PRODUCTION AND PURIFICATION OF SLOW SKELETAL MYOSIN-BINDING PROTEIN C N-TERMINAL FRAGMENTS

by

Rashi Ramesh Singh

A Thesis Submitted to the Faculty of the

DEPARTMENT OF CELLULAR AND MOLECULAR MEDICINE

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2025

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by Rashi Ramesh Singh, titled "Production and purification of slow skeletal myosin-binding protein C N-terminal fragments" and recommend that it be accepted as fulfilling the dissertation requirement for the Master's Degree.

Brett Colson Date: 08/15/2025

[Dr. Brett Colson]

Date: 08/15/2025

[Dr. Christopher Pappas]

John Konhilas (Aug 26, 2025 12:44:46 PDT) Date: 08/15/2025

[Dr. John Konhilas]

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

Brett Colson Date: 08/15/2025

Dr. Brett Colson

Master's Thesis Committee Chair

Department of Cellular and Molecular Medicine

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Myosin binding protein-C (MyBP-C) is a sarcomeric protein that plays an essential role in regulating muscle contraction and maintaining structural integrity in striated muscle. It exists as 3 paralogs expressed in separate genes: slow skeletal (sMyBP-C, MYBPC1), fast skeletal (fMyBP-C, MYBPC2), and cardiac MyBP-C (cMyBP-C, MYBPC3). Mutations in the gene encoding sMyBP-C are implicated in debilitating disorders such as distal arthrogryposis type 1 (DA-1), yet the molecular mechanisms underlying this pathology remain poorly understood. Moreover, sMyBP-C is alternatively spliced in many variants, including a short and long form. Little is known about the differences in function of the skeletal paralogs and these splice variants. Interestingly, sMyBP-C long contains a PKA-mediated phosphorylation site in its N-terminal linker that precedes the C1 domain. This is different from the location of the PKA site in cMyBP-C, which resides in M-domain in between C1 and C2 domains. Therefore, it is likely that the mechanisms of skeletal MyBP-C in health and disease are distinct from cMyBP-C. To enable mechanistic studies, this thesis focuses on producing and purifying human N-terminal domain fragments (C1-C2 domains) of s- and f- MyBP-C using an expression and purification protocol originally developed for c-MyBP-C. The initial constructs of sC1-C2 short, sC1-C2 long, and fC1-C2 each contained a cleavable C-terminal His-tag for purification that contained an 11-residue linker of MyBP-C sequence before the tag termed, 11cHT. These constructs were codon-optimized for bacteria, cloned into the pET45b vector, and expressed in E. coli BL21(DE3) cells. Autoinduction media and a reduced growth temperature (25°C) were used to enhance protein solubility and yield. Following lysis and immobilized metal affinity chromatography (IMAC), proteins were successfully purified and subjected to TEVp cleavage and dialysis. SDS-PAGE analysis confirmed high expression and tag removal, with purity levels ranging from 84% to 94%. Final protein yields were robust (~19 mg per 500 mL culture). These results validate the applicability of the cardiac MyBP-C expression strategy to skeletal paralogs, supporting future studies of actin binding, phosphorylation effects, and pathogenic mutations in skeletal MyBP-C. We note that most

recently, optimizations of skeletal C1-C2 protein design and production by the Colson lab resulted in using an N-terminal cleavable His-tags and even lower growth temperatures of 18ºC. The purified constructs lay the foundation for biophysical assays such as FRET to explore the structural and functional consequences of DA-associated mutations and the broader role of skeletal MyBP-C in muscle physiology. Rationales and considerations for use of strategies, protocols, and reagents used for production and purification of recombinant MyBP-C are described. Predictions of how such studies previously done in the Colson lab using cMyBP-C may be expected for future studies applied to skeletal MyBP-C biology are discussed.

1. INTRODUCTION

Striated muscle contraction, which includes both cardiac and skeletal muscle types, is fundamentally driven by the highly coordinated interaction between actin and myosin within the sarcomere. The organization of thick and thin filaments within the sarcomere is illustrated in Figure 1. The initiation of this process is tightly regulated by calcium ions (Ca^{2+}), which serve as the primary intracellular signal for contraction. Upon release from the sarcoplasmic reticulum, Ca^{2+} binds to the calcium-sensing subunit, troponin C, situated on the actin-containing thin filament. This binding induces a structural shift in the troponin–tropomyosin complex, causing tropomyosin to move away from the myosin-binding sites on actin (McKillop & Geeves, 1993; Gordon et al., 2000). Once these sites are exposed, the myosin heads, which are part of the thick filaments, can attach to actin, initiating the cross-bridge cycle.

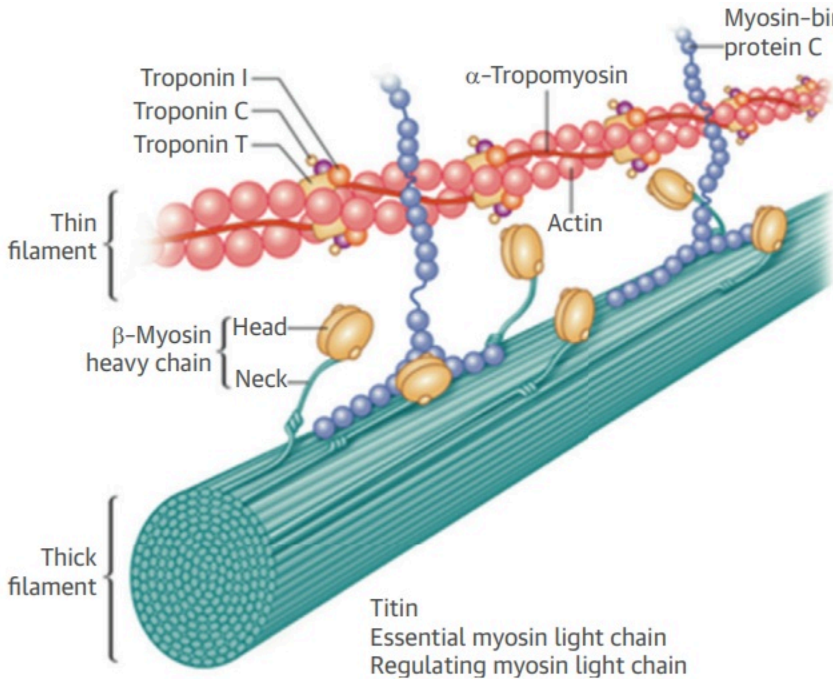

Figure 1: Sarcomere Structure Depicting Thick and Thin Filament Proteins Relevant to MyBP-C Function

Schematic representation of the sarcomere, illustrating the arrangement of thick and thin filament proteins. The thick filament includes β-myosin heavy chain, titin, essential and regulatory myosin light chains, and myosin-binding protein C (MyBP-C), while the thin filament consists of actin, α-tropomyosin, and the troponin complex (TnI, TnT, and TnC). Although the image originates from studies on cardiac muscle, the structural organization and MyBP-C localization are conserved in skeletal muscle. Mutations in the skeletal isoform of MyBP-C (sMyBP-C) are implicated in congenital myopathies such as distal arthrogryposis. Adapted from Semsarian et al. (2015).

The actomyosin ATPase cycle involves a sequence of molecular events powered by ATP hydrolysis. A simplified representation of this process is illustrated in Figure 2. It begins with the binding of ATP to the myosin head, leading to detachment from actin. ATP is then hydrolyzed to ADP and inorganic phosphate (Pi), which primes the myosin head into a high-energy state. Upon reattachment to actin, the release of Pi triggers the power stroke, during which the myosin head pivots and pulls the actin filament, generating contractile force. The subsequent release of ADP completes the cycle, allowing a new ATP molecule to bind and initiate the next round (Rayment et al., 1993; Lehman et al., 1994). While Ca^{2+} does not directly influence the ATP hydrolysis step, it plays a critical role in accelerating Pi release when both Ca^{2+} and strongly bound myosin heads are present. This step is tightly linked to force production and is significantly enhanced through activation of the thin filament, positioning Pi release as a central regulatory point in contraction kinetics (Rosenfeld & Taylor, 1987).

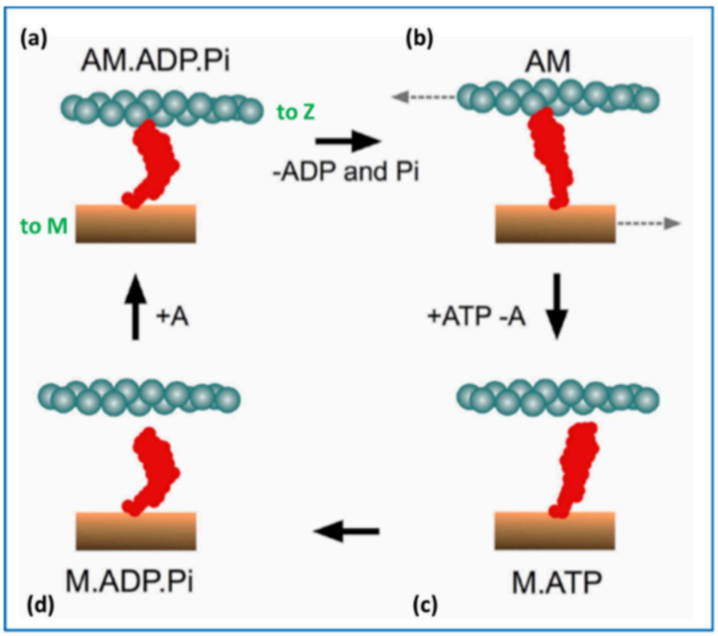

Figure 2: Representation of the Actomyosin ATPase Cycle Driving Muscle Contraction

Simplified depiction of the ATP-driven contractile cycle between actin and myosin filaments, following the classical Lymn-Taylor scheme. (a) The myosin head (red) binds actin (blue/green) in the presence of ADP and Pi, forming a strong-binding, pre-power stroke state. (b) Release of Pi and ADP initiates the power stroke, resulting in filament sliding. (c) Binding of ATP induces dissociation of the myosin head from actin. (d) ATP hydrolysis re-cocks the myosin head into a high-energy state, ready for the next contraction cycle. This mechanochemical cycle is the fundamental basis of force generation in striated muscle. Adapted from (Squire, 2019).

Robinett et al. (2019) explored this interplay using isolated rat skeletal muscle fibers and systematically varying Ca^{2+} concentrations to mimic different levels of physiological activation. Their experiments, conducted on both slow- and fast-twitch muscle fibers, utilized permeabilized fibers and a calcium range from pCa 6.4 to 4.5. They demonstrated that contractile force and velocity are not only dependent on Ca^{2+}-induced activation of the thin filament but also on cooperative recruitment of myosin heads to actin. This cooperative activation suggests that initial myosin–actin interactions enhance subsequent cross-bridge binding events, amplifying the overall contractile response. In this context, structural regulatory proteins such as MyBP-C provide an additional layer of control.

MyBP-C in striated muscle biology

MyBP-C is located in the C-zone of the sarcomere, a region situated in the central portion of each half of the A-band where thick filaments are present and MyBP-C is specifically localized. Interacting with both actin and myosin, MyBP-C modulates myofilament dynamics. A proposed model of how MyBP-C coordinates contraction and relaxation through filament interactions is illustrated in Figure 3. Robinett et al. further demonstrated that phosphorylation of slow skeletal MyBP-C by protein kinase A (PKA) significantly increases the rate of cross-bridge cycling (k_{tr}), likely by relieving structural constraints on myosin heads and enhancing their interaction with actin. These findings emphasize the integrated roles of Ca^{2+} signaling, thin filament regulation, and ATP hydrolysis in coordinating the force, speed, and efficiency of skeletal muscle contraction.

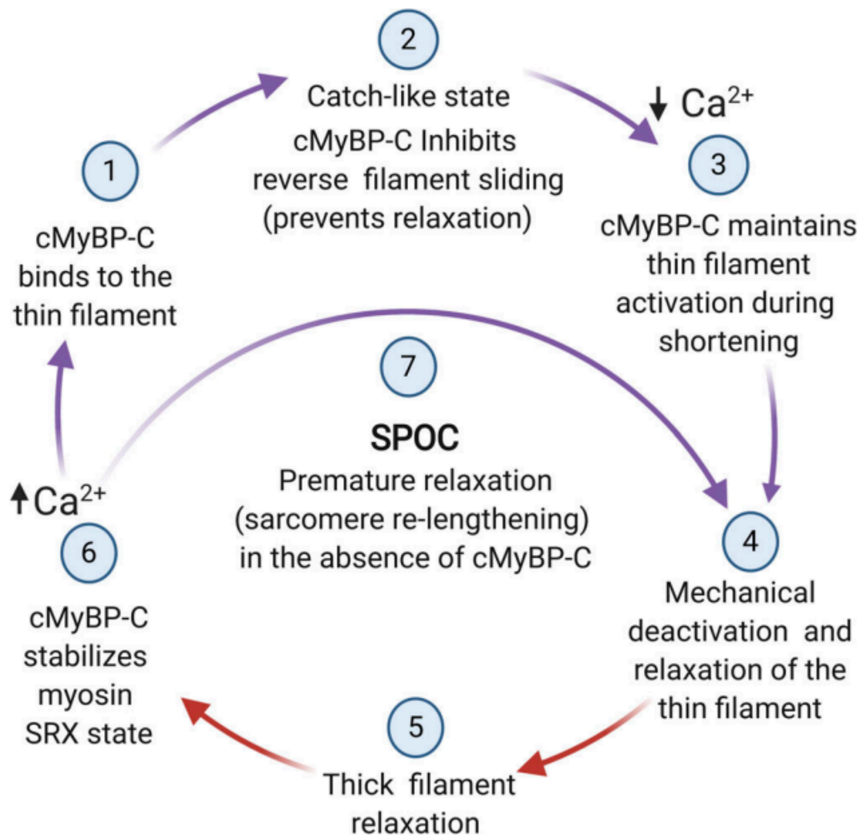

Figure 3: Model of MyBP-C–Mediated Regulation of Sarcomere Contractility (Contraction and Relaxation) in Striated Muscle

This schematic illustrates a proposed model in which myosin-binding protein C (MyBP-C) dynamically coordinates contraction and relaxation through alternating interactions with the thin and thick filaments, regulated by calcium levels and mechanical signaling. Upon calcium binding, MyBP-C associates with the thin filament (Step 1), promoting a catch-like state that prevents premature filament sliding (Step 2). During continued sarcomere shortening, MyBP-C stabilizes thin filament activation via tropomyosin repositioning (Step 3). Relaxation is initiated by a drop in calcium and mechanical deactivation (Step 4), followed by thick filament relaxation

(Step 5). MyBP-C then promotes the super-relaxed (SRX) state of myosin through thick filament interactions (Step 6). In the absence of MyBP-C, sarcomere relengthening occurs prematurely (Step 7), a phenomenon described as SPOC (sarcomere premature onset of relaxation). Although originally described in cardiac muscle, this regulatory framework is increasingly applied to skeletal muscle, where MyBP-C mutations are associated with congenital contracture disorders such as distal arthrogryposis. Adapted from Harris (2021).

Actin-myosin interactions are modulated by other sarcomeric proteins including MyBP-C to fine-tune contraction. Specifically, it is thought that the key interactions of MyBP-C's N-terminal domains, particularly the C1-C2 region containing M-domain, regulates actin-myosin interactions by binding to both actin and myosin. While the cardiac paralog of MyBP-C N-terminal domains have been studied in detail, largely due to MyBP-C variants causing hypertrophic cardiomyopathy, relatively little is known about the N-terminal domains of fast and slow paralogs of MyBP-C. However, skeletal MyBP-C has now been implicated in skeletal muscle disorders, underscoring the importance of understanding these paralogs (Bunch et al., 2019; Gurnett et al., 2010; Wong et al., 2024)

To better understand skeletal MyBP-C N-terminal domains in health and disease, I will discuss key aspects and considerations for adapting current methodologies optimized for the construct design, protein expression, and protein purification of c-MyBP-C N-terminal fragments (Bunch et al., 2018) for future research using slow and fast skeletal MyBP-C N-terminal fragments. I also present my experiment to express and purify these skeletal MyBP-C proteins using established cardiac MyBP-C protein methods and how these can be used for future optimization of constructs, protocols, and mechanistic studies.

MyBP-C domain organization

There are three major isoforms or paralogs of MyBP-C, cardiac, slow skeletal, and fast skeletal, referred to as cMyBP-C, sMyBP-C, and fMyBP-C, respectively. cMyBP-C is expressed in the heart. While sMyBP-C is expressed as numerous alternatively spliced variants that are expressed in both slow-and fast-twitch skeletal muscles, whereas fMyBP-C is expressed in fast-twitch skeletal muscles. The skeletal MyBP-C domain organization includes an extended chain of 7 immunoglobulin (Ig)-like domains and 3 fibronectin (Fn)-like domain, referred to starting from the N-terminus as C1–C10 (Ackermann & Kontrogianni-Konstantopoulos, 2011, 2013). The domain structure of the three MyBP-C isoforms, including isoform-specific features and phosphorylation sites, is illustrated in Figure 4. The C1 domain is flanked by two distinct regions: a proline- and alanine-rich linker (PAL) on one side and a conserved linker region, referred to as the MyBP-C motif (M-domain), on the other. The PAL, located upstream of C1, consists of approximately 50 amino acids, while M-domain, positioned downstream, spans around 100 amino acids. Although the three paralogs exhibit considerable structural and sequence similarity, the cardiac MyBP-C isoform possesses three additional features not found in skeletal isoforms. These include an Ig domain at the extreme N-terminus, termed C0, a unique 9-residue insertion within the M-domain containing phosphorylation sites crucial for its regulatory role in muscle contraction, and a 28-amino acid loop within the central region of the C5 domain (Ackermann & Kontrogianni-Konstantopoulos, 2010, 2013). In contrast, human sMyBP-C long form contains a PKA site in the PAL.

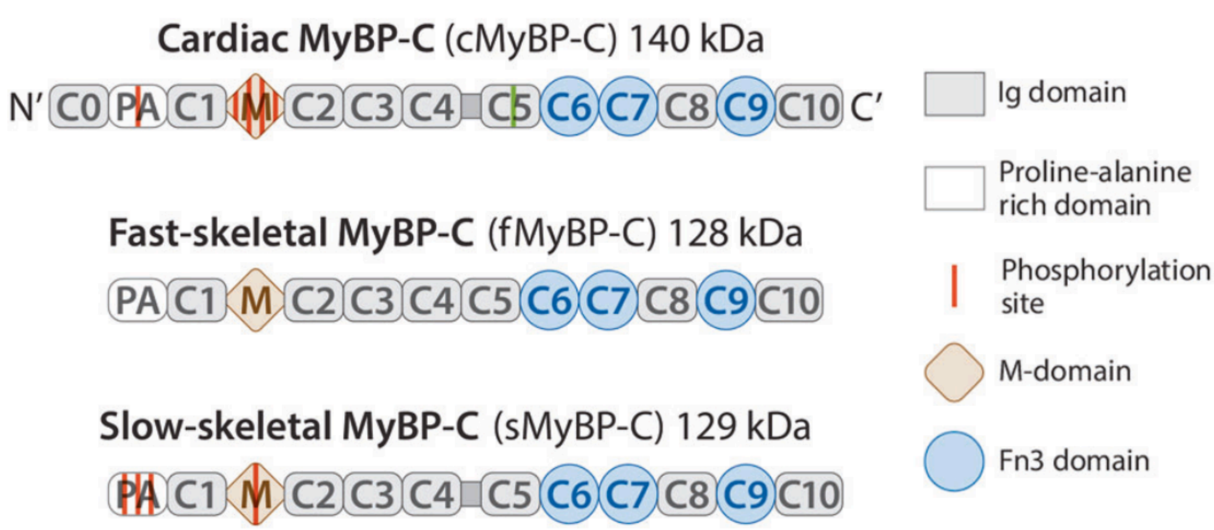

Figure 4: Domain Organization of Cardiac and Skeletal Isoforms of Myosin-Binding Protein C

Schematic representation of the domain structures of the three MyBP-C isoforms: cardiac (cMyBP-C), fast-skeletal (fMyBP-C), and slow-skeletal (sMyBP-C), encoded by the genes MYBPC3, MYBPC2, and MYBPC1, respectively. All isoforms share a conserved modular architecture composed of seven immunoglobulin-like (Ig) domains (gray), three fibronectin type III (Fn3) domains (blue), an M-domain (tan diamond), and an N-terminal proline-alanine–rich region (white box). Distinct features include: an additional N-terminal Ig domain (C0) and phosphorylation sites (orange lines) in cMyBP-C; absence of phosphorylation in fMyBP-C; and unique phosphorylation patterns in sMyBP-C, including three sites in the PA region and one in the M-domain. A linker between C4 and C5 has been identified in both cMyBP-C and sMyBP-C. Additionally, the cardiac isoform contains a 28-residue insertion within the C5 domain (indicated by a green vertical line), further distinguishing them structurally. Understanding these isoform-specific structural and regulatory differences is critical for elucidating the molecular basis of skeletal myopathies, such as distal arthrogryposis, associated with mutations in MYBPC1 and MYBPC2. Adapted from Song et al. (2023).

Protein level studies of MyBP-C in health and disease

Protein-level investigations into cMyBP-C have revealed critical insights into how its N-terminal domains (C0–C2) modulate sarcomeric function through interactions with actin and myosin. Quantitative binding assays, including cosedimentation and fluorescence lifetime (FLT), demonstrate that C0–C2 binds filamentous actin (F-actin) with micromolar affinity, a process enhanced by tropomyosin (Tm) and significantly reduced upon phosphorylation by protein kinase A (PKA) (Bunch et al 2021a, Bunch et al 2021b). Fluorescence resonance energy transfer (FRET) and FLT-detected FRET biosensor assays have validated these interactions in high-throughput screening formats, enabling identification of small-molecule inhibitors such as suramin, NF023, and aurintricarboxylic acid that block unphosphorylated and phosphorylated C0–C2 binding to actin (Bunch et al., 2021b; Bunch et al., 2023). Mass spectrometry revealed robust phosphorylation at Ser275, Ser284, and Ser304 within the M-domain, with negligible modification at Ser311(Bunch et al.,2019, Supplementary Figure 2). Structural and functional studies further demonstrate that phosphorylation disrupts the tri-helix bundle (THB) within the M-domain, altering its environment and weakening binding to actin and myosin (Wong et al.,2024). Cooperative binding of C0–C2 across multiple actin monomers was observed via transient phosphorescence anisotropy (TPA), with phosphorylation or specific HCM mutations such as R282W diminishing this effect (Bunch et al., 2019). Tropomyosin presence augments C0–C2 affinity for actin and reveals mutation-specific alterations in this cooperative interaction (Bunch et al., 2021a). Furthermore, phosphorylation not only reduces actin binding but also disrupts interactions with myosin S2, especially through conformational shifts in the THB domain (Wong et al.,2024). This means that MyBP-C N-terminal domains can bind either actin or myosin in a phosphorylation-sensitive manner, although it remains unclear what is the distribution of these dynamic interactions in muscle cells. The dual regulatory roles of MyBP-C in activating and inhibiting thin filament interactions under different conditions are illustrated in Figure 5. Together, these findings delineate a dynamic regulatory mechanism by which phosphorylation and disease mutations modulate cMyBP-C's role in contraction and relaxation, establishing C0–C2 as both a functional modulator and a promising therapeutic target in cardiomyopathy and heart failure. Given the successful application of these cardiac constructs to mechanistic studies in health and disease, the long-term goals stemming from the current project aim to apply a similar strategy to study skeletal MyBP-C with an emphasis on the sMyBP-C due to its involvement in the DA disorder.

Seminal work from the Harris lab has contributed extensively to our understanding of cardiac MyBP-C, particularly its N-terminal domains, phosphorylation regulation, and disease-associated mutations findings that can be effectively leveraged to inform skeletal MyBP-C research. Risi et al. (2021) demonstrated that the C2 domain of cMyBP-C interacts directly with filamentous actin (F-actin), providing molecular evidence for thin filament regulation beyond the classic myosin-centric model. This interaction is modulated by phosphorylation events within the M-domain, with specific serine residues such as Ser275 and Ser284 playing pivotal roles in switching actin affinity (Shaffer et al., 2009). Further structural studies using cryo-electron microscopy revealed how the M-domain interacts with both actin and tropomyosin in a calcium-sensitive manner, emphasizing its regulatory flexibility (Risi et al., 2022). Mutations within the tri-helix bundle, such as those studied by van Dijk et al. (2018), impair actomyosin dynamics, delaying cardiac relaxation and mirroring pathophysiological changes seen in hypertrophic cardiomyopathy (HCM). While the sequence and domain architecture of cardiac and skeletal MyBP-C paralogs is largely conserved, unique regions in each isoform likely contribute to their specialized functional roles. These structural and mechanistic insights from cMyBP-C studies can thus be thoughtfully adapted to elucidate the role of sMyBP-C in muscle disorders such as distal arthrogryposis (DA), while also accounting for paralog-specific distinctions.

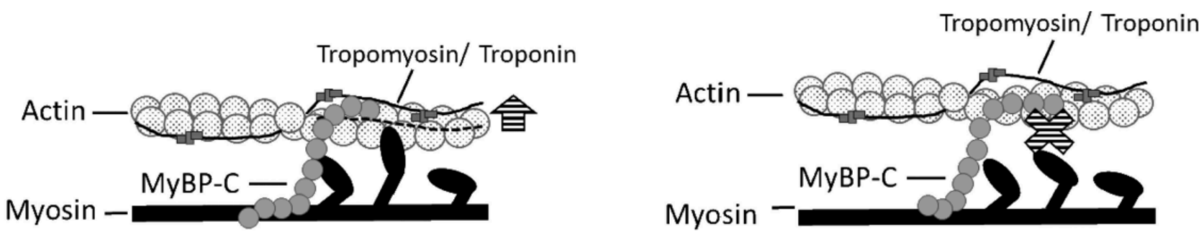

Figure 5: Dual Role of MyBP-C in Thin Filament Regulation During Muscle Contraction

Illustration of two proposed mechanisms by which myosin-binding protein C (MyBP-C) modulates actin–myosin interactions via thin filament regulation. (a) MyBP-C directly activates the thin filament under low calcium conditions by displacing tropomyosin, promoting actin exposure and facilitating cross-bridge formation with myosin. (b) Conversely, MyBP-C may sterically block myosin S1 binding sites on actin, preventing cross-bridge formation even when tropomyosin is in the permissive (open) state. This dual functionality, activation versus inhibition, highlights the context-dependent regulatory role of MyBP-C. In skeletal muscle, spatial and temporal restriction of MyBP-C localization within the sarcomere may balance these opposing actions to fine-tune contractility. Adapted from van Dijk et al. (2014).

sMyBP-C missense mutations cause DA muscle disorder

Distal arthrogryposis type 1 (DA-1) is a congenital musculoskeletal disorder characterized by contractures of the distal limbs, particularly affecting the hands and feet. Clinical features include clubfoot, vertical talus, camptodactyly, overriding fingers, and ulnar deviation. Genetic studies have identified mutations in the MYBPC1 gene, which encodes sMyBP-C, as a cause of DA. Representative clinical manifestations observed in affected individuals from a multigenerational family with MYBPC1 mutations are illustrated in Figure 6. Two specific missense mutations, W236R and Y856H, were found in these individuals, implicating functionally significant regions of the sMyBP-C protein in disease pathology. The W236R mutation is located in the unique M-domain motif between the C1 and C2 immunoglobulin (Ig) domains, a region known to interact with the S2 segment of myosin. The Y856H mutation lies within the C8 domain, which is part of the protein's C-terminal region and is known to associate with both the light meromyosin (LMM) region of myosin and with titin (Gurnett et al., 2010). Thus, mutations can occur both in the N-terminal and C-terminal ends of the sMyBP-C molecule that result in DA.

Figure 6: Clinical Manifestations of Distal Arthrogryposis Type 1 Associated with MYBPC1 Mutations

Photographic documentation of lower and upper limb contractures in affected members of a multigenerational family with autosomal dominant distal arthrogryposis type 1 (DA-1) caused by MYBPC1 mutations. (A) Lower limb deformities include bilateral and unilateral congenital clubfoot in the proband and his great aunt, shown before and after corrective surgery. (B) Hand contractures feature camptodactyly, ulnar deviation, and finger extension abnormalities. These clinical phenotypes reflect the distal joint contractures characteristic of DA-1 and underscore the pathological relevance of MYBPC1 missense mutations (W236R and Y856H) in skeletal muscle function. Adapted from Gurnett et al. (2010).

While the genetic basis of DA is well established, the mechanistic consequences of these mutations at the molecular and cellular levels remain poorly understood. sMyBP-C is a sarcomeric protein that contributes to the structural stability of thick filaments and plays a regulatory role in modulating actin-myosin interactions during muscle contraction (Song et al., 2023; Geist et al., 2018). The protein also serves as a substrate for post-translational modifications, such as phosphorylation by PKA and protein kinase C (PKC), which may fine-tune its function during contractile activity (Ackermann & Kontrogianni-Konstantopoulos, 2011). However, it remains unclear how mutations such as W236R and Y856H affect these functional properties and lead to the specific contractile defects observed in DA.

To elucidate these mechanisms, various biochemical and biophysical approaches can be employed using purified wild-type and mutant sMyBP-C fragments. For instance, protein–protein interaction assays, such as co-sedimentation, isothermal titration calorimetry (ITC), and microscale thermophoresis (MST), can quantify potential differences in binding affinity between sMyBP-C and its sarcomeric partners (Bunch et al., 2021b). Additionally, FRET can detect intermolecular binding or intramolecular conformational changes in the protein structure caused by mutations (Wong et al,2024; Kanassatega et al.,2022). Given the regulatory role of phosphorylation in sMyBP-C function, in vitro kinase assays using PKA or PKC could reveal whether these mutations impair phosphorylation efficiency or modify the accessibility of phosphorylation sites (Ackermann & Kontrogianni-Konstantopoulos, 2011). Techniques such as circular dichroism (CD) spectroscopy can also provide insights into structural stability and folding differences between wild-type and mutant proteins. Furthermore, cell-based expression systems, including skeletal muscle cell lines or iPSC-derived myotubes, can be utilized to evaluate the impact of these mutations on sarcomere assembly, protein localization, and contractile behavior (Song et al., 2023).

Testing established cardiac MyBP-C expression and purification methods on skeletal MyBP-C

This study leverages the existing knowledge and methodologies developed for cMyBP-C to design a comparable protein construct for sMyBP-C. The goal is to eventually investigate how mutations and phosphorylation status influence actin binding and regulation, potentially illuminating the molecular mechanisms underpinning DA. Such insights may lay the groundwork for therapeutic screening approaches or molecular interventions targeting sMyBP-C dysfunction. Therefore, I hypothesize that the established methodologies that have been successful for expression and purification of cardiac MyBP-C N-terminal fragments will be an effective starting point for using skeletal MyBP-C N-terminal fragment constructs. Below, key strategies and considerations for effective recombinant protein expression and purification are reviewed, followed by Results of an initial testing of the expression and purification of fast skeletal, slow skeletal long, and slow skeletal short MyBP-C N-terminal fragments and Discussion of this experiment with implications for future directions of protocol optimizations and studies using these proteins.

Strategies and considerations for effective recombinant protein expression and purification

The production of skeletal MyBP-C N-terminal fragment “C1-C2” containing PAL-C1-M-C2 domains involves cloning this region into an expression vector for use with an E. coli bacterial expression system. Following the culture growth, the C1-C2 fragment is purified from the cell homogenate by its His-tag using nickel affinity column. The tag is then cleaved using tobacco etch virus protease (TEVp) to leave only the human MyBP-C N-terminal protein. The purification process, yield, and final purity is monitored by visualizing the migration and separation of protein content by SDS-PAGE. While each one of these steps had been optimized for cMyBP-C, there are considerations for improvement on the protein quality control and reproducibility for downstream experiments that are highly sensitive to small perturbations in structure and/or function. For these considerations, I will review and discuss these processes and options for alternative strategies to address potential pitfalls and areas to improve protein yield, purity, and quality for biophysical studies of skeletal MyBP-C structural dynamics, post-translational modifications, and DA mutations.

A. MyBP-C fragment expression vector

The C1-C2 fragment of the human MyBP-C sequence is cloned into the pET45b vector and expressed in a bacterial host cell of Escherichia coli. The plasmid map highlighting the vector elements used in this construct is shown in Figure 7. pET45b is a plasmid vector commonly used in molecular biology research for the expression of recombinant proteins in bacterial hosts, particularly E. coli. Protein expression in bacteria is tightly regulated and modulated by sophisticated genetic mechanisms utilizing components of the lac operon and T7 RNA polymerase derived from bacteriophage T7.

pET45b contains a bacterial origin of replication, which allows it to replicate autonomously in E. coli cells. This feature ensures the maintenance of the plasmid during bacterial growth and division. It carries a selectable marker gene, such as an antibiotic resistance gene, in this case to carbenicillin. This gene allows for the selection of bacterial cells that have successfully taken up the plasmid when the bacteria are grown in media containing carbenicillin. pET45b provides a multiple cloning site (MCS) downstream of the promoter. The MCS contains unique restriction enzyme recognition sites where the gene of interest can be inserted. Specifically, DNA encoding the protein to be expressed, in this case, the C1-C2 domain, can be inserted into the plasmid at these sites. This insertion allows for the expression of the recombinant protein under the control of the T7lac promoter.

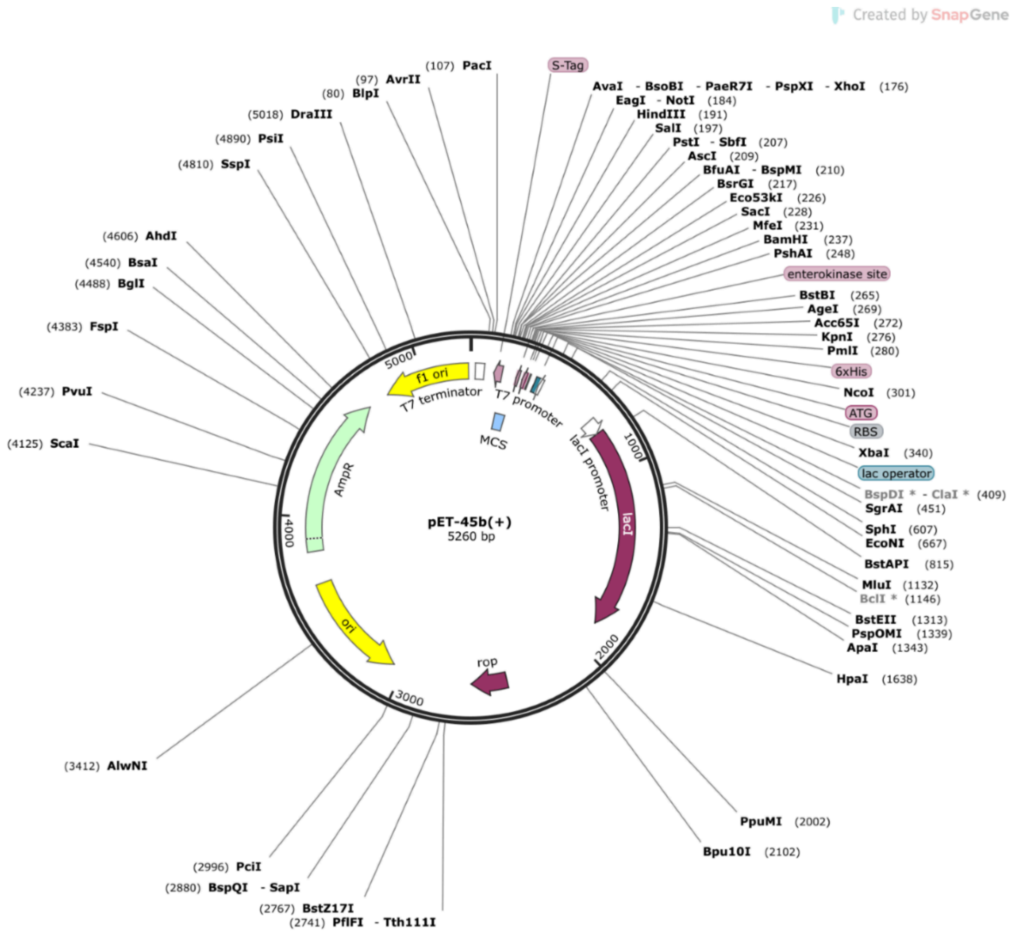

Figure 7: Map of pET-45b(+) Expression Vector

Schematic representation of the pET-45b(+) plasmid used for protein expression. The vector includes a T7 promoter for high-level expression, a multiple cloning site (MCS) for gene insertion. Additional elements such as the ampicillin resistance gene (AmpR) for selection and the origin of replication (Ori) ensure plasmid maintenance in host cells. This figure provides an overview of the genetic elements facilitating controlled expression of recombinant proteins. Adapted from SnapGene (n.d.).

pET45b also contains a T7lac promoter upstream of cloning sites used to incorporate the MyBP-C fragment coding sequences. The vector utilizes the T7lac promoter system for strong and tightly controlled gene expression. In this system, there is a T7 promoter that can be acted upon by T7 RNA polymerase to drive high-level expression of the gene of interest. This promoter drives the transcription of the MyBP-C fragment inserted into the cloning site in bacterial cells, allowing for the expression of the recombinant protein. There is a lac operator (LacO) sequence just downstream of the T7 promoter that can be acted upon by the lac repressor (LacI) protein to block transcription of the T7 promoter. The plasmid also carries the LacI gene. The LacI protein binds to LacUV5 promoter repressing expression of the T7 RNA polymerase gene. It also binds to the T7lac promoter on the PET vector blocking transcription of the MyBP-C fragment gene by any T7 RNA polymerase that may be made due to leaky expression. Therefore, expression of C1-C2 is tightly controlled at two levels, the T7 RNA polymerase and the C1-C2 expression genes, by lac repressor.

One way to induce protein (C1-C2) expression from the pET45b vector is using isopropyl β-D-1-thiogalactopyranoside (IPTG). It is a chemical compound commonly used to induce the expression of proteins. IPTG is a molecular mimic of lactose. It resembles lactose but cannot be metabolized by E. coli. When IPTG is added to the growth medium of E. coli, it enters the bacterial cells. Inside the cell, IPTG binds to the LacI repressor protein, preventing it from binding to the T7lac and the LacUV5 promoters. As a result, T7 RNA polymerase is expressed and can now transcribe the gene expressing MyBP-C. Another method for inducing MyBP-C fragment expression utilizes autoinduction medium, designed to achieve high levels of protein production. This medium includes key components to optimize the overall expression process. Glycerol serves as the primary carbon source, promoting bacterial growth to high density. It also contains minimal levels of glucose and amino acids, which initially suppress lactose uptake and allolactose production until the culture reaches the stationary phase. At this point, lactose is utilized and converted into allolactose. Allolactose then act on the repressor, relieving repression of both the T7 RNA polymerase and MyBP-C fragment genes within the bacteria. Here, we selected to use the pET45b vector, although alternative DNA plasmids area also suitable, and the autoinduction media approach for convenience.

B. Cell line

BL21(DE3) is the specific E. coli cell line utilized for protein expression, wherein the T7 RNA polymerase gene is encoded by the lambda DE3 prophage integrated into the chromosome, regulated by the lacUV5 promoter under lac operon control. This cell line is also deficient in two proteases, Lon protease and OmpT protease, and this further enhances the production of proteins derived from cloned genes. Other lines are available to deal with challenges to express proteins, although this is not expected for skeletal C1-C2 given the success with the similar cardiac paralog. The nucleotide and protein sequences for the skeletal MyBP-C constructs are presented in Figures 8–13. The slow short construct (hsskC1C2_S_11cHT) is shown in Figures 8 and 9, while the fast skeletal construct (hfskC1C2_11cHT) appears in Figures 10 and 11. The long slow variant (hsskC1C2_L_11cHT), included for comparative expression analysis, is depicted in Figures 12 and 13.

For plasmid maintenance, NEB5 alpha cells are employed. These cells possess three beneficial traits: they are recombination deficient, preventing DNA rearrangements involving repeated sequences; they are endonuclease deficient, ensuring the quality of plasmid preparations by preventing plasmid degradation during handling; and they carry a mutation (hsdR17) that eliminates endonuclease activity against foreign DNA, facilitating the manipulation and stability of cloned genetic material. These cells should also be suitable for these purposes without expected issues.

C. Growth Temperature

Expressing proteins at lower temperatures during autoinduction helps improve folding and solubility, especially for complex or multi-domain proteins. At higher temperatures, such as 37°C, cellular processes accelerate, leading to faster protein synthesis. However, this rapid pace can overwhelm the cell's folding machinery, resulting in increased misfolding and aggregation. By reducing the temperature to around 25°C or less, protein synthesis slows down, allowing more time for proper folding and giving the cell's chaperones a better opportunity to assist in stabilizing complex proteins. Additionally, lower temperatures reduce the activity of proteases, which can degrade misfolded proteins, further enhancing the yield of stable, functional proteins. Proteins are also less prone to forming insoluble inclusion bodies at lower temperatures, resulting in a higher proportion of soluble, functional protein in the final extract. While 25°C as performed here is sufficient for soluble skeletal MyBP-C, we have since found optimized conditions with 18°C growth.

D. Optimizing codon sequences for E. coli expression

Optimizing codon usage is a strategy for enhancing protein expression in E. coli. In E. coli, the efficiency of protein synthesis can be significantly influenced by the codon usage of the gene being expressed. The process of codon optimization involves modifying the DNA sequence of a gene, in this case encoding human MyBP-C, to align its codon usage with that preferred and most used for native proteins by E. coli. This approach aims to reduce issues related to rare codons, which can impede translation by causing delays or errors due to insufficient availability of corresponding tRNAs.

This strategy can lead to substantial improvements in protein expression. This is achieved by replacing rare codons with those that are more frequently used by E. coli. By doing so, the translation machinery can more efficiently read the mRNA and produce the desired protein. Moreover, optimizing codon usage often results in enhanced protein solubility, as it facilitates smoother translation, which reduces the risk of incomplete or misfolded proteins that tend to aggregate. Properly folded proteins are more likely to remain soluble, which is critical for their functional and structural integrity.

Additionally, codon usage optimization not only impacts the efficiency of translation but also affects mRNA stability and folding. Changes in the mRNA sequence can influence its secondary structure, which in turn can affect how well it is translated. This dual impact on both translation efficiency and mRNA stability can lead to increased yields of soluble protein. For human sC1-C2 and fC1-C2, the original pET45b vectors containing their E. coli optimized sequences were designed and ordered from

GenScript. It is expected that the codon optimization will improve protein expression for solubility, yield, and aid in optimized purification.

E. Protein Purification

A protein purification process was used to isolate the MyBP-C protein fragment from other proteins produced by the E. coli expression system, ensuring high protein purity for downstream applications. The first step was to resuspend the bacterial pellet in lysis buffer, followed by lysing the cells using a high-pressure homogenizer. This device is operated by pushing the biological suspension through a tight space and exposing it to a strong shear force. The cells break apart due to the high-pressure impact of the fluid against the valve surface, which disrupts the cellular membrane without damaging the protein. The homogenization process in the Emulsiflex (C3 model by Avestin, Inc.) instrument, in our lab's hands, provides more homogenous mixture, reproducible protein quality, and yield than other more labor-intensive methods such as cycles of cell freeze-thaw and manual homogenization in lysis buffer.

Following homogenization in the Emulsiflex, the lysed mixture is centrifuged to remove cell debris, organelles, insoluble proteins, and other particulate matter, leaving a crude mixture of soluble proteins including the MyBP-C fragment and nucleic acids in the supernatant. To purify the MyBP-C fragment, poly-histidine (6xHis) affinity tags were used via immobilized metal-affinity chromatography (IMAC). IMAC is based on the interaction between a Ni^{2+} metal ion immobilized on a resin matrix and the histidine residues of the affinity tag. Nickel-nitrilotriacetic acid (Ni^{2+}-NTA) was used as the matrix, which securely coordinates the nickel ion through four sites while leaving two sites exposed to interact with His residues. This enabled selective binding of the His-tagged protein, allowing for typically 90-95% purity for cMyBP-C with high yield (~20-30 mg per 500 ml of culture). For the MyBP-C fragments that were purified, a 6x-His tag was added to the C-terminus of the construct. However, we note that subsequent work has determined that placing the His-tag at the N-terminus of skeletal MyBP-C results in better protein quality that behaves more reproducibly in downstream functional assays, such as fluorescence-based actin-binding of skeletal MyBP-C. Another consideration is to include some length of a native linker between the domain and tag. Typically, lengths of ~8-11 residues have been used to ensure availability of the His tag.

Two IMAC buffers are prepared. The IMAC Equilibration Buffer contains 50 mM Na_2HPO_4, 200 mM NaCl, 25 mM imidazole, and 10% glycerol, adjusted to pH 8.0 with HCl. In this buffer, sodium phosphate (Na_2HPO_4) acts as a buffering agent to maintain a stable pH, sodium chloride (NaCl) provides ionic strength to minimize nonspecific electrostatic interactions, imidazole at a low concentration (25 mM) reduces weak binding of contaminant proteins, and glycerol stabilizes the protein during purification. The IMAC Elution Buffer is formulated similarly but contains 250 mM imidazole instead of 25 mM. The increased imidazole concentration competes with the His-tag for binding to the Ni^{2+} resin, effectively eluting the bound protein.

Binding of the poly-histidine-tagged proteins is performed using a column-based procedure. The 6x-His-tagged protein has a higher affinity for the Ni^{2+}-NTA resin than proteins that bind nonspecifically. By using the minimum effective volume of resin, most of the available binding sites are occupied by the target protein, which reduces nonspecific interactions. Sodium chloride up to 200 mM and low concentrations of imidazole (up to 25 mM) in the binding buffer further minimize nonspecific protein binding by increasing solubility and weakening low-affinity interactions. Following binding of the tagged protein, the column is washed with low imidazole to remove nonspecific proteins that bind weakly to the column.

To elute the His-tagged MyBP-C fragment from the nickel-affinity column, a high concentration of imidazole (250 mM) is used. Imidazole, a structural analog of histidine, competes with the His residues of the 6x-His tag for binding to the Ni^{2+} ions on the column, thereby releasing the bound protein into the elution fractions. This method is preferred over low pH elution, as it minimizes the risk of protein denaturation or activity loss.

Following elution, tobacco etch virus protease (TEVp) is added to cleave the 6x-His tag located at the C-terminal end of the MyBP-C fragment. To facilitate efficient cleavage by TEVp, dithiothreitol (DTT) and ethylenediaminetetraacetic acid (EDTA) are added to the reaction. DTT acts as a reducing agent, maintaining cysteine residues in a reduced state and preventing the formation of disulfide bonds that could interfere with protease activity or cause protein aggregation. EDTA acts as a chelating agent, binding divalent metal ions such as Mg^{2+} or Ca^{2+} that might otherwise inhibit protease activity or promote protein degradation through metal-dependent proteases.

After cleavage, the protein sample undergoes dialysis to remove excess imidazole, DTT, EDTA, cleaved His-tags, and TEVp protease. Dialysis is performed using a semi-permeable membrane, allowing small molecules and buffer components to diffuse out while retaining the larger protein molecules. This process enables buffer exchange into a suitable final storage buffer and contributes to obtaining a purified, tag-free MyBP-C fragment ready for downstream structural or biochemical assays. As noted earlier, we have now found better protein quality with N- rather C-terminal tags.

Finally, gel electrophoresis is a technique used to separate mixtures of DNA, RNA or proteins based on molecular size using an electrical field through a gel. Here, Coomassie blue staining is used to detect the abundance of proteins loaded on the gel after electrophoresis to evaluate protein overexpression, purity, yield as well as potential issues with correct size of proteins, breakdown, and contaminating proteins to the skeletal MyBP-C protein of interest.

Gene expression from plasmid_

TCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGG
GCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTG
GTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTG
TCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTA
ATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATG
CCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTT
CCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCG
CCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGA
TGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGG
TCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATC
CTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCAC
CGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAG
TTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGG
AGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGA
ATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGG
CCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATA
ACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACC
GCGAAAGGTTTTGCGCCATTCGATGGTGTCAAGGAGATAATACGACTCACTATAGGGGAATTG
TGAGCGGATAACAATTCCAAGGAGAATGCCAGAGCCGACCAAGAAAGAGGAAAACGAAGTG
CCGGCGCCGGCTCCGCCGCCGGAGGAACCGAGCAAGGAGAAAGAAGCGGGCACCACCCCG
GCGAAGGATTGGACCCTGGTTGAGACCCCGCCGGGTGAAGAGCAGGCGAAACAAAACGCGA
ACAGCCAGCTGAGCATCCTGTTCATTGAAAAGCCGCAAGGTGGCACCGTGAAAGTTGGCGAG
GACATCACCTTTATTGCGAAGGTGAAAGCGGAGGACCTGCTGCGTAAGCCGACCATCAAATG
GTTCAAGGGTAAATGGATGGACCTGGCGAGCAAGGCGGGCAAACACCTGCAGCTGAAAGAA
ACCTTCGAGCGTCACAGCCGTGTTTACACCTTTGAGATGCAAATCATTAAGGCGAAAGATAAC
TTTGCGGGTAACTACCGTTGCGAAGTGACCTATAAGGACAAATTCGATAGCTGCAGCTTTGAC
CTGGAAGTTCACGAGAGCACCGGCACCACCCCGAACATCGATATTCGTAGCGCGTTCAAGCG
TAGCGGTGAAGGCCAGGAAGACGCGGGCGAGCTGGATTTTAGCGGCCTGCTGAAGCGTCGTG
AAGTTAAACAGCAAGAAGAGGAACCGCAGGTGGACGTTTGGGAGCTGCTGAAGAACGCGAA
ACCGAGCGAGTACGAAAAAATCGCGTTCCAATATGGTATTACCGACCTGCGTGGTATGCTGAA
GCGTCTGAAACGTATGCGTCGTGAGGAAAAGAAAAGCGCGGCGTTTGCGAAGATCCTGGACC
CGGCGTACCAGGTGGATAAAGGTGGCCGTGTTCGTTTTGTGGTTGAGCTGGCGGACCCGAAG
CTGGAAGTGAAGTGGTACAAAAACGGTCAAGAAATCCGTCCGAGCACCAAGTATATTTTCGA
GCACAAAGGCTGCCAGCGTATCCTGTTTATTAACAACTGCCAAATGACCGACGATAGCGAGTA
CTATGTTACCGCTGGTGACGAAAAGTGCAGCACCGAGCTGTTCGTGCGTGAACCGCCGATTAT
GGTTACCAAACAGCTGGAGAACCTGTATTTTCAAGGCCACCACCACCACCACCACTAA

His-tag TEVp Start codon Stop codon MyBP-C T7 promoter Lac operator Lac I promoter Lac I
Ribosome Binding Site (RBS)

Figure 8: Nucleotide Sequence for Human Slow Short Skeletal With a C-terminal His-tag (hsskC1C2_S_11cHT) Within the Plasmid Vector

Highlighted in color are (see key above) His-tag, TEVp site, Start codon, Stop codon, MyBP-C fragment C1-C2, T7 promoter, Lac operator, Lac I promoter, RBS).

PEPTKKEENEVPAPAPPPEEPSKEKEAGTTPAKDWTLVETPPGEEQAKQNANSQLSILFIEKPQGGT
VKVGEDITFIAKVKAEDLLRKPTIKWFKGKWMDLASKAGKHLQLKETFERHSRVYTFEMQIIKA
KDNFAGNYRCEVTYKDKFDSCSFDLEVHESTGTTPNIDIRSAFKRSGEGQEDAGELDFSGLLKRR
EVKQQEEEPQVDVWELLKNAKPSEYEKIAFQYGITDLRGMLKRLKRMRREEKKSAAFAKILDPA
YQVDKGGRVRFVVELADPKLEVKWYKNGQEIRPSTKYIFEHKGCQRILFINNCQMTDDSEYYVT
AGDEKCSTELFVREPPIMVTKQLENLYFQG

Figure 9: Protein Sequence of hsskC1C2_S_11cHT

Orange is the PAL region, pink is C1 domain, yellow is M-domain, blue is C2 domain, black is the end linker and green is the TEV recognition site.

TCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGG
GCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTG
GTTCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCTGCCCCAGCAGGC
GAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTA
TCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGC
CCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTT
GCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAA
TTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTA
ATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTC
GCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAA
ATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGAT
AGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTT
CGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCGCGAGATT
TAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCAATC
AGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCC
ATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGGCCTGGTTCACCACGCGG
GAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACTGGTTTCACA
TTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCGCGAAAGGTTTTGCGCC
ATTCGATGGTGTCTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCAAGGAG
AATGCCAGAAGCGAAGCCGGCGGCGAAGAAAGCGCCGAAGGGTAAAGATGCGCCGAAGGG
TGCGCCGAAAGAAGCGCCGCCGAAGGAAGCGCCGGCGGAGGCTCCTAAAGAAGCGCCGCCG
GAGGATCAAAGCCCGACCGCTGAGGAACCGACCGGCGTGTTCCTGAAGAAGCCGGACAGCG
TTAGCGTGGAGACCGGCAAAGATGCGGTGGTTGTGGCGAAGGTGAACGGCAAAGAACTGCC
GGACAAGCCGACCATCAAATGGTTTAAGGGTAAATGGCTGGAGCTGGGTAGCAAGAGCGGCG
CGCGTTTCAGCTTTAAAGAGAGCCACAACAGCGCGAGCAACGTTTACACCGTGGAACTGCAC
ATCGGCAAGGTTGTGCTGGGTGATCGTGGCTACTACCGTCTGGAAGTGAAGGCGAAAGACAC
CTGCGATAGCTGCGGTTTTAACATTGATGTGGAAGCGCCGCGTCAGGATGCGAGCGGTCAAA

GCCTGGAGAGCTTCAAACGTACCAGCGAAAAGAAAAGCGACACCGCTGGTGAGCTGGATTT
TAGCGGCCTGCTGAAGAAACGTGAAGTTGTGGAGGAAGAGAAGAAAAAGAAAAAGAAAGA
CGATGACGATCTGGGTATCCCGCCGGAAATTTGGGAGCTGCTGAAGGGCGCGAAGAAAAGCG
AATACGAGAAAATCGCGTTCCAGTATGGTATTACCGACCTGCGTGGTATGCTGAAGCGTCTGA
AGAAAGCGAAAGTTGAGGTGAAGAAAAGCGCGGCGTTTACCAAGAAACTGGACCCGGCGTA
TCAAGTTGATCGTGGTAACAAGATCAAACTGATGGTGGAAATTAGCGACCCGGATCTGACCCT
GAAGTGGTTCAAAAACGGTCAGGAGATCAAGCCGAGCAGCAAATACGTTTTTGAAAACGTG
GGCAAGAAACGTATCCTGACCATTAACAAGTGCACCCTGGCGGACGATGCGGCGTATGAGGT
TGCGGTGAAGGACGAAAAATGCTTCACCGAGCTGTTTGTTAAAGAACCGCCGGTTCTGATTG
TGACCCCGCTGGAAAACCTGTACTTCCAGGGTCACCACCACCACCACCACTAA

Histag TEVp Start codon Stop codon MyBP-C T7 promoter Lac operator Lac I promoter Lac I Ribosome Binding Site (RBS)

Figure 10: Nucleotide Sequence for Human Fast Skeletal With a C-terminal His-tag (hfskC1C2_11cHT) Within the Plasmid Vector

Highlighted in color are (see key above): His-tag, TEVp site, Start codon, Stop codon, MyBP-C fragment C1-C2, T7 promoter, Lac operator, Lac I promoter, RBS.

PEAKPAAKKAPKGKDAPKGAPKEAPPKEAPAEAPKEAPPEDQSPTAEEPTGVFLKKPDSVSVETG
KDAVVVAKVNGKELPDKPTIKWFKGKWLELGSKSGARFSFKESHNSASNVYTVELHIGKVVLGD
RGYYRLEVKAKDTCDSCGFNIDVEAPRQDASGQSLESFKRTSEKKSDTAGELDFSGLLKKREVVE
EEKKKKKKDDDDLGIPPEIWELLKGAKKSEYEKIAFQYGITDLRGMLKRLKKAKVEVKKSAAFT
KKLDPAYQVDRGNKIKLMVEISDPDLTLKWFKNGQEIKPSSKYVFENVGKKRILTINKCTLADDA
AYEVAVKDEKCFTELFVKEPPVLIVTPLENLYFQG

Figure 11: Protein Sequence of hfskC1C2_11cHT

Orange is the PAL region, pink is C1 domain, yellow is M-domain, blue is C2 domain, black is the end linker and green is the TEV recognition site.

TCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCACCAGTGAGACGG
GCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTG
GTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTG
TCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCCGGACTCGGTA
ATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATG
CCCTCATTCAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTT
CCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCG
CCGAGACAGAACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGA
TGCTCCACGCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGG

TCAGAGACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATC
CTGGTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCAC
CGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAG
TTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGG
AGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGA
ATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGAAACGTGGCTGG
CCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATA
ACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACC
GCGAAAGGTTTTGCGCCATTCGATGGTGTCTAATACGACTCACTATAGGGGAATTGTGAGCGG
ATAACAATTCCAAGGAGAATGCCAGAGCCGACCAAGAAAGAGGAAAACGAAGTGCCGGCGC
CGGCTCCGCCGCCGGAGGAACCGAGCAAGGAGAAAGAAGCGGGCACCACCCCGGCGAAGG
ATGAGGAAGAGGTGAGCCCGCCGAGCGCGCTGCCGCCGGGTCTGGGCAGCCGTGCGCTGGA
ACGCAAGGACAGCGATTGGACCCTGGTTGAGACCCCGCCGGGTGAAGAGCAGGCGAAACAA
AACGCGAACAGCCAGCTGAGCATCCTGTTCATTGAAAAGCCGCAAGGTGGCACCGTGAAAGT
TGGCGAGGACATCACCTTTATTGCGAAGGTGAAAGCGGAGGACCTGCTGCGTAAGCCGACCA
TCAAATGGTTCAAGGGTAAATGGATGGACCTGGCGAGCAAGGCGGGCAAACACCTGCAGCTG
AAAGAAACCTTCGAGCGTCACAGCCGTGTTTACACCTTTGAGATGCAAATCATTAAGGCGAA
AGATAACTTTGCGGGTAACTACCGTTGCGAAGTGACCTATAAGGACAAATTCGATAGCTGCAG
CTTTGACCTGGAAGTTCACGAGAGCACCGGCACCACCCCGAACATCGATATTCGTAGCGCGTT
CAAGCGTAGCGGTGAAGGCCAGGAAGACGCGGGCGAGCTGGATTTTAGCGGCCTGCTGAAG
CGTCGTGAAGTTAAACAGCAAGAAGAGGAACCGCAGGTGGACGTTTGGGAGCTGCTGAAGA
ACGCGAAACCGAGCGAGTACGAAAAAATCGCGTTCCAATATGGTATTACCGACCTGCGTGGTA
TGCTGAAGCGTCTGAAACGTATGCGTCGTGAGGAAAAGAAAAGCGCGGCGTTTGCGAAGATC
CTGGACCCGGCGTACCAGGTGGATAAAGGTGGCCGTGTTCGTTTTGTGGTTGAGCTGGCGGA
CCCGAAGCTGGAAGTGAAGTGGTACAAAAACGGTCAAGAAATCCGTCCGAGCACCAAGTATA
TTTTCGAGCACAAAGGCTGCCAGCGTATCCTGTTTATTAACAACTGCCAAATGACCGACGATA
GCGAGTACTATGTTACCGCTGGTGACGAAAAGTGCAGCACCGAGCTGTTCGTGCGTGAACCG
CCGATTATGGTTACCAAACAGCTGGAGAACCTGTATTTTCAAGGCCACCACCACCACCACCAC
TAA

Histag TEVp Start codon Stop codon MyBP-C T7 promoter Lac operator Lac I promoter Lac I RBS

Figure 12: Nucleotid e Sequence for Human Slow Long Skeletal with C-terminal His-tag (hsskC1C2_L_11cHT) Within the Plasmid Vector

Highlighted in color are (see key above): His-tag, TEVp site, Start codon, Stop codon, MyBP-C fragment C1-C2, T7 promoter, Lac operator, Lac I promoter, RBS.

PEPTKKEENEVPAPAPPPEEPSKEKEAGTTPAKDEEEVSPPSALPPGLGSRALERKDSDWTLVETPP
GEEQAKQNANSQLSILFIEKPQGGTVKVGEDITFIAKVKAEDLLRKPTIKWFKGKWMDLASKAG
KHLQLKETFERHSRVYTFEMQIIKAKDNFAGNYRCEVTYKDKFDSCSFDLEVHESTGTTPNIDIRS
AFKRSGEGQEDAGELDFSGLLKRREVKQQEEEPQVDVWELLKNAKPSEYEKIAFQYGITDLRGM
LKRLKRMRREEKKSAAFAKILDPAYQVDKGGRVRFVVELADPKLEVKWYKNGQEIRPSTKYIFE
HKGCQRILFINNCQMTDDSEYYVTAGDEKCSTELFVREPPIMVTKQLENLYFQG

Figure 13: Protein Sequence of hsskC1C2_L_11cHT

Orange is the PAL region, pink is C1 domain, yellow is M-domain, blue is C2 domain, black is the end linker and green is the TEV recognition site.

2. METHODS

In the previous section, the rationale and considerations for methods and protocol choices were described. The following section provides the methodological details of the initial protocol used to test the expression of the skeletal MyBP-C fragments using the established cMyBP-C approach (Bunch et al., 2018; Wong et al., 2024). More recently, two key optimizations have been made to place the cleavable His-tag at the N-terminus of the skeletal C1-C2 and grow proteins at 18ºC.

Transformation of DNA into competent cells

BL21(DE3) competent E. coli cells were thawed, gently mixed and 50 ml of cells was pipetted into a tube for transformation on ice. 5 ml containing 1pg-100ng of skeletal C1-C2 plasmid DNA was added to the cell mixture. Tubes were gently mixed by manual flicking (rather than vortex). The mixture was placed in ice for 30 min and then heat shocked at 42°C for 10 seconds and again placed on ice for 5 min. 950 ml of Super Optimal broth with Catabolite repression (SOC) medium at room temperature (23-25ºC) was added to the mixture and incubated at 37°C for 60 min on a shaker at 250 rpm. Small aliquots of the mixture in the presence of 100 ml carbenicillin were spread on the LB plate and incubated overnight at 37℃.

Bacterial growth to overexpress protein of interest

Bacterial colonies were inoculated in 2 ml of LB media with carbenicillin by shaker incubation at 37℃ at 250rpm. Once the culture turned cloudy due to protein expression, the 2 ml culture was transferred into 50 ml of LB media with carbenicillin and incubated like the 2 ml culture. When the optical density (OD) at 600 nm reached 0.5-1, the bacterial culture was diluted to an OD_{600} of between 0.02-0.05 in 500ml LB. The bacterial culture was subsequently diluted to an OD_{600} of 0.04 in 500 mL of overnight TB medium and incubated at 28 °C overnight until it reached the stationary phase, typically corresponding to an OD_{600} of between 3-10 at 600nm. The next morning the culture was centrifuged for 5 min at 6000xg in a JA-10 rotor at 4°C. The pellet was collected and suspended in 30 ml of phosphate buffered saline (PBS) supplemented with phenylmethylsulphonyl fluoride (PMSF), a protease inhibitor to prevent protein breakdown. The mixture was centrifuged again for 15 min at 4000 rpm and 4°C. The resulting supernatant was immediately discarded. The pellet was saved, weighed, and stored for later use -80°C.

To monitor protein content of cell homogenates during expression protocol, sample aliquots were prepared and diluted in 2X Protein Solubilization Buffer (PSB from Bio-Rad, Inc.). Samples were heated for 10 min at 90℃. Meanwhile, 900 ml of MOPS running buffer was prepared for adding to the gel box. 20 ml of a protein molecular weight ladder and each sample was loaded on the 8% SDS-PAGE gel, which was run for 50 min at 200V. After the run was complete, the gel was transferred from the casting box to a staining box. The gel was stained with 100 ml Coomassie stain for 20 min and 3 times destained with 100 ml for 20 min each and each time rinsed with DI water. The gel is then imaged by Image Studio software (LICORbio).

Cell homogenization

Bacterial pellets were removed from -80℃ storage and resuspended in IMAC equilibration buffer (50 mM Na_2HPO_4, 10% glycerol, 200 mM NaCl, 25 mM imidazole, pH 8 with HCl) supplemented with protease inhibitors (0.1mM PMSF and an EDTA-free protease inhibitor tablet per 50 ml) to a final volume of 40 ml. This bacterial suspension was then homogenized through an Emulsiflex homogenizer to lyse the bacterial cells and release the proteins. 5% polyethyleneimine (PEI) was then added to precipitate nucleic acids. 1.5 ml of protein sample was added to 98 ml of 2x PBS for monitoring protein purity and yield using gel electrophoresis. The remaining protein sample was clarified to remove insoluble proteins by centrifugation at 30,000xg for 20 min at 4°C using a JA-17 rotor.

Protein purification by IMAC

All skeletal MyBP-C C1–C2 constructs used in this study are of human origin and were cloned with a C-terminal 11-residue His-tag (His_6 plus linker residues) and a TEVp (Tobacco Etch Virus protease) cleavage site. A short flexible linker was included between the protein and tag to improve protease accessibility and prevent steric hindrance. These features enabled affinity purification via Ni^{2+}-IMAC and efficient removal of the tag during TEVp cleavage. This standardized design was used for all three constructs: the short slow (sC1C2_S), fast (fC1C2), and long slow (sC1C2_L) to ensure consistent expression and purification performance across experiments.

For column chromatography, 5 ml His_{60} beads (10 ml of a 50% suspension in ethanol) were equilibrated using IMAC equilibration buffer. The beads were spun at 2000 rpm for 4 min at 4°C, and this process was repeated at least three times. Once the sample was centrifuged, the supernatant was poured into the His_{60} beads equilibrated with an IMAC equilibration buffer. To protect against protein degradation, 0.5x protease inhibitor (PI) and 0.1 mM PMSF (a serine protease inhibitor) were added to the samples. Samples were then incubated overnight in a tube with 1 mM sodium azide to prevent bacterial contamination.

The next day the sample was centrifuged for 3 min at 2000 rpm at 4℃. The supernatant was transferred to a new tube, and the pellet was resuspended in an IMAC equilibration buffer. The pellet was then centrifuged repeatedly until the OD of the solution was reduced to less than 0.05. The sample was then eluted using IMAC elution buffer in the column and then OD and final concentration (mg/ml) was measured on the UV/vis spectrophotometer. EDTA and DTT were added to the sample to a final concentration of 0.5 mM and 1 mM, respectively. Then, 20 mg/ml of TEVp was added to cleave the C-terminal His-tag from the purified skeletal MyBP-C fragments, removing extraneous sequences and yielding the native form of the protein for downstream studies. The mixture was kept on a shaker at room temperature until the dialysis was prepared.

For protein dialysis, an IMAC equilibration buffer was added at 10x the volume of the sample to a beaker with 1 mM DTT and 0.5 mM EDTA. The protein sample was transferred to the cellulose dialysis membrane bag and dialyzed in the beaker of the buffer overnight.

A second dialysis was performed on the next day with change of a fresh volume of IMAC equilibration buffer without DTT and EDTA and incubated for another 2 hours. This second dialysis step was necessary to remove residual DTT and EDTA which can interfere with metal affinity chromatography as well as cleaved His-tag peptides and TEVp protease from the sample. Additionally, this step equilibrated the protein into a fresh buffer suitable for final purification on His60 affinity resin, free of chelating agents and reducing agents, which is essential to ensure efficient binding of remaining His-tagged impurities while allowing the untagged MyBP-C fragment to flow through during the final purification step. Following dialysis, the sample was transferred from the tubing into a clean conical tube and then loaded onto a column containing approximately 5 mL of a 50% suspension of His60 Ni^{2+} affinity beads. This volume was selected based on the amount used during the initial purification to ensure sufficient binding capacity for any remaining His-tagged contaminants, such as TEVp and cleaved His-tag peptides. The native MyBP-C fragment, which lacks the His-tag after TEV cleavage, flows through without binding to the resin. The sample was passed over the column multiple times to maximize removal of His-tagged species. Protein concentration was monitored by measuring optical density (OD), and the final purified sample was stored in 60% glycerol at –20°C. Aliquots were collected at each step of the purification process for later analysis by SDS-PAGE to assess purity and track the presence of the target protein.

3. RESULTS

Skeletal MyBP-C N-terminal fragment constructs

Using a purification pipeline originally developed for cardiac MyBP-C (cMyBP-C), we adapted and optimized this workflow to establish initial conditions to efficiently isolate skeletal MyBP-C fragments. The sequences for the slow short, fast skeletal, and slow long constructs used in this study are shown in Figures 8, 10, and 12. This starting protocol allowed for the successful purification of recombinant proteins with high yield and purity, ensuring minimal contamination from non-target proteins.

Protein expression and tag cleavage

SDS-PAGE analysis provides further validation of protein expression and purification (Figure 14). The gel images reveal distinct protein bands corresponding to the expressed fragments, with samples taken at multiple stages, including pre-TEVp (before TEVp cleavage of the His-tag), Post-TEVp (after cleavage of the His-tag), and final purification (F). At the Pre-TEVp stage, densitometry was used to estimate the proportion of the sC1–C2_S fragment band relative to total visible protein. The slow short construct (sC1–C2_S) initially appeared to exhibit nearly 100% target band intensity; however, this was likely an overestimation due to baseline distortion in the densitometry trace (including a slightly negative value in the lower region of the gel lane). Adjusted estimates suggest high expression approximately 90–95% but not complete purity.

In comparison, the fast skeletal construct (hfskC1C2) showed 88% target band proportion, and the slow long construct (hsskC1C2_L) exhibited 81%. These values reflect visual estimates of relative band intensity compared to background and adjacent bands (see Table 1).

Notably, in the Post-Beads fraction for fC1 C2, a visible amount of target protein remained unbound, suggesting that not all of the expressed protein interacted with the His60 resin. This may be due to limited accessibility of the C-terminal His-tag potentially obstructed by local folding or steric hindrance or partial saturation of the resin during binding. The unbound fraction was not reloaded onto the column in this experiment, so the amount of recoverable protein lost at this step remains uncertain. Future optimizations could include altering incubation time, tag positioning, or column capacity to improve binding efficiency.

Following TEVp cleavage (Post-TEVp), a downward shift in band migration confirmed successful His-tag removal, and target band proportions were estimated at 78% for hsskC1C2_S, 92% for hfskC1C2, and 89% for hsskC1C2_L, indicating efficient cleavage without significant protein loss. Notably, hfskC1C2 and hsskC1C2_L displayed improved solubility post-cleavage, whereas hsskC1C2_S showed a slight reduction, potentially due to precipitation or minor degradation. This variability will be closely monitored in future optimization efforts and is illustrated in Figure 15.

Protein Purity

As shown in Table 2, the final protein purity levels confirm the effectiveness of this approach. Among the three constructs, hfskC1C2 exhibited the highest purity at 94%, followed by hsskC1C2_L at 91%, and hsskC1C2_S at 84%. While all three constructs were successfully purified, the variations in purity levels suggest potential differences in expression efficiency, solubility, or stability of the respective proteins. The SDS-PAGE gel in Figure 16 is that many unwanted contaminants were efficiently removed, leading to pure recombinant proteins suitable for downstream applications. Note that further purification by size exclusion chromatography (SEC) is often done in experiments to further reduce contaminants of the incorrect molecular weight to further increase MyBP-C N-terminal domain fragment purity. Additionally, the final protein yield for each construct is reported in Table 3. The observed differences in purity levels could reflect variations in the efficiency of TEVp cleavage, protein stability, or residual non-specific binding during purification. Additionally, the final protein yield further supports the reproducibility and suitable yield of the purification process. Among the three constructs, hsskC1C2_S exhibited the highest final yield at 19.9 mg, followed by hsskC1C2_L at 18.8 mg, and hfskC1C2 at 17.9 mg. The minor variations in protein yield, despite similar culture conditions, suggest that differences in protein expression levels, solubility, and recovery efficiency may have influenced the final output. The correlation between purity and yield highlights the balance between maximizing recovery and maintaining high sample integrity during purification.

Conclusion

Overall, these results validate the robustness of our optimized purification pipeline, demonstrating its ability to yield highly purified sMyBP-C fragments with excellent recovery and efficiency. The effectiveness of the expression and purification workflow is further supported by the SDS-PAGE analysis, summarized across Figures 14–16. We note that following downstream studies of protein structure, function, and quality control, this protocol has been updated and optimized with the two major differences being a change in the His_6-TEVp site to the N-terminus of skeletal C1-C2 and grow-ups were done at 18ºC.

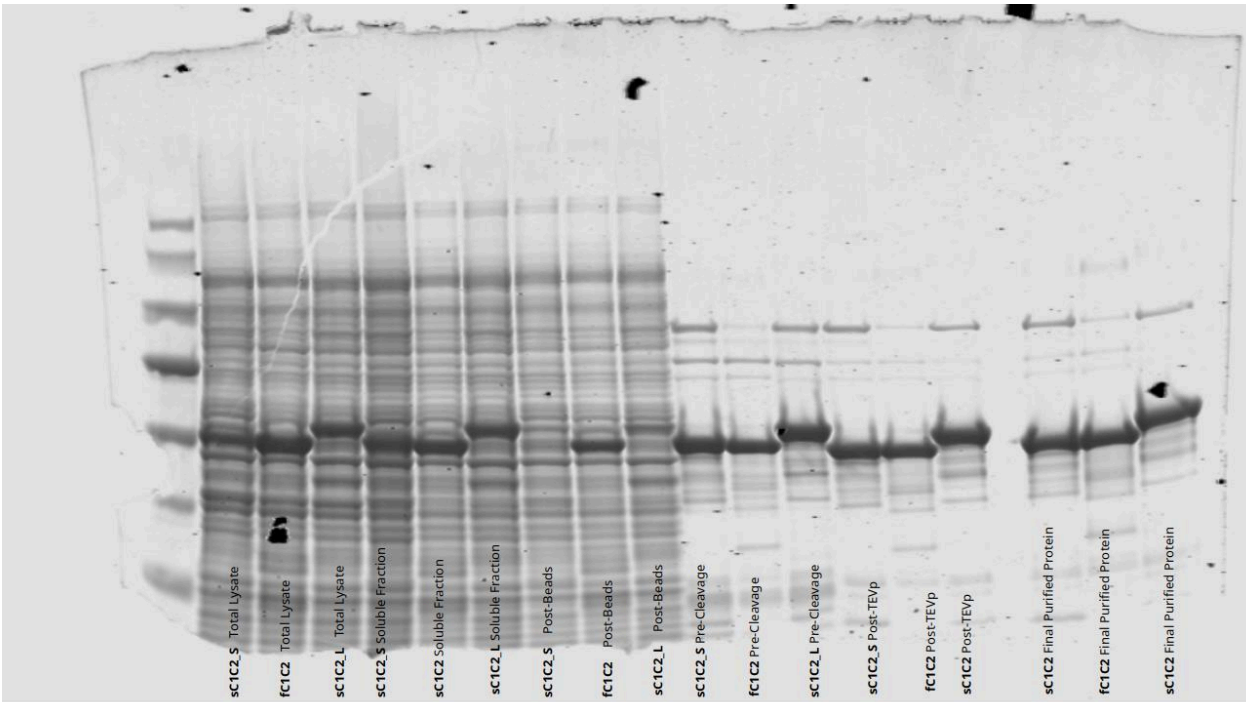

Figure 14: SDS-PAGE Analysis of sMyBP-C C1-C2 Construct Expression and Purification

SDS-PAGE gel showing the expression and purification workflow for three skeletal MyBP-C C1–C2 fragments: slow short (sC1–C2_S), fast skeletal (fC1–C2), and slow long (sC1–C2_L). Each construct was sampled at six key purification stages: Total Lysate, Soluble Fraction, Post-Beads, Pre-Cleavage (His-tagged), Post-TEVp (tag cleaved), and Final Purified Protein. Band intensity across lanes illustrates expression levels, solubility, tag cleavage efficiency, and final purity. All constructs are human and include an 11-residue C-terminal His-tag and TEVp cleavage site.

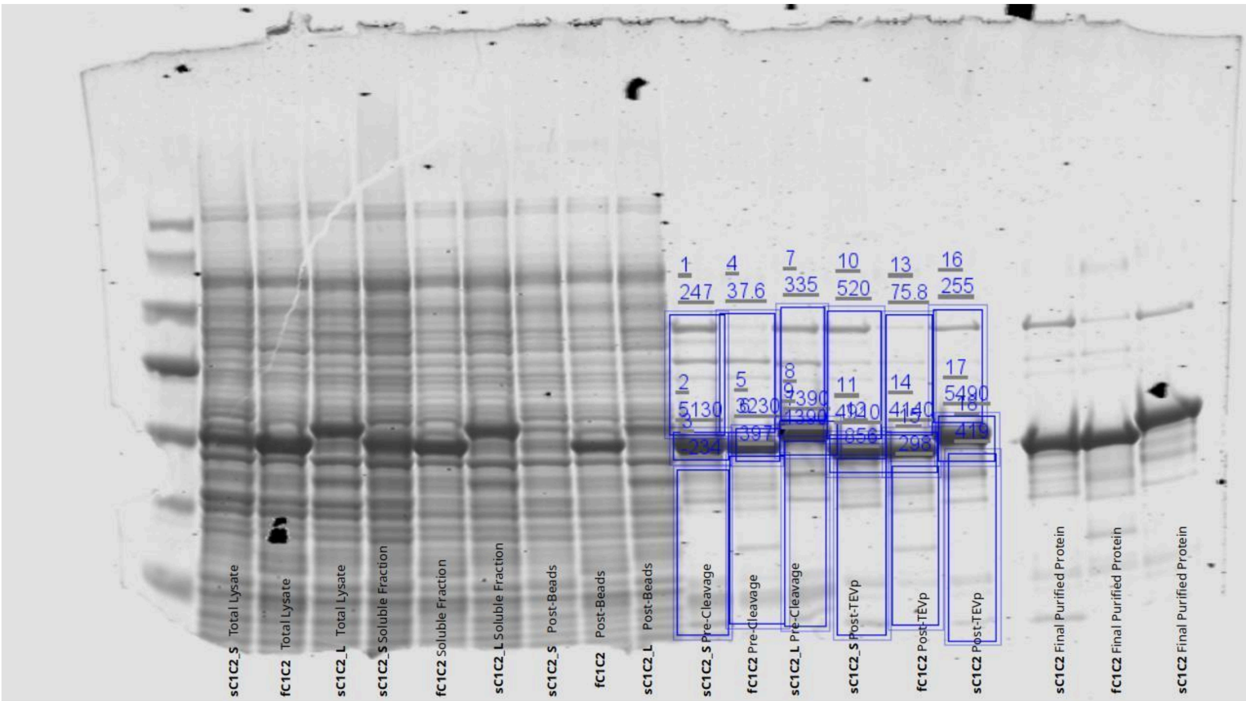

Figure 15: SDS-PAGE Analysis of sMyBP-C C1–C2 Constructs Before and After TEVp Cleavage

Coomassie-stained SDS-PAGE gel showing the purification of three skeletal MyBP-C C1–C2 fragments slow short (sC1–C2_S), fast skeletal (fC1–C2), and slow long (sC1–C2_L) at two purification stages: Pre-Cleavage (His-tagged) and Post-TEVp (tag cleaved). Each lane is annotated with estimated band intensities. The major band (“Pro”) represents the target protein, with minor upper and lower bands labeled “Top” and “Bottom,” respectively, to capture non-target species or partial degradation products. Percent values reflect the approximate proportion of the total lane signal attributed to the desired construct. Across most constructs, TEVp cleavage effectively reduced tag-related species and maintained or improved target protein purity.

Gel 1	Pro	Top	bottom	total	% band
1 hsskC1C2_S PreTEVp	5130	247	-234	5143.0	100%

2 hfskC1C2 PrevTEVp	3230	37.6	397	3664.6	88%
3 hsskC1C2_L PreTEVp	7390	335	1390	9115.0	81%
4 hsskC1C2_S PostTEVp	4910	520	856	6286.0	78%
5 hfskC1C2 PostTEVp	4140	75.8	298	4513.8	92%
6 hsskC1C2_L PostTEVp	5490	255	419	6164.0	89%

Table 1: Protein Expression Levels Pre-TEVp and Post-TEVp

Relative expression levels were determined by estimating the intensity of the target protein band ("Pro") as a proportion of the total lane intensity on Coomassie-stained SDS-PAGE gels. At the Pre-TEVp stage, hsskC1C2_S showed 100% relative target band proportion (i.e., no detectable contaminant bands), while hfskC1C2 and hsskC1C2_L exhibited 88% and 81% target band proportions, respectively. Post-TEVp cleavage, the target band proportions were 78% for hsskC1C2_S, 92% for hfskC1C2, and 89% for hsskC1C2_L, reflecting efficient tag removal with minimal loss of protein integrity.

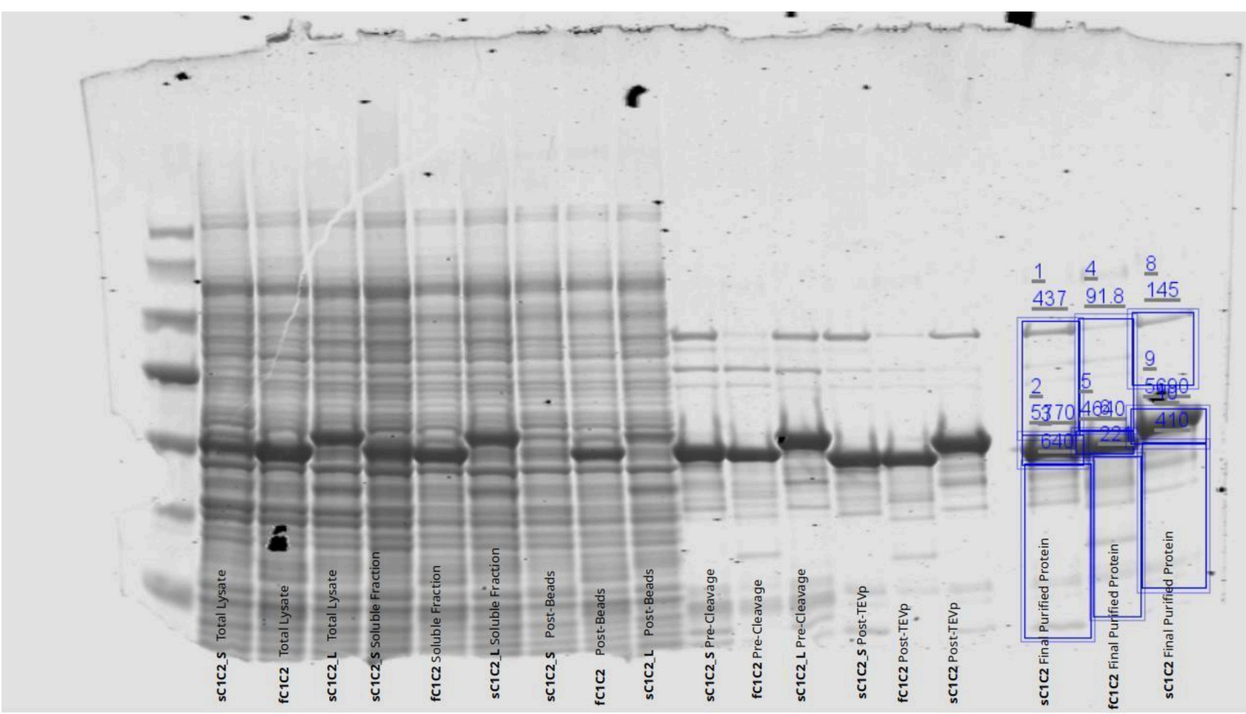

Figure 16: SDS-PAGE Analysis of Final Purified sMyBP-C C1–C2 Fragments

Coomassie-stained SDS-PAGE gel showing the final purified protein samples of slow short (sC1–C2_S), fast skeletal (fC1–C2), and slow long (sC1–C2_L) MyBP-C C1–C2 fragments. The dominant band in each lane corresponds to the target protein ("Pro"), while minor bands above and below ("Top" and "Bottom") represent residual contaminants or degradation products. Visual estimation of band intensities yielded final purities of 84% for sC1–C2_S, 94% for fC1–C2, and 91% for sC1–C2_L, indicating efficient purification. Blue annotation boxes highlight bands and percentages used in these calculations.

Gel 1	Pro	Top	bottom	total	% band
1 hsskC1C2_S Final	5770	437	640	6847.0	84%
2 hfskC1C2 Final	4640	91.8	221	4952.8	94%
3 hsskC1C2_L Final	5690	145	410	6245.0	91%

Table 2: Final Purity Levels of Purified Proteins

The percentage of the total protein corresponding to the target band (Pro) of purified protein at the final stage was 84% for hsskC1C2_S, 94% for hfskC1C2, and 91% for hsskC1C2_L, indicating successful purification with high yields of all fragments.

Construct	Pellet weight (g)	Culture volume (mL)	Protein yield (mg)
hsskC1C2_S	8.24	500	19.9
hfskC1C2	8.27	500	17.9
hsskC1C2_L	7.95	500	18.8

Table 3: Final Purified Protein Yields

The final purified protein yield for hsskC1C2_S, hfskC1C2, and hsskC1C2_L was measured after purification. hsskC1C2_S exhibited the highest yield at 19.9 mg, followed by hsskC1C2_L at 18.8 mg, and hfskC1C2 at 17.9 mg. This was further validated by SDS-PAGE.

4. DISCUSSION

Building upon the success of expressing and purifying skeletal MyBP-C fragments, the insights gained from previous studies in the lab using cardiac MyBP-C (cMyBP-C) offer a valuable framework for understanding how phosphorylation and disease-associated mutations might similarly influence skeletal paralogs. For instance, Colson et al. (2016) demonstrated that phosphorylation of the N-terminal C0–C2 fragment in cMyBP-C results in increased structural compaction and reduced molecular disorder, an effect that enhances actin binding and modulates contractility. Given the structural conservation in the MyBP-C Ig-like and M- domains across isoforms, it is plausible that phosphorylation of sMyBP-C would also induce conformational changes that regulate its interaction with actin and myosin, albeit with specific differences due to the absence of the cardiac-specific C0 domain. However, in human sMyBP-C, the phosphorylation site is in the PAL rather than the M-domain, unique structural changes from cMyBP-C may be involved related to the PAL region. These studies and those described below are currently ongoing in the Colson lab.

Moreover, the cooperative, phosphorylation-sensitive modulation of actin dynamics observed in cardiac systems (Bunch et al., 2019, 2021b) suggests that unphosphorylated skeletal MyBP-C may similarly stabilize actin filaments, promoting strong contractile states, while phosphorylation could weaken this interaction to allow for dynamic cycling. These regulatory mechanisms, if conserved, could explain how mutations in sMyBP-C particularly those linked to DA might mimic or interfere with phosphorylation effects, thereby disrupting filament interactions and muscle performance. Wong et al. (2024) further revealed that mutations in the M-domain of cMyBP-C could enhance myofilament binding even under phosphorylated conditions, supporting the idea that skeletal mutations may similarly increase binding affinity or rigidity, contributing to disease phenotypes such as stiffness or impaired relaxation.

Critically, the lack of certain phosphorylation hotspots in skeletal isoforms may indicate a different threshold or sensitivity to regulation, emphasizing the need for skeletal-specific studies using high-throughput fluorescence lifetime assays, as developed in Bunch et al. (2021a); Dvornikov et al. (2023). Applying such approaches to the newly purified sMyBP-C constructs will help determine whether similar structural transitions govern actin and myosin interactions and whether skeletal isoforms possess alternative compensatory mechanisms for dynamic regulation. These predictions not only justify the production of sMyBP-C fragments but also lay a strong foundation for future studies into the biochemical consequences of phosphorylation and mutation, which remain understudied in skeletal muscle contexts. By drawing from the cardiac MyBP-C paradigm, this underscores the potential of skeletal variants to exhibit conserved yet distinct regulatory behaviors, advancing our understanding of isoform-specific modulation in health and disease.

Considering the findings from Bunch et al. (2018), the cosedimentation assays planned for skeletal MyBP-C fragments should consider how N-terminal elements and isoform-specific sequences might influence filament interactions. In that study, the presence of an N-terminal extension (NTE) in mouse cMyBP-C significantly inhibited myosin binding, without affecting actin binding, highlighting a modulatory role specific to myosin interfaces. This suggests that when assessing actin and myosin interactions with skeletal MyBP-C fragments, careful attention should be paid to the presence or absence of analogous structural features such as the PAL region or phosphorylation motifs. Furthermore, differential binding stoichiometry between cardiac isoforms with and without the NTE provides a framework for interpreting how skeletal MyBP-C fragments particularly those differing in N-terminal composition might exhibit varying affinities or regulatory sensitivities in cosedimentation experiments (Bunch et al., 2018).

Objective and impact

This study aimed to successfully express and purify fragments of slow and fast skeletal myosin binding protein-C (sMyBP-C), specifically the constructs for fast, slow short and slow long with C-terminal TEVp cleavable His-tags (hfskC1C2, hsskC1C2_S, and hsskC1C2_L). These protein fragments represent functionally relevant regions of the sMyBP-C isoform, important for actin/myosin binding and PKA phosphorylation, and serve as a critical foundation for future investigations into their biochemical properties, structural characteristics, and potential roles in muscle-related disorders such as distal arthrogryposis (DA).

The successful expression and purification of these skeletal MyBP-C constructs demonstrate that a protocol originally developed for cardiac C0-C2 can be effectively applied to skeletal isoforms. The ability to cleave C-terminal tags efficiently, achieve high purity (84–94%), and maintain good yields (~18–20 mg) confirms the robustness of this workflow and its suitability for downstream studies. This finding is impactful because it establishes the feasibility of using recombinant skeletal constructs for actin-binding assays, structural analysis, and disease modeling. Compared to cardiac C0-C2, the skeletal constructs expressed and purified with comparable efficiency. Among the skeletal variants, hfskC1C2 had the highest purity, while hsskC1C2_S achieved the highest yield, reflecting minor differences in solubility or folding efficiency. However, given recent variable results between independent preparations skeletal C1-C2 we have now found more reproducible and high-quality protein by moving this His-TEVp tag to the N-terminus and also improved folding by further reducing growth temperatures to 18°C.

What worked particularly well in this study was the codon optimization of the gene sequences for E. coli, which enhanced translation efficiency. The use of the pET45b vector, BL21(DE3) strain, and autoinduction media at a reduced temperature (25°C) collectively contributed to improved solubility and reduced inclusion body formation. The C-terminal His-tag and TEVp cleavage site functioned as expected, with minimal non-specific binding and effective cleavage. This analysis focused on the soluble protein fractions; insoluble fractions or inclusion bodies were not characterized in this study. However, some variability in yield and purity suggests opportunities for improvement. For instance, optimizing TEVp protease conditions or adding an additional size-exclusion chromatography step could improve purity, particularly for the hsskC1C2_S construct. Expression consistency might also be enhanced by exploring chaperone co-expression or alternate expression strains.

Additionally, partial loss of fskC1-C2 during the initial affinity purification step, evident by its presence in the unbound Post-Beads fraction, indicates suboptimal binding efficiency. This outcome likely stems from structural factors that may limit His-tag accessibility, such as local folding or orientation effects. It is also possible that column saturation or insufficient incubation contributed to incomplete capture. To minimize protein loss in future experiments, strategies like extending binding time, modifying tag placement, or increasing resin volume should be considered.

The core objective of this work was to generate structurally intact and biochemically pure sMyBP-C fragments that can be used in downstream functional and structural studies. In doing so, this study lays the groundwork for a broader understanding of skeletal MyBP-C isoform function and its implications in muscle physiology and disease. Overall, the high purity of these sMyBP-C fragments confirms the efficiency of the purification protocol, making them suitable candidates for further structural and functional characterization. Future experiments should focus on confirming structural integrity, interaction dynamics, and potential physiological relevance to gain deeper insights into their role in cardiac function and related muscular disorders such as distal arthrogryposis.

Future Directions

To further verify the quality of the purified proteins, additional analytical techniques can be employed. Mass spectrometry (MS) can be used to confirm the molecular weight and integrity of the purified fragments, ensuring that the expected sequences are retained without degradation or modifications (Figure 17). Circular dichroism (CD) spectroscopy could provide insights into the secondary structure, confirming proper folding. Furthermore, dynamic light scattering (DLS) can assess protein aggregation, ensuring monodispersity.

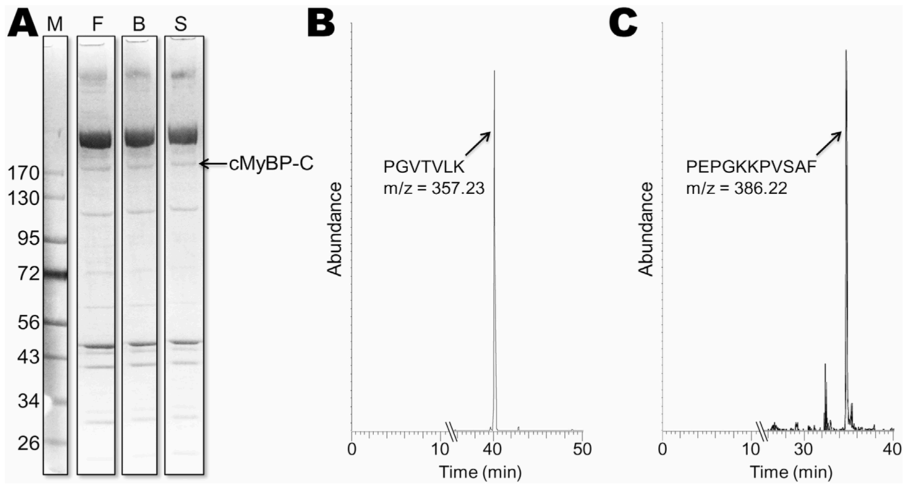

Figure 17: Mass Spectrometry-Based Validation of cMyBP-C for Protein Quality Assessment (representative example)

(A) SDS-PAGE analysis of myofibrillar proteins isolated from three mouse strains (FVB/N, Black 6, and SVE129), with the cMyBP-C band indicated. (B, C) Representative LC-MS/MS ion chromatograms identifying tryptic (PGVTVLK, m/z = 357.23) and α-chymotryptic (PEPGKKPVSAF, m/z = 386.22) peptides from cMyBP-C, confirming protein identity of the N-terminal extension present in mouse but not human cMyBP-C discovered by the Colson lab. This figure illustrates the type of MS based analysis proposed for verifying the identity and integrity of purified skeletal MyBP-C fragments in future studies. Adapted from Bunch et al., 2018.

For functional validation, these purified protein fragments can be utilized in biochemical assays to assess their interaction with binding partners. Surface plasmon resonance (SPR) or isothermal titration calorimetry (ITC) can quantify binding affinities, while in vitro phosphorylation assays can determine post-translational modification patterns. Additionally, cell-based assays could be employed to explore their functional role in relevant biological contexts.

Moreover, targeted mutagenesis can be performed on these fragments to study their role in disease models, particularly in conditions such as distal arthrogryposis. Specific mutations associated with this disorder can be introduced to assess their impact on protein stability, structure, and interactions. Functional assays could help elucidate the molecular mechanisms underlying the disease and identify potential therapeutic targets. Structural studies, including X-ray crystallography or cryo-electron microscopy, may further clarify the conformational changes induced by these mutations and their biological implications.

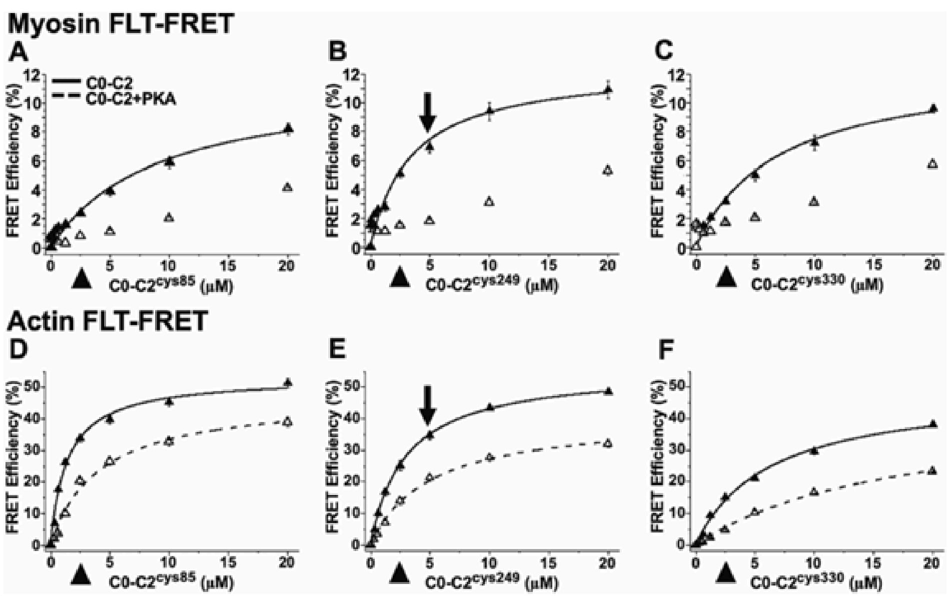

Figure 18: FLT-FRET Binding Curves of Cardiac MyBP-C C0–C2 Domains with Actin and Myosin

FLT-FRET assays showing binding efficiency of unphosphorylated (solid lines) and phosphorylated (dashed lines) C0–C2 domains of cardiac MyBP-C to actin (D–F) and myosin

(A–C). Different FRET acceptor positions (Cys85, Cys249, and Cys330) allow domain-specific analysis of interaction strength and regulation by phosphorylation. These binding profiles offer a model framework for future comparative studies between cardiac and skeletal MyBP-C isoforms. Applying similar FLT-FRET assays to skeletal MyBP-C constructs will allow evaluation of phosphorylation-dependent interaction dynamics and help determine if disease-associated mutations affect unphosphorylated and phosphorylated states, as seen in cardiac proteins. Adapted from Wong et al., 2024.

Given the extensive body of work on cardiac MyBP-C in the lab, future directions should also include comparative studies between cardiac and skeletal MyBP-C isoforms to uncover isoform-specific differences in regulation and function. Using the same fluorescence lifetime-based assays, FLT-FRET, and actin/myosin-binding assays previously established for cardiac proteins (Figure 18), the newly purified skeletal fragments can be evaluated under phosphorylated and unphosphorylated conditions to determine how phosphorylation modulates actin or myosin interactions in skeletal muscle. Performing parallel phosphorylation studies will help clarify whether skeletal MyBP-C undergoes similar structural transitions or exhibits alternative regulatory mechanisms due to the absence of the C0 domain or differences in phosphorylation sites. These comparative experiments could also validate whether disease-associated mutations in skeletal MyBP-C mimic unphosphorylated states, as observed in cardiac systems, providing a direct link between structural dysregulation and disease phenotypes like distal arthrogryposis. Ultimately, these studies would not only deepen our understanding of skeletal muscle contractility but also support the development of isoform-specific therapeutic strategies.

References

Ackermann, M. A., & Kontrogianni-Konstantopoulos, A. (2010). Myosin binding protein-C slow: An intricate subfamily of proteins. Journal of Biomedical Biotechnology, 2010, 652065. https://doi.org/10.1155/2010/652065

Ackermann, M. A., & Kontrogianni-Konstantopoulos, A. (2011). Myosin binding protein-C slow is a novel substrate for protein kinase A (PKA) and C (PKC) in skeletal muscle. Journal of Proteome Research, 10(10), 4547–4555. https://doi.org/10.1021/pr200355w

Ackermann, M. A., & Kontrogianni-Konstantopoulos, A. (2013). Myosin binding protein-C slow: A multifaceted family of proteins with a complex expression profile in fast and slow twitch skeletal muscles. Frontiers in Physiology, 4, 391. https://doi.org/10.3389/fphys.2013.00391

Bunch, T. A., Lepak, V. C., Bortz, K. M., & Colson, B. A. (2021). A high-throughput fluorescence lifetime-based assay to detect binding of myosin-binding protein C to F-actin. Journal of General Physiology, 153(3), e202012707. https://doi.org/10.1085/jgp.202012707

Bunch, T. A., Guhathakurta, P., Lepak, V. C., Thompson, A. R., Kanassatega, R. S., Wilson, A., Thomas, D. D., & Colson, B. A. (2021). Cardiac myosin-binding protein C interaction with actin is inhibited by compounds identified in a high-throughput fluorescence lifetime screen. Journal of Biological Chemistry, 297(1), 100840. https://doi.org/10.1016/j.jbc.2021.100840

Bunch, T. A., Guhathakurta, P., Thompson, A. R., Lepak, V. C., Carter, A. L., Thomas, J. J., Thomas, D. D., & Colson, B. A. (2023). Drug discovery for heart failure targeting myosin-binding protein C. Journal of Biological Chemistry, 299(12), 105369. https://doi.org/10.1016/j.jbc.2023.105369

Bunch, T. A., Kanassatega, R. S., Lepak, V. C., & Colson, B. A. (2019). Human cardiac myosin-binding protein C restricts actin structural dynamics in a cooperative and phosphorylation-sensitive manner. Journal of Biological Chemistry, 294(44), 16228–16240. https://doi.org/10.1074/jbc.RA119.009543

Bunch, T. A., Lepak, V. C., Kanassatega, R. S., & Colson, B. A. (2018). N-terminal extension in cardiac myosin-binding protein C regulates myofilament binding. Journal of Molecular and Cellular Cardiology, 125, 140–148. https://doi.org/10.1016/j.yjmcc.2018.10.009

Colson, B. A., Thompson, A. R., Espinoza-Fonseca, L. M., & Thomas, D. D. (2016). Site-directed spectroscopy of cardiac myosin-binding protein C reveals effects of phosphorylation on protein structural dynamics. Proceedings of the National Academy of Sciences, 113(12), 3233–3238. https://doi.org/10.1073/pnas.1521281113

Dvornikov, A. V., Bunch, T. A., Lepak, V. C., & Colson, B. A. (2023). Fluorescence lifetime-based assay reports structural changes in cardiac muscle mediated by effectors of contractile regulation. Journal of General Physiology, 155(3), e202113054. https://doi.org/10.1085/jgp.202113054

Geist, J., Ward, C. W., & Kontrogianni-Konstantopoulos, A. (2018). Structure before function: Myosin binding protein-C slow is a structural protein with regulatory properties. FASEB Journal, 32(12), fj201800624R. https://doi.org/10.1096/fj.201800624R

Gordon, A. M., Homsher, E., & Regnier, M. (2000). Regulation of contraction in striated muscle. Physiological Reviews, 80(2), 853–924. https://doi.org/10.1152/physrev.2000.80.2.853

Gurnett, C. A., Desruisseau, D. M., McCall, K., Choi, R., Meyer, Z. I., Talerico, M., Miller, S. E., Ju, J. S., Pestronk, A., Connolly, A. M., Druley, T. E., Weihl, C. C., & Dobbs, M. B. (2010). Myosin binding protein C1: A novel gene for autosomal dominant distal arthrogryposis type 1. Human Molecular Genetics, 19(7), 1165–1173. https://doi.org/10.1093/hmg/ddp587

Harris, S. P. (2021). Making waves: A proposed new role for myosin-binding protein C in regulating oscillatory contractions in vertebrate striated muscle. Journal of General Physiology, 153(3), e202012729. https://doi.org/10.1085/jgp.202012729

Kanassatega, R. S., Bunch, T. A., Lepak, V. C., Wang, C., & Colson, B. A. (2022). Human cardiac myosin-binding protein C phosphorylation- and mutation-dependent structural dynamics monitored by time-resolved FRET. Journal of Molecular and Cellular Cardiology, 166, 116–126. https://doi.org/10.1016/j.yjmcc.2022.02.005

Lehman, W., Hatch, V., Korman, V., Rosol, M., Thomas, L., Maytum, R., Geeves, M. A., & Tobacman, L. S. (1994). Tropomyosin and the steric mechanism of muscle regulation. Advances in Experimental Medicine and Biology, 332, 15–22. https://doi.org/10.1007/978-1-4615-2862-3_2

McKillop, D. F., & Geeves, M. A. (1993). Regulation of the interaction between actin and myosin subfragment 1: Evidence for three states of the thin filament. Biophysical Journal, 65(2), 693–701. https://doi.org/10.1016/S0006-3495(93)81110-X

Rayment, I., Holden, H. M., Whittaker, M., Yohn, C. B., Lorenz, M., Holmes, K. C., & Milligan, R. A. (1993). Structure of the actin–myosin complex and its implications for muscle contraction. Science, 261(5117), 58–65. https://doi.org/10.1126/science.8316857

Risi, C. M., Patra, M., Belknap, B., Harris, S. P., White, H. D., & Galkin, V. E. (2021). Interaction of the C2 Ig-like domain of cardiac myosin binding protein-C with F-actin. Journal of Molecular Biology, 433(19), 167178. https://doi.org/10.1016/j.jmb.2021.167178

Risi, C. M., Villanueva, E., Belknap, B., Sadler, R. L., Harris, S. P., White, H. D., & Galkin, V. E. (2022). Cryo-electron microscopy reveals cardiac myosin binding protein-C M-domain interactions with the thin filament. Journal of Molecular Biology, 434(24), 167879. https://doi.org/10.1016/j.jmb.2022.167879

Robinett, J. C., Hanft, L. M., Geist, J., Kontrogianni-Konstantopoulos, A., & McDonald, K. S. (2019). Regulation of myofilament force and loaded shortening by skeletal myosin binding protein C. Journal of General Physiology, 151(5), 645–659. https://doi.org/10.1085/jgp.201812200

Rosenfeld, S. S., & Taylor, E. W. (1987). The mechanism of regulation of actomyosin subfragment 1 ATPase. Journal of Biological Chemistry, 262(21), 9984–9993. https://doi.org/10.1016/S0021-9258(18)48042-1

Semsarian, C., Ingles, J., Maron, M. S., & Maron, B. J. (2015). New perspectives on the prevalence of hypertrophic cardiomyopathy. Journal of the American College of Cardiology, 65(12), 1249–1254. https://doi.org/10.1016/j.jacc.2015.01.019

Shaffer, J. F., Kensler, R. W., & Harris, S. P. (2009). The myosin-binding protein C motif binds to F-actin in a phosphorylation-sensitive manner. Journal of Biological Chemistry, 284(18), 12318–12327. https://doi.org/10.1074/jbc.M808850200

SnapGene. (n.d.). pET-45b(+). SnapGene. https://www.snapgene.com/plasmids/pet_and_duet_vectors_(novagen)/pET-45b(%2B)

Song, T., Landim-Vieira, M., Ozdemir, M., Gott, C., Kanisicak, O., Pinto, J. R., & Sadayappan, S. (2023). Etiology of genetic muscle disorders induced by mutations in fast and slow skeletal MyBP-C paralogs. Experimental & Molecular Medicine, 55(3), 502–509. https://doi.org/10.1038/s12276-023-00953-x

Squire, J. (2019). Special issue: The actin-myosin interaction in muscle: Background and overview. International Journal of Molecular Sciences, 20(22), 5715. https://doi.org/10.3390/ijms20225715

van Dijk, S. J., Bezold, K. L., & Harris, S. P. (2014). Earning stripes: Myosin binding protein-C interactions with actin. Pflügers Archiv - European Journal of Physiology, 466(3), 445–450. https://doi.org/10.1007/s00424-013-1432-8

van Dijk, S. J., Kooiker, K. B., Napierski, N. C., Touma, K. D., Mazzalupo, S., & Harris, S. P. (2018). Point mutations in the tri-helix bundle of the M-domain of cardiac myosin binding protein-C influence systolic duration and delay cardiac relaxation. Journal of Molecular and Cellular Cardiology, 119, 116–124. https://doi.org/10.1016/j.yjmcc.2018.05.001

Wong, F. L., Bunch, T. A., Lepak, V. C., Steedman, A. L., & Colson, B. A. (2024). Cardiac myosin-binding protein C N-terminal interactions with myosin and actin filaments: Opposite effects of phosphorylation and M-domain mutations. Journal of Molecular and Cellular Cardiology, 186, 125–137. https://doi.org/10.1016/j.yjmcc.2023.11.010