Journal of Molecular and Cellular Cardiology 107 (2017) 13–21

Contents lists available at ScienceDirect

Journal of Molecular and Cellular Cardiology

journal homepage: www.elsevier.com/locate/yjmcc

Distinct sequences and post-translational modifications in cardiac atrial and ventricular myosin light chains revealed by top-down mass spectrometry

Zachery R. Gregorich [a,b], Wenxuan Cai [a,b], Ziqing Lin [b], Albert J. Chen [b], Ying Peng [b], Takushi Kohmoto [c], Ying Ge [a,b,d,e,*]

[a] Molecular and Cellular Pharmacology Training Program, University of Wisconsin-Madison, Madison, WI 53705, USA
[b] Department of Cell and Regenerative Biology, University of Wisconsin-Madison, Madison, WI 53705, USA
[c] Department of Surgery, University of Wisconsin-Madison, Madison, WI 53705, USA
[d] Human Proteomics Program, University of Wisconsin-Madison, Madison, WI 53705, USA
[e] Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA

ARTICLE INFO

Article history:
Received 23 March 2017
Accepted 15 April 2017
Available online 17 April 2017

Keywords:
Myosin light chain
Top-down mass spectrometry
Post-translational modification
Phosphorylation
Acetylation
Methylation

ABSTRACT

Myosin is the principal component of the thick filaments that, through interactions with the actin thin filaments, mediates force production during muscle contraction. Myosin is a hexamer, consisting of two heavy chains, each associated with an essential (ELC) and a regulatory (RLC) light chain, which bind the lever-arm of the heavy chain and play important modulatory roles in striated muscle contraction. Nevertheless, a comprehensive assessment of the sequences of the ELC and RLC isoforms, as well as their post-translational modifications, in the heart remains lacking. Herein, utilizing top-down high-resolution mass spectrometry (MS), we have comprehensively characterized the sequences and N-terminal modifications of the atrial and ventricular isoforms of the myosin light chains from human and swine hearts, as well as the sites of phosphorylation in the swine proteins. In addition to the correction of disparities in the database sequences of the swine proteins, we show for the first time that, whereas the ventricular isoforms of the ELC and RLC are methylated at their N-termini, which is consistent with previous studies, the atrial isoforms of the ELC and RLC from both human and swine are N^{α}-methylated and N^{α}-acetylated, respectively. Furthermore, top-down MS with electron capture dissociation enabled localization of the sites of phosphorylation in swine RLC isoforms from the ventricles and atria to Ser14 and Ser22, respectively. Collectively, these results provide new insights into the sequences and modifications of myosin light chain isoforms in the human and swine hearts, which will pave the way for a better understanding of their functional roles in cardiac physiology and pathophysiology.

1. Introduction

The thick filaments of muscles, including cardiac, skeletal, and smooth muscles, are predominantly composed of myosin that, through interactions with actin (the principal component of thin filaments), mediates force production during muscle contraction [1,2]. Myosin is a hexamer consisting of two heavy chains (MHCs), each of which is associated with an essential (or alkali) light chain (ELC) and a regulatory (or phosphorylatable) light chain (RLC) [3]. In striated muscles both the ELC and RLC, which bind and stabilize the lever-arm of the MHC [4,5], play critical roles in the modulation of contractile function [6–11]. In support of this, mutations in both the ELC and RLC are associated with the development of cardiac and skeletal muscle myopathies [12]. Additionally, phosphorylation of the RLC by the Ca^{2+}/calmodulin-dependent myosin light chain kinase (MLCK) represents a critical mechanism regulating contractility, especially in the heart [6,11,13–16]. In particular, recent studies have convincingly demonstrated that loss of RLC phosphorylation leads to pathological cardiac hypertrophy and heart failure [16–19]. These findings have generated substantial interest in the roles played by myosin light chains and their modifications in cardiac physiology and pathophysiology; thus, a comprehensive assessment of

Abbreviations: MHCs, myosin heavy chains; ELC, essential light chain; RLC, regulatory light chain; MLCK, myosin light chain kinase; PTMs, post-translational modifications; ELC_v, ventricular isoform of ELC; ELC_a, atrial isoform of ELC; RLC_v, ventricular isoform of RLC; RLC_a, atrial isoform of RLC; MS, mass spectrometry; MS/MS, tandem MS; ECD, electron capture dissociation; CID, collision induced dissociation; LC, liquid chromatography; LTQ, linear ion trap; FT-ICR, Fourier transform ion cyclotron resonance; $_pRLC_a$, mono-phosphorylated RLC_a; $_{pp}RLC_a$, bis-phosphorylated RLC_a; $_pELC_a$, mono-phosphorylated ELC_a; $_pRLC_v$, mono-phosphorylated RLC_v; $_{pp}RLC_v$, bis-phosphorylated RLC_v; $_pELC_v$, mono-phosphorylated ELC_v.

* Corresponding author at: 1111 Highland Ave., WIMR II 8551, Madison, WI 53705, USA.
E-mail address: ge2@wisc.edu (Y. Ge).

http://dx.doi.org/10.1016/j.yjmcc.2017.04.002
0022-2828/

myosin light chain isoforms and their post-translational modifications (PTMs) in the heart is apropos.

A number of different myosin light chain isoforms have been identified to date, all of which belong to the EF-hand protein super-family of Ca^{2+}-binding proteins [20]. Four of these genes are expressed in the heart, with expression patterns that vary by chamber, developmentally, and in response to pathological stimuli [7,10,15]. The *MYL3* and *MYL4* genes encode the ventricular (ELC_v) and atrial (ELC_a) isoforms of the ELC, while the ventricular (RLC_v) and atrial (RLC_a) isoforms of the RLC are encoded by the *MYL2* and *MYL7* genes, respectively. Although expression of ELC_v is primarily restricted to the ventricles of the heart, ELC_a is expressed in both the atria and ventricles during normal embryonic development [7,10]. In adulthood, however, expression of the ELC_a is restricted to the atria, although re-expression in the ventricles occurs in response to pressure overload and heart failure [7,10,21]. On the other hand, RLC_v expression is restricted to the ventricles both in the developing and adult heart [15]. Conversely, RLC_a, like ELC_a, is expressed throughout the heart early in development, and becomes restricted to the atria later in development [15].

Top-down mass spectrometry (MS) has gained considerable popularity as the premier approach for comprehensively characterizing proteins [22–24]. Unlike in conventional bottom-up MS, in which proteins are digested and the resulting peptides are analyzed by MS, intact proteins are analyzed in top-down MS, providing a global or "bird's eye" view of all protein species, including those containing sequence variations (due to mutations/polymorphisms or alternative splicing) and/or PTMs [22–30]. Following intact protein analysis, specific protein species of interest can be isolated and fragmented using a variety of tandem MS (MS/MS) techniques, including, but not limited to, electron capture dissociation (ECD) and collision induced dissociation (CID), to obtain sequence information and localize PTMs [22–30]. In particular, top-down MS with ECD represents a powerful method for the comprehensive characterization of proteins; especially those containing labile PTMs such as phosphorylation, which are frequently lost when proteins are fragmented using energetic dissociation methods such as CID [25]. Additionally, the use of high-resolution mass spectrometers in top-down MS studies offers unparalleled mass accuracy, which not only increases confidence in protein identification, but also in the identification of protein PTMs [31]. Herein, utilizing top-down high-resolution MS, we have characterized the sequences and N-terminal modifications of cardiac myosin light chain isoforms from human and swine, as well as the sites of phosphorylation in the swine proteins, towards a better understanding of the functional roles of these proteins in cardiac physiology and pathophysiology. Interestingly, we found that, whereas the ventricular ELC and RLC are N^{α}-tri-methylated, the atrial ELC is methylated at its N-terminus while the atrial RLC in both swine and human is N^{α}-acetylated; making the atrial RLC unique among cardiac myosin light chain isoforms. Importantly, we have also precisely localized the sites of phosphorylation in swine RLC isoforms from the ventricles and atria to Ser14 and Ser22, respectively. Although prior studies have reported atrial RLC phosphorylation, this represents the first study to definitively localize a site of phosphorylation in this isoform.

2. Methods

Detailed methods are found in Supporting information. To comprehensively characterize the sequences and PTMs of swine myosin light chain isoforms, myofilament-enriched extracts were prepared from the atrial and ventricular myocardium of 1–3 healthy adult Yorkshire domestic swine (*Sus scrofa*) (approximately 3 months of age) using the two-step extraction procedure described by Van Eyk et al. [32,33] with modifications. Subsequently, protein extracts prepared from swine atrial or ventricular myocardium were separated by 1D reverse phase liquid chromatography (LC) coupled directly online with a linear ion trap (LTQ) mass spectrometer (Thermo Scientific, Bremen, Germany) (Fig. S1), and eluting myosin light chain isoforms were fraction collected. MS/MS analysis of the fraction collected swine myosin light chain isoforms was conducted on a 7T LTQ/FT Ultra high-resolution mass spectrometer (Thermo Scientific) (Fig. 1). Moreover, to characterize the sequences and N-terminal modifications of human cardiac myosin light chain isoforms, myofilament-enriched extracts were prepared from donor atrial and ventricular myocardium as above, and online liquid chromatography (LC)-MS/MS analysis was carried out on a maXis II mass spectrometer (Bruker, Billerica, MA, USA). Accession numbers for all myosin light chain sequences utilized in this study can be found in Table S1.

3. Results

3.1. Inhibition of phosphatase activity towards the RLC

To quantify RLC phosphorylation we employed quantitative online top-down LC-MS. This method provided robust and highly-reproducible measurement of RLC phosphorylation in cardiac tissue extracts prepared using the two-step extraction procedure described by Van Eyk et al. [32,33] with modifications (Fig. S2, Supplemental results). By varying the concentration of phosphatase inhibitors used in the HEPES-based extraction buffer [32,33], we determined that supplementation with 600 mM NaF (in combination with 5 mM sodium pyrophosphate and 100 mM β-glycerophosphate) provided maximal inhibition of phosphatase activity towards the RLC in myocardial extracts (Figs. S3–S4, Supplemental results). Therefore, HEPES-based buffer containing 600 mM NaF, 5 mM sodium pyrophosphate, and 100 mM β-glycerophosphate was used for all subsequent experiments to prevent de-phosphorylation of RLC during the preparation of myofilament-enriched protein extracts from atrial and ventricular myocardium.

3.2. Sequencing and N-terminal modification characterization of swine and human RLC_a

Top-down MS profiling of RLC_a protein species in extracts prepared from the atrial myocardium of swine revealed the presence of major and minor protein species presumably corresponding to un-phosphorylated RLC_a (M_r 19,372.59 Da), mono-phosphorylated RLC_a ($_pRLC_a$; M_r 19,452.55 Da), and bis-phosphorylated RLC_a ($_{pp}RLC_a$; 19,532.51 Da) (Fig. 2A). Top-down LC-MS-based quantification yielded a mean value of 0.16 ± 0.03 mol of Pi/mol of RLC for swine RLC_a (n = 5) (Fig. 2B).

However, the expected mass of swine RLC_a, predicted based on the current unconfirmed sequence in the UniProtKB/Swiss-Prot database (M_r 16,694.24 Da), was substantially different than the experimentally determined mass for this protein (Fig. 2A), indicating potential sequence discrepancies in the unconfirmed sequence. Therefore, to correct potential sequence discrepancies in the database sequence, we first verified the sequence of the human protein using online top-down LC-MS/MS (Fig. S5), and, subsequently, utilized the confirmed human RLC_a sequence as a template to characterize the swine protein. Interestingly, although previous studies have shown that ventricular myosin light chain isoforms are N^{α}-tri-methylated [34], high accuracy mass measurement revealed that human RLC_a is N-terminally acetylated (+42.010 Da) rather than tri-methylated (+42.046 Da) (Fig. S5A). Similarly, swine RLC_a was also N^{α}-acetylated rather than N-terminally methylated (Fig. 2C).

Comprehensive top-down MS analysis permitted the confirmation of 6 amino acid differences between the swine and human proteins (Fig. 2D; Figs. S6–S7; Supplementary results). Moreover, swine RLC_a contained a 26 amino acid N-terminal stretch not present in the unconfirmed swine sequence in the UniProtKB/Swiss-Prot database (Fig. 2D; Fig. S8). After sequencing, a total of 41 *a*•, 89 *c*, 28 *y*, and 87 *z*• ions could be matched to the sequence of swine RLC_a (from a combined 2 ECD tandem mass spectra), representing cleavage of 144 out of 173 inter-residue bonds (~83% sequence coverage) (Fig. 2D). Note that *a*•, *b*, and *c* ions contain the N-terminus, while *x*, *y*, and *z*• ions contain