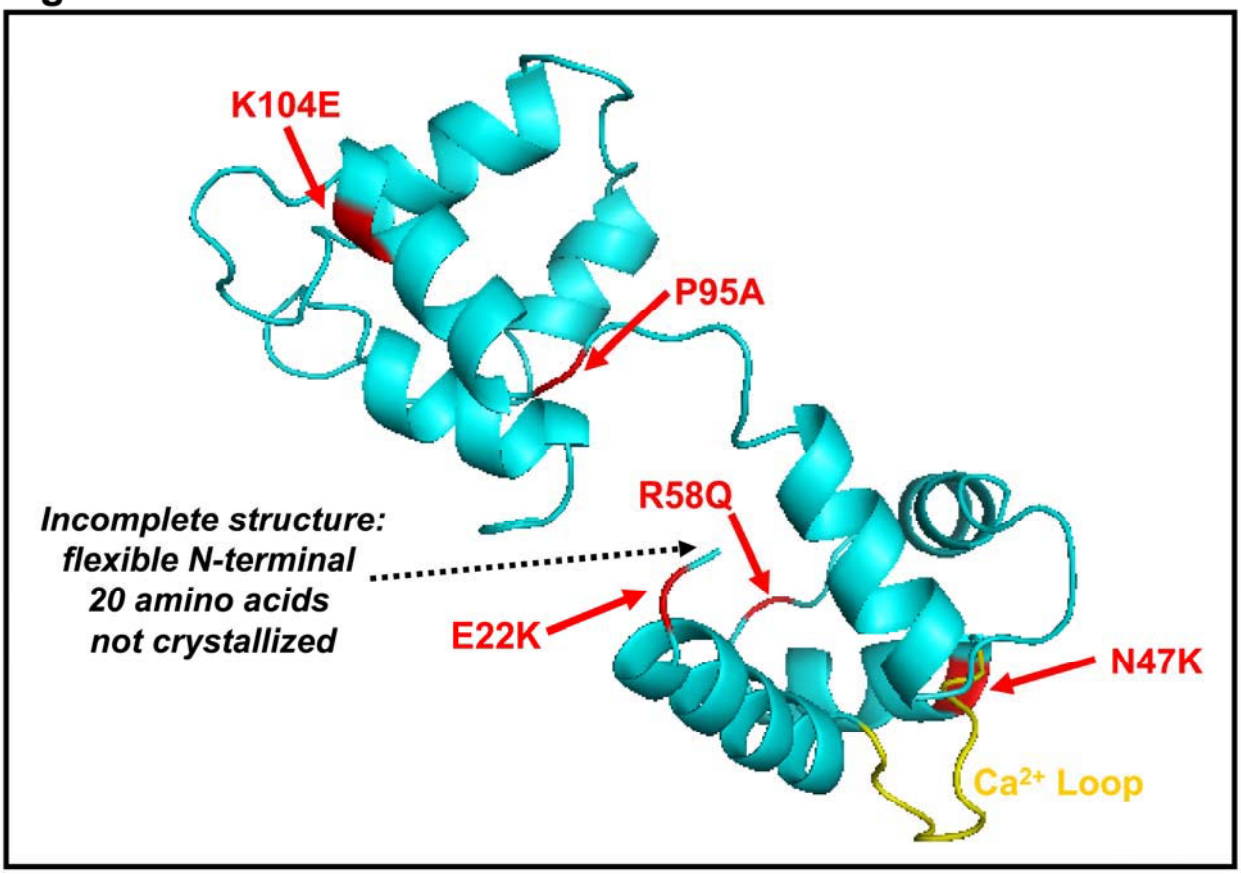

(14/16)NS15
(14/16)nS15
(14/16)Ns15
(14/16)ns15
Integrated Transition Area
400000
300000
200000
100000
15000
10000
5000
0
*
*
830.4
830.72
857.06
857.38
m/z
Control
DCM
n=2 biological replicates
n=3 technical replicates
* p < 0.05

Figure 8


K104E
P95A
R58Q
Incomplete structure:
flexible N-terminal
20 amino acids
not crystallized
E22K
N47K
Ca2+ Loop