Structural Analysis of Human Cardiac Troponin C and Myosin Binding Protein C

by

Xiaolu Linda Zhang

B.Sc., University of British Columbia, 2009

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name: **Xiaolu Linda Zhang**

Degree: **Doctor of Philosophy (Molecular Biology and Biochemistry)**

Title of Thesis: ***Structural Analysis of Human Cardiac Troponin C and Myosin Binding Protein C***

Examining Committee: **Chair:** Lisa Craig
Associate Professor

Mark Paetzel
Senior Supervisor
Professor

Jack N. Chen
Supervisor
Professor

Glen F. Tibbits
Supervisor
Professor

Thomas Claydon
Internal Examiner
Assistant Professor
Department of Biomedical Physiology and Kinesiology

Filip Van Petegem
External Examiner
Associate Professor
Department of Biochemistry and Molecular Biology
University of British Columbia

Date Defended: July 25, 2014

Partial Copyright Licence

Abstract

Sarcomeric proteins are essential for the proper structural assembly and functioning of the sarcomere, the basic contractile unit in striated muscles. When mutations are present in the genes that encode for these proteins, it may lead to cardiac diseases, such as hypertrophic cardiomyopathy (HCM), the leading cause of death in young athletes. Hundreds of mutations within the genes that encode sarcomeric proteins have been shown to cause HCM. The L29Q mutation in cardiac troponin C (cTnC) and the R502W mutation in cardiac myosin binding protein C (cMyBP-C) are two of these mutations. cTnC senses the cytosolic Ca^{2+} concentration and transduces this signal to allow for cross-bridging, leading to muscle contraction. cMyBP-C regulates muscle contractility by interacting with myosin and actin. How the L29Q and the R502W mutations, respectively, affect the cTnC and cMyBP-C and cause disease is unclear.

To advance our knowledge of how these two mutations affect the cardiac proteins' structures and functions, I first generated high resolution structures of the wild type (WT) and mutant regulatory domains of cTnC (cNTnC) using X-ray crystallography and then, determined the WT and R502W mutant structures of the C3 domain of cMyBP-C using nuclear magnetic resonance (NMR). The WT cNTnC was discovered to have coordinating Cd^{2+} ions at both of its calcium-binding sites. This is true for the mutant cNTnC as well. In the WT cNTnC, the vestigial site (EF1) coordinated Cd^{2+} in a noncanonical 'distorted' octahedral geometry, while the functional calcium-binding site (EF2) coordinated Cd^{2+} in the canonical pentagonal bipyramidal geometry. A subtle structural change was observed in the region near the L29Q mutation, and it may play a role in the increased Ca^{2+} affinity of the mutant. The R502W mutation in cMyBP-C did not change the protein's global structure. The dynamics and thermal stabilities of the protein were also not affected by the mutation, as shown by techniques such as amide ^{15}N relaxation and circular dichroism spectroscopy. The mutation does, however, alter the surface charge on the C3 domain and, like other HCM-related mutations found within the same domain, may disrupt the protein's interactions with other sarcomeric proteins, such as actin. The data acquired from this thesis project contributes to a better understanding of the structures of sarcomeric proteins and the pathophysiology of hypertrophic cardiomyopathy.

Keywords: cardiac troponin C; calcium-binding proteins; cadmium; myosin binding protein C; immunoglobulin fold; hypertrophic cardiomyopathy

To my beloved family

Acknowledgements

Foremost, I would like to express my sincere gratitude to my senior supervisor, Dr. Mark Paetzel, who has inspired and motivated me to overcome many obstacles in my research. His guidance, patience and encouragement enabled me to complete my Ph.D thesis with the breadth of knowledge in my area of interest.

I am greatly thankful to my supervisory committee members, Dr. Jack Chen and Dr. Glen Tibbits, who have provided remarkable insights and advices throughout my studies. I am grateful to my internal examiner, Dr. Thomas Claydon, external examiner, Dr. Filip Van Petegem and chair, Dr. Lisa Craig, for serving on my examining committee. I would also like to thank our collaborator, Dr. Lawrence McIntosh, Dr. Soumya De, Mark Okon and Jacob Brockerman for their help and guidance for my NMR work at the University of British Columbia.

I owe it to our lab manager, Deidre de Jong-Wong, for her amazing work in managing the lab and her constant thoughtful support in my life as a Ph.D. student. I could not have done this without her.

Sincere appreciation goes to the past and present members of the Paetzel lab: Dr. Jaeyong Lee, Dr. Apollos Kim, Dr. Kelly Kim, Dr. Ivy Chung, Dr. Charles Stevens, Dr. Sung-Eun Nam, Chuanyun Luo, Alison Li, Suraaj Aulakh, Daniel Chiang, Minfei Fu, Michael Ungerer, Zohreh Sharafianardekani, Eugean Shin, Lauretta Yu and Cindy Li. Their encouragement and friendship made working in the lab a fun and rewarding experience.

From the bottom of my heart, my gratitude goes to my Mom and Dad for their unconditional love and support. And thank-you to my beloved husband, Yu, for always believing in me and helping me to achieve many goals in life.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ASU	Asymmetric Unit
ATP	Adenosin Triphosphate
CD	Circular Dichroism
DXC	Deoxycholic Acid
HCM	Hypertrophic Cardiomyopathy
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LMM	Light Meromyosin
LV	Left Ventricle
NIQD	D2N, V28I, L29Q and G30D
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PKA	Protein Kinase A
RMSD	Root-mean-square Deviation
SAD	Single-wavelength Anomalous Diffraction
SDS	Sodium Dodecyl Sulphate
SL	Sarcomere Length
SPR	Surface Plasmon Resonance
SR	Sarcoplasmic Reticulum
TFP	Trifluoroperazine
WT	Wild-type

Glossary

Å	Ångströms (10^{-10} m)
Asymmetric Unit	The smallest unit that can be rotated and translated to generate one unit cell using only the symmetry operators. The asymmetric unit may be one molecule, one subunit of a multimeric protein, or more than one molecule.
B-factor	Also called 'temperature-factor'. A factor that describes the degree to which the electron density is spread out. It indicates the static or dynamic mobility of an atom. Higher B-factor values indicate higher disorder or mobility.
Chemical Shift	Resonant frequency of a nucleus relative to a standard. It is influenced by the chemical environment around the nucleus. It is expressed in ppm (parts per million).
Column Volume	The total volume of a chromatography column (the sum of the void volume and the matrix volume).
Completeness	The average number of independent measurements of each reflection in a crystallographic data set. Completeness is calculated as (number of measured reflections)/(number of unique reflections).
Crystal	A regular three-dimensional repeat of molecules, with internal symmetry.
Crystal Lattice	The regular array of points about which molecules composing a crystal are centered.
Crystallographic Refinement	A cyclic process of improving R-factor (i.e. agreement between the molecular model and the crystallographic data) by adjustment of the model to obtain highly precise structural model that matches the measured data.
HSQC	Heteronuclear Single Quantum Correlation, an NMR experiment that results in a 2D spectrum with one axis for the chemical shift of ^{1}H and the other for non-hydrogen nucleus (most often ^{15}N or ^{13}C for protein NMR).
in vitro	Latin for "within glass". Describes experiments that are performed with cells or biological molecules outside their normal biological context.
in vivo	Latin for "within the living". Describes experiments that are performed within a living organism or natural setting.
kDa	Kilodalton, non-SI unit for molecular mass
Matthews Coefficient	The crystal volume per unit of protein molecular mass and has a unit of $Å^3Da^{-1}$. It is also called the specific volume (Vm). The average is 2.4, in a range of 1.9-4.2.

Molecular Replacement	A method for deriving initial phases using a known homologous structure.
NMR	Nuclear Magnetic Resonance, a phenomenon in which magnetic nuclei in a magnetic field absorb and re-emit electromagnetic radiation.
NOESY	Nuclear Overhauser Effect Spectroscopy. In a NOESY experiment, nuclei are correlated that are close in space (<5Å). Therefore, peaks in a NOESY contain distance information that can be used to determine the structure of a molecule.
PCR	Polyemerase Chain Reaction, a molecular biology technique to amplify selected DNA sequence by several orders of magnitude.
Penetrance	The portion of genotypes that show expected phenotypes.
Proband	An individual who is affected with a genetic disorder and through whom his/her family with the genetic disorder is ascertained.
R_{cryst} and R_{free}	A measure of agreement between the crystal structure model and the original X-ray diffraction data.
Resonance Assignment	A process of identifying which resonances in the NMR spectrum originate from which atom in the protein that is being examined.
R_{merge}	A measure of agreement among multiple measurements of the same reflections.
RMSD	Root-mean-square deviation, a measure of the differences between several measured values. In structural biology, RMSD is used to describe how well two or more structures align with each other. A lower RMSD value indicates higher structural similarity.
SAD	Single-wavelength Anomalous Diffraction, a method for deriving initial phases that involves measuring diffraction data at a single wavelength near the absorption edge of a heavy-atom such as selenium.
Space Group	A mathematic description of a crystal lattice with a certain type of symmetry and a unit cell. Symmetry type is defined by a set of crystallographic symmetry operations, such as rotation, translation, and screw axis. A Hermann-Mauguin space-group symbol specifies the Bravais lattice and a list of symmetry directions. For instance, $P6_122$, is a hexagonal space group, with primitive lattice (P) and six-fold screw axes parallel to the x-axis (6_1), and a normal two-fold rotational axis (2) along the y and z axes.
Structure Factor	A mathematical function describing the amplitude and phase of a wave diffracted from crystal lattice planes characterised by Miller indices h,k,l.
Synchrotron	A particle accelerator that produces very bright light (electromagnetic waves) in the region from infrared through to gamma rays.

Tris	tris(hydroxymethyl)aminomethane, widely used as a component of buffer solutions. It has a pKa value of 8.06 at 25 °C.
Unit Cell	The smallest unit that repeats itself by translation through a crystal.
Wild-Type	The typical form of an organism, strain, gene, or characteristic as it occurs in nature.
X-ray	A form of electromagnetic radiation that has a wavelength ranging between 0.01 to 10 nm.

Chapter 1.

Introduction

This chapter describes myocardial contractility and the function of sarcomeres in cardiac muscle. The background of hypertrophic cardiomyopathy (HCM), a genetic disease caused by mutations within the genes that encode sarcomeric proteins, is then provided. Cardiac troponin C (cTnC) and cardiac myosin-binding protein C (cMyBP-C) are two sarcomeric proteins that are essential for proper cardiac contractile function. HCM-related mutations are found in both proteins. The respective roles of cTnC and cMyBP-C are illustrated and their previously determined crystal and NMR derived structures are summarized.

1.1. Human cardiac muscle

1.1.1. Myocardial contractility and its regulation

The human heart is the organ that circulates blood throughout the human body via involuntary muscle contraction. The heart goes through cycles of systole and diastole, during which the muscle contracts and relaxes, respectively. The proper functioning and regulation of cardiac muscle is therefore of great importance.

Cardiac muscle contraction follows an excitation-contraction coupling process, meaning that an electrical stimulus is converted to a mechanical response. An action potential, initiated by pacemaker cells in either the sino-atrial or atrio-ventricular node, depolarizes the cell membrane and is then conducted to all cells in the heart via gap junctions. The action potential results in an increase in cytosolic Ca^{2+} via L-type calcium channels. This increase in Ca^{2+} is detected by ryanodine receptors in the membrane of the sarcoplasmic reticulum (SR), which then releases more Ca^{2+} in a process called calcium-induced calcium release. Within the cardiomyocyte, troponin senses the Ca^{2+}

and induces cross-bridge formation between myosin and actin, resulting in muscle contraction. As the muscle relaxes, the SR ATPase, which resides in the SR, transfers Ca^{2+} from the cytosol of the cell to the lumen of the SR at the expense of ATP hydrolysis. In the meantime, a sodium/calcium exchanger ejects Ca^{2+} from the cell. The drop in intracellular Ca^{2+} concentration is again sensed by troponin, which then abolishes the cross-bridging, causing the muscle to relax. The entire myocardial contraction process is regulated by various means, including controlling of the cellular Ca^{2+} flux and controlling the response of the myofilaments to Ca^{2+}.

Phosphorylation is one example of how this regulation is achieved. To control the cellular Ca^{2+} flux, Ca^{2+}-regulating proteins such as those found in the L-type calcium channels and ryanodine receptors are phosphorylated to increase the inward Ca^{2+} current and the SR's Ca^{2+}-loading/release into the cytosol, thereby enhancing contractile force. To alter the myofilaments' sensitivity to Ca^{2+}, protein kinase A phosphorylates human cardiac troponin I at serines 23 and 24. This causes an increase in the cross-bridge cycling rate and a decrease in the myofilaments' Ca^{2+} sensitivity, accelerating muscle relaxation.

1.1.2. Sarcomere – basic contractile unit

The muscular layer that lines the heart chambers and provides structure to the organ is composed of cardiomyocytes. Each of these cells can physically contract to shorten the heart muscle. Inside each cardiomyocyte are hundreds of rod-shaped myofibrils (Figure 1.1) running parallel to each other along the long axis of the cardiomyocyte. These myofibrils are composed of proteins which are organized into thin and thick filaments. Along the length of the myofibrils are repeating sections called sarcomeres (Figure 1.1), the smallest unit of the cardiac muscle responsible for contraction. The sarcomeres of one myofibril are aligned with those on the myofibrils next to it giving cardiomyocytes a striated appearance (Figure 1.1), like that in the skeletal muscle. In smooth muscle cells, this alignment is absent.

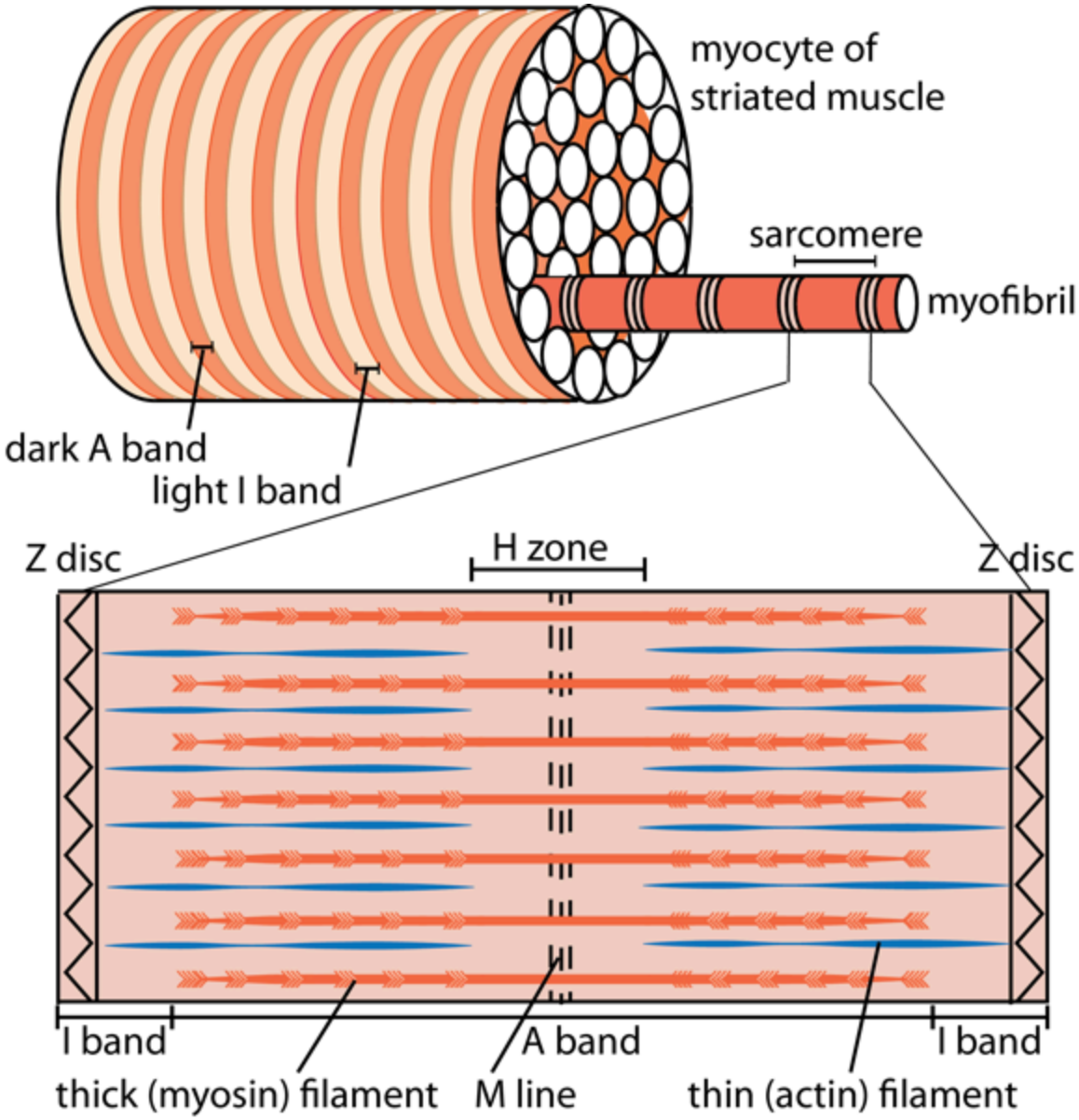

Figure 1.1 The schematics of striated myocyte and sarcomere.
Each striated myocyte contains myofibrils, which are long chains of sarcomeres. The sarcomeres are composed of thin and thick filaments. The Z discs set boundaries of sarcomeres and the M line defines the sarcomere centre. A-bands represents regions that contains myosin and H-zones are regions of only myosin. I-bands are regions of only actin.

The Z discs serve as boundaries for each sarcomere and provide attachment for the titin and actin filaments. At the midline of each sarcomere is the M line, providing attachment for the myosin filament. The myosin filaments are thicker and thus appear darker than the actin filament. As a result, the dark A band (made up of myosin filaments) and the light I band (made up of actin filaments) which alternate, create a

striated appearance on each cardiomyocyte. When muscle contraction occurs, the sarcomere shortens as myosin form cross bridges with actin, the action of which brings the Z discs closer to the M line and reduces the width of the H zone.

The sarcomere is assembled with hundreds of proteins and the correct function of each is essential for proper cardiac behaviour. Cardiac diseases, such as hypertrophic cardiomyopathy, highlight how heart functions can go wrong when these sarcomeric proteins are affected by genetic mutations.

1.2. Hypertrophic cardiomyopathy and sarcomeric protein mutations

1.2.1. Symptoms, diagnosis and treatments

Hypertrophic cardiomyopathy (HCM) is a common genetic cardiac disease, which affects 1 in 500 people worldwide, and is the leading cause of sudden death for athletes and people under the age of 35 (Maron et al., 1995). HCM is characterized by unexplained left ventricular (LV) hypertrophy associated with nondilated ventricular chambers, defective diastolic filling, and ventricular outflow obstruction (Kumar et al., 2013). Adult patients recognized clinically with HCM have LV walls greater than 15mm in thickness, based on echocardiography or cardiovascular magnetic resonance (Gersh et al., 2011). For children, the increased LV wall thickness is defined as wall thickness more than two standard deviations above the mean for age, sex or body size. Systolic function is usually preserved in HCM, but the myocardium does not relax and therefore exhibits primary diastolic dysfunction.

There is also a group of patients with pathogenic mutations but lacking the HCM phenotype. For these genotype-positive/phenotype-negative patients, electrocardiogram, transthoracic echocardiogram and clinical assessment at periodic intervals (12 to 18 months in children and adolescents and every five years in adults) are performed (Gersh et al., 2011).

Beta-blocking drugs and Verapamil therapy are two examples of pharmacological management for HCM patients. Beta-blocking drugs inhibit beta-blockers, inducing

negative inotropic and chronotropic effects (ie. decreased myocardial contractility and heart rate). Verapamil is an L-type calcium channel blocker. Since these calcium channels are concentrated in the atrio-ventricular nodes, Verapamil can be used to decrease impulse conduction through the nodes, thus protecting the ventricles from atrial tachyarrhythmias. There are also invasive therapies: such as septal reduction therapy and alcohol septal ablation. These procedures remove or kill the area of muscle in a hypertrophied heart, which obstructs blood being ejected from the heart, and cause the treated muscle to become less thick. In some patient, a pacing device or an implantable cardioverter-defibrillator may be implanted to relieve symptoms (Gersh et al., 2011).

1.2.2. Genetics of hypertrophic cardiomyopathy

HCM is a genetically heterogeneous disorder caused by mutations that encode sarcomeric proteins and is inherited in an autosomal dominant manner. The majority of the causative variants are missense (Hershberger et al., 2009; Richard et al., 2003).

When patients have HCM symptoms, genetic testing is recommended. Genetic sequencing is also used to identify asymptomatic patients with HCM mutations. There are more than 400 HCM-related mutations identified in the genes of at least nine different sarcomere proteins (Kumar et al., 2013). These include β-cardiac myosin heavy chain (*MYH7*) (Geisterfer-Lowrance et al., 1990; Watkins et al., 1992), cardiac myosin-binding protein C (*MyBPC3*) (Watkins et al., 1995), cardiac troponin T (*TnT2*) (Thierfelder et al., 1994), cardiac troponin I (*TnI3*) (Kimura et al., 1997), α-tropomyosin (*Tpm*) (Watkins et al., 1995), titin (*TTN*) (Satoh et al., 1999), α-cardiac actin (*ACTC*) (Mogensen et al., 1999) and the essential and regulatory myosin light chains (*Myl3/Myl2*) (Poetter et al., 1996). The diverse mutations underlying HCM have one unifying feature: they all increase myofilament activation (Kumar et al., 2013). This results in myocyte hypercontractility with a concomitant increase in energy used and a net negative energy balance.

A recent “1000 Genomes” project sequenced approximately 2500 individuals from 25 different populations, and it provides an overview of all human genetic variation (Abecasis et al., 2012). Further studies checking the presence of the HCM-related

mutations in the “1000 Genomes” database will provide insights into the prevalence of these disease-related mutations in general populations.

The role of some of the HCM-related mutations in disease mechanisms remains unclear. To fill these knowledge gaps, my research investigates the L29Q mutation in cardiac troponin C (cTnC) and the R502W mutation in cardiac myosin-binding protein C (cMyBP-C) from a structural biology point of view.

1.3. Cardiac troponin C

1.3.1. Structural and functional role of cTnC

While studying the Ca^{2+}-dependent contractile behaviour of actomyosin, Ebashi and colleagues discovered troponin and recognized its essential role in the contraction/relaxation process (Ebashi, 1963; Ebashi and Kodama, 1965; Ebashi and Kodama, 1966; Ebashi and Endo, 1968). Further research revealed that troponin is a complex of three proteins: one with a high affinity for calcium, one that inhibits the MgATPase of actomyosin and one that forms a viscous complex with tropomyosin(Hartshorne et al., 1966; Hartshorne and Mueller, 1968; Hartshorne et al., 1967; Schaub and Perry, 1969). Each troponin complex sits on the thin filament at a regular interval distance of 400 Å and its binding stoichiometry with tropomyosin and actin is 1:1:7 (Ebashi et al., 1969; Otsuki et al., 1967). In 1972, at the Cold Spring Harbor symposium, the nomenclature of troponin C (calcium binding protein), troponin I (inhibitory protein) and troponin T (tropomyosin binding protein) was adopted (Greaser and Gergely, 1971). The troponin complex is thought to accommodate faster contraction and relaxation, because to date it has only been found in the striated muscles: skeletal and cardiac muscle.

Cardiac troponin C (cTnC), the Ca^{2+}-sensor, together with cardiac troponin I (cTnI) and cardiac troponin T (cTnT), forms the heterotrimeric troponin complex. Like its skeletal isoform, skeletal troponin C (sTnC), the cTnC protein has two structural domains: the N-terminal regulatory domain (cNTnC) and the C-terminal structural domain (cCTnC) with each of these domains containing two EF-hands, a classical helix-loop-helix motif that binds Ca^{2+}. In total, there are 4 ion-binding sites (Sites I, II, III and

IV, from N-terminus to C-terminus) within each cTnC. Site I, is inactive in cTnC, as compared to sTnC, due to sequence differences. An insertion of Val28 and substitutions of Asp30/Asp32 in sTnC with Leu29/Ala31 in cTnC eliminate key ion-coordinating ligands in cTnC Site I.

The cTnC senses the cytosolic Ca^{2+} concentration as the Ca^{2+} ion binds within its N-terminal regulatory domain, which then induces the binding of switch peptide of cTnI (residues 147-163) to the hydrophobic cavity of cNTnC, consequently causing detachment of the inhibitory peptide of cTnI (residues 128-146) from actin. This detachment results in exposure of the myosin binding site on actin which allows for the cross-bridging between myosin and actin, leading to muscle contraction (Gordon et al., 2000; Li et al., 1999).

1.3.2. Known structures of cTnC

The first structure of cTnC was determined by NMR almost two decades ago (Table 1.1). It showed that the N-terminal regulatory domain (cNTnC) exists in a “closed” conformation, even in the Ca^{2+}-bound state (Sia et al., 1997). This contrasts with what have been observed in Ca^{2+}-saturated sTnC, which has a large conformational change upon Ca^{2+}-binding at its N-terminal EF-hands. This observation may imply that cTnC can vary its degree of openness in response to muscle contraction events, by mechanisms such as the phosphorylation of cTnI, which has phosphorylation sites (ie. Ser23/Ser24) at its cardiac-specific N-terminus.

Structures of cTnC in complex with cTnI provide insight into the conformational changes within cardiac troponin during muscle contraction (Table 1.1). For example, the structure of the Ca^{2+}-bound cNTnC and cTnI switch peptide complex shows how the cTnI peptide (comprising residues Arg147-Lys163) takes on an alpha-helical conformation and forms hydrophobic interactions with the cavity of cNTnC (Li et al., 1999) (Figure 1.2a). In skeletal troponin I (sTnI), the same region has been shown to bind the exposed hydrophobic pocket of Ca^{2+}-saturated regulatory domain of sTnC (sNTnC) (McKay et al., 1997). The cNTnC renders an open conformation, similar to that observed in Ca^{2+}-saturated sTnC (Slupsky and Sykes, 1995). The binding affinity between cNTnC and the cTnI peptide, however, is six times weaker than that between

their skeletal counterparts. This suggests that the difference in the cardiac and skeletal muscle contraction can be understood in terms of different thermodynamics and kinetics equilibria. The inhibitory peptide of cTnI (comprising residues Thr128-Arg147) dissociates from actin and binds to cTnC upon Ca^{2+}-induced muscle contraction. To illustrate the details of this conformational change, the complex structure of cCTnC and the inhibitory peptide was determined (Lindhout and Sykes, 2003) and showed that the peptide adopts a helical conformation between residues 134 and 139 and interacts with the cCTnC electrostatically (Figure 1.2b).

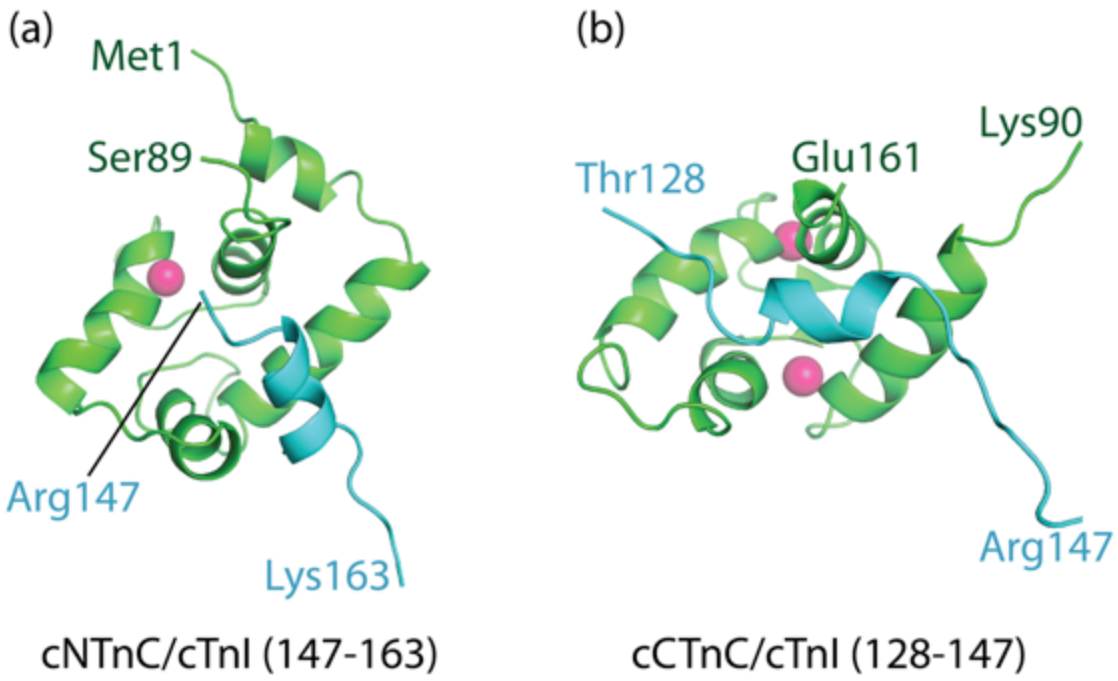

Figure 1.2 Structures of cTnC in complex with cTnI.
Structures of (a) Ca^{2+}-bound cNTnC and cTnI switch peptide complex (PDB: 1MXL) and (b) Ca^{2+}-bound cCTnC and cTnI inhibitory peptide complex (PDB:1OZS) are shown in ribbon diagrams. The cNTnC and cCTnC are colored in green and the cTnI peptides are in cyan. The Ca^{2+} ions are shown as pink spheres. The start and end residues of each cTnC and cTnI chains are labelled.

Structures of various chemical compounds in complex with cTnC have been resolved (Figure 1.3). Many of the chemical compounds are Ca^{2+}-sensitizing/desensitizing drugs and their effects on cTnC are clearly illustrated in the structures. Compounds such as bepridil and trifluoperazine have been found within the hydrophobic cavity of cNTnC and are thought to enhance cTnI binding to cNTnC and to increase the Ca^{2+} affinity of cTnC (Li et al., 2000).

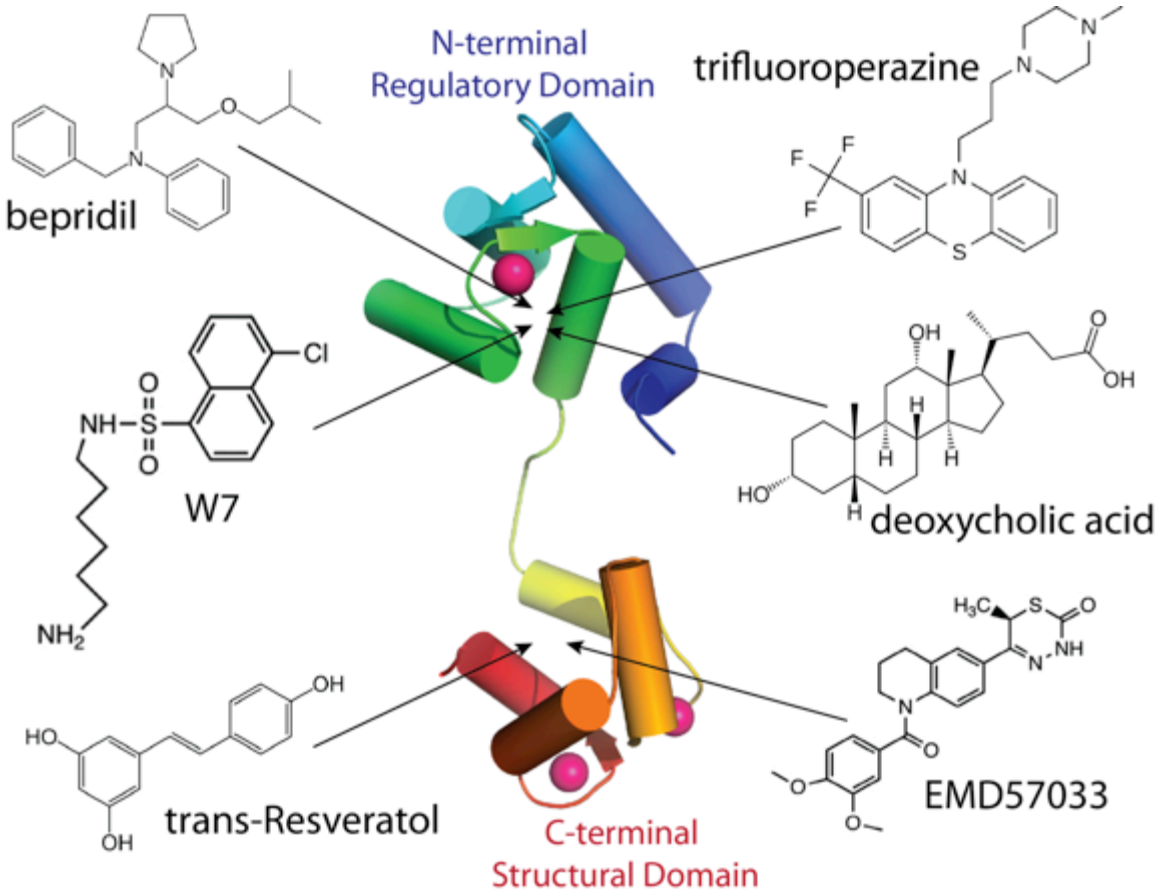

Figure 1.3 Structures of various chemical compounds found in complex with cTnC.

Cartoon diagram of solution structure of chicken cTnC (C35S/C84S) (Sia et al., 1997) is shown with various chemical compounds found in complex with cTnC. The helices are shown as cylindars and the β-strands as arrows. Ca^{2+} ions are showns as pink spheres. The binding site of each compound on cTnC is shown by an arrow.

In 2003, Takeda *et al.* published the core domain structure of the cardiac troponin complex (Figure 1.4). Although parts of cTnI and cTnT are not well defined in the structure, the overall architecture of the troponin molecule is clear. The complex is divided into subdomains: the regulatory head, the IT arm, TnT1, c-TnT and the regulatory segment of TnI (Takeda et al., 2003). The regulatory head consists of cTnC residues 3-84 and cTnI residues 150-159, and the IT arm consists of cTnC residues 93-161, cTnI residues 42-136 and cTnT residues 203-271. These subdomains are connected by flexible linkers, allowing for major changes both in position and conformation within the cardiac troponin complex. It is suggested that the flexible nature of the troponin molecule is relevant to its physiological function.

During calcium regulation, the regulatory segment of TnI undergoes major changes in position. At lower Ca^{2+} concentrations, the regulatory segment of TnI attaches to actin-tropomyosin. At higher Ca^{2+} concentrations, this segment is detached from actin-tropomyosin and associates with cNTnC as shown in the determined core domain structure (Figure 1.4). The IT arm is rigid and has no direct interaction with actin-tropomyosin. This region, however, bridges TnT1 and C-TnT subdomains, which are two distinct portions of TnT that anchor troponin to the thin filament through tropomyosin binding. The IT arm is also upstream from the regulatory segment of TnI, which changes position in a Ca^{2+}-dependent manner. Thus the IT arm may rotate in this process and leads to the movement of tropomyosin strand on the thin filament. This structure and the other mentioned previously provide a framework for understanding the interactions within the troponin complex and their effects on heart contraction.

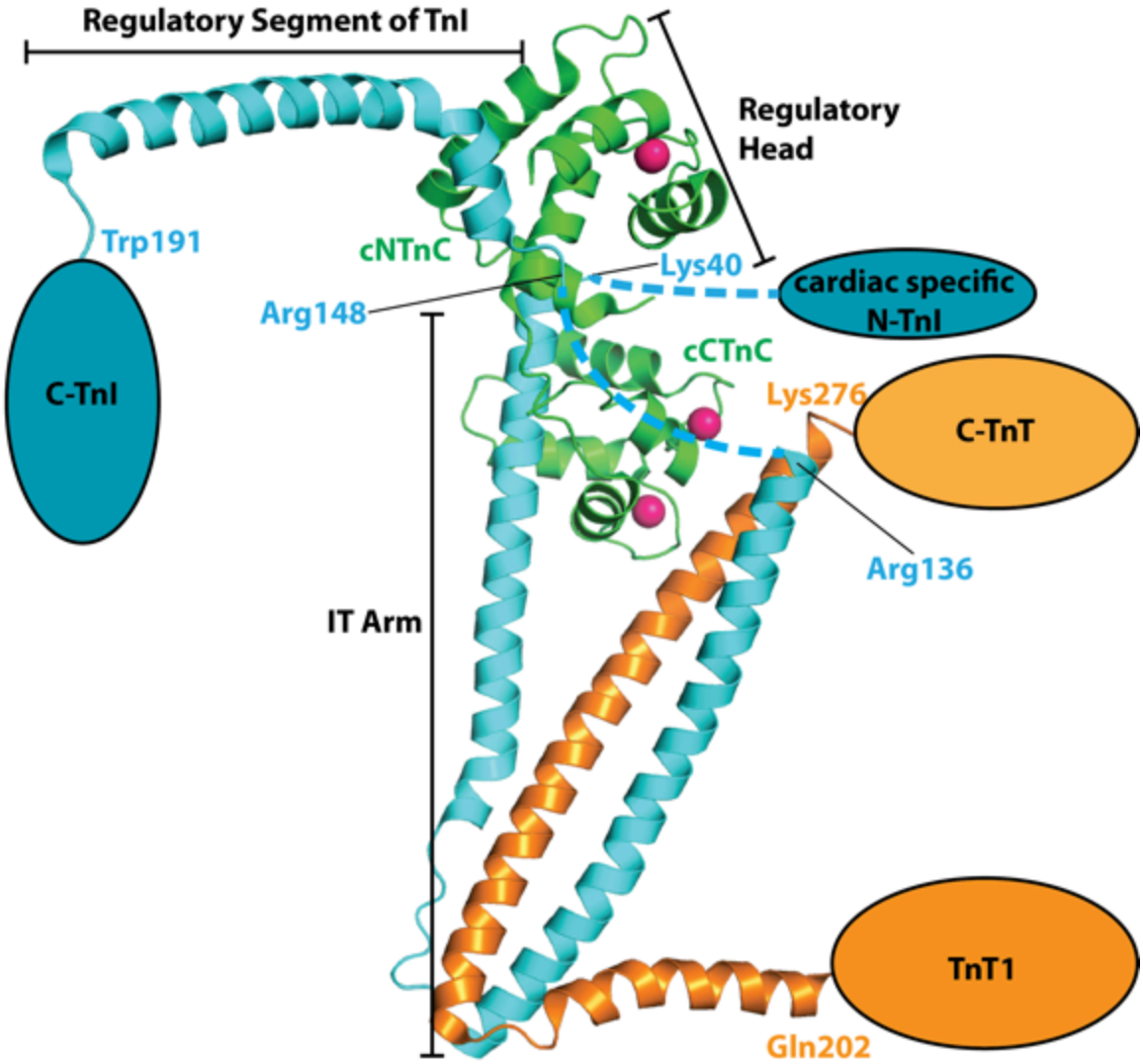

Figure 1.4 Crystal structure of the core domain of the cardiac troponin complex.

The core domain structure of the cardiac troponin complex (PDB: 1J1E) is shown in cartoon diagram. The N-terminal and C-terminal domains cTnC (green) are labelled and the Ca^{2+} ions are shown as pink spheres. Residues 137-147 (dashed lines) of cTnI (cyan) are not shown in the structure as the associated electron densities are not well defined. The cardiac specific domain of cTnI (residues 1-30, shown as an oval) is not included in the cTnI that was used to generate the crystal structure, and it is connected to the Lys40 of cTnI by a dashed line. The C-terminus of cTnI (residues 192-210), the TnT1 (residues 183-201) and the C-terminus of TnT (residues 277-288) are not well defined and are shown as ovals. The structure is divided into five subdomains: the regulatory head, IT arm, c-TnT, TnT1 and the regulatory segment of TnI.

Table 1.1 Structures of cTnC deposited in Protein Data Bank.

Constructs	Species	Method	PDB Code
cNTnC (D2A/C35S/C84S)	*Gallus gallus*	NMR	2CTN (Sia et al., 1997)
cCTnC	*Gallus gallus*	NMR	3CTN (Sia et al., 1997)
cTnC (C35S/C84S)	*Gallus gallus*	NMR	1AJ4 (Sia et al., 1997)
cCTnC / N-cTnI (33-80)*	*Gallus gallus*	NMR	1FI5 (Gasmi-Seabrook et al., 1999)
cTnC (in cTnC-cTnI complex)	*Gallus gallus*	NMR	1LA0 (Dvoretsky et al., 2002)
cCTnC / N-cTnI (33-80)*	*Gallus gallus*	NMR	1SCV (Finley et al., 2004)
cCTnC / N-cTnI (33-80)*	*Gallus gallus*	NMR	1SBJ (Finley et al., 2004)
cNTnC	*Homo sapiens*	NMR	1AP4 (Spyracopoulos et al., 1997)
cNTnC without Ca^{2+}	*Homo sapiens*	NMR	1SPY (Spyracopoulos et al., 1997)
cNTnC / cTnI switch peptide	*Homo sapiens*	NMR	1MXL (Li et al., 1999)
cCTnC / EMD57033	*Homo sapiens*	NMR	1IH0 (Wang et al., 2001)
cNTnC / cTnI switch peptide / bepridil	*Homo sapiens*	NMR	1LXF (Wang et al., 2002)
cCTnC / cTnI inhibitory peptide	*Homo sapiens*	NMR	1OZS (Lindhout and Sykes, 2003)
cNTnC (F77W/V82A)	*Homo sapiens*	NMR	2JXL (Julien et al., 2008)
cTnC (F153W/C35S/C84S)	*Homo sapiens*	NMR	2JT3 (Wang et al., 2005)
cTnC (F153(FTR)/C35S/C84S)	*Homo sapiens*	NMR	2JT8 (Wang et al., 2005)
cTnC (F104(FTR)/C35S/C84S)	*Homo sapiens*	NMR	2JTZ (Wang et al., 2005)
cTnC (F104W/C35S/C84S)	*Homo sapiens*	NMR	2JT0 (Wang et al., 2005)
cNTnC (C35S/C84S) / W7	*Homo sapiens*	NMR	2KFX (Hoffman and Sykes, 2009)
cNTnC (F77W) / cTnI switch peptide	*Homo sapiens*	NMR	2KGB (Julien et al., 2009)
cCTnC / Green Tea Polyphenol	*Homo sapiens*	NMR	2KDH (Robertson et al., 2009)
cNTnC / cTnI switch peptide / W7	*Homo sapiens*	NMR	2KRD (Oleszczuk et al., 2010)
cNTnC / cTnI switch peptide / dfbp-o	*Homo sapiens*	NMR	2L1R (Robertson et al., 2010)
cCTnC / trans-Resveratol	*Homo sapiens*	NMR	2L98 (Pineda-Sanabria et al., 2011)
cCTnC / cTnI (34-71).	*Homo sapiens*	NMR	2MLE (Baryshnikova et al., 2008)
cNTnC at 30 °C	*Onchorhynchus mykiss*	NMR	1R2U (Blumenschein et al., 2004)

Constructs	Species	Method	PDB Code
cNTnC at 7 °C	*Onchorhynchus mykiss*	NMR	1R6P (Blumenschein et al., 2004)
cTnC (C35S/C84S) / bepridil	*Gallus gallus*	X-Ray	1DTL (Li et al., 2000)
cTnC (C35S/C84S) / cTnI (T31M/C80A/C97A) (31-210) / cTnT (183-288)	*Homo sapiens*	X-Ray	1J1E (Takeda et al., 2003)
cTnC (C35S/C84S) / cTnI (T31M/C80A/C97A) (31-163) / cTnT (183-288)	*Homo sapiens*	X-Ray	1J1D (Takeda et al., 2003)
cNTnC (C35S/C84S) / trifluroparazine	*Homo sapiens*	X-Ray	1WRK (To be published)
cNTnC (C35S/C84S)/ trifluroparazine	*Homo sapiens*	X-Ray	1WRL (To be published)
cNTnC / deoxycholic acid / Cd^{2+}	*Homo sapiens*	X-Ray	3RV5 (Li et al., 2011)

cNTnC: residues 1-89 of cTnC; cCTnC: residues 90-161 of cTnC; cTnI switch peptide: residues 147-163 of cTnI; cTnI inhibitory peptide: residues 128-147 of cTnI; FTR: 5-flurotryptophan; W7: N-(6-aminohexyl)-5-chloro-1-naphthalenesulfinamide; dfbp-o: 2',4'-difluoro(1,1'-biphenyl)-4-yloxy acetic acid

* The N-cTnI (33-80) peptides are in complex with cCTnC, but only cCTnC is presented in the determined structures.

1.3.3. Classical Ca^{2+}-binding EF-hand

The cNTnC domain contains two EF-hand motifs (EF1 and EF2). The EF-hand is a common functional motif found in many Ca^{2+}-binding proteins. It has a 12-residue long ion-binding loop flanked by two perpendicularly positioned α-helices (Kretsinger and Nockolds, 1973). Canonically, residues at loop positions 1, 3, 5, 7 and 12 participate in coordinating Ca^{2+}, and the coordinating ligands are positioned in a pentagonal bipyramidal geometry (Strynadka and James, 1989). The equatorial plane is constructed from ligands at positions Y, Z, –Y and –Z. The –Z position is normally a glutamate residue that contributes its two carboxylate oxygen atoms as a bidentate ligand. The ligands orientated at the X and –X positions make up the vertical axis (Figure 1.5). The ligand at the –X position is often a water molecule.

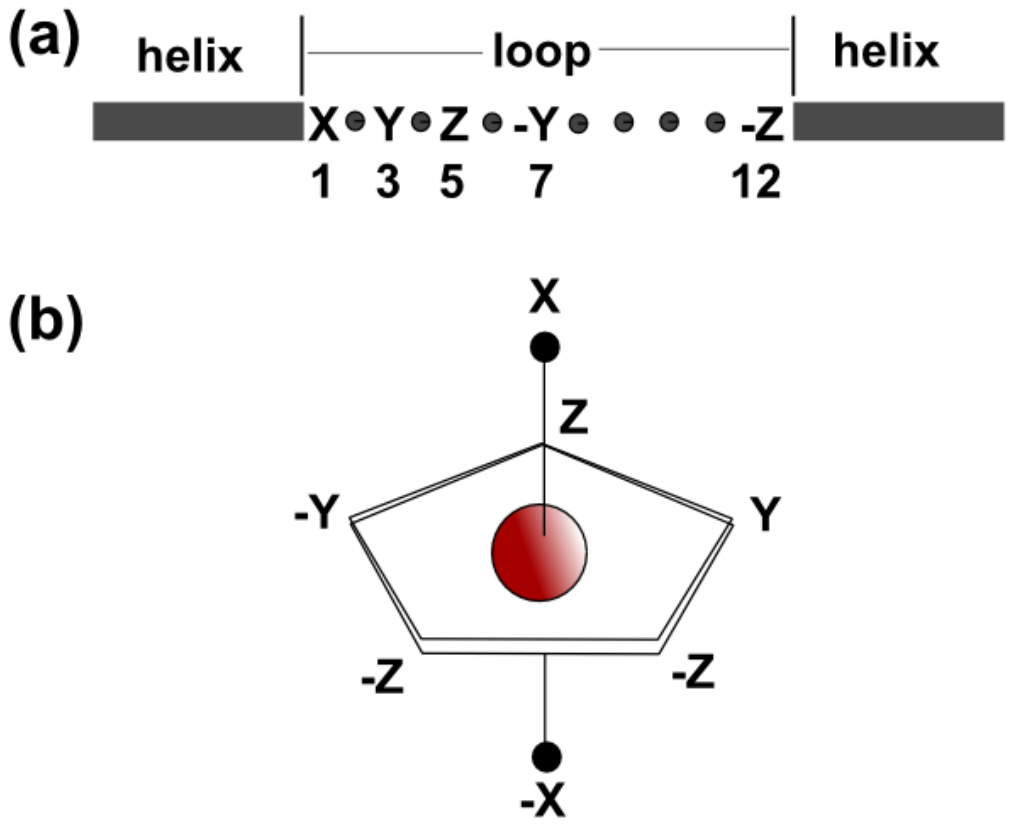

Figure 1.5 Canonical pentagonal bipyramidal geometry coordination displayed by EF-hand ion-binding proteins.

Canonical pentagonal bipyramidal geometry coordination displayed by EF-hand ion-binding proteins. (a) Flanked by two helices, the EF-hand ion-coordinating loop contains 12 residues. The residues at positions 1, 3, 5, 7 and 12 are labelled with their spatial orientation in pentagonal bipyramidal geometry. (b) Spatial orientation of coordinating ligands in canonical pentagonal bipyramidal geometry.

Only one of the two EF-hands in cNTnC binds Ca^{2+} under physiological conditions (Li et al., 1997,van Eerd & Takahashi, 1975). By comparing the EF1 loop sequence to the canonical EF-hand loop sequence, it can be seen that the residues at positions 1 and 3 are replaced by two nonpolar residues that are unable to coordinate Ca^{2+} ions (van Eerd & Takahashi, 1975). Consequently, EF2 serves as the sole Ca^{2+}-binding site in cNTnC under physiological Ca^{2+} concentrations and acts as the lone Ca^{2+} sensor in the troponin complex for cardiac muscle function (Li et al., 1997). This is supported by NMR studies of Ca^{2+} bound human cNTnC which show that Ca^{2+} is observed only in EF2 (Spyracopoulos et al., 1997). To reveal the atomic details of ion coordination within the EF-hands of the protein, a high-resolution crystal structure of cNTnC is needed.

1.3.4. HCM-related L29Q mutation in cTnC

The first mutation in cardiac troponin C (cTnC) was found in a German patient with hypertrophic cardiomyopathy (HCM) and reported in *Mutation and Polymorphism Report* in 2001 (Hoffmann et al., 2001). A thymine nucleotide was mutated to an adenine in the *TNNC1* gene at chromosomal location 3p21.3-p14.3, causing a leucine-to-glutamine exchange at amino acid 29. The L29Q mutation has been documented with only one HCM case so far and the frequency of its occurrence in the general population is unknown.

The L29Q mutation has been shown to enhance the Ca^{2+}-binding characteristics of cTnC and alter myocyte contractility when incorporated into cardiac myocytes (Liang et al., 2008). In experiments, the Ca^{2+}-binding and release characteristics of L29Q in mammalian cTnC were monitored using the F27W substitution with fluorescence spectrometry (Liang et al., 2008). The Ca^{2+}-binding affinity of site II of L29Q cTnC was 1.3-fold higher than that of wild type (WT), and the Ca^{2+}-disassociation rate from the same site was not significantly different. The rate of Ca^{2+}-binding to site II was 1.5-fold higher in L29Q cTnC. The mutational effects on Ca^{2+} activation of force generation in single mouse cardiac myocytes were also studied, and it is 1.4-fold higher than myocytes reconstituted with L29Q cTnC.

In single skinned cardiomyocytes, the L29Q mutation increased Ca^{2+}-senstivity in a highly sarcomere length (SL)-dependent manner (Li et al., 2013). Ca^{2+}-sensitivity is normally reduced by phosphorylated cTnI (Robertson et al., 1982), and this reduction was significantly diminished in L29Q cTnC. L29Q also depressed the SL-dependent increase in myofilament Ca^{2+}-sensitivity. In the thin filament, and in the presence of phosphomimetic cTnI, the L29Q cTnC reduced the Ca^{2+}-affinity by 27% in steady-state measurements and increased the Ca^{2+}-dissociation rate by 20% in kinetic studies (Li et al., 2013).

1.3.5. L29Q mutation and NIQD in exothermic species

In an evolutionary study of cardiac troponin C (cTnC) by Gillis et al (Gillis et al., 2007), ectothermic species (organisms that regulates their body temperature by exchanging heat with external environments; eg. rainbow trout [*Oncorhynchus mykiss*])

were found to have a glutamine at position 29, the missense mutation in cTnC that relates to hypertrophic cardiomyopathy in humans (HCM), as well as three other conserved residue differences from the human cTnC (D2N, V28I, and G30D) (Figure 1.6). These four residues, Asn2, Ile28, Gln29 and Asp30 (NIQD) have functional importance in Ca^{2+}-sensing at lower temperatures (which is the physiological temperature of most ectothermic species) (Gillis et al., 2007). This has been supported by observations of a higher Ca^{2+}-affinity for the human NIQD mutant of cTnC (Gillis et al., 2005). Spyracopoulos *et al.* had suggested previously that EF1 and EF2 loops are structurally linked (Spyracopoulos et al., 1997). Based on these findings, it is possible that the primary sequence differences in site I may have an effect on the structure and thus, on the ion-binding ability of site II (Gillis et al., 2007). Determination of the L29Q and NIQD cNTnC structures may provide greater insight into the conformational changes and their ramifications.

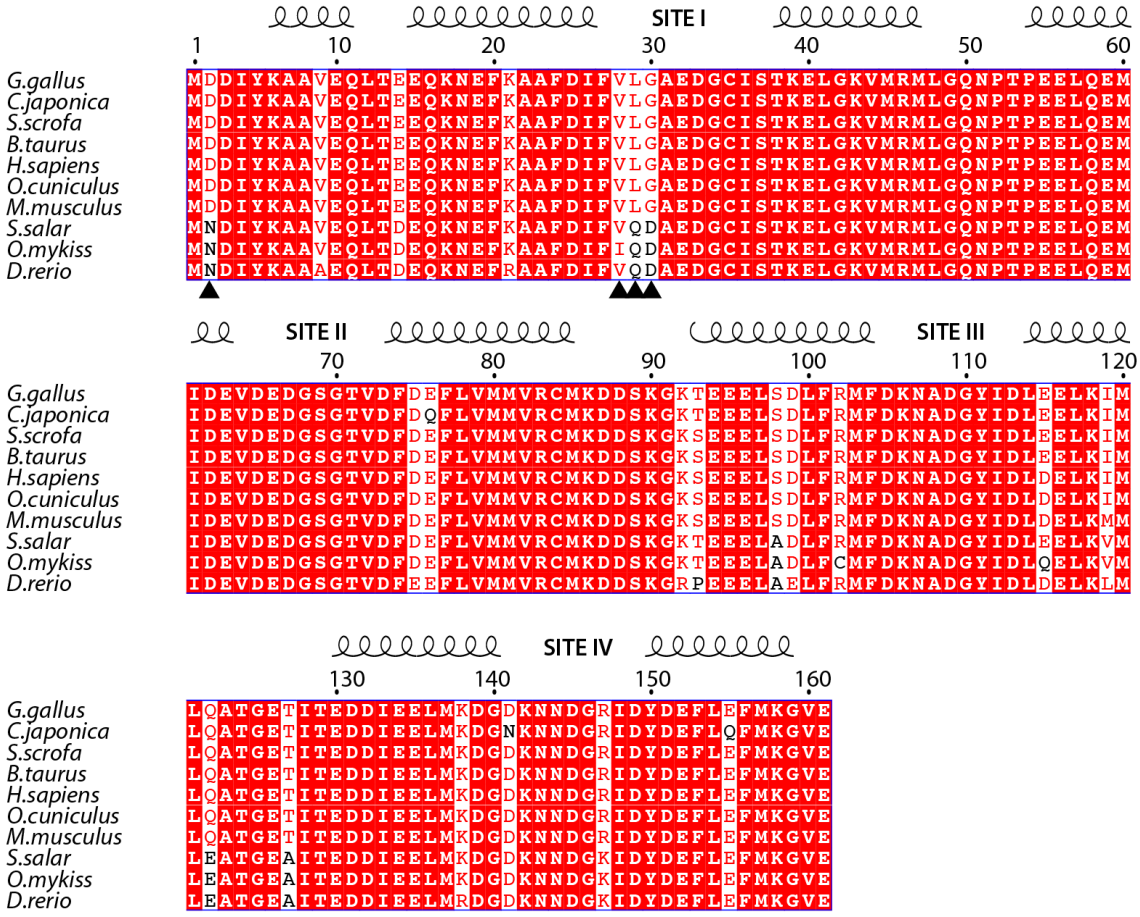

Figure 1.6 Cross-species sequence alignment of cTnC.
Invariant residues are shown in red boxes, similar residues in red text, and stretches of amino acids that are similar across the group of sequences in blue boxes. The protein sequences were acquired from the Swiss-Prot database: *G.gallus* (P09860); *C.japonica* (P05936); *S.scrofa* (P63317); *B.taurus* (P63315); *H.sapiens* (P63316); *O.cuniculus* (P02591); *M.musculus* (P19123); *S.salar* (B5X8Q3); *O.mykiss* (Q7ZZB9); *D.rerio* (Q800V7). Residues at position 2, 28, 29 and 30 have functional importance in Ca^{2+}-sensing at lower temperatures, which is the physiological temperature of most ectothermic species, and they are shown with black arrows. The secondary structure of full-length cTnC (Li et al., 2000) is shown above the alignment. The helices are shown as coils and the four Ca^{2+}-binding sites, including the vestigial Site I, are labelled at their respective positions.

1.4. Cardiac myosin binding protein C

1.4.1. Structural and functional role of cMyBP-C

Cardiac myosin binding protein C (cMyBP-C), a 130kDa protein located on the thick filaments of vertebrate cardiac muscle, manifests as seven to nine stripes approximately 43nm apart in the crossbridge-containing C-zones of the sarcomere (Bennett et al., 1986; Offer et al., 1973; Sjöström and Squire, 1977). Structurally, cMyBP-C is comprised of eight immunoglobulin-like (Ig) domains and three fibronectin-like domains (Gruen et al., 1999; Squire et al., 2003) (Figure 1.7). A (Pro-Ala)-rich sequence is also found at the N-terminus of the protein. The cardiac isoform of myosin-binding protein C contains an N-terminal domain C0, four phosphorylation sites in the mybpc domain and a 28-residue insert in the C5 domain (Gautel et al., 1995).

The protein, cMyBP-C, is essential for the regulation of contractility in cardiac muscle (Flashman et al., 2004). In an *MYBPC3* (the gene encoding cMyBP-C) knocked-out mouse, the sarcomere assembles normally but the mouse shows cardiac hypertrophy and impaired contractile function (Harris et al., 2002). As shown by co-sedimentation experiments and immunofluorescence microscopy of recombinant proteins, the C-terminus (C7-C10) of cMyBP-C is tethered to the myosin backbone (Flashman et al., 2007; Gilbert et al., 1996; Okagaki et al., 1993). The N-terminus of cMyBP-C modulates contractility by interacting with the S2 region of myosin in a phosphorylation-controlled manner (Gruen et al., 1999; Kunst et al., 2000). Studies have shown that the (Pro-Ala)-rich sequence within cMyBP-C interacts with actin (Gruen et al., 1999; Squire et al., 2003) and the mybpc motif may also provide a platform for actin binding (Howarth et al., 2012). A three-dimensional electron microscopy reconstruction has confirmed that N-terminal fragments of cMyBP-C (C0 to C3) lie alongside F-actin (Mun et al., 2011).

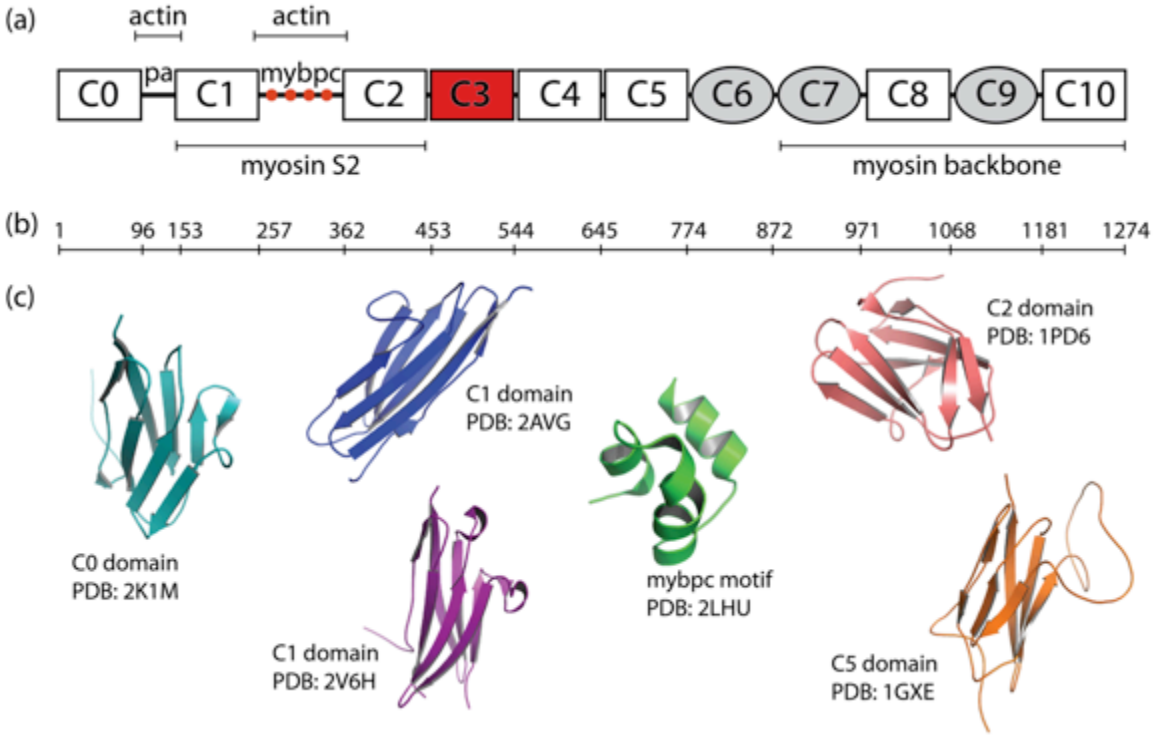

Figure 1.7 Schematic representation of the modular architecture of cMyBP-C. (a) The fibronectin-like domains are shown in gray ovals and immunoglobulin-like domains in white rectangles, except for C3 (red rectangles). The 4 phosphorylation sites within the mybpc motif are shown as red dots and the (Pro-Ala)-rich sequence as 'pa'. Regions of interaction on cMyBP-C as reported in the literature are indicated with horizontal bars and labelled with the name of interacting proteins. (b) The sequence numbers for the beginning and end of each domains are displayed on a horizontal scale and is aligned with the schematic map in (a). (c) Previously determined structures of cMyBP-C are shown in cartoon with their representing region labelled: C0 domain (cyan, PDB: 2K1M); C1 domain (blue, PDB: 2AVG); C1 domain (purple, PDB: 2V6H); mybpc motif (green, PDB: 2LHU); C2 domain (pink, PDB: 1PD6) and C5 domain (orange, PDB:1GXE).

1.4.2. Known structures of cMyBP-C

While the C-terminal region has a structural role in positioning cMyBP-C on the myosin backbone, the N-terminal region of cMyBP-C has an important role in the modulation of actomyosin activities. As a result, the functions of the various domains within the N-terminal cMyBP-C have received more attention and many of their structures (ie. C0, C1, mybpc motif and C2) have been determined.

As predicted from their sequence, C0, C1 and C2 have immunoglobulin-like topology (Figure 1.7). Using NMR interaction experiments, the binding site of C1 has been located in the immediate vicinity of the S1-S2 hinge of myosin (Ababou et al.,

2008). The C2 domain binds to S2 of myosin with low affinity but exhibits a high binding specificity (Ababou et al., 2007)

Effects of HCM-related mutations have also been speculated upon based on the known structures. For example, HCM-related mutations in C1 are clustered at one end towards the C-terminus and they may alter the structural integrity of this region and its interactions with myosin S2 (Ababou et al., 2008). In C2, the HCM-related E301Q mutation is within the binding interface between the C2 domain and myosin S2. As most key interactions are between polar amino acids, the E301Q mutation is thought to disrupt these interactions and lead to disease-causing alterations. However, no structure of any of the HCM-related mutants of cMyBP-C had been resolved before the start of my thesis project.

Table 1.2 Structures of cMyBP-C in the Protein Data Bank.

Constructs	Species	Method	PDB Code
C0 domain (2-96)	*Homo sapiens*	NMR	2K1M (To be published)
C1 domain (151-260)	*Homo sapiens*	NMR	2AVG (Ababou et al., 2008)
C2 domain (355-451)	*Homo sapiens*	NMR	1PD6 (Ababou et al., 2007)
C5 domain (641-770)	*Homo sapiens*	NMR	1GXE (Idowu et al., 2003)
mybpc motif (255-357)	*Mus musculus*	NMR	2LHU (Howarth et al., 2012)
C1 domain (151-258)	*Homo sapiens*	X-Ray	3CX2 (Fisher et al., 2008)
C1 domain (151-258)	*Homo sapiens*	X-Ray	2V6H (Govada et al., 2008)

1.4.3. Immunoglobulin-like protein fold

The C3 domain of cMyBP-C renders an immunoglobulin-like (Ig-like) protein fold, which is found in a variety of proteins including many non-immunoglobulins. The immunoglobulin fold has been shown to interact with multiple proteins and can withstand force when incorporated in a filament, such as those in titin, a large multi-domain protein located within the sarcomere.

Structurally, the Ig-like domain contains seven to nine antiparallel β-strands (A, B, C, [C', C",] D, E, F, G) that fold into a Greek-key β-sandwich of two sheets (Bork et al., 1994). The β-sandwich has a core domain, composed of strands B, C, E and F.

When compared with other Ig-like proteins, the lengths of the non-core strands and surrounding loops are highly variable, but the core strands are well-conserved (Kabsch and Sander, 1983).

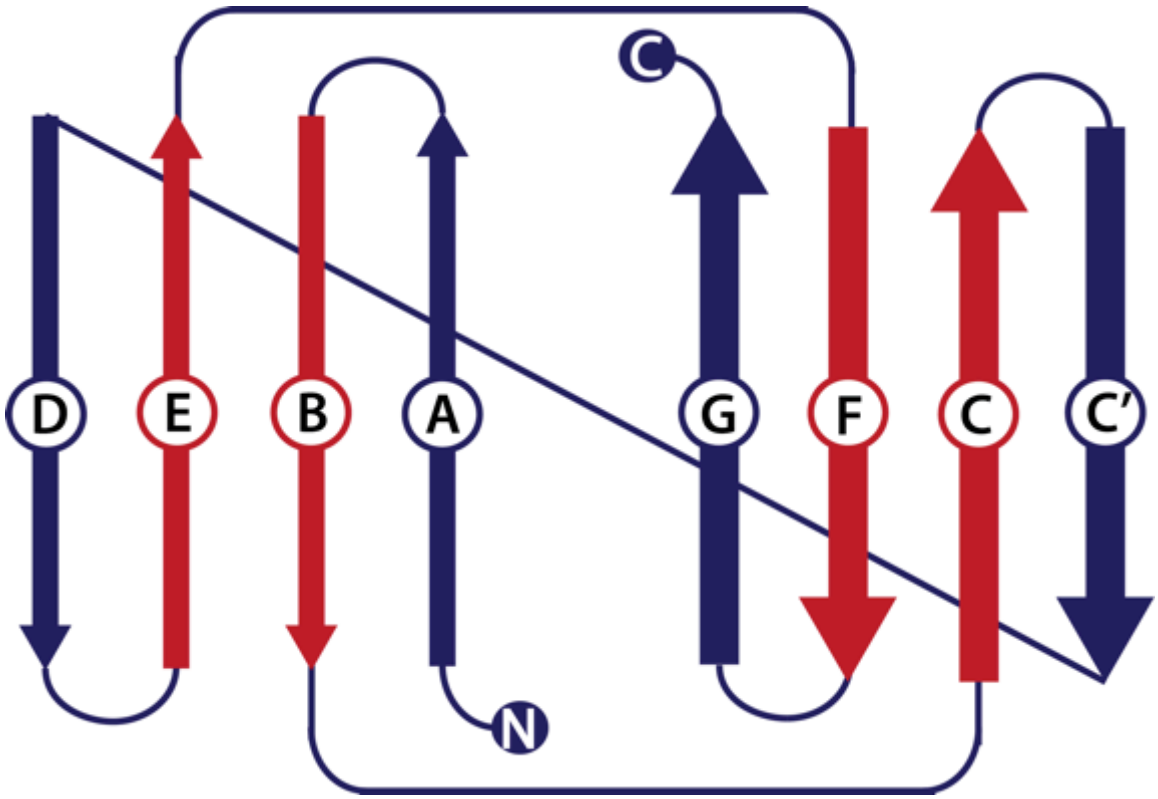

Figure 1.8 Two dimensional topology diagram of the Ig-like protein fold.
The strands (A, B, C, C', D, E, F, G), depicted in arrows, fold into a sandwich of two sheets, distinguished by thick and thin arrows. The core strands (B, C, E and F) are shown in red.

1.4.4. HCM-related R502W mutation in cMyBP-C C3

The gene encoding cMyBP-C in humans is located on chromosome 11p11.2 and contains 35 exons (Gautel et al., 1995). Approximately 42% of the HCM-related mutations are found in the *MYBPC3* gene that encodes cMyBP-C (Bonne et al., 1995; Richard et al., 2003; Watkins et al., 1995).

The majority of *MYBPC3* mutations are splice donor, splice acceptor, deletion or insertion mutations, encoding for truncated proteins lacking the C-terminus, which contains the myosin-binding site and the titin-binding site (Carrier et al., 1998). These truncated proteins will be proteolyzed rapidly by the proteasome complex within the ventricular tissue if they are not incorporated in the myofibrillar assembly (Tanaka et al.,

1992). cMyBP-C is essential for the normal myofibrillar assembly (Freiburg and Gautel, 1996; Furst and Gautel, 1995) and a reduced amount of functional protein, together with the inclusion of truncated cMyBP-C could lead to myofibrillar disassembly and sarcomeric disarray.

The mutation R502W, in the C3 domain of cMyBP-C, has been found in a patient with HCM and was first reported in 2003 by Richards *et al.* (Richard et al., 2003). This mutation is caused by a C10951T nucleotide change within Exon17 of the *MYBPC3* gene (Richard et al., 2003). It is later found that the R502W mutation occurs in 2.4% of unrelated probands of European descent(Saltzman et al., 2010). In a recent study, 25 individuals from 9 unrelated families share the R502W mutations(Page et al., 2012).

1.5. Research objectives

There were no resolved structures for the L29Q mutations in cTnC and the R502W mutations in cMyBP-C at the beginning of this PhD project. My research project, therefore, aims to investigate the structural effects of the HCM-related mutations, L29Q and R502W on cTnC and cMyBP-C, respectively, and how they may affect the functions of each protein.

Chapter 2.

Crystal structure of the cTnC regulatory domain

Note regarding contributions:

Portions of this chapter were published previously in the Acta Crystallographica Section D: Biological Crystallography.

Zhang, X.L., Tibbits, G.F. & Paetzel, M. (2013). Crystal structure of cardiac troponin C regulatory domain with bound Cd^{2+} reveals a closed conformation and unique ion coordination. *Acta Crystallographica Section D Biological Crystallography* **D69**, 722-734

2.1. Overview

In order to study the structure of cTnC and the mutational effect of L29Q on the protein, various constructs were created (detailed list in Appendix A) and cTnC was over-expressed in three different lengths: full-length (residues 1-161), N-terminus (residues 1-89) and C-terminus (residues 90-161). Various residues were also mutated to characterize the protein's function and to make the protein suitable for specific experiments. cTnI and cTnT were also over-expressed to reconstruct the troponin complex for co-crystallization purposes (Appendix B). Purification of these proteins using ion exchange chromatography and size exclusion chromatography was successful and numerous crystallization trials were set up for different cTnC proteins and troponin complexes.

Several hits from the crystallization screenings of different cTnC proteins were observed, but no hits were obtained from the crystallization screenings of the troponin complexes, despite the range of efforts made so far (Appendix C). Of the hits for cTnC, one condition, which contained Cd^{2+}, yielded crystals that diffracted well. From this crystallization condition, we obtained crystals for the wild-type (WT), L29Q and NIQD cNTnC and thus determined their high-resolution structures.

In this chapter, the highest-resolution structure of WT human cNTnC is described. The structures of the L29Q and NIQD mutants are compared with their WT counterparts, and the effect of the mutations are discussed. Since the structure was obtained with Cd^{2+} ions present, several interesting features are observed. For example, Cd^{2+} ions replaced Ca^{2+} in the classical calcium-binding site of the protein and intermolecular Cd^{2+} ions promoted crystallization. A refolded cNTnC was also crystallized to study the effect of refolding on the protein.

2.2. Materials and methods

2.2.1. WT cNTnC over-expression and purification

The plasmid pET21a_WT_cNTnC_1-89 previously described (Li et al., 2011) was transformed into *Escherichia coli* expression host BL21(DE3) competent cells. The expressed protein (cNTnC) is 89 residues long (residues 1-89 of human troponin C, UniProt database accession number P63316) with a calculated molecular mass of 10,062 Da and a theoretical isoelectric point of 4.0. Cells containing the plasmid were grown at 37°C in LB media (100 µg/mL ampicillin) until the OD600 reached 0.6. The cells were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and were incubated for an additional 4 hours at 37°C. Cell pellets were collected by centrifugation at 6000 g for 7 minutes at 4 °C and resuspended in resuspension buffer [10 mM Tris-HCl pH 8.0, 1 mM dithiothreitol (DTT)] (3 times the volume of the pellet). Cells were sonicated for 45 seconds at 30% amplitude using a Model 500 Dismembrator (Fisher Scientific) and lysed using an Avestin EmulsiFlex-C3 high-pressure homogenizer set at 15,000 - 20,000 psi for 5 minutes. The lysate was centrifuged at 28,964 g for 35 minutes at 4 °C and then the supernatant was loaded onto a Q-sepharose anion exchange column (5 mL column volume; GE HealthCare) pre-equilibrated with resuspension buffer. After washing the column with 10 mL of resuspension buffer, elution was performed in five steps with increasing NaCl concentrations (100 mM to 500 mM in 100 mM increments). Samples from the elution fractions were run on a 15% SDS-PAGE gel. Fractions containing cNTnC were combined and concentrated to a 5mL volume using an Amicon centrifugal filter unit (5-kDa cutoff, Millipore). The protein was then further purified using a size exclusion chromatography column (HiPrep 26/60 Sephacryl S-100 High-Resolution) on an AKTA Prime system (Pharmacia Biotech). The column was equilibrated with 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1mM DTT and run at a flow rate of 1.0 mL/min. All fractions containing cNTnC were identified using 15% SDS-PAGE and concentrated to 20 mg/mL using an Amicon centrifugal filter unit (5kDa cutoff, Millipore). The concentration was determined by UV_{280} absorption with a NanoDrop Spectrophotometer ND-100 (Thermo Scientific) using an extinction coefficient of 1490 $M^{-1}cm^{-1}$ (ProtParam) (Wilkins et al., 1999).

2.2.2. L29Q and NIQD cNTnC over-expression and purification

The L29Q and NIQD cNTnC constructs (residues 1-89) were made from full-length L29Q and NIQD cTnC constructs by converting Lys90 to a termination codon by site-directed mutagenesis (SDM). The SDM primers used were described previously (Li et al., 2011). The over-expression and purification procedures were the same as the ones described for WT cNTnC.

2.2.3. Protein refolding

For one of the crystal structures (PDB: 4GJE), the purified WT cNTnC was subjected to refolding before being used for crystallization. The protein was first denatured in buffer that contained 4 M guanidine-HCl, 10 mM Tris-HCl, pH 8.0 and 1 mM DTT. It was then dialyzed against buffer that contained 10 mM Tris-HCl, pH 8.0 and 1 mM DTT to remove the denaturant. The protein stayed soluble throughout the refolding process.

2.2.4. Crystallization

The sitting-drop vapor diffusion method was used to grow the crystals in this study. Each crystallization drop was prepared by mixing 1 µL of purified protein (20 mg/mL) with 1 µL of reservoir solution and equilibrated against 1 mL of reservoir solution. The reservoir solution used for the crystals that generated the high-resolution WT cNTnC and refolded WT cNTnC data contained 0.05 M cadmium sulfate octahydrate, 1.0 M sodium acetate and 0.1 M HEPES, pH 7.5, while the crystals for the SAD dataset, L29Q and NIQD cNTnC, were obtained from conditions containing 0.02 M cadmium sulfate octahydrate, 0.6 M sodium acetate and 0.1 M Tris-HCl, pH 8.0. In both cases, crystals were grown for three days at room temperature (298 K). The crystals were flash cooled in liquid nitrogen in a 0.2 mm CrystalCap copper magnetic loop (Hampton Research) using a cryoprotectant solution that had the same chemical components as the respective reservoir solutions but with glycerol replacing 30% of the water.

2.2.5. Anomalous data collection

The SAD dataset for the WT cNTnC was collected at Beamline X4A at the National Synchrotron Light Source using a wavelength of 0.97910 Å. The detector was an ADSC QUANTUM 4r CCD. A total of 720 images were collected with a 0.5-degree oscillation angle and 10-second exposure time while the crystal-to-detector distance was 160 mm. Collected diffraction images were integrated and scaled with HKL2000 (Otwinowski & Minor, 1997). The reflections were detectable to 1.7 Å resolution and the Matthew's Coefficient and % solvent were calculated using the CCP4 suite (Collaborative Computational Project, Number 4., 1994). Crystal parameters and data collection statistics are given in Table 2.1.

2.2.6. High-resolution dataset data collection

The high-resolution diffraction data set for WT cNTnC was collected at Beamline 08ID-1 of the Canadian Light Source, Canadian Macromolecular Crystallography Facility, using a wavelength of 0.97949 Å. A MARMOSAIC 300mm CCD detector was used with a crystal-to-detector distance of 180 mm and a total of 201 images were collected with a 0.5-degree oscillation angle and a 0.6-second exposure time. The diffraction data were processed with HKL2000 (Otwinowski & Minor, 1997). The reflections were detectable to 1.4 Å resolution and the Matthew's coefficient and % solvent were calculated using the CCP4 suite (Collaborative Computational Project, Number 4., 1994). Crystal parameters and data collection statistics are given in Table 2.1.

2.2.7. Refolded WT cNTnC data collection

The refolded WT cNTnC dataset was collected at Beamline 08B1-1 of the Canadian Light Source, Canadian Macromolecular Crystallography Facility, using a wavelength of 0.97921 Å. A Rayonix MX300HE CCD detector was used with a crystal-to-detector distance of 180 mm and a total of 180 images were collected with a 1-degree oscillation angle and a 5-second exposure time. The diffraction data were processed with HKL2000 (Otwinowski & Minor, 1997). The reflections were detectable to 1.5 Å resolution. Crystal parameters and data collection statistics are given in Table 2.1.

2.2.8. L29Q and NIQD cNTnC data collection

The L29Q and NIQD cNTnC datasets were collected at Beamline 08ID-1 of the Canadian Light Source, Canadian Macromolecular Crystallography Facility, using a wavelength of 0.97949 Å. A MARMOSAIC 300mm CCD detector was used with a crystal-to-detector distance of 210mm for L29Q cNTnC. A total of 120 images were collected with a 0.75-degree oscillation angle and a 1-second exposure time. For NIQD cNTnC, the detector distance was 200mm and 90 images were collected with a 1-degree oscillation and 1-second exposure time. The diffraction data were processed with HKL2000 (Otwinowski & Minor, 1997). The reflections were detectable to 1.7 Å resolution for L29Q cNTnC and 1.8 Å for NIQD cNTnC. Crystal parameters and data collection statistics are given in Table 2.1.

2.2.9. SAD phasing and structure solution

The program HKL2MAP (Pape & Schneider, 2004) was used to search for cadmium sites at 2.2 Å resolution and five sites with an occupancy higher than 0.5 were found. The electron density map generated by HKL2MAP (Pape & Schneider, 2004) was then used for model building in PHENIX (Adams et al., 2010). The final model was optimized by manual building in Coot (Emsley & Cowtan, 2004) and rounds of restrained refinement in REFMAC5.5.0109 (Murshudov et al., 1997). The data collection and refinement statistics are presented in Table 2.1.

2.2.10. Structure determination and refinement of the high-resolution dataset

Phases for the high-resolution data of WT cNTnC were obtained by molecular replacement using the program PHASER (McCoy et al., 2007). The coordinates from the SAD dataset structure were used as a search model. Manual adjustments to the initial model were performed in the program Coot (Emsley & Cowtan, 2004) and the structure was refined using restrained refinement in REFMAC5 (Murshudov et al., 1997). Seven Cd^{2+} ions and two Ca^{2+} ions are observed in the structure. The two different types of ions were distinguished by fitting Cd^{2+} or Ca^{2+} into the difference electron density maps. Because of the significant difference in the number of electrons between these ions, only

one of the two ions would fit the electron density without creating positive or negative difference density after refinement.

2.2.11. Structure determination and refinement of the L29Q, NIQD and refolded WT cNTnC

Phases for these three datasets were obtained by molecular replacement using the program MOLREP (Vagin & Teplyakov, 2010), where the coordinates from the high-resolution dataset structure were used as a search model. Manual adjustments to the initial model were performed using the program Coot (Emsley & Cowtan, 2004) and the structure was refined using restrained refinement in REFMAC5 (Murshudov et al., 1997). The complete data collection and refinement statistics are summarized in Table 2.1.

2.2.12. Structural analysis

The root-mean-square deviation (RMSD) values reported for all-atoms protein superpositions were calculated using PyMOL (DeLano, 2002) and atomic distances were measured using the program Coot (Emsley & Cowtan, 2004). To visualize central hydrophobic cavity exposure, the structures were displayed in inverted surface cavity mode using PyMOL (DeLano, 2002).

2.3. Results

2.3.1. Protein purification, crystallization and structure solution

Human cNTnC (residues 1-89) with WT primary sequence was successfully expressed, purified (Figure 2.1) and crystallized (Figure 2.2) and its structure was determined by anomalous diffraction methods. This structure was then used as a molecular replacement search model to obtain phases for a higher-resolution data set. The 1.4 Å resolution structure contains one molecule in the asymmetric unit (ASU) and includes 86 of the 89 residues within the protein. Three residues (residues 67-69) are not modeled due to lack of electron density. The final refined model has an R_{work} of 13.5 % and an R_{free} of 14.8 %. Along with the 86 residues from the single protein chain, 7 cadmium, 2 calcium, 2 acetate ions and 49 water molecules are modeled. The average

B-factor of the structure is 25.6 $Å^2$. Ramachandran plot statistics indicate no residue outliers present in the final model (Table 2.1). The atomic coordinates and structure factors have been deposited in the Protein Data Bank and are assigned the following PDB codes: 3SD6 (high-resolution WT cNTnC); 3SWB (SAD phasing WT cNTnC); 4GJE (refolded WT cNTnC); 4GJF (L29Q cNTnC); 4GJG (NIQD cNTnC).

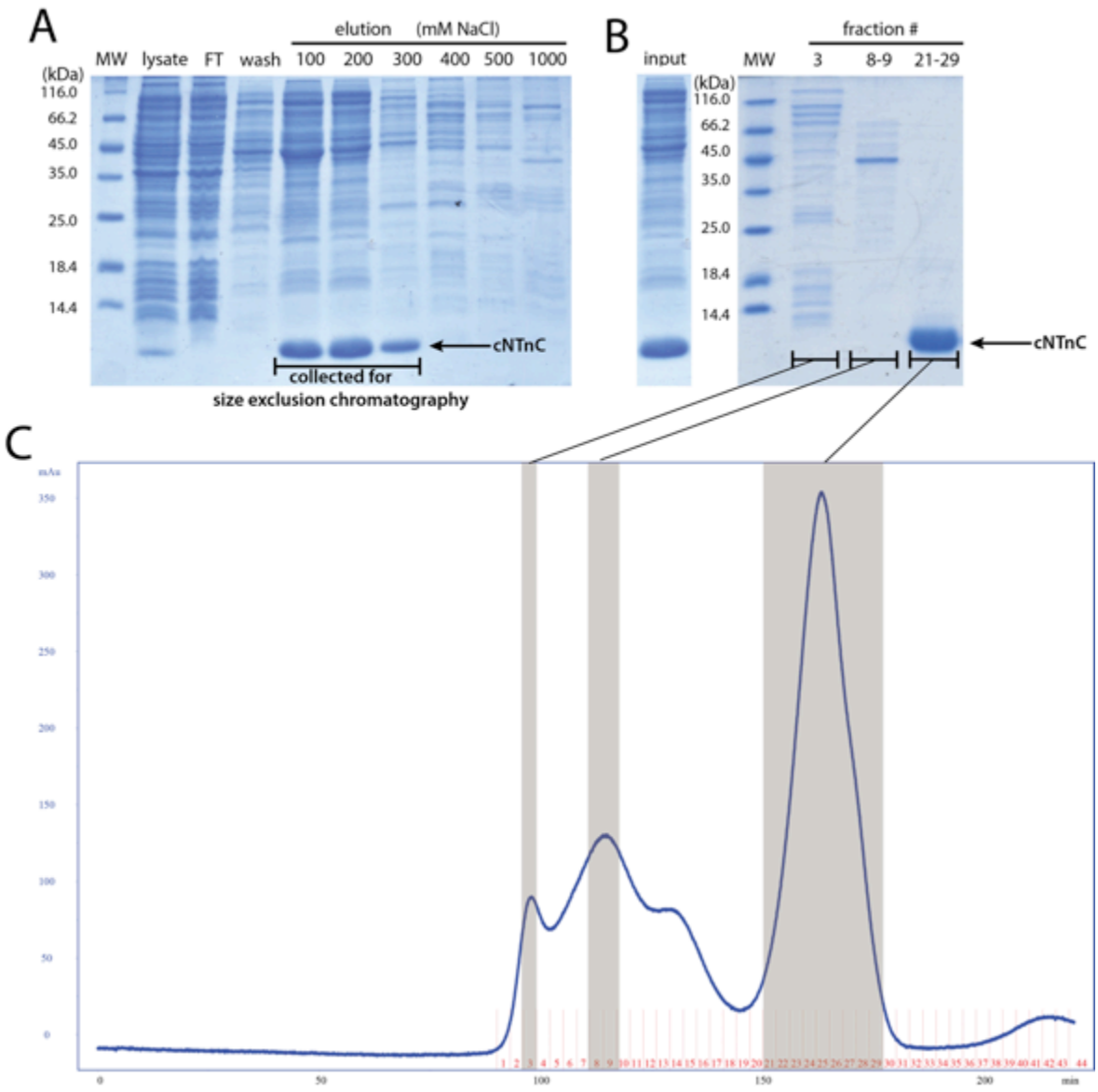

Figure 2.1 Q-sepharose anion exchange chromatography and size exclusion chromatography of WT cNTnC.

A. Elution fractions of WT cNTnC from Q-sepharose anion exchange chromatography were run on a 15% SDS-PAGE gel. The NaCl concentrations used to elute the sample are shown on top of each lane. The molecular marker is shown on the far left. **B.** Elution fractions of WT cNTnC from size exclusion chromatography Sephacryl S-100 were run on a 15% SDS-PAGE gel. The fraction numbers corresponding to those in C are shown on top of each lane. The molecular marker is shown on the far left. **C.** Chromatogram of WT cNTnC from Sephacryl S-100 size exclusion chromatography. Flow rate was 1 mL/min.

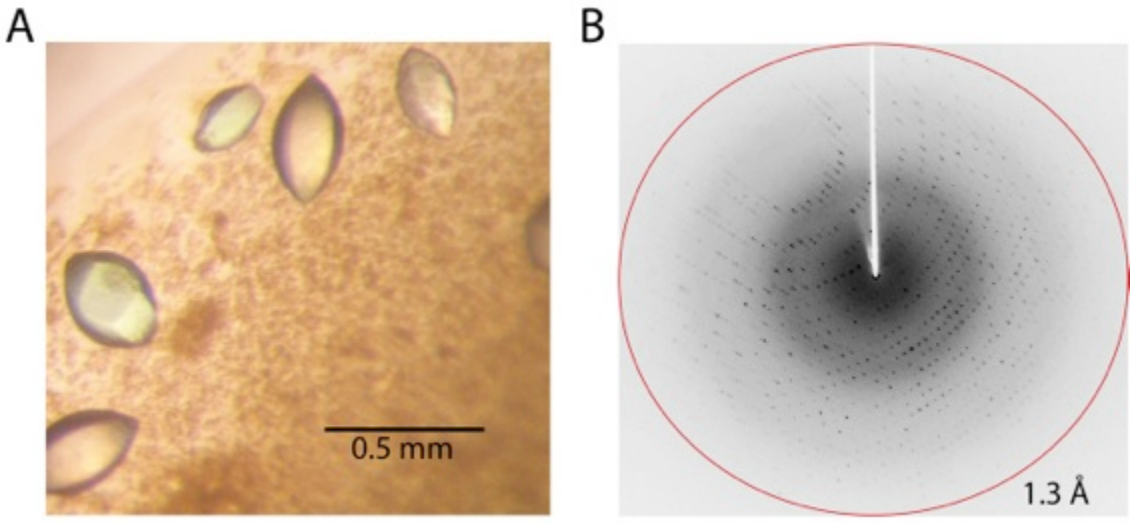

Figure 2.2 Crystals of WT cNTnC and their diffraction pattern.
A. Picture of WT cNTnC crystals. The scale is indicated with a ruler. **B.** Diffraction pattern of a WT cNTnC crystal shown in A. The outer edge of the diffraction is at 1.3 Å.

The L29Q, NIQD and refolded WT cNTnC structures were all determined by molecular replacement using the high-resolution dataset as a search model. Refinement statistics of these structures are summarized in Table 2.1.

Table 2.1. Data collection and refinement statistics.

	High resolution (PDB: 3SD6)	**SAD (PDB: 3SWB)**	**Refolded WT (PDB: 4GJE)**	**L29Q (PDB: 4GJF)**	**NIQD (PDB: 4GJG)**
Beamline	CLS 08ID-1	NSLS X4A	CLS 08B1-1	CLS 08ID-1	CLS 08ID-1
Crystal Parameters					
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P6_122$
a,b,c (Å)	49.6, 49.6, 116.0	49.8, 49.8, 117.4	49.6, 49.6, 116.3	49.3, 49.3, 113.4	49.6, 49.6, 115.1
Matthews coefficient	2.11	2.15	2.06	1.98	2.02
Solvent content (%)	41.72	42.94	40.45	37.87	39.21
Data Collection Statistics					
Wavelength (Å)	0.97949	0.97910	0.97921	0.97949	0.97949
Resolution (Å)	42.93 – 1.37 (1.42 – 1.37)[a]	43.12 – 1.67 (1.73 – 1.67)	42.98 – 1.60 (1.66 – 1.60)	42.66 – 1.90 (1.97 – 1.90)	42.91 – 2.00 (2.07 – 2.00)
Total reflections	190197	383749	242530	66732	57016
Unique reflections	17420 (1251)	10488 (982)	11912 (1155)	6990 (662)	6190 (594)
R_{merge} [b]	0.050 (0.442)	0.128 (0.432)	0.058 (0.382)	0.062 (0.327)	0.091 (0.409)
Mean $(I)/\sigma(I)$	64.07 (2.283)	102.75 (11.632)	84.85 (11.967)	55.09 (8.650)	46.91 (8.441)
Completeness (%)	97.0 (79.8)	99.9 (99.9)	99.9 (100.0)	99.7 (99.8)	99.8 (100.0)
Multiplicity	10.5 (4.3)	35.8 (17.8)	20.4 (20.6)	9.5 (9.8)	9.2 (9.1)
Phasing Statistics					
Number of Cd sites	-	5	-	-	-
FOM	-	0.713	-	-	-
Refinement Statistics					
Protein molecules in A.U.	1	1	1	1	1
Residues	86	86	86	86	86
Waters	49	45	67	45	34
Cadmium	7	6	7	8	7
Calcium	2	3	2	1	3

Acetate	2	1	3	1	1
Glycerol	-	-	-	1	1
Total number of atoms	745	738	766	742	731
R_{cryst} [c] / R_{free} [d] (%)	13.5 / 14.8	13.5/ 16.5	12.1/ 15.0	14.2/ 17.8	14.4/ 19.8
Average B-factor (Å²)					
All atoms	25.6	24.7	28.3	40.5	38.7
Protein	25.3	24.1	26.8	38.8	37.2
Water	38.6	36.2	43.5	63.6	62.4
Cadmium	19.6	19.5	22.2	35.4	39.7
Calcium	28.8	27.0	27.9	175.4	76.8
Acetate ion	29.5	27.7	35.1	36.9	41.7
Glycerol	-	-	-	50.6	50.2
RMSD on angles (º)	1.833	1.908	1.735	1.640	1.585
RMSD on bonds (Å)	0.018	0.018	0.018	0.016	0.015
Ramachandran plot statistics (Percentage of residues in regions):					
Preferred regions	80 (97.56%)	80 (97.56%)	80 (97.56%)	80 (97.56%)	81 (98.78%)
Allowed regions	2 (2.44%)	2 (2.44%)	2 (2.44%)	2 (2.44%)	1 (1.22%)
Outliers	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)

2.3.2. Overall protein fold of WT cNTnC

The WT cNTnC structure contains five α-helices (αN, αA, αB, αC and αD) and two short β-strands (β1 and β2) (Figure 2.3). The EF1 motif is comprised of αA (residues 13-27), loop 1 (residues 28-37), and αB (residues 38-48), while the EF2 motif is comprised of αC (residues 53-64), loop 2 (residues 65-72), and αD (residues 73-86) (Figure 2.3). A short two-stranded anti-parallel β-sheet, sometimes referred as the 'EF β-scaffold' (Grabarek, 2006), formed by β1 (residues 36-37) and β2 (residues 71-72) within the EF1 and EF2, brings the two ion binding loops into close proximity to each other.

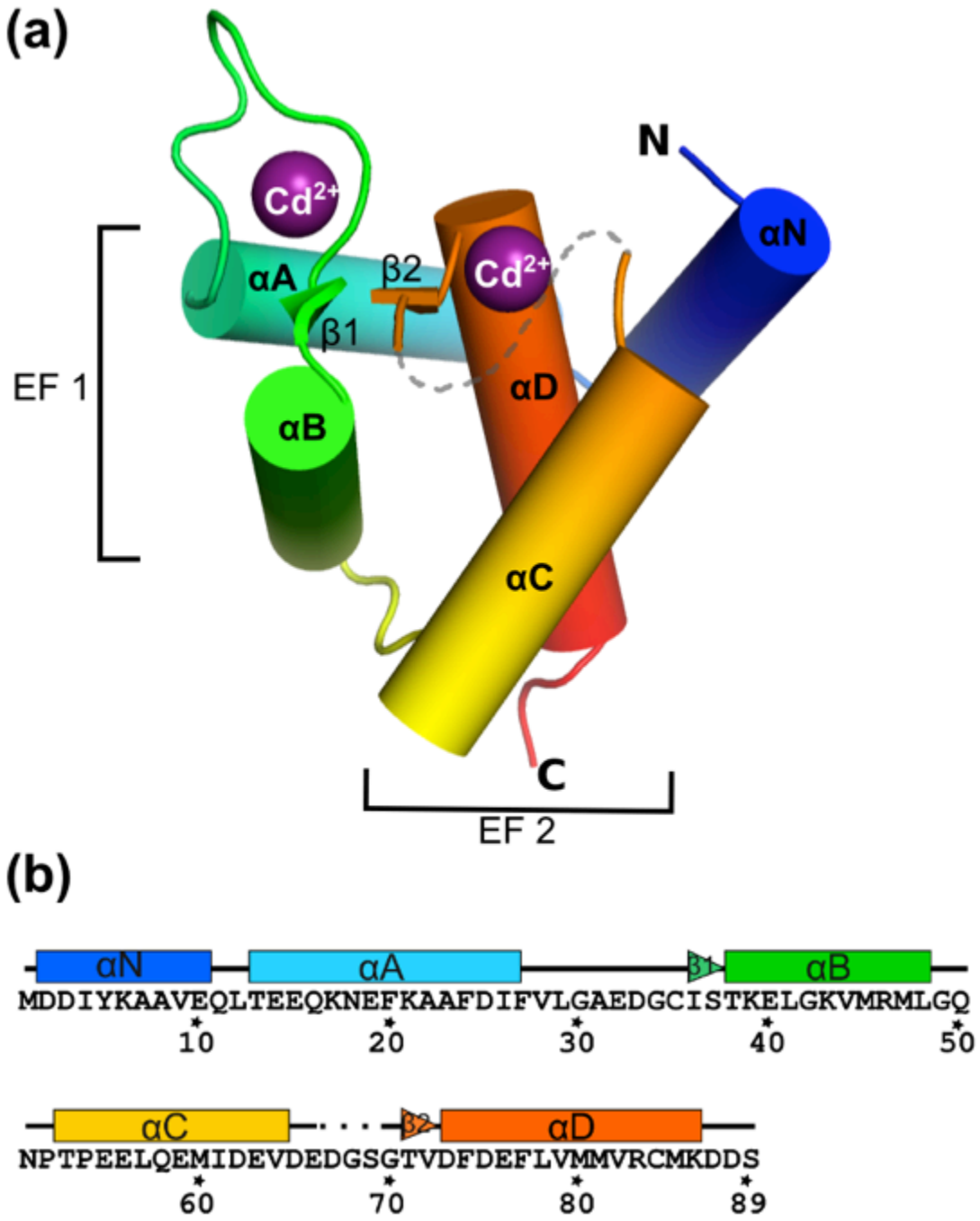

Figure 2.3 Overall protein fold of Cd^{2+}-bound human cNTnC.
(a) WT human cNTnC contains five α-helices (αN, αA, αB, αC and αD) and two β-strands (β1 and β2). Two EF-hand motifs are present in the structure: EF1 and EF2. Each EF loop coordinates a Cd^{2+} ion. The three residues missing from the EF2 loop region owing to a lack of electron density are represented by a grey dashed line. (b) Secondary structure of Cd^{2+}-bound human cNTnC aligned with the primary sequence (UniProt accession No. P63316). The α-helices and β-strands are presented in the same colour scheme as in (a).

Clear and sharp electron density was observed for every residue within the protein chain except for the three residues, Asp67, Gly68 and Ser69, which are within the EF2 loop (Figure 2.4). An SDS-PAGE gel of the protein crystal that generated the structure showed a single band at ~10kDa, which corresponds to the approximate molecular mass of cNTnC (Figure 2.4). This suggests that the lack of electron density for the three EF2 loop residues is a result of thermal motion rather than a result of proteolytic cleavage or radiation damage at this site. Investigation of the crystalline lattice shows that this region of the loop faces a solvent channel.

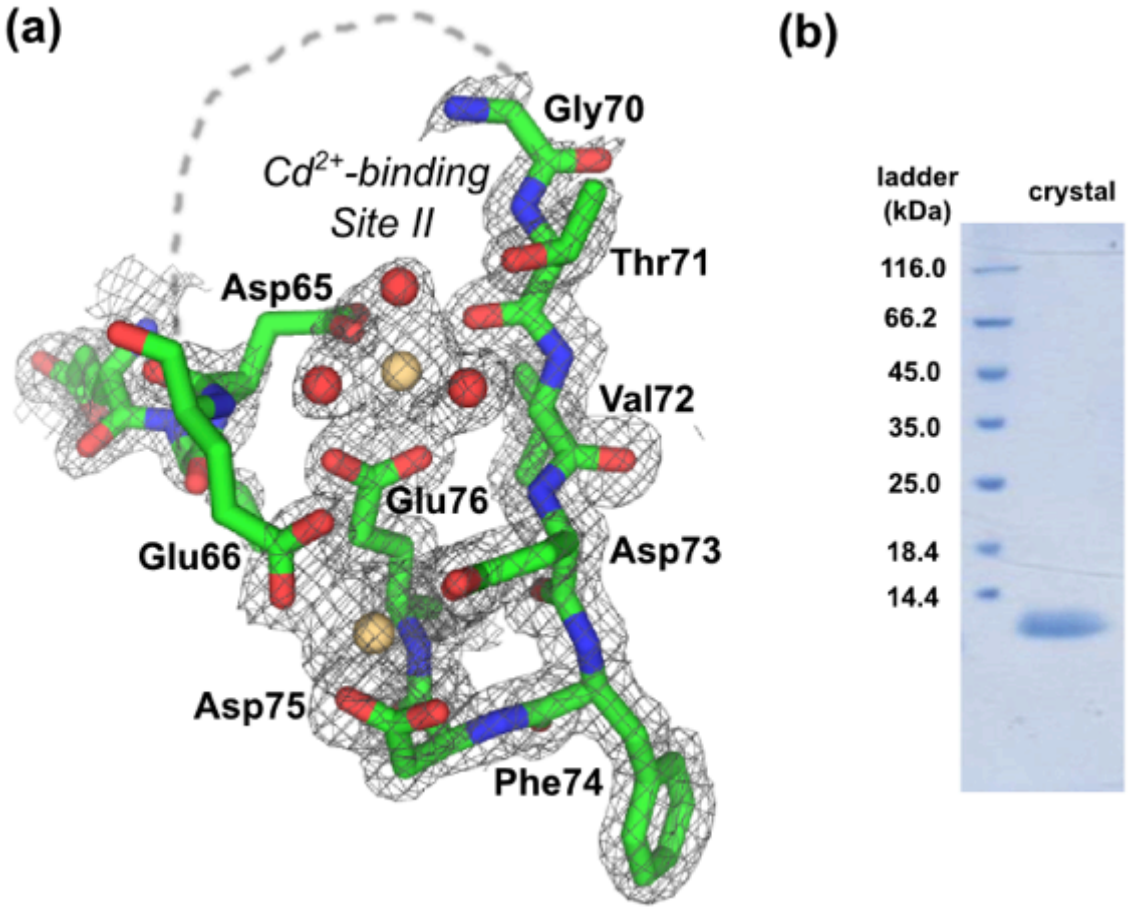

Figure 2.4 Three residues within EF2 are disordered in the Cd^{2+}-bound cNTnC structure.

(a) Electron density ($2F_o$-F_c, at the 1.0σ level) is shown for the region near the Cd^{2+} bound in EF2. No clear electron density is visible for residues 67, 68 and 69. The missing residues are represented by a dashed line. The cNTnC residues are shown as sticks (carbon, green; nitrogen, blue; oxygen, red; sulfur, yellow). The Cd^{2+} ion and water molecules are represented by light orange and red spheres, respectively. (b) An SDS-PAGE gel (stained with PageBlue) of the cNTnC crystal used in data collection shows a single protein band at ~10kDa.

2.3.3. L29Q and NIQD mutations caused a subtle main chain shift in the EF1 loop of cNTnC

The electron density maps confirmed the mutations in the L29Q and NIQD structures. In Figure 2.5a, the L29Q cNTnC dataset was refined with a WT leucine at position 29. Positive difference density (Fo - Fc, at sigma level 3.0) appears at the end, and along the side of the leucine side chain. This suggests that the side chain of residue 29 should be more extended than leucine (i.e. branching point at Cδ instead of Cγ), and glutamate is a better fit to the density.

In Figure 2.5b, the NIQD cNTnC dataset was refined with WT residues: valine at position 28, leucine at position 29 and glycine at position 30. For V28I, positive difference density (sigma level 3.0) showed up at one of the Cγ ends of valine. This is a clear indication for isoleucine instead of valine. Also, positive difference densities for both L29Q and G30D render distinct shapes for the side chains of mutation residues Gln29 and Asp30 (Figure 2.5). This strongly supports the presence of L29Q and G30D in NIQD cNTnC. The superimposition of WT, L29Q and NIQD cNTnC (Figure 2.5) showed no significant overall structural difference among the three. The all-atom RMSD value between L29Q and WT cNTnC is 0.187 Å and 0.123 Å between NIQD and WT cNTnC (DeLano, 2002).

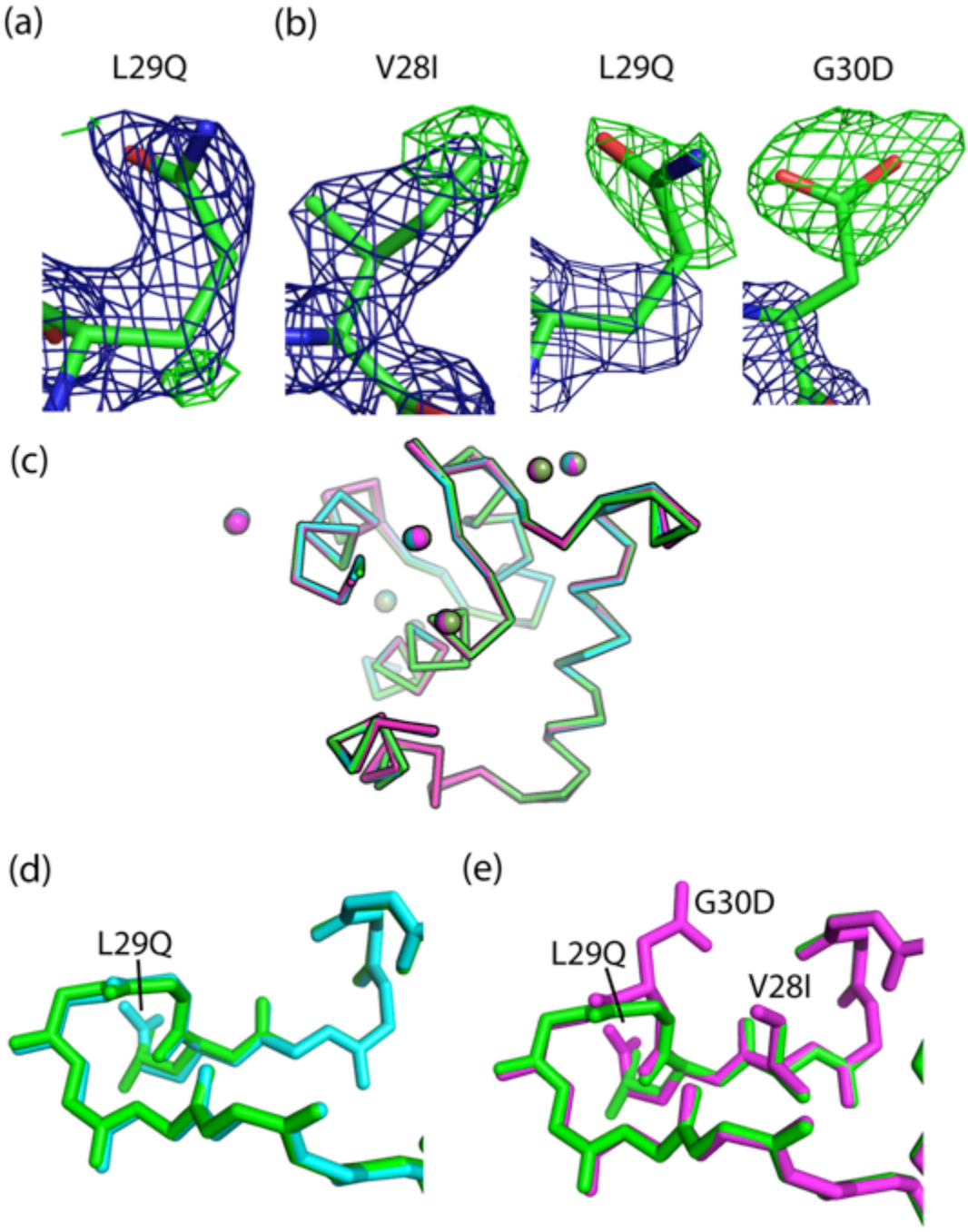

Figure 2.5 Crystal structures of L29Q and NIQD cNTnC.
Electron density (2Fo-Fc, at the 1.0σ level, blue mesh) and difference density (Fo-Fc, at the 3.0σ level, green mesh) are shown for the region of cNTnC with (a) an L29Q mutation in L29Q cNTnC and (b) V28I, L29Q and D30G mutations in NIQD cNTnC. Each of these figures shows the electron densities when refined with the WT sequence residues (i.e. Val28, Leu29 and Gly30). The mutated residues are shown as sticks (carbon, green; nitrogen, blue; oxygen, red). (c) Superimposition of WT (green), L29Q (cyan) and NIQD (magenta) cNTnC (shown as ribbons). The Cd^{2+} ion in each structure is shown as a sphere. Superimposition of (d) L29Q (cyan) and WT (green) cNTnC and (e) NIQD (magenta) and WT cNTnC with a focus on the EF1 loop region. The main-chain shift is visible in both cases. Protein chains are shown as sticks.

L29Q is located at the beginning of the EF1 loop of cNTnC. This non-polar leucine to polar glutamine mutation extends the side chain of residue 29 by one atom. Its side chain points away from the protein molecule and into the solvent channel between neighboring protein molecules. The presence of Oε1 and Nε2 of Gln29 did not create extra hydrogen bonds at residue position 29 in the structure. However, the main chain between residues 29 and 32 shifted towards the solvent channel by up to 0.5 Å in L29Q cNTnC compared to WT. The presence of the L29Q mutation did not change the Cd^{2+} coordination in EF1 compared to its WT counterpart.

As well, in the NIQD cNTnC mutant, which contains V28I, L29Q and G30D mutations in the EF1 loop, the Cd^{2+} coordination in EF1 is not altered compared to its WT counterpart. The main chain between residues 29 and 32 is also shifted towards the solvent channel in the NIQD mutant by up to 1.0 Å compared to WT. The change from glycine at position 30 to aspartic acid introduces a new hydrogen bond between the Oδ1 of the Asp30 side chain and the main chain N of Asp30. The two Oδ's of Asp30 also formed hydrogen bonds with three separate water molecules.

It was proposed previously that since EF1 and EF2 loops are structurally linked, the sequence mutation in the EF1 loop may alter the structure of the EF2 loop in cNTnC and thus affect the ion-binding ability of EF2 (Spyracopoulos et al., 1997, Gillis et al., 2007). However, superimposition of WT, L29Q and NIQD cNTnC did not show any differences in the EF2 loop among the three structures (Figure 2.5) nor is the coordination of Cd^{2+} in EF2 changed. Therefore, what accounts for the increased Ca^{2+} affinity (Liang et al., 2008) in L29Q and NIQD cNTnC is likely to be the non-polar to polar mutation (Leu29 to glutamine, and Gly30 to aspartic acid).

Shown in the mutant structures here (Figure 2.5), both residues 29 and 30 are located on the surface of the protein with their side chains facing the solvent. Non-polar to polar mutations at these positions tend to increase the surface hydrophilicity of the protein. It has been demonstrated previously that an increase in surface hydrophilicity in proteins (such as trypsin and subtilisin) results in proteins being more easily activated (Genicot et al., 1996, Smalås et al., 1994). The increased interaction between the protein and solvent decreases the free energy needed to be overcome for conformational change to occur that leads to protein activation (Gillis et al., 2007).

2.3.4. Overall conformation of Cd^{2+}-bound cNTnC resembles Ca^{2+}-bound cNTnC

The overall conformation of our Cd^{2+}-bound cNTnC structure resembles that of the Ca^{2+}-bound structures. An all-atom superimposition of the Cd^{2+}-bound cNTnC crystal structure with a representative chain of the Ca^{2+}-bound cNTnC NMR structure (PDB ID: 1AP4 (Spyracopoulos et al., 1997)) results in an RMSD of 1.65 Å (Figure 2.6).

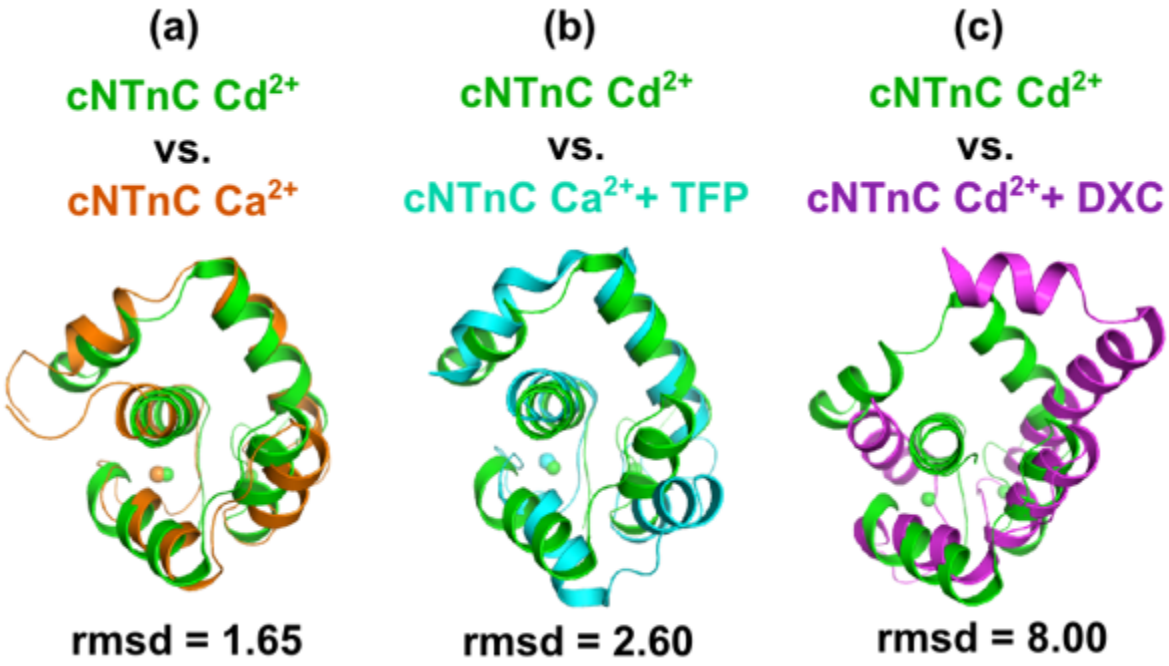

Figure 2.6 Superimposition of the Cd^{2+}-bound cNTnC structure with other cNTnC structures.

All structures are presented as ribbon diagrams. The Cd^{2+}-bound cNTnC is in green and is superimposed with (a) a representative chain from the NMR structure of Ca^{2+}-bound cNTnC (orange, PDB entry 1AP4), (b) chain A of Ca^{2+}-bound cNTnC in compex with TFP (cyan, PDB entry 1WRK) and (c) chain A of Cd^{2+}-bound cNTnC in complex with DXC (magenta, PDB entry 3RV5). The RMSD is shown below each superimposition.

The two other available crystal structures of cNTnC adopt a different conformation compared to the structure presented here. The crystal structures of Ca^{2+}-bound human cNTnC, with Cys35 mutated to serine and with bound trifluoroperazine (TFP) (unpublished data, PDB ID: 1WRK), superimposes on to our structure with a RMSD of 2.60 Å. The helix B position is the major difference between the structures (Figure 2.7b), as it moves away from the centre of the Ca^{2+}-bound cNTnC that is in complex with TFP. The crystal structure of WT human cNTnC coordinating Cd^{2+} in both

EF-loops and in complex with deoxycholic acid (DXC) (Li et al., 2011) reveals a significantly different conformation than the Cd^{2+}-bound cNTnC in the absence of DXC. Helices A in the two structures are oriented approximately 90 degrees apart from each other, and all-atom superposition of the two structures results in a RMSD of 8.00 Å (Figure 2.6).

Examination of the three cNTnC crystal structures reveals that there are differences in the degree of 'openness' (Figure 2.7). The term 'openness' refers to a measure of how far apart helices A/B are positioned relative to helices C/D. The farther apart the two groups of helices are, the more exposed is the central hydrophobic core. Our Cd^{2+}-bound cNTnC represents the most closed (or lowest degree of "openness") structure among the three, while the Cd^{2+}-bound cNTnC in complex with DXC has the highest degree of "openness". The three structures are placed on the close-to-open scale in Figure 2.7.

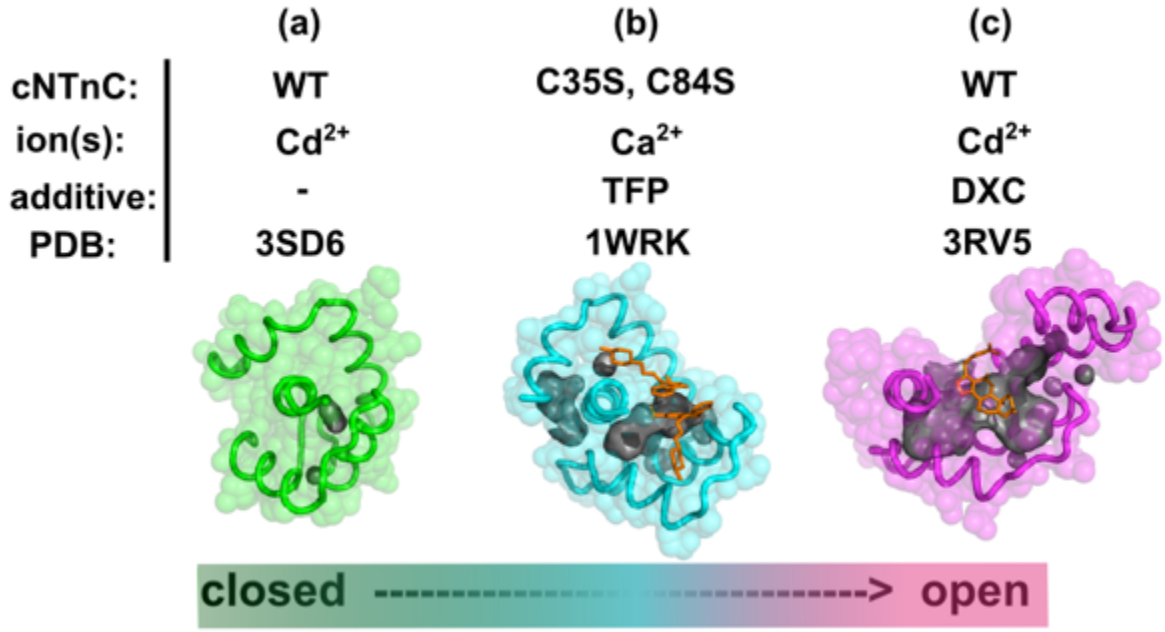

Figure 2.7 Crystal structures of cNTnC placed on a 'closed-to-open' scale.
(a) Cd^{2+}-bound cNTnC (green), (b) Ca^{2+}-bound cNTnC in complex with TFP (trifluoroperazine, orange; protein, cyan) and (c) Cd^{2+}-bound cNTnC in complex with DXC (deoxycholic acid, orange; protein, magenta) are displayed as ribbons with semitransparent surfaces. Dark grey blobs represent the void area of the central hydrophobic cavity on the protein surface in each cNTnC protein.

While TFP is a Ca^{2+}-sensitizing drug that increases the Ca^{2+} affinity of troponin C and muscle contractility (Kass & Solaro, 2006, Endoh, 2007), DXC is a derivative of bile acid (Hofmann, 1999). The two chemicals have similar structural characteristics in that they both contain a hydrophobic ring system and a hydrophilic group at the end. The hydrophobicity of the ring system allows these chemical compounds to bind within the central hydrophobic core of cNTnC, and stabilize the open conformation of cNTnC. Figure 2.7 shows how TFP and DXC are bound at the hydrophobic core of cNTnC.

Before the structure presented here was determined, it was unclear what contributed more to the observed wide open conformation: the Cd^{2+} binding at the EF-loops or the presence of DXC (Li et al., 2011). It is now clear that Cd^{2+} binding to the EF-loops of cNTnC does not by itself induce the wide-open conformation. Instead, it is likely that the presence of compounds such as DXC and TFP in the hydrophobic core of the protein contributes to the stabilization of the open conformation of cNTnC (Figure 2.7). There are examples of other Ca^{2+}-sensitizing drugs that also stabilize the open conformation of cNTnC (Robertson et al., 2010,Wang et al., 2002).

2.3.5. EF1 coordinates Cd^{2+} in a 'distorted' octahedral geometry

The Cd^{2+} ion located in EF1 is coordinated by six ligands (Figure 2.8). The main chain carbonyl oxygen of Cys35 and the oxygen from a water take up the axial positions X and –X respectively. The equatorial plane is constructed by Asp33 Oδ2 (Y), Asp33 Oδ1 (Z), Cys35 Sγ (-Y) and acetate OXT (-Z) (Figure 2.8). This makes the Cd^{2+} coordination at the EF1 loop non-canonical, as the ion is coordinated in a 'distorted' octahedral geometry. This is the first time this coordination geometry has been observed in an EF-loop of cNTnC, but octahedral coordination has been observed in EF-loops of other Ca^{2+}-binding proteins (Gifford et al., 2007, Grabarek, 2011); for example, the Ca^{2+} binding apoptosis-linked gene-2 protein (Jia et al., 2001). Ca^{2+} is coordinated in an octahedral geometry using only residues from the N-terminal half of the EF-loop. The octahedral coordination is also utilized by EF-loops to coordinate ions that have a smaller ionic radius than calcium (Grabarek, 2011). For example, in calbindin D9k, the C-terminal EF-loop utilizes canonical pentagonal bipyramidal coordination when bound to Ca^{2+} (Svensson et al., 1992) but octahedral coordination when bound to Mg^{2+} (Andersson et al., 1997). The octahedral coordination allows the EF-loop to achieve optimum coordination angles with the smaller ion (Grabarek, 2011). In all cases of octahedral coordination described above, including the structure presented here, a common feature is observed: the bidentate ligand at position 12 (-Z), usually a glutamate, is too distant from the coordination sphere and is therefore replaced by a water molecule or an acetate ion. Interestingly, a metalloprotease that does not contain an EF-hand motif has been observed to bind Cd^{2+} (replacing the native Zn^{2+} ion) via 'distorted' octahedral geometry using waters and histidine side chains as ligands (Huang et al., 2002).

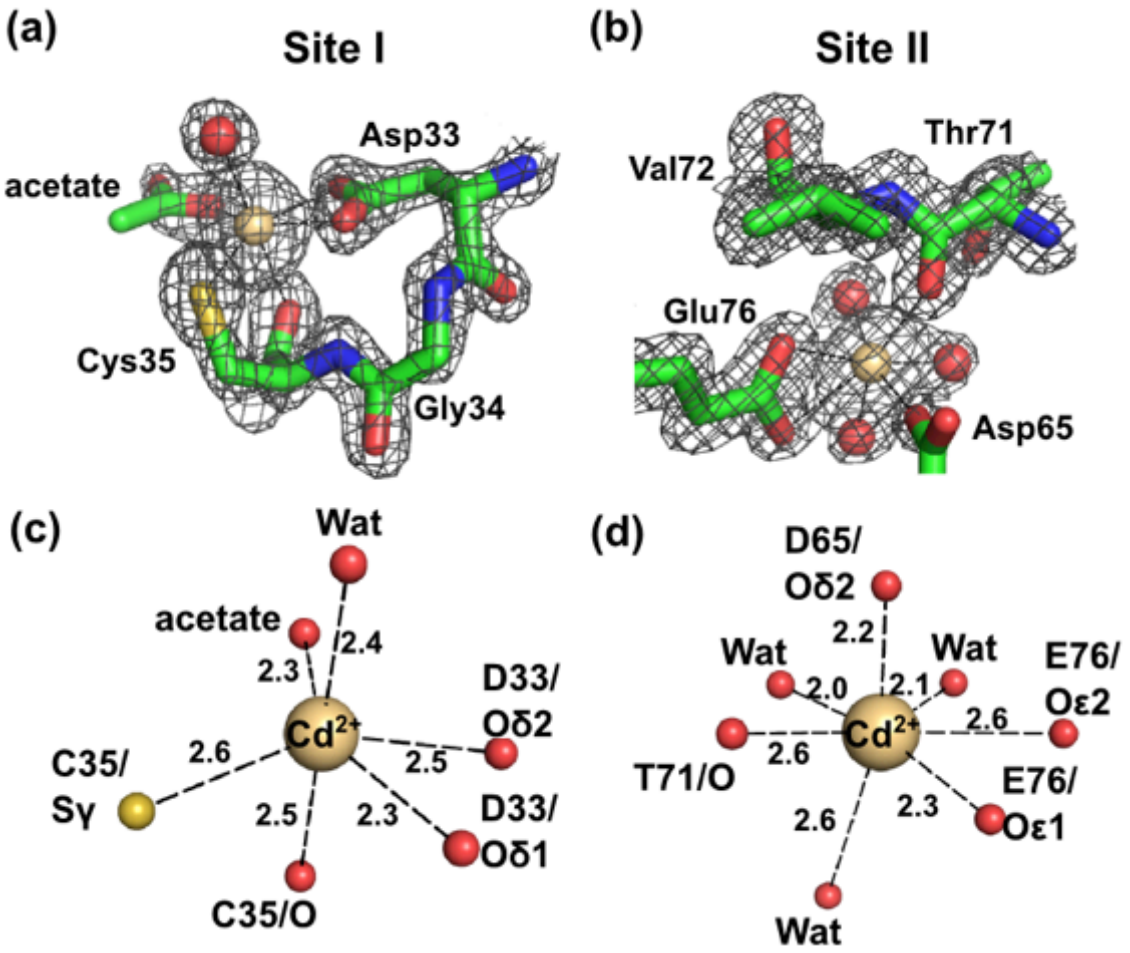

Figure 2.8 Electron density and coordination geometry of Cd^{2+} in the EF1 and EF2 loops of cNTnC.

(a) The electron density for residues within proximity of the Cd^{2+} ion at EF1. (b) The electron density for residues within proximity of the Cd^{2+} ion at EF2. The residues are rendered as sticks (carbon, green; nitrogen, blue; oxygen, red; sulfur, yellow). The Cd^{2+} ion and water molecules are represented by light orange and red spheres, respectively. The σ level of the $2F_o$-F_c electron density is 1.5σ in (a) and 1.0σ in (b). (c) The atoms involved in Cd^{2+} coordination and their atomic distances are shown for EF1. (d) The atoms involved in Cd^{2+} coordination and their atomic distances are shown for EF2. The Cd^{2+} ions are depicted as light orange spheres. Atomic distances are indicated in Å. The coordinating atoms are shown as small spheres (oxygen, red; sulfur, yellow).

2.3.6. EF2 coordinates Cd^{2+} in a canonical pentagonal bipyramidal geometry using only three residues. Occupancy of Cd^{2+} in EF1 and EF2 supports the Ca^{2+}-binding nature of the EF-hand motif

Our structure shows that EF2 coordinates Cd^{2+} in a canonically pentagonal bipyramidal geometry (Figure 2.8). Four of the ligands come from Asp65 Oδ2 (X), Thr71 O (-Y) and Glu76 Oε1/Oε2 (-Z) and three others from water molecules (-X, Y and Z) (Figure 2.8). Since two of the conserved residues, Asp67 (Y) and Ser69 (Z), are not observed in the electron density (presumably they are in a dynamic state), water molecules substitute for the disordered ligands. To our knowledge, this is the first time such an ion coordination has been observed in cNTnC or any other EF-hand protein. This reveals a tremendous flexibility in this protein's ion coordination repertoire.

The EF-hand motif is normally present in proteins for the purpose of Ca^{2+}-binding. The Ca^{2+}-binding loop in the EF-hand motif has evolved to be very specific in sequence and spatial arrangement for the accommodation of the Ca^{2+} ion. The coordinating ligands come together to coordinate Ca^{2+} in a pentagonal bipyramidal geometry. Any ion with a bigger or smaller ionic radius would not fit as well in the EF loop.

In contrast to the full occupancy and sharp electron density of Cd^{2+} coordination in EF1, the occupancy of Cd^{2+} in EF2 is partial (0.5 in WT cNTnC, 0.6 in L29Q and NIQD cNTnC) and the electron density of three EF2 loop residues was missing. The B-factor values of Cd^{2+} in EF2, in all the structures presented here, are higher than the Cd^{2+} in their corresponding EF1 loops. This is the first structural evidence for a disordered region in the EF2 loop of troponin C. For the functional EF2 that is designed for Ca^{2+}-binding with its conserved coordination residues, the presence of Cd^{2+} likely destabilizes and disorders this region which results in the missing electron densities for the three loop residues.

2.3.7. Refolded cNTnC

The 1.6 Å resolution structure of the *in vitro* refolded WT cNTnC rendered the same conformation as the WT cNTnC (RMSD = 0.057) that was natively folded within the cytoplasm of the expression host. This suggests that refolding of cNTnC does not alter the fold of the protein.

The troponin complex has been studied for decades and troponin C has been purified in various ways; some purified from muscle tissue directly and others expressed using *Escherichia coli*. Regardless of the initial source of the protein, the majority of the previous studies used refolded troponin C. The WT, L29Q and NIQD cNTnC structures presented here were expressed in *Escherichia coli* and purified under non-denaturant conditions. To investigate any possible structural difference between refolded and non-refolded cNTnC, refolded cNTnC was crystallized and the same crystal form was generated as for the non-refolded WT cNTnC. Therefore, regardless of the refolding status of cNTnC, it can be seen that the protein renders the same conformation, as indicated by the crystal structures presented here.

2.4. Discussion

2.4.1. Cys35 plays an important role in coordinating the Cd^{2+} in EF1

Although it did not evolve to bind Ca^{2+} at physiological conditions, the vestigial EF1 site in the cNTnC structure shown here coordinates a Cd^{2+} ion in full occupancy with strong electron density that clearly reveals the coordination geometry (Figure 2.8). While the similarity of ionic radius between Cd^{2+} and Ca^{2+} allows Cd^{2+} to fit in the EF1 loop, the presence of cysteine in the EF1 loop is proposed to play a major role in coordinating Cd^{2+} at this vestigial site. Comparing EF1 in the cNTnC to skeletal troponin C, which has a functional EF1 that coordinates Ca^{2+}, the unique presence of cysteine (Cys35) in the vestigial EF1 of cNTnC makes this loop more adaptable for Cd^{2+} binding. The cysteine Sγ atom tends to form a thiolate bond with the soft metal Cd^{2+} rather than with the hard metal Ca^{2+}(Jalilehvand et al., 2009). Although the amino acid sequence of EF1 makes the classical pentagonal bipyramidal coordination of Ca^{2+} difficult, the presence of cysteine and the flexibility of Cd^{2+} coordination in the protein makes the EF1 loop a suitable niche for Cd^{2+}, as shown in the structure presented here. Similar examples are cysteine-rich proteins, such as metallothioneins (MT), which can actively remove Cd^{2+} *in vivo* by thiolate coordination with cysteine residues (Boulanger et al., 1982, Melis et al., 1983).

The cysteine at the EF1 loop, Cys35, also participates in coordinating Cd^{2+} intermolecularly. This connection between cysteine of one molecule and Cd^{2+} from the neighbouring molecule plays an important role in crystal packing. Parallel crystallization trials with identical experimental conditions were set up for both WT and C35S cNTnC. While the WT cNTnC crystallized and generated the high-resolution structure, the C35S cNTnC yielded no crystals.

2.4.2. Intermolecular Cd^{2+} promotes crystallization

Cd^{2+} ions are bound at both the EF1 and EF2 loops (Figure 2.3). As well as these two Cd^{2+} ions, five other Cd^{2+} and two Ca^{2+} ions are found in each ASU, where they are positioned between the protein and its symmetry-related molecules. Crystallization does not occur in these conditions with less than 20 mM cadmium sulfate octahydrate. This high concentration of Cd^{2+} and the large number of Cd^{2+} ions coordinated between neighboring protein molecules suggest that these ions play an important role in crystal packing. A similar situation has been shown for leucine/isoleucine/valine binding protein (LIVBP) (Trakhanov & Quiocho, 1995). Various divalent ions were screened to optimize crystallization conditions for LIVBP and Cd^{2+} was found to generate the most ordered and highest diffraction crystals (Trakhanov & Quiocho, 1995). The addition of Cd^{2+} has also improved the crystal morphology of histidine-binding protein (Trakhanov et al., 1998). In all the cases mentioned above, Cd^{2+} concentration needs to reach a certain level for crystallization to occur (for example, 1 mM for LIVBP and 20 mM for cNTnC). The concentration needed for crystallization correlates to the abundance and distribution of negative charged residues on the protein surfaces (Trakhanov et al., 1998). Since 25.9% of cNTnC residues are acidic, this protein is a good candidate for utilizing Cd^{2+} to promote crystallization.

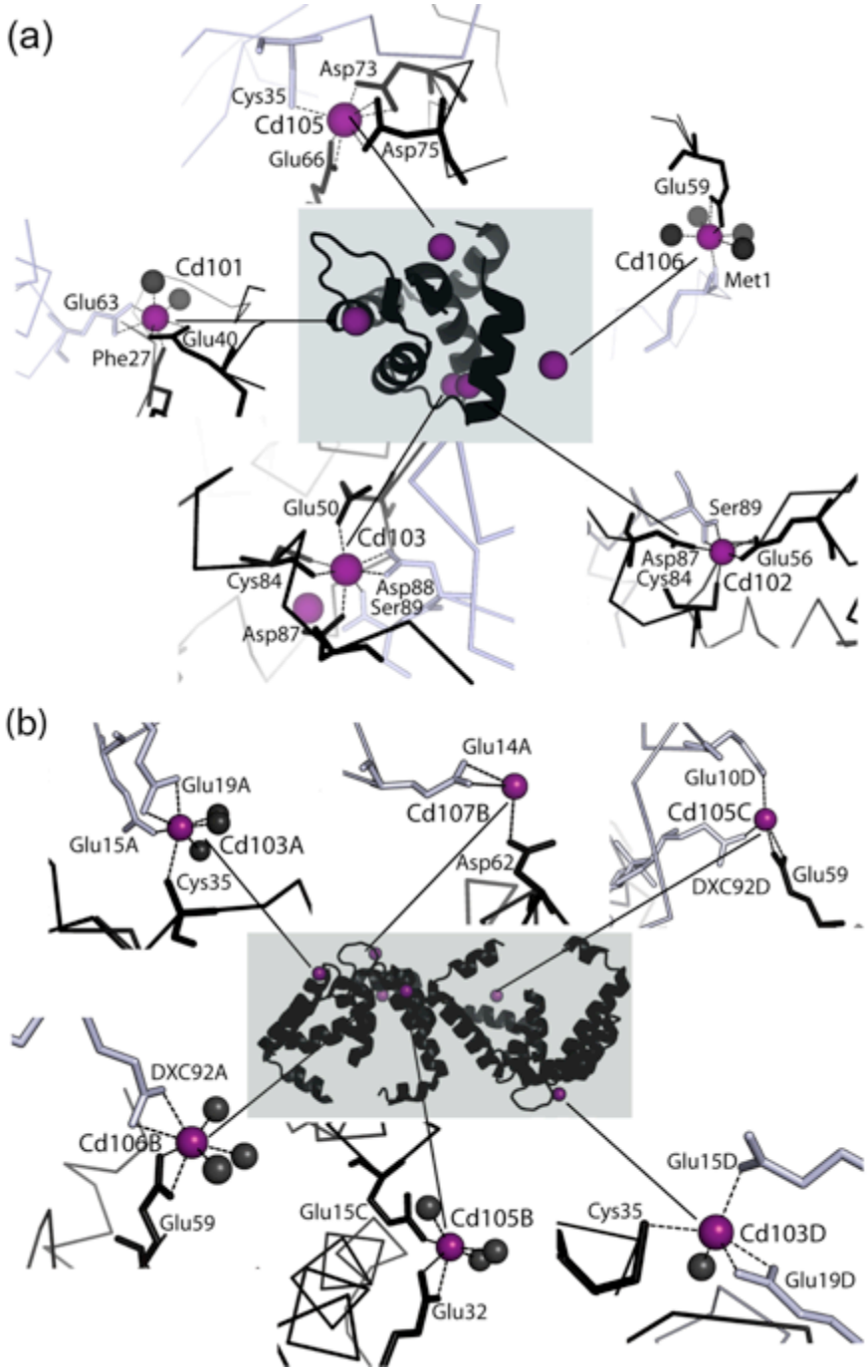

Figure 2.9 Intermolecular Cd^{2+} coordination in cNTnC and cNTnC complexed with DXC.

Intermolecular Cd^{2+} ions (purple spheres) and their spatial arrangement around both (a) cNTnC (PDB: 3SD6) and (b) cNTnC complexed with DXC (PDB: 3RV5; deoxycholic acid; protein rendered as a black cartoon) are shown in the shaded area. The coordination details of each intermolecular Cd^{2+} ion are also presented around the shaded area. The protein chain and its neighbouring symmetry-related molecule that coordinate the Cd^{2+} ion are shown as black and white ribbons, respectively. Coordinating residues are labelled and rendered in sticks in the corresponding chain colours. Grey spheres are water molecules that coordinate Cd^{2+}.

Figure 2.9 shows all the intermolecular Cd^{2+} ions for cNTnC and cNTnC bound to DXC. The Cd^{2+} ions are coordinated in various arrangements (for example, tetrahedral, trigonal bipyramidal, octahedral) intermolecularly in both structures. The majority of the coordinating residues are negatively-charged glutamates and aspartates. The only residue that is conserved between the two structures to coordinate intermolecular Cd^{2+} is Cys35. This again stresses the important relationship between Cys35 and Cd^{2+} in this structure, as discussed in the previous section. Table 2.2 summarizes the coordination of intermolecular Cd^{2+} in the two cNTnC structures shown in Figure 2.9.

Table 2.2. Coordination of intermolecular Cd^{2+} in two cNTnC structures.

	Open conformation (PDB:3RV5) **4 molecules/ASU** **21 Cd's total in ASU: 6 coordinated intermolecularly,** **8 in EF-hands**		**Closed conformation (PDB:3SD6)** **1 molecule/ASU** **7 Cd's total in ASU: 5 coordinated intermolecularly,** **2 in EF-hands**	
Intermolecular Cd	Cd103A	Cys35A (Sγ)	Cd101 Occupancy = 1.0	Phe27 (O)
		HOH115A, 118A, 119A		Glu40 (Oε1, Oε2)
		Glu15A (Oε2) [Symm]		HOH125, 145
		Glu19A (Oε1, Oε2) [Symm]		Glu63 (Oε1, Oε2) [Symm]
	Cd105B	Glu32B (Oε1, Oε2)	Cd102 Occupancy = 1.0	Glu56 (Oε1, Oε2)
		HOH142B, 148B		Cys84 (Sγ)
		Glu15C (Oε2)		Asp87 (Oδ2)
		HOH138C		Ser89 (O, OXT) [Symm]
	Cd106B	Glu59B (Oε1, Oε2)	Cd103 Occupancy = 1.0	Gln50 (Oε1)
		HOH133B, 134B, 135B		Cys84 (Sγ, O)
		DXC92A (O3, O4) [Symm]		Asp87 (Oδ1)
	Cd107B	Asp62B (Oδ2)		Asp88 (Oδ1, Oδ2) [Symm]
		Glu14A (Oε1, Oε2) [Symm]		Ser89 (O) [Symm]
	Cd105C	Glu59C (Oε1, Oε2)	Cd105 Occupancy = 1.0	Glu66 (Oε1, Oε2)
		Glu10D (Oε2) [Symm]		Asp73 (Oδ1, Oδ2)
		DXC92D (O4) [Symm]		Asp75 (Oδ1, Oδ2)
	Cd103D	Cys35D (Sγ)		Cys35 (Sγ) [Symm]
		HOH128D	Cd106 Occupancy = 0.6	Glu59 (Oε1, Oε2)
		Glu15D (Oε2) [Symm]		HOH131, 132, 133, 146
		Glu19D (Oε1, Oε2) [Symm]		Met1 (O) [Symm]

2.4.3. Cd^{2+} coordination in EF-hand proteins

Cadmium bound to EF-hand motif containing proteins other than cNTnC have been structurally investigated previously (Li et al., 2011, Rao et al., 1996, Leinala et al., 2003, Swain et al., 1989). Figure 2.10 compares the ion coordinating residues in EF-loops from five Cd^{2+} bound EF-hand structures. These structures belong to parvalbumin (Swain et al., 1989), calpain (Leinala et al., 2003), cardiac troponin C in complex with DXC (Li et al., 2011), skeletal troponin C (Rao et al., 1996) and cardiac troponin C. Detailed spatial orientations of the coordinating ligands and atomic distances are listed in Table 2.3. All of these proteins are well-known Ca^{2+}-binding proteins and except for EF1 in our cardiac troponin C structure, all EF-loops coordinate Cd^{2+} in the canonical pentagonal bipyramidal geometry, as shown by the schematic in Figure 2.10.

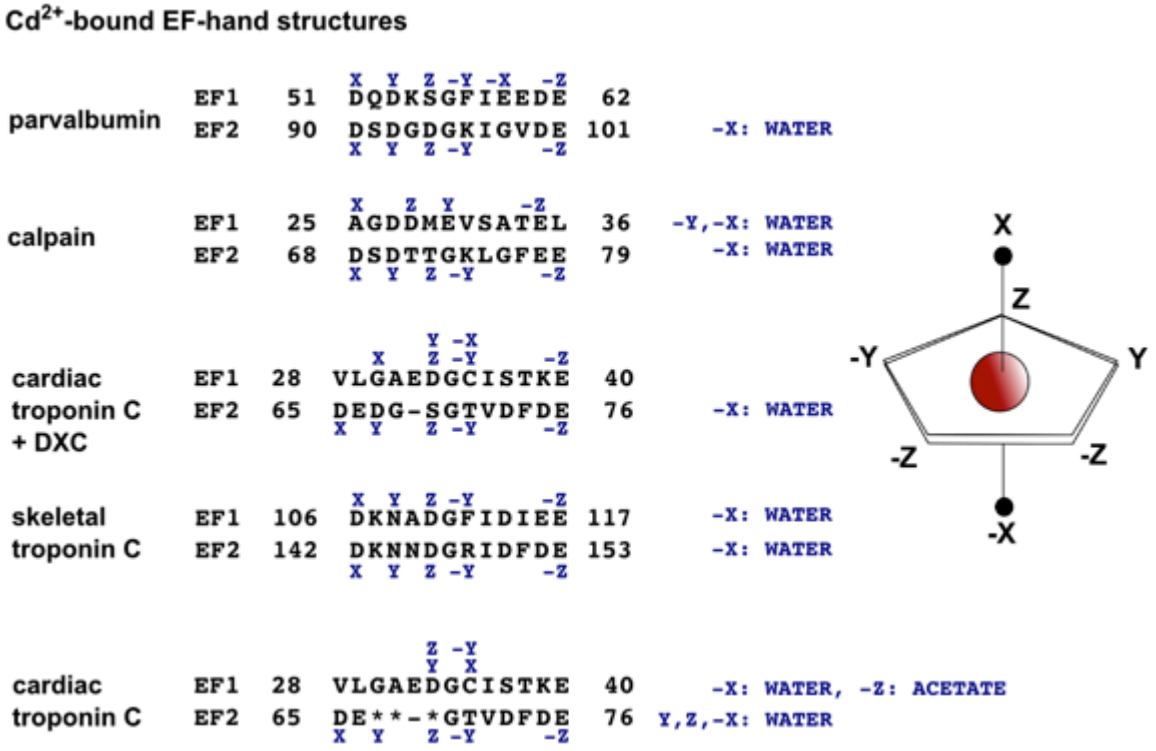

Figure 2.10 The residues used to coordinate Cd^{2+} in EF-hand protein structures. EF-hand loop sequences of parvalbumin (PDB entry 1CDP; Swain et al., 1989), calpain (PDB entry 1NP8; Leinala et al., 2003), cardiac troponin C complexed with DXC (PDB entry 3RV5; Li et al., 2011), skeletal troponin C (PDB entry 1NCX; Rao et al., 1996) and cardiac troponin C (PDB entry 3SD6; Zhang et al., 2013) are shown. Spatial orientation assignments of the coordinating ligands are labelled either above or below the residues. The canonical pentagonal bipyramidal geometry labels are schematically shown for reference.

Canonically, five EF-loop residues at positions 1, 3, 5, 7 and 12 contribute coordinating ligands. In EF1 of the structure presented here, only two residues, Asp33 and Cys35 (positions 5 and 7), contribute to the ion coordination. As mentioned previously, the EF1 loop in cNTnC normally does not bind ions (vestigial site); however, the structure presented here shows that EF1 manages to coordinate Cd^{2+} by changing the loop conformation. This also causes the side chain of the residue at position 12, glutamate, to point away from the coordination site so it is not available for coordinating the Cd^{2+} in canonical fashion. This feature can be observed by comparing the EF1 coordination (listed in Table 2.3) between cNTnC with bound DXC and the cNTnC structure presented here.

EF2 in cNTnC-DXC and EF2 in the cNTnC structure presented here both coordinate Cd^{2+} in canonical fashion. The two EF-loops in skeletal troponin C also coordinate Cd^{2+} canonically (Rao et al., 1996), leading to the conclusion that Cd^{2+} ions are able to substitute for Ca^{2+} in EF-hands in a canonical fashion. Indeed, functional data shows that Cd^{2+} can activate skeletal troponin C (in which site I is functional) by substituting for Ca^{2+} (Chao et al., 1990).

By comparing the EF-loop sequences of the five proteins in Figure 2.11, it can be seen that cardiac troponin C is the only one with cysteine as a coordination residue; keeping in mind that the cysteine is coordinating a Cd^{2+} ion in a vestigial ion binding site, so its significance is not clear in a physiological sense. To our knowledge, there is no other example of an EF-hand protein in which cysteine serves as an ion coordinating ligand. Interestingly, this is not the case in non-EF-hand proteins.

Table 2.3. Cd^{2+} bound to EF-hand motif-containing proteins: distance (Å) between ligand atom and Cd^{2+}.

	EF1		EF2	
Parvalbumin (PDB entry 1CDP; Swain *et al.*, 1989)				
X	Asp51/Oδ1	2.1	Asp90/Oδ2	2.3
Y	Asp53/Oδ1	2.4	Asp92/Oδ1	2.4
Z	Ser55/Oγ	2.6	Asp94/Oδ1	2.2
-Y	Phe57/O	2.3	Lys96/O	2.3
-X	Glu59/Oε1	2.3	HOH116/O	2.4
-Z	Glu62/Oε1	2.8	Glu101/Oε1	2.3
	Glu62/Oε2	2.4	Glu101/Oε2	2.7
Calpain (PDB entry 1NP8; Leinala *et al.*, 2003)				
X	Ala25/O	2.2	Asp68/Oδ1	2.3
Y	HOH726/O	2.5	Asp70/Oδ1	2.3
Z	Asp28/Oδ1	2.2	Thr72/Oγ	2.5
-Y	Glu30/O	2.4	Lys74/O	2.3
-X	HOH772/O	2.3	HOH778/O	2.4
-Z	Glu35/Oε1	2.4	Glu79/Oε1	2.1
	Glu35/Oε2	2.4	Glu79/Oε2	2.9
Cardiac troponin C with DXC (PDB entry 3RV5; Li *et al.*, 2011)				
X	Gly30/O	2.2	Asp65/Oδ2	2.1
Y	Asp33/Oδ1	2.4	Asp67/Oδ1	2.3
Z	Asp33/Oδ2	2.1	Ser69/Oγ	2.3
-Y	Cys35/O	2.5	Thr71/O	2.4
-X	Cys35/S	2.5	HOH110/O	2.3
-Z	Glu40/Oε1	2.4	Glu76/Oε1	2.3
	Glu40/Oε2	2.6	Glu76/Oε2	2.6
Skeletal troponin C (PDB entry 1NCX; Rao *et al.*, 1996)				
X	Asp106/Oδ1	2.0	Asp142/Oδ1	2.1
Y	Asn108/Oδ1	2.3	Asn144/Oδ1	2.4
Z	Asp110/Oδ1	2.3	Asp146/Oδ1	2.3
-Y	Phe112/O	2.3	Arg148/O	2.3

	EF1		EF2	
-X	HOH176/O	2.2	HOH209/O	2.3
-Z	Glu117/Oε1	2.5	Glu153/Oε1	2.4
	Glu117/Oε2	2.5	Glu153/Oε2	2.6
Cardiac troponin C (PDB entry 3SD6; Zhang *et al.*, 2013)				
X	Cys35/O	2.5	Asp65/Oδ2	2.2
Y	Asp33/Oδ2	2.5	HOH147/O	2.1
Z	Asp33/Oδ1	2.3	HOH148/O	2.0
-Y	Cys35/Sγ	2.6	Thr71/O	2.6
-X	HOH97/O	2.4	HOH151/O	2.6
-Z	ACT90/OXT	2.3	Glu76/Oε1	2.3
			Glu76/Oε2	2.6

ACT: actetate; HOH: water.

2.4.4. Cd^{2+} coordination in non-EF-hand proteins

Cadmium with a filled d10 orbital shell can take on a number of different coordination geometries. Examination of non-EF-hand proteins in complex with Cd^{2+} reveals that Cd^{2+} is most often coordinated in either tetrahedral or trigonal bipyramidal geometries, with histidine and cysteine residues most frequently involved in coordinating the Cd^{2+}. For example, in cadmium carbonic anhydrase (Xu et al., 2008) and horse liver alcohol dehydrogenase (Meijers et al., 2001), Cd^{2+} is coordinated by two cysteines, one histidine and two water molecules in a trigonal bipyramidal geometry. Histidines are normally involved in tetrahedral coordination (Benning et al., 2001, Ferraroni et al., 1999), and in one case, octahedral coordination (Huang et al., 2002). The ionic radius of Cd^{2+} is similar to that of Ca^{2+} and this similarity in size and charge make the two ions competitors for the same protein ion binding sites (Mastrángelo et al., 2011, Adiele et al., 2012). As shown in the examples above, Cd^{2+} appears to be more flexible in its coordination number and geometry than Ca^{2+}. Cd^{2+} is also able to replace Zn^{2+} ions, in many proteins (Xu et al., 2008, Meijers et al., 2001).

2.4.5. Cd^{2+} in the cNTnC vestigial site, a potential means of cardiac toxicity

Cadmium is a soft metal that can form a covalent bond with the thiol sulfur atom from cysteine (Dokmanić et al., 2008). Therefore, it is possible that Cd^{2+} cannot dissociate from the EF1 loop once it has been coordinated. Although EF1 is a vestigial site in cNTnC, it has been shown previously to have a functional effect on the protein (Gillis et al., 2007, Gillis et al., 2005). Cadmium, which is not normally present in mammals, is a toxic metal and can disrupt physiological function (Houston, 2007, Peters et al., 2010). For example, intake of Cd^{2+} in rats has been shown to cause irregular cardiac function (Kopp et al., 1983). The formation of a potential covalent cadmium-sulfur bond leading to the inability of Cd^{2+} to dissociate from EF1 of cNTnC could suggest a potential contributor to cadmium's toxicity via cardiac dysfunction.

Chapter 3.

Structural characterization of immunoglobulin domain C3 of cMyBP-C by nuclear magnetic resonance

Note regarding contributions:

Portions of this chapter were published previously in Biochemistry.

Zhang, X.L., Soumya, D., McIntosh, L.P. & Paetzel, M. (2014). Structural characterization of the C3 domain of cardiac myosin binding protein-C and its hypertrophic cardiomyopathy-related R502W mutant. *Biochemistry* **53(32)**, 5332-5342

The work summarized in this chapter was a collaborative effort with Dr. McIntosh's laboratory at the University of British Columbia. I made all constructs, expressed and purified proteins, assigned resonance, calculated and refined the structures, made all figures, and analyzed all data. Jacob Brockerman and Soumya De helped me with running the instruments for data collection and initial data processing. Soumya De trained me to use various NMR softwares for resonance assignment and structure calculation, and made intellectual contributions throughout the structure solving process.

3.1. Overview

In order to study how the R502W mutation may affect the function of the cMyBP-C, the gene of the C3 domain of the cMyBP-C was subcloned into a protein expression vector with an intact N-terminal His_{6x}-tag for Ni^{2+}-NTA affinity purification. R502W mutant construct was also created by site directed mutagenesis. The proteins had high expression level (10 mg/L of culture) and after thrombin cleavage of the N-terminal His_{6x}-tag, the expressed protein contained 95 residues. The abundance of protein and smaller protein size made the C3 domain of cMyBP-C an ideal candidate for structural investigations by NMR spectroscopy.

This chapter presents the structural and dynamic properties of the WT and R502W C3 domain of cMyBP-C. The two structures have similar backbone conformations and backbone amide ^{15}N relaxation profiles. They both rendered the immunoglobulin-like protein fold, a β-sandwich with two anti-parallel β-sheets, forming a Greek-key topology. The R502W mutation on the protein surface and it did not disrupt the protein fold, nor did it change its thermal stability significantly. Comparing the surface electrostatics of the two structures, the R502W mutation reduced the size of a positively charged area on the surface of the protein, which may be important for the protein's interactive functions. The cMyBP-C has shown to interact with a number of proteins, such as myosin and actin. Further investigation of the interactive properties of the protein may reveal the impact of the R502W on the protein's function and its role in the disease mechanism of HCM.

3.2. Materials and methods

3.2.1. Cloning and mutagenesis

A 273 base pair DNA fragment, encoding the C3 region (residues 453-543) of human cMyBP-C, was amplified from the full-length *MYBPC3* cDNA gene. The PCR reaction used the forward primer 5'-GCCATATGCCTGTGCTCATCACGCG-3' and the reverse primer 5'-GCGCGGCCGCTTACTTTTCCTGCACAATGAGCT-3' that contain the restriction sites *Nde*I and *Not*I, respectively. The PCR product was ligated into the pET28a vector (Novagen). The resulting construct, denoted as His_{X6}-C3, encodes the WT C3 domain of cMyBP-C with a cleavable N-terminal hexahistidine affinity tag. Subsequent DNA sequencing (GENEWIZ) confirmed that the C3 insert matched the sequence reported in the Swiss-Prot database (UniProt entry Q14896/MYBPC3_HUMAN). The R502W C3 construct was made from the WT C3 construct by site-directed mutagenesis using the primers 5'-GACCTTCAAATACTGGTTCAAGAAGGACGG-3' and 5'-CCGTCCTTCTTGAACCAGTATTTGAAGGTC-3'.

3.2.2. Isotopic labeling and protein purification

The expression plasmid was transformed into *Escherichia coli* BL21(λDE3) cells. Uniformly ^{15}N-labeled His_{X6}-C3 was expressed in M9T media (200 μg/mL kanamycin) supplemented with 1 g/L $^{15}NH_4Cl$ (Sigma-Aldrich). Uniformly $^{13}C/^{15}N$-labeled His_{X6}-C3 was expressed in M9T media containing 3 g/L [$^{13}C_6$]-glucose (Sigma-Aldrich) and 1 g/L $^{15}NH_4Cl$. For both isotopically labeled samples, cultures were grown at 37°C to an OD_{600} of 0.6 and induced with 1 mM IPTG. After an overnight incubation, cells were harvested by centrifugation at 6,000*g* (4°C) for 7 min, and subsequently lysed using an Avestin Emulsiflex-3C high-pressure homogenizer for 5 min in 50 mM sodium phosphate buffer, pH 6.5. The resulting lysate was further centrifuged at 29,000*g* for 30 min at 4°C. The supernatant containing the His_{X6}-tagged proteins was loaded onto a Ni^{2+} affinity column (5 mL column volume; GE Healthcare), pre-equilibrated with phosphate buffer. Elution was performed in five steps with an increasing imidazole concentration (100-500 mM, in 100 mM increments). Samples of the elution fractions were run on a 15% SDS-PAGE gel. Fractions containing C3 were combined and concentrated to a volume of 5 mL using

an Amicon centrifugal filter unit (5kDa cutoff, Millipore). The His_{X6}-tag was cleaved by adding 10 units (1 unit/uL) of thrombin and incubating for 6 hours. The protein was then further purified using a size-exclusion chromatography column (HiPrep 26/60 Sephacryl S-100 High-Resolution) on an ÄKTAprime system (Phamacia Biotech). The column was equilibrated with 50mM sodium phosphate buffer, pH 6.5, 100 mM NaCl and run at a flow rate of 1.0 mL/min. Fractions containing C3 were identified using 15% SDS-PAGE, pooled and concentrated to 10 mg/mL using an Amicon ultracentrifugal filter device (3 kDa cutoff, Millipore). The final protein (Pro453-Lys543 with four remnant N-terminal residues, Gly-Ser-His-Met, resulting from cloning and thrombin cleavage) is 95 residues in length and has a calculated molecular mass of 10,812 Da and a calculated isoelectric point of 5.1. The concentration was determined by UV_{280} absorption with a NanoDrop ND-100 spectrophotometer (Thermo Scientific) using predicted extinction coefficients of 8480 M^{-1} cm^{-1} and 13980 M^{-1} cm^{-1} for the WT and R502W species, respectively (Wilkins et al., 1999).

3.2.3. NMR data acquisition

NMR experiments were performed on TCI-cryoprobe equipped Bruker Avance III 500 and 600 MHz NMR spectrometers at 25 °C. All samples consisted of 0.8 mM protein in 50 mM sodium phosphate buffer, 100 mM NaCl, pH 6.5 with 10% D_2O. The spectra were processed with NMRPipe (Delaglio et al., 1995) and analyzed using Sparky (Goddard and Kneller, 1999).

3.2.4. Chemical shift assignments and structure calculation

Signals from the backbone ^{1}H, ^{13}C, and ^{15}N nuclei of the $^{13}C/^{15}N$-labeled WT and R502W C3 domains were assigned using the following experiments: $^{15}N/^{13}C$-HSQC, HNCO, HN(CA)CO, HNCACB, HN(CO)CACB (Sattler et al., 1999). Aliphatic side-chain assignments were obtained from 3D (H)CC(CO)NH-TOCSY, H(CC)(CO)NH-TOCSY and HCCH-TOCSY experiments. The aromatic side-chains were assigned using the 2D (Hβ)Cβ(CγCδ)Hδ and (Hβ)Cβ(CγCδCε)Hε experiments (Yamazaki et al., 1993). Backbone dihedral angles were predicted from $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, $^{1}H^{\alpha}$, $^{1}H^{N}$ and ^{15}N chemical shifts with TALOS+ (Shen et al., 2009) and used as restraints for structure calculations. NOE-derived distance restraints were obtained from simultaneous 3D ^{1}H-$^{15}N/^{13}C$-^{1}H

NOESY-HSQC (mixing time = 110 ms) and methyl-NOESY spectra (Zwahlen et al., 1998). The NOESY spectra were partially assigned manually, followed by automated assignment using CYANA 3.0 (Güntert, 2004). Structure calculations were performed in seven iterative cycles. Each cycle yielded 100 structures and the 10 lowest energy structures were taken into the following cycle as the basis for further spectral assignments and structure calculations. The 10 lowest energy structures generated in the final cycle were then refined with CNS using explicit solvent and molecular dynamics simulations (Brunger et al., 1998). Table 3.1 shows a summary of the restraints and structural statistics for both proteins. The chemical shifts and structural coordinates of the WT C3 domain ensemble have been deposited in the BioMagResData Bank (accession number: 25007) and RCSB Protein Data Bank (accession number: 2MQ0), respectively. The chemical shifts and structural coordinates of the R502W C3 domain ensemble have been deposited in the BioMagResData Bank (accession number: 25010) and RCSB Protein Data Bank (accession number: 2MQ3), respectively.

3.2.5. Backbone amide ^{15}N relaxation

Amide ^{15}N relaxation data (T_1, T_2, heteronuclear NOE) were collected at 25°C on the WT and R502W C3 proteins with a 500 MHz NMR spectrometer (Farrow et al., 1994). Relaxation rate constants were determined by fitting each set of peak heights to a single exponential decay. Errors in the rate constants were determined by Monte Carlo simulations. The heteronuclear {^{1}H}-^{15}N NOE values were determined from the ratio of peaks height versus a control reference spectrum and the NOE errors were estimated from the spectral noise. The global tumbling correlation times and anisotropic model-free order parameters (S^2) were calculated from the relaxation data and structural ensembles using TENSOR2 (Dosset et al., 2000).

Amide relaxation compensated ^{15}N-CPMG-HSQC spectra were collected at 600 MHz. Interleaved spectra at different CPMG field strengths (ν_{CPMG} = 50, 100, 200, 400, 600, 800 and 1000 Hz) were collected in random order with a constant time delay T_{relax} = 80 ms. Peak volumes, determined using the autoFit.tcl script of NMRPipe, were used to calculate the effective transverse relaxation rate $R_{2,eff} = (-1/T_{relax})\ln(V_{CPMG}/V_0)$, where V_{CPMG} and V_0 are the peak volumes with and without CPMG pulse train.

3.2.6. Structural and electrostatic analyses

The secondary structural analysis was performed using the MICS program (Shen and Bax, 2012), combined with manual inspection of the backbone hydrogen bonding patterns. The root mean square deviation (RMSD) values reported for protein superimpositions were calculated by pair-wise ensemble comparison in PyMol [37]. The DALI (Holm and Rosenström, 2010) and FATCAT (Ye and Godzik, 2003) servers were used to find proteins with similar folds. The surface electrostatics analyses were performed with the adaptive Boltzmann-Poisson solver plug-in (Baker et al., 2001) within PyMol (DeLano, 2004) using dielectric constants of 2 for protein and 80 for solvent. The solvent radius was set at 1.4 Å. The per-atom charge and radius were calculated using the AMBER forcefield (Wang et al., 2000) within the PDB2PQR program (Dolinsky et al., 2004). Default parameters were used in the PDB2PQR program, except for the protonation states, which were assigned at pH 6.5 to match the NMR experimental conditions.

3.2.7. Circular dichorism data collection and analysis

Circular dichroism (CD) spectra were collected for both the WT and R502W C3 samples as a function of temperature with a Jasco J-810 spectropolarimeter. Temperature was controlled using a Peltier thermoelectric system. Purified proteins (180 µL at 15 µM concentration) in buffer (20 mM sodium phosphate, pH 6.5, 50mM NaCl) were loaded into a 0.05 cm path length quartz cuvette. Spectra, covering a wavelength range of 198 nm to 260 nm, were collected for the protein samples and buffer controls at 5 °C intervals between 20 °C and 70 °C. The samples were incubated at each temperature for 5 minutes before data collection.

The collection of each set of data was repeated three times for both the WT and the R502W mutant. The buffer baseline spectrum was subtracted from each protein spectrum at each corresponding temperature to correct for background signal. The set of averaged spectra were deconvoluted into basis component spectra using the convex constraint algorithm in the program CCA+ (Perczel et al., 1991). The sum of the fractions of the component spectra was normalized to one at each temperature. The change in the fraction of each basis spectrum, contributing to the total CD as a function of temperature, was used to monitor the transition between different folding stages. The

protein unfolding, depicted by transitions between the component spectra, was plotted against temperature. Upon fitting the data points with a Boltzmann sigmoidal curve (Liu et al., 2010; Niesen et al., 2007; Orwig and Lieberman, 2011) using GraphPad Prism version 5.00 (GraphPad Software), the mid-point unfolding temperatures were estimated.

3.2.8. Figure preparation

Figures were prepared using PyMol (DeLano, 2004). The alignment figure was prepared using the programs CLUSTALW (Larkin et al., 2007) and ESPript (Gouet et al., 1999).

3.3. Results

3.3.1. Expression, purification and thrombin digestion

Both the WT and the R502W C3 domain were successfully expressed with N-terminal His-tag and purified using Ni^{2+}-NTA affinity chromatography (Figure 3.1a). Upon digestion with thrombin, the His_{6x}-tag was removed, leaving remnant residues: Gly-Ser-His-Met at the protein's N-terminus. The removal of His_{6x}-tag was confirmed by band shift on SDS-PAGE gel (Figure 3.1a, b). Size exclusion chromatography of the thrombin digested C3 yielded samples of high purity, which was further concentrated to 10 mg/mL for subsequent experiments.

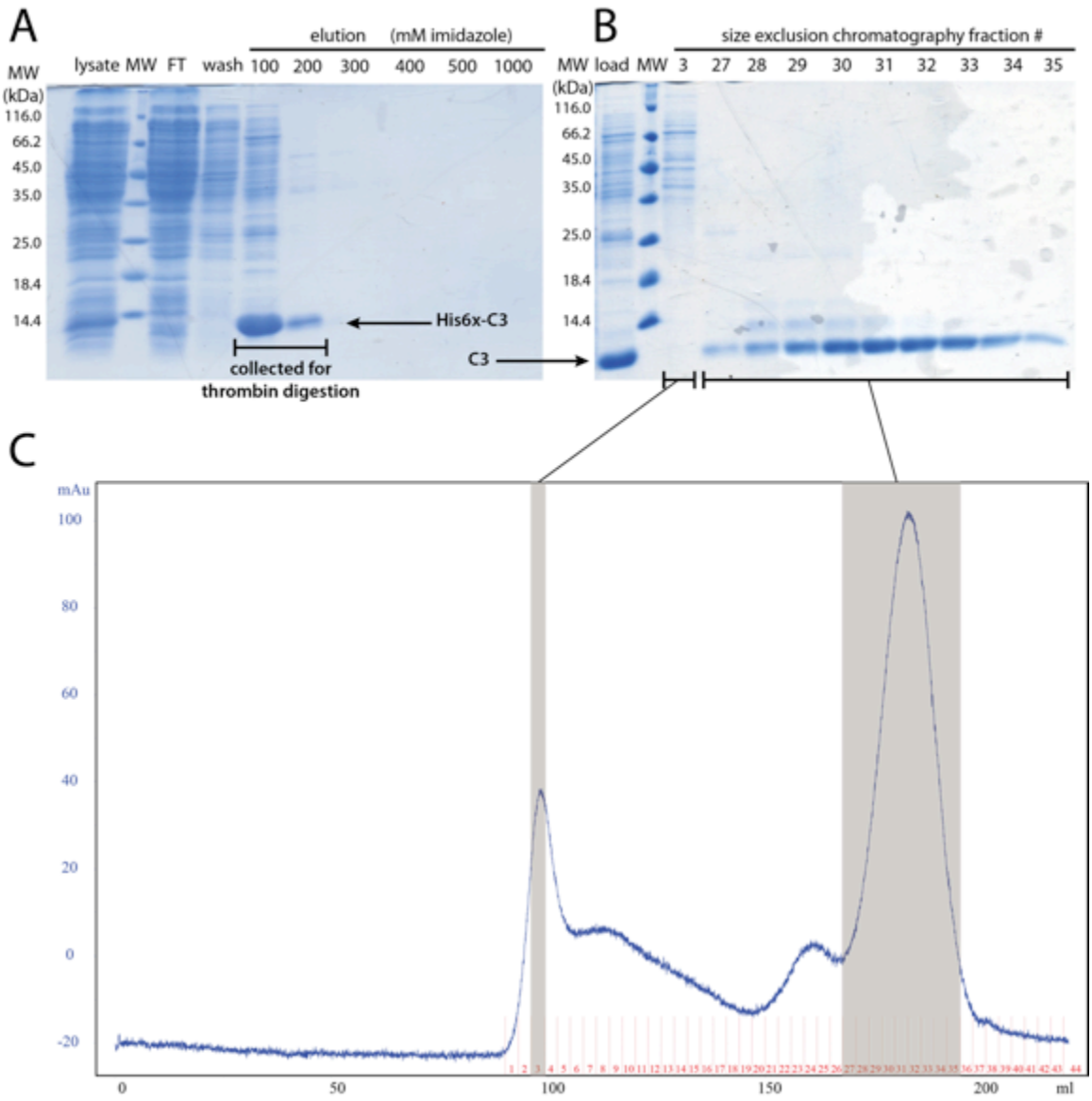

Figure 3.1 Ni^{2+}-NTA affinity chromatography and size exclusion chromatography of WT cMyBP-C C3.

A. Elution fractions of cMyBP-C C3 from Ni^{2+}-NTA affinity chromatography are ran on 15% SDS-PAGE gel. The imidazole concentrations used to elute the sample are shown on top of the elution lanes. Molecular marker is shown on far left. B. Elution fractions of cMyBP-C C3 from size exclusion chromatography Sephacryl S-100 are ran on 15% SDS-PAGE gel. The fraction numbers corresponding to that in C are shown on top of the lanes. C. Size exclusion chromatogram of cMyBP-C C3. The fractions are labeled with numbers in red and those covered by grey box were ran on 15% SDS-PAGE gel (shown in B). The R502W C3 has the same purification profile.

3.3.2. NMR-derived structure of WT and R502W C3 domain

Both the WT and R502W C3 domains yielded well-dispersed ^{15}N-HSQC spectra (Figure 3.2), confirming that they adopt independently folded structures when excised from the full-length protein. Using chemical shift-based dihedral angles restraints and an extensive set of NOE-derived distance restraints, the structural ensembles of the proteins were calculated with CYANA 3.0 (Figure 3.3a and Table 3.1). The resulting ensembles are well defined with backbone atom RMSD's of 0.61 Å and 0.80 Å between residues 454-538 of 10 lowest energy structures in WT and R502W C3, respectively.

As expected from sequence analysis, the C3 domain adopts an Ig-like protein fold. This is a β-sandwich structure composed of two β-sheets with a Greek-key topology (Figure 1.8). The two sheets are formed by strands denoted as ABED and C'CFGG'A' to maintain consistency with Ig-fold nomenclature (Figure 3.3b and c). The β-sheets are anti-parallel, except for the parallel arrangement of strand G' with the short strand A'. Residues 519-522, which are part of the loop between strands E and F, fold into a short helix. This secondary structure description is consistent with that predicted by MICS using chemical shifts.

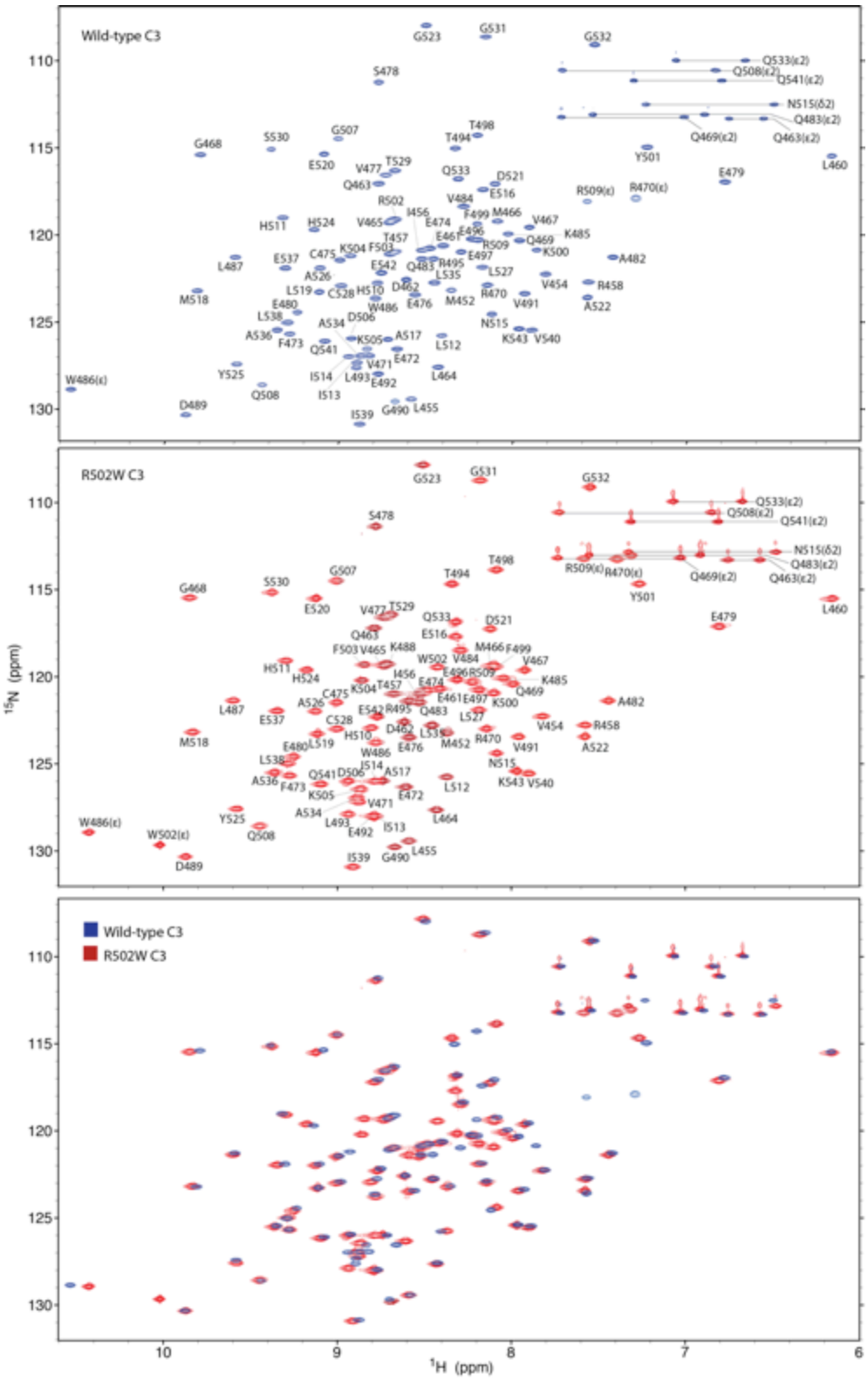

Figure 3.2 HSQC spectra of WT and R502W C3 domains.
Assigned ^{15}N-HSQC spectra of WT (top; blue) and R502W (middle; red) C3 domains. The overlaid spectra are also shown (bottom).

Table 3.1 Summary of the NMR data and structural statistics.

	WT	R502W mutant
Summary of restraints		
NOE distance restraints		
Intra-residue	471	356
Sequential	626	384
Medium range (1< |i-j| < 5)	294	152
Long range (|i-j| ≥ 5)	1103	520
Total	2494	1412
Dihedral angle restraints		
ϕ,ψ	86, 88	83, 85
Ramachandran analyses		
Residues in most favoured regions (%)	92.3	92.6
Residues in allowed regions (%)	7.7	7.4
Residues in generously allowed and disallowed regions (%)	0.0	0.0
Violations (average ± standard deviation)		
Distance restraints (Å)	0.12 ± 0.28	0.06 ± 0.21
Dihedral angle restraints (°)	0.47 ± 1.57	0.31 ± 0.82
Deviations from idealized geometry		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.6	0.6
RMSD[1] (Å)		
Backbone atoms (C, N, O)	0.61±0.09	0.80±0.12
All atoms (C, N, O)	0.84±0.12	1.19±0.12

1. Pairwise RMSD was calculated among 10 refined structures for residues 454-538 of human cMyBP-C C3 domains.

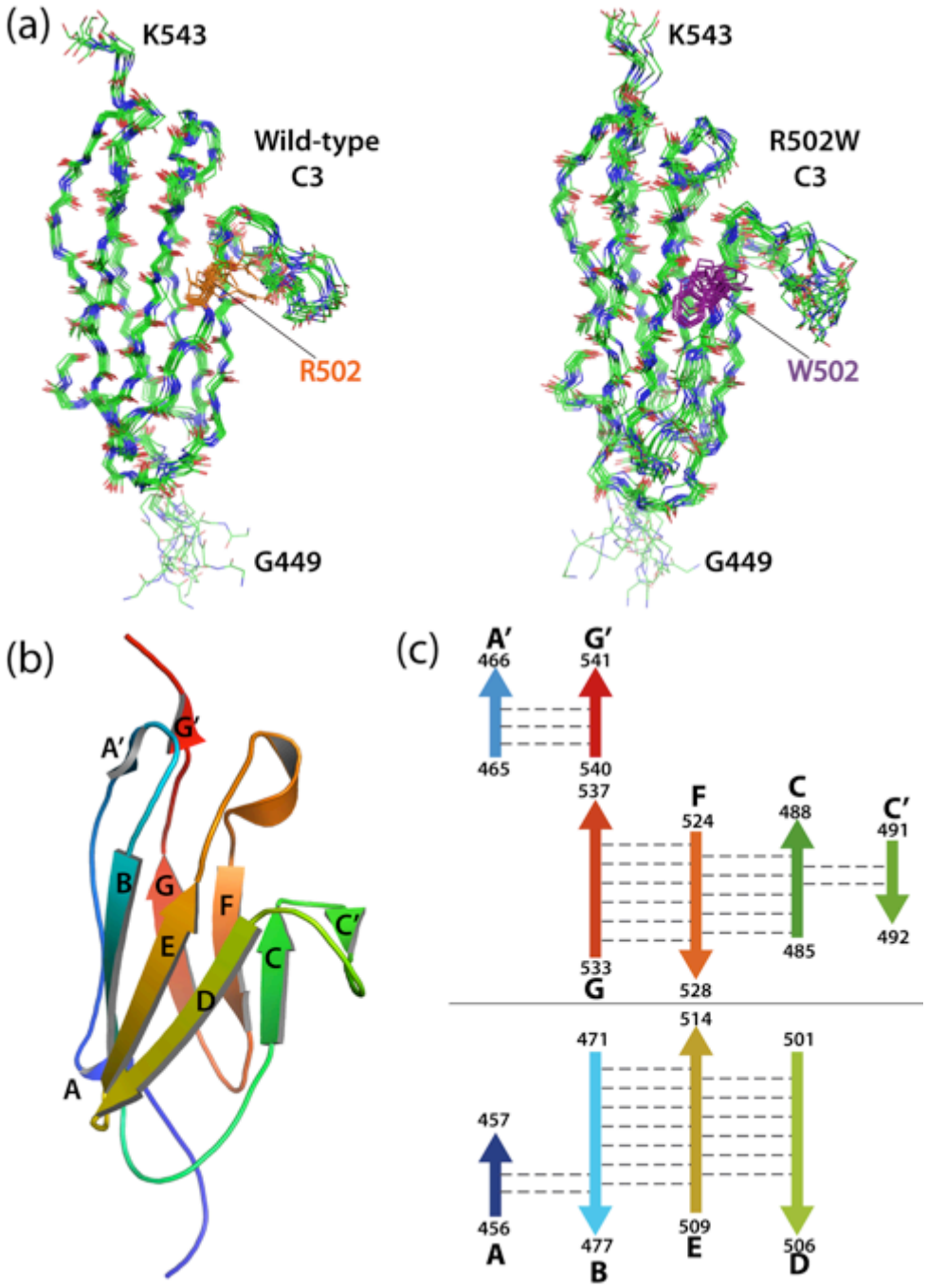

Figure 3.3 The NMR-derived structural ensembles of the WT and R502W C3 domains.

(a) The tertiary structural ensembles of the WT and R502W C3 domains are shown as main chain line diagrams (carbon: green, nitrogen: blue, oxygen: red). The side-chains of R502 (orange) and W502 (purple) are also highlighted. (b) The ribbon diagram presents the overall fold of the WT C3 domains. The β-strands are shown as arrows in rainbow colouring from N-terminus (blue) to C-terminus (red). Two β- sheets are formed by strands ABED and C'CFGG'A'. (c) The hydrogen bonds between the β-strands are shown as gray dash lines in a topology diagram, with the same colouring as in (b). The two β-sheets are separated for clarity. The start and end residues of each β-strand are labelled.

3.3.3. The R502W mutation does not perturb the C3 domain structure

Globally, the structures of the WT and R502W C3 domains are highly similar, with a pair-wise backbone RMSD of 1.37 ± 0.12 Å between their structural ensembles for residues 454-538 (Figure 3.4). Residue 502 is in the second position on strand D, and its backbone carbonyl and amide form hydrogen bonds with Ile513 (second to the last residue on strand E). Locally, substitution of Arg502 with tryptophan also did not perturb the structure of this sheet. In the WT protein ensemble, the side-chain of Arg502 is not well defined, adopting a range of conformations due to the lack of any measurable distance restraints (Figure 3.4). Similarly, In the R502W mutant ensemble, the hydrophobic side-chain of Trp502 also adopts a range of solvent exposed orientations, again due to the lack of measurable restraints. Consistent with these minimal perturbations in the NMR-derived structural ensembles, the amide $^{1}H_{N}$ and ^{15}N chemical shift differences between the WT and R502W C3 domains are localized around position 502 (Figure 3.5). Together, these data indicate that the R502W mutation does not alter the function of cMyBP-C through a large-scale conformational change within the C3 domain itself.

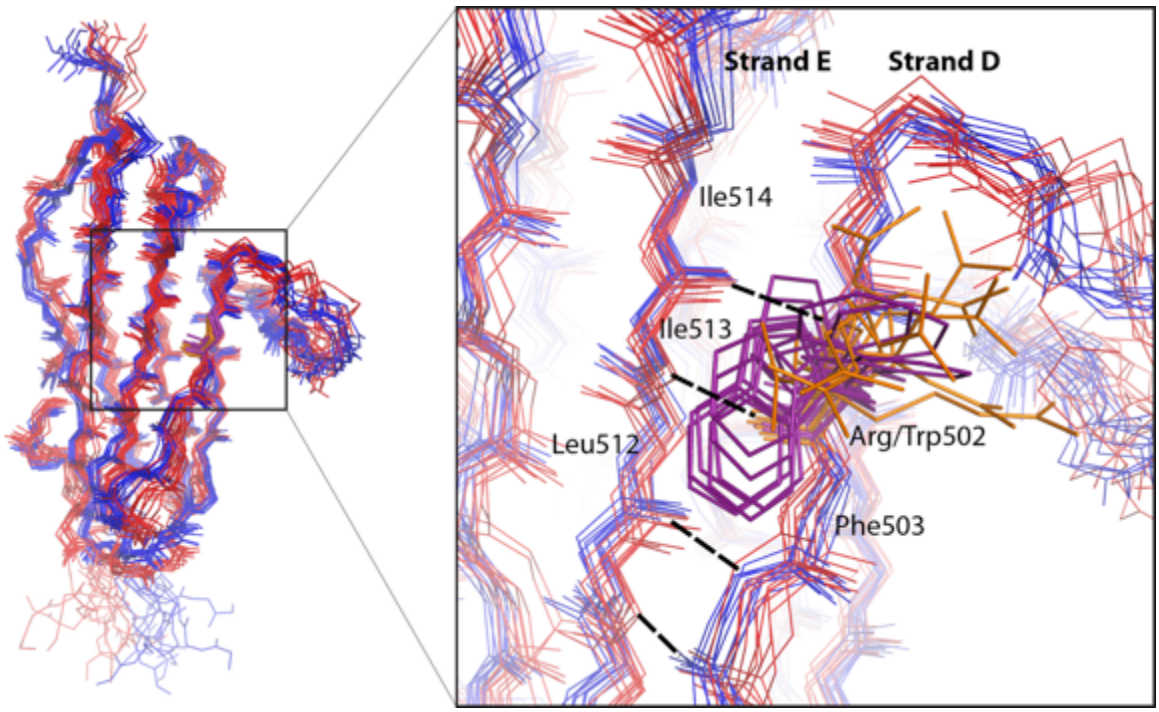

Figure 3.4 Structural superimposition of the WT and R502W C3 domains.
The structural ensemble of the WT C3 domain (blue lines) is overlaid on that of the R502W C3 domain (red lines). The region in the proximity of the R502W mutation is also shown in an expanded view, with the side-chains of R502 (orange) and W502 (purple) shown as sticks. Hydrogen bonds between strands E and D are depicted as black dashed lines.

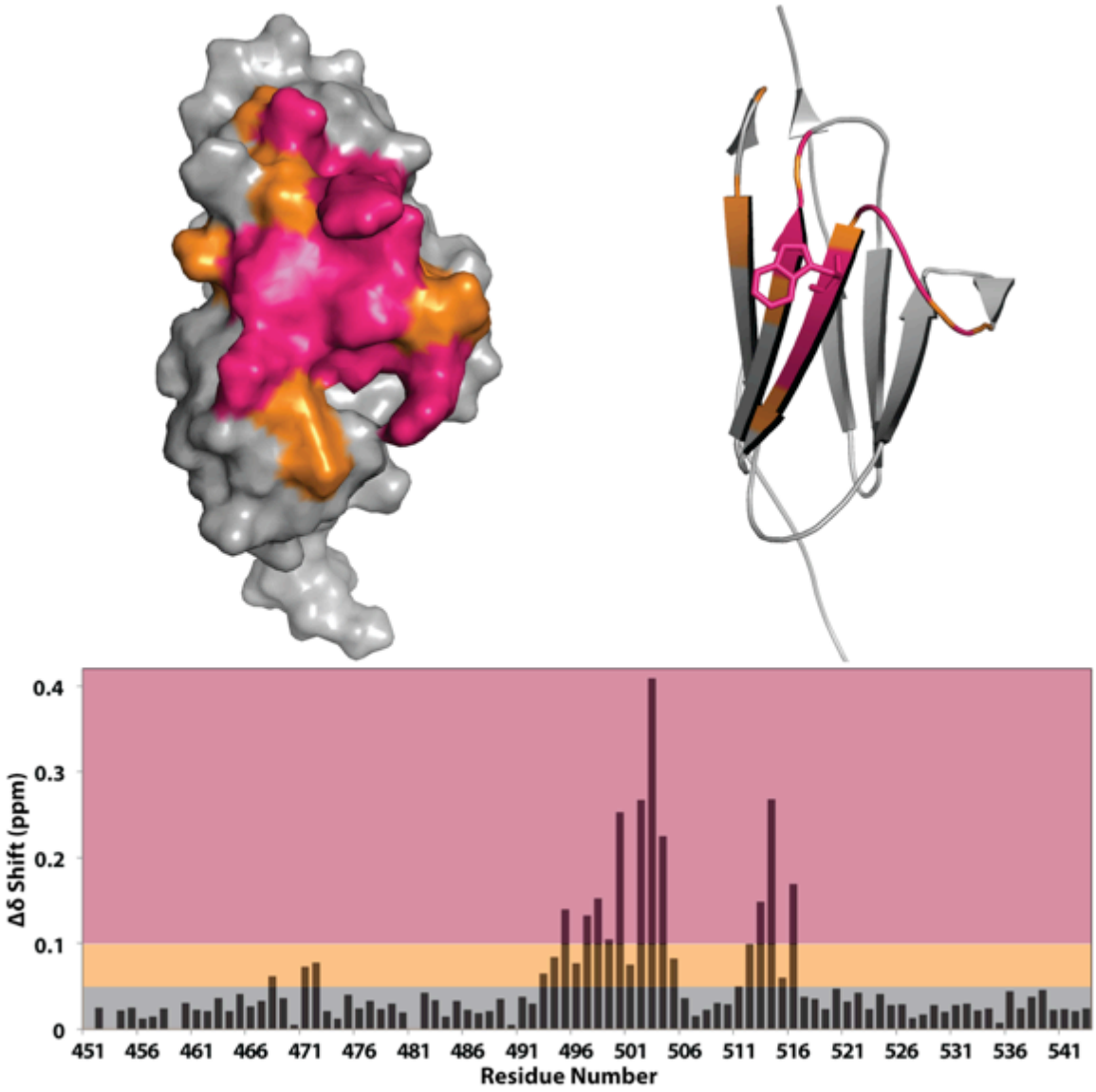

Figure 3.5 Amide chemical shift perturbations due to the R502W mutation. Combined amide chemical shifts perturbations ($\Delta\delta = \{(\Delta\delta_H^2 + (0.2\Delta\delta_N)^2\}^{1/2}$) of the corresponding backbone amides between the WT and R502W C3 domains are calculated and mapped onto a surface and ribbon diagram of the R502W C3 structure. Residues with Δδ greater than 0.1 ppm are shown in pink and those between 0.05 and 0.1 ppm are shown in orange. The side-chain of R502W is shown as sticks.

3.3.4. The R502W mutation does not perturb the C3 domain dynamics

To characterize the nsec-psec dynamic properties of the WT and R502W C3 domain, amide ^{15}N T_1, T_2, and heteronuclear NOE relaxation data were collected for both proteins at 25°C using a 500 MHz NMR spectrometer. From the T_1/T_2 ratios, global

isotropic tumbling correlation times of 6.0 ± 0.03 ns and 5.7 ± 0.02 ns were determined for the WT and R502W mutant C3 domains, respectively. These correlation times are consistent with those expected for an 11 kDa protein (Daragan and Mayo, 1997), and thus confirm that both isolated C3 domains are monomeric.

The ^{15}N-NOE values are very sensitive to the sub-nanosecond timescale motions of the ^{15}N-^{1}H^{N} bond vector and report on the local dynamics of the associated amino acid. In the mutant protein the ^{15}N-NOE value of the Trp502 ^{15}N$^{\varepsilon 1}$ is 0.47 ± 0.01, whereas that of the buried Trp486 is 0.70 ± 0.01. The substantially lower ^{15}N-NOE value indicates that the side-chain of Trp502 in the R502W C3 domain is relatively mobile. This is expected from its solvent exposed position and its multiple conformations in the structural ensemble due to a lack of experimental distance restraints.

The full sets of ^{15}N relaxation data for the main-chain amides of both proteins were fit according to the anisotropic model-free formalism (Lipari and Szabo, 1982). A comparison of these order parameters, as well as heteronuclear NOE values, indicates that both WT and the R502W C3 domains exhibit very similar nsec-psec timescale motions (Figure 3.6). In each case, amides in β-strands are well ordered, whereas those at the termini and in loop regions are somewhat more flexible. The most notable of the latter case are residues in the C'-D loop, which also show higher RMSD values in the structural ensembles calculated for each protein.

As a sensitive probe of msec-μsec timescale conformational exchange, the two proteins were investigated using ^{15}N-amide relaxation dispersion measurements (Tollinger et al., 2001). These studies revealed no conformational exchange for either the WT or R502W C3 domains (data not shown). This is consistent with the model-free analysis where the ^{15}N T_2 data for most amides were fit without the need for additional T_2 exchange contribution. Collectively, these relaxation measurements demonstrate that substitution of Arg502 with tryptophan does not markedly change the dynamic properties of the C3 domains across a wide timescale range.

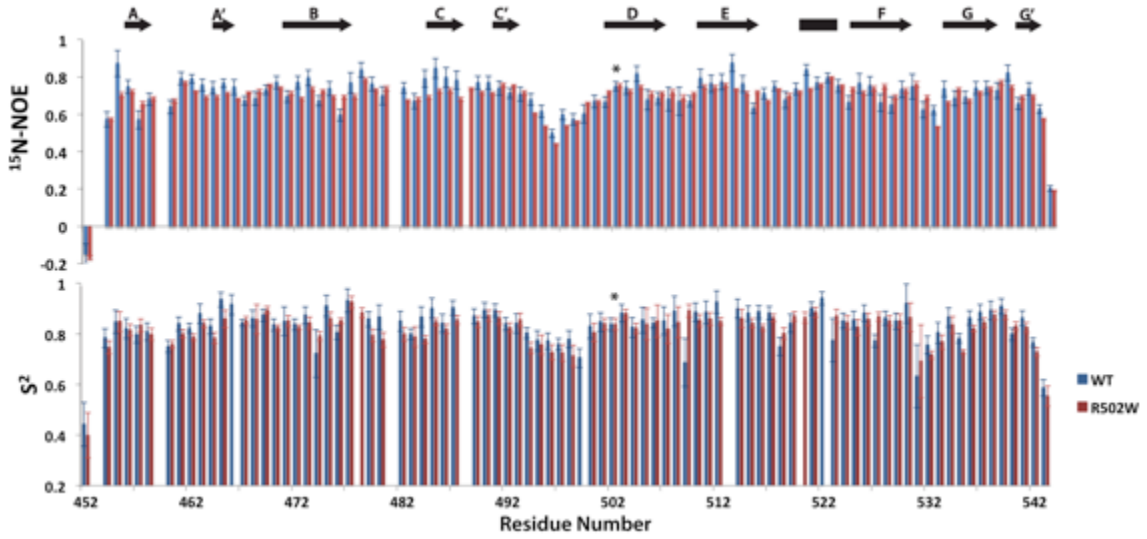

Figure 3.6 Backbone dynamics of cMyBP-C C3 domain from amide ^{15}N relaxation analysis.

Plots of ^{15}N-NOE (upper panel) and order parameter (S^2) values (lower panel) versus residue number are shown for both WT (blue) and R502W (red) mutant C3 domain. Lower ^{15}N-NOE and S^2 values, as seen for residues near the N- and C-termini and in the C'-D loop, are indicative of faster sub-nsec timescale backbone motions. Residue 502 is identified with an asterisk. The secondary structures identified in the 3D structures are mapped on top (arrows: β-strands; rectangles: helix).

3.3.5. The R502W mutation does not affect the C3 domain thermal stability

To investigate potential stability effects induced by the R502W mutation, the heat-induced denaturation of both the WT and mutant were monitored using circular dichroism (CD) spectroscopy. Sets of spectra were collected over a temperature range from 20°C to 70°C (Figure 3.7a). With increasing temperature, the minima of the spectra started at 217 nm (characteristic for β-sheets) and moved towards 200 nm (characteristic for random coils). This provides evidence that the C3 domain, mainly composed of β-sheets, was unfolding and losing its secondary structure as the temperature increased.

Using convex constraint algorithm, the collected sets of spectra were then deconvoluted into basis components. Two basis components (Figure 3.7b) were resulted, which suggests that the WT and R502W mutant both unfold in a predominantly global two-state transition. Upon obtaining the two basis components, the spectrum collected at each temperature can be interpreted as the sum of fractions of each basis components. As the temperature changed from 20°C to 70°C, the fraction of one basis component decreased while the other increased.

The percentage of unfolding, depicted by the percentage of transition from one basis component spectra to the other, as a function of temperature is shown in Figure 3.7c. Both the WT and the R502W mutant data fit well with Boltzmann sigmoidal curves. The unfolding profiles of the two proteins overlap closely (within the error range), with fit mid-point unfolding temperatures of 56.8 ± 2.0°C for the WT C3 and 55.4 ± 1.5°C for the mutant. This demonstrates that the R502W mutation did not have a significant impact on the overall thermal stability of the C3 domain.

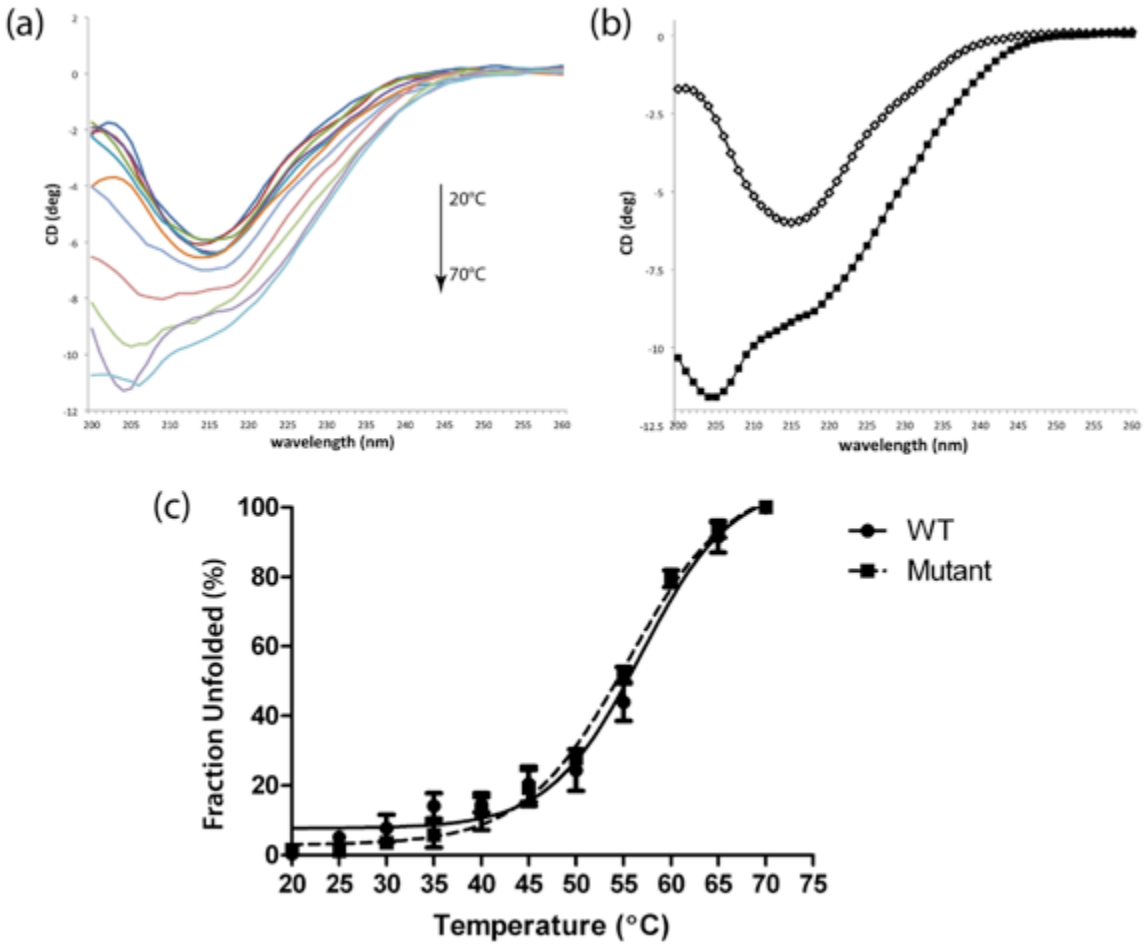

Figure 3.7 Temperature induced unfolding of C3 domain monitored by CD.
(a) The CD spectra of the WT C3 domain collected as a function of temperature. (b) The two deconvoluted basis component curves determined by the convex constraint algorithm for the WT C3 domain. Open squares correspond to the folded state and black squares correspond to unfolded state. The R502W C3 has similar spectra and basis components curves as the WT. (c) The data points of the fraction unfolded versus temperature follow Boltzmann-weighted sigmoidal curves, with R^2 values of 0.97 and 0.98, for the for the WT (round box) and R502W (square box) C3 domains, respectively. Each point on the plot (displayed with error bars) was an average of three experimental replicates.

3.3.6. The R502W mutation alters the predicted C3 domain surface charge

A comparison of the calculated Boltzmann-Poisson electrostatic features of the WT and R502W mutant C3 domains reveals a marked change of the surface charge induced by the R502W mutation (Figure 3.8). In the WT C3, the side chains of Arg470, Lys485, Arg495, Arg502, Lys504, Lys505 contribute to a large positively-charged surface. This stands in contrast to the net negative-charge on the remainder of the protein surface. Upon mutating Arg502 to tryptophan, the size of the positively-charged patch is reduced significantly. Such an electrostatic change could alter the interaction of the cMyBP-C C3 domain with a complementary charged partner protein.

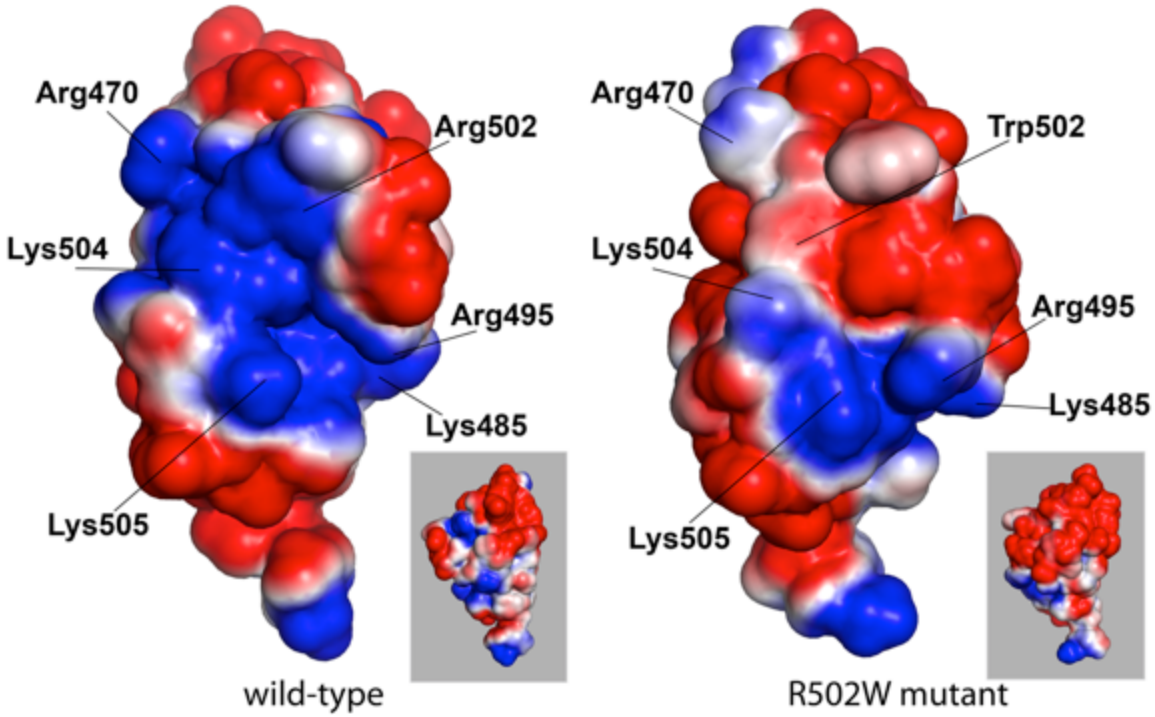

Figure 3.8 Electrostatic properties of the cMyBP-C C3 domain.
Surface representations of the WT and R502W C3, colored by electrostatic potential (negative, red; positive, blue) as calculated using default settings within the APBS (Baker et al., 2001) plug-in in PyMol (DeLano, 2002). The proteins are shown in the same orientation, and rotated 180^{o} along the long axis in the small insets. The residues that contribute to the positively-charged patch are labelled.

3.4. Discussion

3.4.1. The C3 domain belongs to immunoglobulin-like family

Using NMR spectroscopy, we have determined that the isolated C3 domain of cMyBP-C adopts an Ig-like fold. This was not unexpected as, along with seven other domains within cMyBP-C, the C3 domain was predicted by sequence analysis to be Ig-like (Oakley et al., 2004). Indeed, this fold, which is composed of seven to nine antiparallel β-strands arranged in a Greek-key β-sandwich (Lesk and Chothia, 1982), has been reported in numerous proteins, including many non-immunoglobulins (Bork et al., 1994). Strands B, C, E and F are defined as the common core of Ig-like proteins (Bork et al., 1994) (Figure 1.8). When compared with other Ig-like proteins, the lengths of the non-core strands and surrounding loops are highly variable, but these core strands are well-conserved (Kabsch and Sander, 1983). Figure 3.11 shows structural alignment of the C3 domain with six other proteins of high structural similarity scores. In this figure, the similarities of the core strands of the C3 domain with other Ig-like proteins can be clearly seen.

Like other Ig-like proteins, the C3 domain has three buried hydrophobic residues in its protein core, one from each of the core strands B, C, and F (Halaby et al., 1999). In the cMyBP-C C3 domain, these are Phe473 on strand B, Trp486 on strand C and Tyr525 on strand F. Also, in classical Ig domains, a disulfide bridge exists across the β-sandwich, formed between strands B and F (Lesk and Chothia, 1982). However, this feature is not conserved in Ig-like proteins (Halaby et al., 1999). In the case of the cMyBP-C C3 domain, two cysteines are also juxtaposed in strands B (Cys475) and F (Cys528). However, rather than pointing at each other, their side chains point to the same direction, preventing them from forming a disulfide bridge. The reduced states of Cys475 and Cys528 were confirmed from their highly diagnositc $^{13}C^{\beta}$ chemical shifts (Sharma and Rajarathnam, 2000).

3.4.2. R502W mutation may disrupt protein-protein interaction of C3 domain

The WT and R502W C3 domains are highly similar in terms of their structure (Figure 3.3), dynamics (Figure 3.6), and thermal stability (Figure 3.7). This strongly

suggests that the functional consequences of the mutation do not arise by disrupting the structural integrity of the domain, or through conformational rearrangements leading to allosteric effects. Rather, the mutation may directly alter the interactions of the C3 domain with cMyBP-C binding partners.

Cross-species alignment of the cMyBP-C C3 domain shows an overall high level of residue conservation, with Arg502 being invariant (Figure 3.9b). Furthermore, this arginine residue has been considered as a mutation “hot spot” since both R502Q and R502W are related to HCM (Richards *et al.*, 2003). Our structural investigations suggest that Arg502 may play an important role in the proper functioning of the cMyBP-C C3 domain by contributing to an intermolecular interface.

Comparison of the electrostatic features of the WT and the R502W mutant C3 domain shows that a large positively-charged area on the surface of the protein is altered by the R502W mutation (Figure 3.8). Residues Arg470, Lys485, Arg495, Arg502, Lys504, and Lys505 form this positively-charged patch on an otherwise negatively charged protein. The R502W mutation reduces this positively-charged patch, and thus could disrupt electrostatically-guided protein interactions between C3 domain and its protein binding partners. In addition, the mutation could eliminate direct interactions of the domain with such a partner via arginine-mediated salt bridges and hydrogen bonds.

The cMyBP-C has been shown to interact with various proteins within the sarcomere. In its C-terminal region, domains C7 to C10 bind to titin and light meromyosin (LMM) (Gilbert et al., 1996). At its N-terminal region (C0-C2 domains), the mybpc motif reaches out to interact with S2 of myosin in a phosphorylation-dependant manner (Gruen et al., 1999), while the C0 and C1 domains and the (Pro-Ala)-rich motif may bind to myosin and/or actin (Kulikovskaya et al., 2003; Mun et al., 2011; Squire et al., 2003; Witt et al., 2001). The domain of interest in this study, the C3 domain, lies in the central region (C3 to C6), which may be required for the flexibility of the N-terminal region, allowing it to interact with myosin S2 or actin. Because of its proximity to the N-terminal domains of cMyBP-C, the C3 domain may also be in contact with either of these potential partners, interacting with myosin or actin, at different binding interfaces than that of the N-terminal region. The C3 domain may also interact with myosin or actin during a different stage of the contraction cycle than that of the N-terminal region. The

neighbouring C2 domain has high structural similarities to C3 and it has been shown to bind to myosin S2 in NMR monitored titration experiments and molecular modeling studies (Ababou et al., 2007). Similar to what we have proposed for the R502W mutation in the C3 domain, the modelling suggests that the key interactions are between polar amino acids, and that HCM-related mutations (i.e E301Q in C2 and R870H in S2) would disrupt these electrostatic interactions.

Actin is a negatively-charged protein at physiological pH, and positively-charged residues from a number of actin-binding partners have been implicated in actin-binding interfaces (Amann et al., 1998; de Arruda et al., 1992; Hüttelmaier et al., 1999; Janssen et al., 2006; Lee et al., 2004; Yarmola et al., 2001). For examples, one of the Ig-like domains in the C-terminal region of paladin has been shown to bind to actin in co-sedimentation assays (Dixon et al., 2008). We thus hypothesize that the positively-charge patch on the C3 domain, which Arg502 and five other residues contribute to, may enhance an electrostatic interaction between cMyBP-C and actin. If so, the R502W mutation would disrupt this interaction, which could lead to altered regulation of contractility in cardiac muscle and ultimately HCM. Our preliminary studies using isothermal titration calorimetry and surface plasmon resonance did not detect any binding between the isolated C3 domain and actin. However, it is possible that additional components of cMyBP-C are required for the postulated association with actin.

3.4.3. Other HCM-related mutations in C3 domain

Eleven known mutations associated with HCM are in the C3 domain of cMyBP-C: R458H, G490R, G490V, R495G, R495Q, R502Q, R502W, Lys504 deletion, G507R, G523W and E542Q. The positions of these mutations are mapped on the structure of the C3 domain and indicated within the cross-species sequence alignment (Figure 3.9). In each case, the side-chains of all these residues are on the protein surface exposed to solvent. As elaborated below, these eleven HCM-related mutations likely perturb electrostatically-guided protein-protein interactions or, in the case of glycine substitutions, the proper folding of the C3 domain.

Several of these HCM-related mutations change conserved residues (Arg495, Arg502 and Lys504) that are positioned within the positively-charged patch (Figure 3.8). This further highlights the importance of this surface in mediating protein-protein interactions and suggests a common mechanism for disruption of function by these mutations. R458H and E542Q, which are not in the positively-charged patch, may also contribute to the charge-mediated interaction of C3. Similar examples can be seen with mutations such as E258K and D228N in C1 domain (Govada et al., 2008) and E301Q in C2 domain (Ababou et al., 2007), all of which are HCM-related mutations and result in a change in protein surface electrostatics.

The other observed HCM-linked mutations occur at glycine residues that are located within loop regions of the protein (e.g., G490 in the CC' loop, G507 in the DE loop, and G523 in the EF loop, Figure 3.10). Gly490 is the third residue within a Type II β-turn, in which the third residue is generally a glycine. Gly507 is part of a Type I β-turn, sitting at the second residue position. Flanked by a short helix (residues 519-522) and the strand F (residues 524-529), Gly523 is within a region with high curvature (defined as 'bend'), which means that the angle between the backbone of the residues before and after the bend has a directional change of greater than 70 degrees (Kabsch and Sander, 1983). These glycines are presumably required for proper folding, and thus the HCM-linked mutations at these sites may disrupt the structure of the C3 domain.

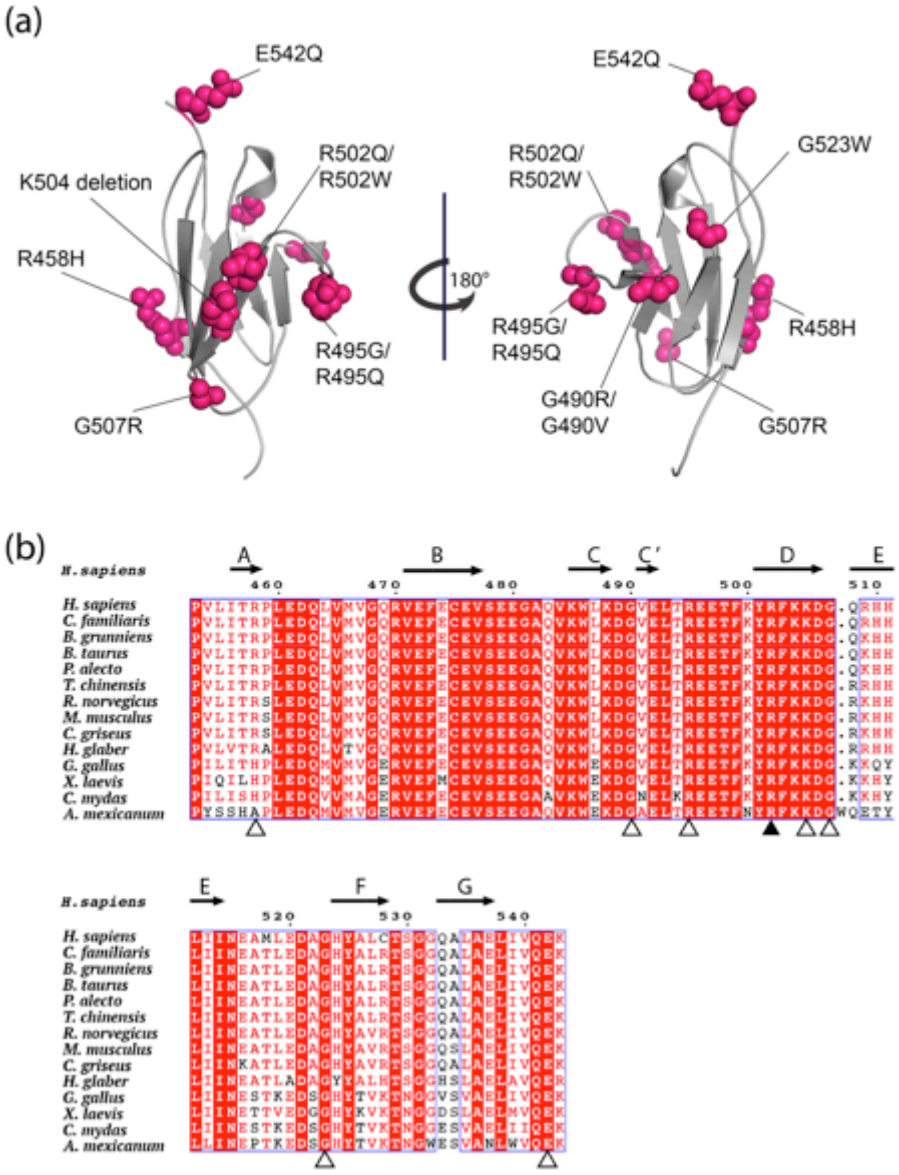

Figure 3.9 HCM-related mutations mapped on to the cMyBP-C C3 domain and cross-species alignment.

(a) The positions of HCM-related mutations (R458H, G490R, G490V, R495G, R495Q, R502Q, R502W, Lys504 deletion, G507R, G523W and E542Q) within the cMyBP-C C3 domain are mapped onto the WT C3 structure (gray ribbon) as pink spheres. (b) Sequence alignment of C3 domains, with invariant residues in red boxes, similar residues in red text, and stretches of amino acids that are similar across the group of sequences in blue boxes. The DSSP classified NMR-derived secondary structure of WT cMyBP-C C3 domain is shown above the alignment, which follows the *Homo sapiens* C3 sequence numbering. Arg502 is identified with a black arrowhead and the remaining mutated residues in (a) are marked with white arrowheads. The protein sequences and the list of HCM-related mutations were acquired from the Swiss-Prot database (The UniProt Consortium, 2014): *H. sapiens* (Q14896); *Canis familiaris* (Q2Q1P6); *Bos grunniens* (L8IEY5); *Bos taurus* (Q0VD56); *Pteropus alecto* (L5KZ62); *Tsuga chinensis* (L9JEW3); *Rattus norvegicus* (P56741); *Mus musculus* (Q3TF37); *Ceanothus griseus* (G3GYK6); *Heterocephalus glaber* (G5BPC6); *Gallus gallus* (E1C7T3); *Xenopus laevis* (Q90X86); *Chelonia mydas* (M7BYN8); *Ambystoma mexicanum* (Q90233).

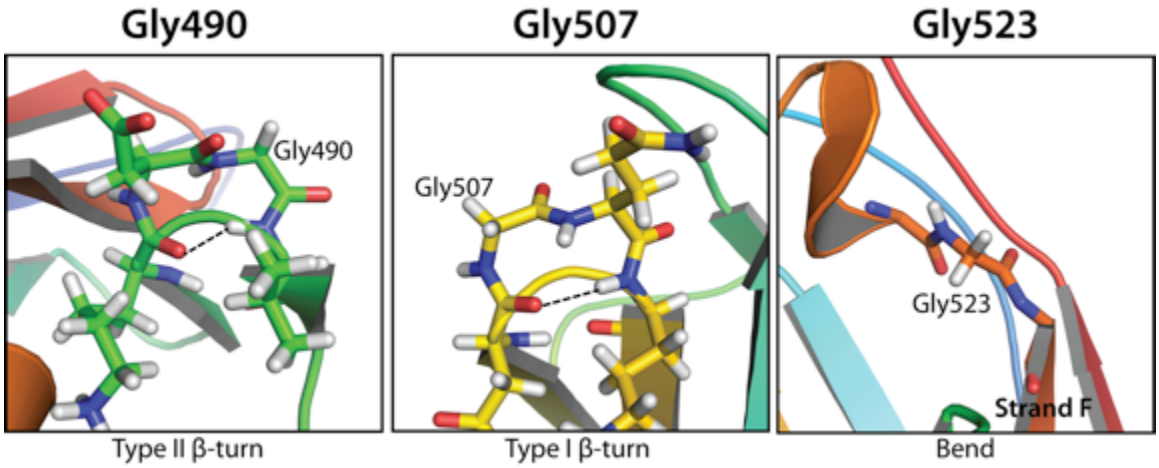

Figure 3.10 Glycine residues in C3 domain associated with HCM-related mutations.

The glycine residues in C3 domain where HCM-related mutations (G490R, G490V, G507R and G523W) occur are shown on the WT C3 domain structure (rainbow cartoon). Gly490, part of the CC' loop, is the third residue in a Type II β-turn (shown in sticks with hydrogen bonds as dashed lines). Gly507, part of the DE loop, is the second residue in a Type I β-turn (shown in sticks). The intraturn hydrogen bonds are shown as dashed lines. Gly523 (shown in sticks) is part of the EF loop and is in the middle of a bend, which has backbone curvature of at least 70 degrees.

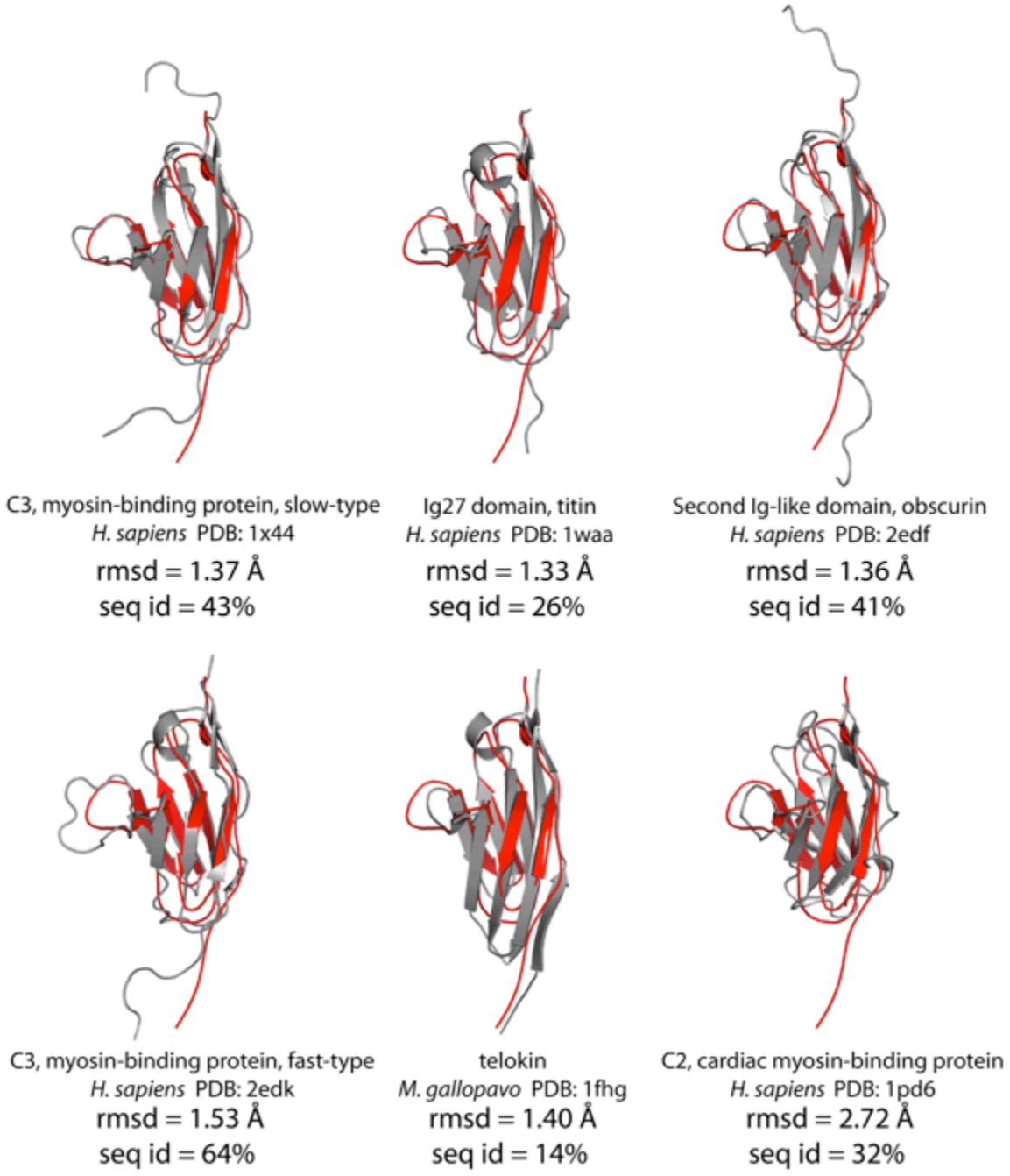

Figure 3.11 Structural homologues of cMyBP-C C3 domain.
The cMyBP-C C3 domain (red) is superposed on the structures of the 6 proteins with the most similar topology and architecture (gray), as identified using the servers DALI (Holm and Rosenström, 2010) and FATCAT (Ye and Godzik, 2003). The RMSD values were calculated against the backbone atoms of the lowest energy C3 domain structure. These proteins are the C3 domains of myosin-binding protein in both fast-type (PDB: 2EDK) and slow-type (PDB: 1X44) muscles, the C2 domain of cMyBP-C (PDB: 1PD6) (Ababou et al., 2007), Ig27 domain of titin (PDB: 1WAA) (Stacklies et al., 2009), telokin (PDB: 1FHG) (Holden et al., 1992), and the second Ig-like domain of obscurin (PDB: 2EDF).

Chapter 4.

Concluding remarks

4.1. Conclusions

HCM is a genetic disease caused by mutations within various sarcomeric proteins. To illustrate the possible structural effects on proteins induced by two of the HCM-related mutations (the L29Q mutation in cTnC and the R502W mutation in cMyBP-C), crystal structures of WT and L29Q cNTnC and NMR solution structures of cMyBP-C were determined.

The WT cNTnC structure was solved at 1.4 Å, the highest resolution structure to date for cTnC while the L29Q cNTnC structure was solved at 1.9 Å. A NIQD cNTnC structure was also solved at 2.0 Å. This structure represents the mutations (D2N, V28I, L29Q and G30D) found in exothermic species that have functional importance in Ca^{2+}-sensing at lower temperatures, which is the physiological temperature of most ectothermic species. All three cNTnC structures have Cd^{2+} coordinated within their Ca^{2+}-binding EF-hand sites, both at the vestigial Site I and the functional Site II. The vestigial Site I coordinated Cd^{2+} in a 'distorted' octahedral geometry and the functional Site II coordinated Cd^{2+} in the canonical pentagonal bipyramidal geometry. The presence of cysteine in the EF1 loop is proposed to play a major role in coordinating Cd^{2+} at the vestigial Site I. Due to a lack of electron density, three residues in the Site II coordinating loop are missing from the structures. However, water molecules were found in place of the missing ligands and were coordinating the Cd^{2+} ions in Site II.

Superimposition of the three structures did not show significant structural differences. Neither the L29Q mutation nor the NIQD mutations altered the overall structure or protein fold of cNTnC. Close examination of the superimposition at the L29Q mutation region showed subtle main-chain shifts in the EF1 loop of cNTnC. In L29Q cNTnC, the main chain between residues 29 and 32 shifted towards the solvent channel

by up to 0.5 Å compared with the WT which in NIQD cNTnC, the same residues shifted by up to 1.0 Å. As these mutations are located on the surface of the protein and their side chains face the solvent, nonpolar-to-polar mutations at these positions would increase the surface hydrophilicity of the protein (Figure 4.1). This increased interaction between the protein and solvent decreases the free energy that needs to be overcome for the conformational change that leads to protein activation.

The overall conformation of the Cd^{2+}-bound cNTnC structure resembles that of the Ca^{2+}-bound cNTnC and both structures represent a closed cNTnC conformation. This is in contrast with the open conformation observed in both the Ca^{2+}-bound cNTnC/TFP complex and the Cd^{2+}-bound cNTnC/DXC complex. Both TFP and DXC are bound at the hydrophobic core of cNTnC and are likely the factors that stabilize the open conformation of cNTnC.

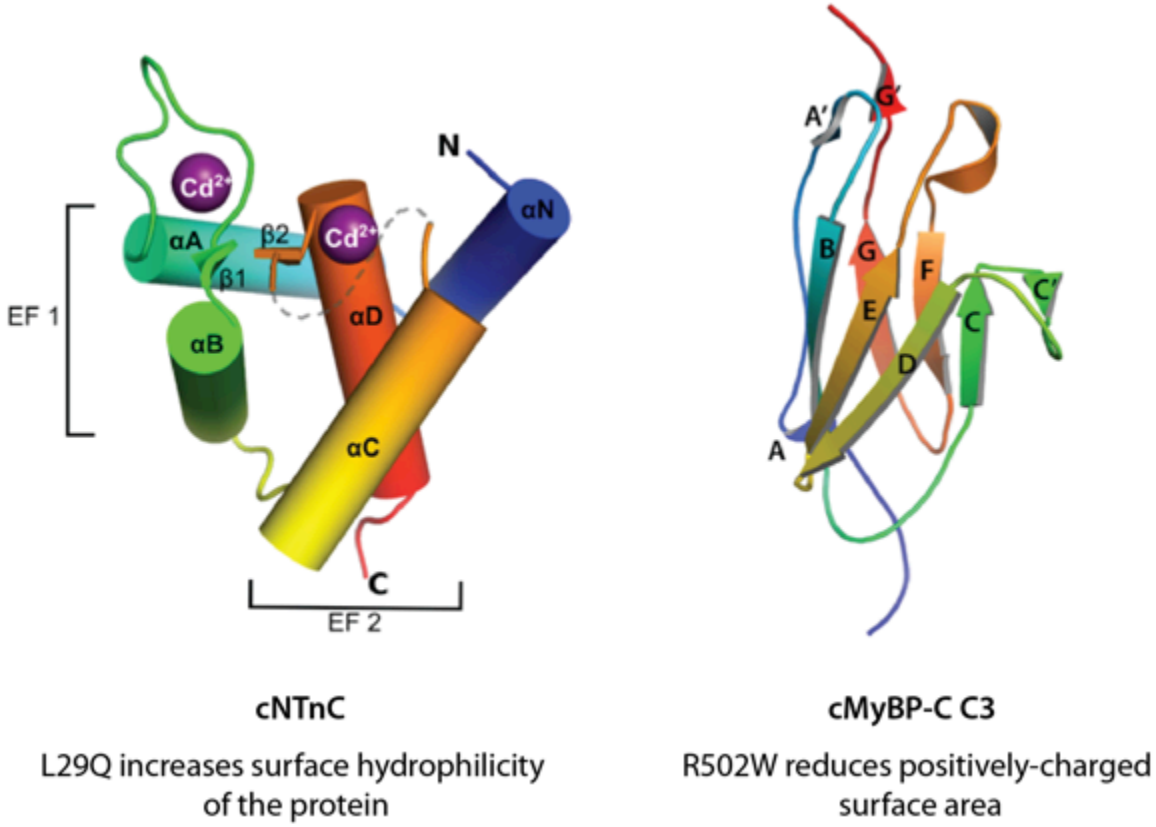

Figure 4.1 Summary of the mutational effect of L29Q on cNTnC and R502W on cMyBP-C C3 domain.

The cartoon diagram of cNTnC and the ribbon diagram of cMyBP-C C3 domain are presented in rainbow coloring (N-terminal: blue; C-terminal: red). Neither L29Q nor R502W causes global conformational change within the proteins, but they both alter protein surface properties.

The first structures of the WT and R502W C3 domain of human cMyBP-C were determined using NMR spectroscopy. Both structures adopt the Ig-like protein fold: a β-sandwich structure composed of two anti-parallel β-sheets with a Greek-key topology. The R502W mutation is in the second position on strand D and does not perturb the C3 domain structure. This suggests that the R502W mutation does not alter the function of cMyBP-C through a large-scale conformational change within the C3 domain itself.

Data from amide ^{15}N T_1, T_2, and heteronuclear NOE relaxation experiments and CD spectroscopy showed that R502W does not perturb either the protein dynamics or the protein's thermal stability. However, R502W greatly altered the surface charge of the protein by reducing the size of a large positively charged surface area (Figure 4.1). Such an electrostatic change could alter the interaction of the cMyBP-C C3 domain with a complementary charged partner protein.

Several of the other HCM-related mutations within the C3 domains are positioned within this positively-charged patch. This further highlights the importance of the positively-charged surface in mediating protein-protein interactions and suggests a common mechanism for the disruption of function by these mutations. Three other HCM-related mutations in C3 occur at glycine residues that are located within loop regions of the protein, which presumably are required for proper folding. Two of the glycines are part of β-turns and the third glycine is within a region with high curvature, which creates a directional change of more than 70 degrees in the backbone.

4.2. Future directions

Despite the determination of both the L29Q cNTnC and R502W cMyBP-C structures, many questions remain unanswered on how these two HCM-related mutations affect the function of each protein. In this section, I propose several future experiments that may result in a greater understanding of the mutational effect of L29Q and R502W on the structures and functions of cTnC and cMyBP-C, respectively.

4.2.1. Co-crystallization of the cTnC/cTnI and troponin complexes

The resolved structures of WT and L29Q cNTnC did not show changes in the protein's secondary or tertiary structure as a result of the L29Q mutation. Further efforts could be focused on the possible effect of the mutation on the interaction of cTnC with its neighbouring protein, cTnI. This structural investigation could be put into the context of the cTnC/cTnI and the troponin complexes to give a more holistic view, as cTnC does not function by itself but within a heterotrimeric complex.

Attempts have been made to co-crystallize the cTnC/cTnI and troponin complexes (Appendix B), but there has been no success to date. Various crystallization strategies could be employed to achieve this goal. This includes, but is not limited to, protein engineering to remove any flexible regions that may hinder crystal formation, optimization of purification procedures to improve the homogeneity of the protein sample and screening for additives that may stabilize the complex and promote crystallization.

4.2.2. NMR solution structures of L29Q cTnC

The determination of a crystal structure relies on having the protein molecules aligned in exactly the same orientation. While this helps to generate an atomic level resolution structure, it also imposes limitations, as the protein might not be crystallized in its physiological conformation. In this thesis project, the cNTnC structures were crystallized with Cd^{2+} ions, and majority (5 out of 7, Table 2.2) of which are coordinated intermolecularly. Crystals could not be generated with Cd^{2+} concentrations lower than 0.02 M. This suggests that the proteins' conformations were rigidly fixed by the Cd^{2+} ions, forming highly ordered crystals that diffracted to 1.4 Å. This may not the true under physiological conditions. Therefore, how the L29Q mutation may alter the structures of cTnC could be further examined with other experimental techniques. Solution structures of L29Q cTnC obtained using NMR spectroscopy might provide additional information supplementary to the crystal structure data, since the protein conformation from the NMR technique would more closely resemble that in the physiological state.

4.2.3. Interaction studies of cMyBP-C with other protein partners

The HCM-related R502W mutation reduces the positively-charged surface area in the cMyBP-C C3 domain and is suggested to impact electrostatic protein-protein interactions between C3 and its protein partners. Future experiments could be focused on this possible impact. Actin is negatively charged at physiological pH and has been shown to interact with the N-terminal of C3. However, preliminary studies using surface plasma resonance (SPR) and isothermal titration calorimetry (ITC) did not detect any binding between the isolated C3 domain and actin. However, it is possible that additional components of cMyBP-C are required for the postulated association with actin. Therefore, different constructs of cMyBP-C (ie. regions spanning various domains in the N-terminal cMyBP-C) could be prepared for the interaction studies. Another interaction partner, such as the S2 region of myosin, could also be a candidate in the interaction studies, as binding is observed between S2 and N-terminal cMyBP-C. The interactions could be studied using techniques such as SPR, ITC and co-sediment assays.

4.2.4. Identification of the binding interface on C3

If a binding partner with C3 or any N-terminal constructs that contain C3 is identified in the experiments suggested above, NMR spectroscopy could be employed to determine the binding interface. This would also confirm whether the positively-charged surface area identified in C3 is of importance in charge-mediated protein-protein interactions. By comparing spectra of C3 with and without the binding partner, the amide chemical shift perturbation could be obtained, where the residues that involved in binding would likely to show greater perturbation in the chemical shift. Co-crystallization of the two binding partners would generate further structural evidence for the binding interface.

References

Ababou, A., Gautel, M., and Pfuhl, M. (2007). Dissecting the N-terminal myosin binding site of human cardiac myosin-binding protein C. Structure and myosin binding of domain C2. J. Biol. Chem. *282,* 9204-9215.

Ababou, A., Rostkova, E., Mistry, S., Le Masurier, C., Gautel, M., and Pfuhl, M. (2008). Myosin binding protein C positioned to play a key role in regulation of muscle contraction: structure and interactions of domain C1. J. Mol. Biol. *384,* 615-630.

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491,* 56-65.

Amann, K.J., Renley, B.A., and Ervasti, J.M. (1998). A cluster of basic repeats in the dystrophin rod domain binds F-actin through an electrostatic interaction. J. Biol. Chem. *273,* 28419-28423.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U. S. A. *98,* 10037-10041.

Baryshnikova, O.K., Robertson, I.M., Mercier, P., and Sykes, B.D. (2008). The dilated cardiomyopathy G159D mutation in cardiac troponin C weakens the anchoring interaction with troponin I. Biochemistry *47,* 10950-10960.

Bennett, P., Craig, R., Starr, R., and Offer, G. (1986). The ultrastructural location of C-protein, X-protein and H-protein in rabbit muscle. J. Muscle Res. Cell. Motil. *7,* 550-567.

Blumenschein, T.M.A., Gillis, T.E., Tibbits, G.F., and Sykes, B.D. (2004). Effect of temperature on the structure of trout troponin C. Biochemistry *43,* 4955-4963.

Bonne, G., Carrier, L., Bercovici, J., Cruaud, C., Richard, P., Hainque, B., Gautel, M., Labeit, S., James, M., Beckmann, J.*, et al.* (1995). Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. Nat. Genet. *11,* 438-440.

Bork, P., Holm, L., and Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. J. Mol. Biol. *242,* 309-320.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54,* 905-921.

Carrier, L., Bonne, G., and Schwartz, K. (1998). Cardiac Myosin-Binding Protein C and Hypertrophic Cardiomyopathy. TCM *8 (4),* 151-157.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y., and McPherson, A. (1994). Screening and optimization strategies for macromolecular crystal growth. Acta Crystallogr. D Biol. Crystallogr. *50,* 414-423.

Daragan, V.A., and Mayo, K.H. (1997). Motional model analyses of protein and peptide dynamics using 13C and 15N NMR relaxation. Progress in Nuclear Magnetic Resonance Spectroscopy *31(1),* 63-105.

de Arruda,M.V., Bazari, H., Wallek, M., and Matsudaira, P. (1992). An actin footprint on villin. Single site substitutions in a cluster of basic residues inhibit the actin severing but not capping activity of villin. J. Biol. Chem. *267,* 13079-13085.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6,* 277-293.

DeLano, W.L. (2004). Use of PYMOL as a communications tool for molecular science. Abstracts of Papers of the American Chemical Society *228,* U313-U314.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System.

Dixon, R.D.S., Arneman, D.K., Rachlin, A.S., Sundaresan, N.R., Costello, M.J., Campbell, S.L., and Otey, C.A. (2008). Palladin is an actin cross-linking protein that uses immunoglobulin-like domains to bind filamentous actin. J. Biol. Chem. *283,* 6222-6231.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. *32,* W665-W667.

Dosset, P., Hus, J.C., Blackledge, M., and Marion, D. (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J. Biomol. NMR *16,* 23-28.

Dvoretsky, A., Abusamhadneh, E.M., Howarth, J.W., and Rosevear, P.R. (2002). Solution structure of calcium-saturated cardiac troponin C bound to cardiac troponin I. J. Biol. Chem. *277,* 38565-38570.

Ebashi, S. (1963). Third Component Participating in the Superprecipitation of 'Natural Actomyosin'. Nature *200,* 1010-1010.

Ebashi, S., and Endo, M. (1968). Calcium ion and muscle contraction. Prog. Biophys. Mol. Biol. *18,* 123-183.

Ebashi, S., Endo, M., and Otsuki, I. (1969). Control of muscle contraction. Q. Rev. Biophys. *2,* 351-384.

Ebashi, S., and Kodama, A. (1966). Interaction of troponin with F-actin in the presence of tropomyosin. J Biochem *59,* 425-426.

Ebashi, S., and Kodama, A. (1965). A new protein factor promoting aggregation of tropomyosin. J Biochem *58,* 107-108.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J., and Kay, L.E. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry *33,* 5984-6003.

Finley, N.L., Howarth, J.W., and Rosevear, P.R. (2004). Structure of the Mg2+-loaded C-lobe of cardiac troponin C bound to the N-domain of cardiac troponin I: comparison with the Ca2+-loaded structure. Biochemistry *43,* 11371-11379.

Fisher, S.J., Helliwell, J.R., Khurshid, S., Govada, L., Redwood, C., Squire, J.M., and Chayen, N.E. (2008). An investigation into the protonation states of the C1 domain of cardiac myosin-binding protein C. Acta Crystallogr. D Biol. Crystallogr. *64,* 658-664.

Flashman, E., Redwood, C., Moolman-Smook, J., and Watkins, H. (2004). Cardiac myosin binding protein C: its role in physiology and disease. Circ. Res. *94,* 1279-1289.

Flashman, E., Watkins, H., and Redwood, C. (2007). Localization of the binding site of the C-terminal domain of cardiac myosin-binding protein-C on the myosin rod. Biochem. J. *401,* 97-102.

Freiburg, A., and Gautel, M. (1996). A molecular map of the interactions between titin and myosin-binding protein C. Implications for sarcomeric assembly in familial hypertrophic cardiomyopathy. Eur. J. Biochem. *235,* 317-323.

Furst, D.O., and Gautel, M. (1995). The anatomy of a molecular giant: how the sarcomere cytoskeleton is assembled from immunoglobulin superfamily molecules. J Mol Cell Cardiol *27,* 951-959.

Gasmi-Seabrook, G., Howarth, J.W., Finley, N., Abusamhadneh, E., Gaponenko, V., Brito, R.M., Solaro, R.J., and Rosevear, P.R. (1999). Solution structures of the C-terminal domain of cardiac troponin C free and bound to the N-terminal domain of cardiac troponin I. Biochemistry *38,* 8313-8322.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy

Server. In The Proteomics Protocols Handbook, Walker, John M. ed., Humana Press) pp. 571-607.

Gautel, M., Zuffardi, O., Freiburg, A., and Labeit, S. (1995). Phosphorylation switches specific for the cardiac isoform of myosin binding protein-C: a modulator of cardiac contraction? EMBO J. *14,* 1952-1960.

Geisterfer-Lowrance, A., Kass, S., Tanigawa, G., Vosberg, H.P., McKenna, W., Seidman, C.E., and Seidman, J.G. (1990). A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. Cell *62,* 999-1006.

Gersh, B.J., Maron, B.J., Bonow, R.O., Dearani, J.A., Fifer, M.A., Link, M.S., Naidu, S.S., Nishimura, R.A., Ommen, S.R., Rakowski, H.*, et al.* (2011). 2011 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J. Thorac. Cardiovasc. Surg. *142,* e153-e203.

Gilbert, R., Kelly, M.G., Mikawa, T., and Fischman, D.A. (1996). The carboxyl terminus of myosin binding protein C (MyBP-C, C-protein) specifies incorporation into the A-band of striated muscle. J. Cell. Sci. *109 (Pt 1),* 101-111.

Gillis, T.E., Liang, B., Chung, F., and Tibbits, G.F. (2005). Increasing mammalian cardiomyocyte contractility with residues identified in trout troponin C. Physiol Genomics *22,* 1-7.

Gillis, T.E., Marshall, C.R., and Tibbits, G.F. (2007). Functional and evolutionary relationships of troponin C. Physiol Genomics *32,* 16-27.

Goddard, T.D., and Kneller, D.G. (1999). SPARKY 3.

Gordon, A.M., Homsher, E., and Regnier, M. (2000). Regulation of contraction in striated muscle. Physiol. Rev. *80,* 853-924.

Gouet, P., Courcelle, E., Stuart, D.I., and Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics *15,* 305-308.

Govada, L., Carpenter, L., da Fonseca, P.,C.A., Helliwell, J.R., Rizkallah, P., Flashman, E., Chayen, N.E., Redwood, C., and Squire, J.M. (2008). Crystal structure of the C1 domain of cardiac myosin binding protein-C: implications for hypertrophic cardiomyopathy. J. Mol. Biol. *378,* 387-397.

Govada, L., Carpenter, L., da Fonseca, P.,C.A., Helliwell, J.R., Rizkallah, P., Flashman, E., Chayen, N.E., Redwood, C., and Squire, J.M. (2008). Crystal structure of the C1 domain of cardiac myosin binding protein-C: implications for hypertrophic cardiomyopathy. J. Mol. Biol. *378,* 387-397.

GraphPad Software. GraphPad Prism version 5.00 for Windows.

Greaser, M.L., and Gergely, J. (1971). Reconstitution of troponin activity from three protein components. J. Biol. Chem. *246,* 4226-4233.

Gruen, M., Prinz, H., and Gautel, M. (1999). cAPK-phosphorylation controls the interaction of the regulatory domain of cardiac myosin binding protein C with myosin-S2 in an on-off fashion. FEBS Lett. *453,* 254-259.

Güntert, P. (2004). Automated NMR structure calculation with CYANA. Methods Mol. Biol. *278,* 353-378.

Halaby, D.M., Poupon, A., and Mornon, J. (1999). The immunoglobulin fold family: sequence analysis and 3D structure comparisons. Protein Eng. *12,* 563-571.

Harris, S.P., Bartley, C.R., Hacker, T.A., McDonald, K.S., Douglas, P.S., Greaser, M.L., Powers, P.A., and Moss, R.L. (2002). Hypertrophic cardiomyopathy in cardiac myosin binding protein-C knockout mice. Circ. Res. *90,* 594-601.

Hartshorne, D.J., and Mueller, H. (1968). Fractionation of troponin into two distinct proteins. Biochem. Biophys. Res. Commun. *31,* 647-653.

Hartshorne, D.J., Perry, S.V., and Davies, V. (1966). A factor inhibiting the adenosine triphosphatase activity and the superprecipitation of actomyosin. Nature *209,* 1352-1353.

Hartshorne, D.J., Perry, S.V., and Schaub, M.C. (1967). A protein factor inhibiting the magnesium activated adenosine triphosphatase of desensitised actomyosin. Biochem. J. *104,* 907-913.

Hershberger, R.E., Cowan, J., Morales, A., and Siegfried, J.D. (2009). Progress with genetic cardiomyopathies: screening, counseling, and testing in dilated, hypertrophic, and arrhythmogenic right ventricular dysplasia/cardiomyopathy. Circ Heart Fail *2,* 253-261.

Hoffman, R.M.B., and Sykes, B.D. (2009). Structure of the inhibitor W7 bound to the regulatory domain of cardiac troponin C. Biochemistry *48,* 5541-5552.

Hoffmann, B., Schmidt-Traub, H., Perrot, A., Osterziel, K.J., and Gessner, R. (2001). First mutation in cardiac troponin C, L29Q, in a patient with hypertrophic cardiomyopathy. Hum Mutat *17,* 524.

Holden, H.M., Ito, M., Hartshorne, D.J., and Rayment, I. (1992). X-ray structure determination of telokin, the C-terminal domain of myosin light chain kinase, at 2.8 A resolution. J. Mol. Biol. *227,* 840-851.

Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids Res. *38,* W545-W549.

Howarth, J.W., Ramisetti, S., Nolan, K., Sadayappan, S., and Rosevear, P.R. (2012). Structural insight into unique cardiac myosin-binding protein-C motif: a partially folded domain. J. Biol. Chem. *287,* 8254-8262.

Hüttelmaier, S., Harbeck, B., Steffens, O., Messerschmidt, T., Illenberger, S., and Jockusch, B.M. (1999). Characterization of the actin binding properties of the vasodilator-stimulated phosphoprotein VASP. FEBS Lett. *451,* 68-74.

Idowu, S.M., Gautel, M., Perkins, S.J., and Pfuhl, M. (2003). Structure, stability and dynamics of the central domain of cardiac myosin binding protein C (MyBP-C): implications for multidomain assembly and causes for cardiomyopathy. J. Mol. Biol. *329,* 745-761.

Janssen, M.E.W., Kim, E., Liu, H., Fujimoto, L.M., Bobkov, A., Volkmann, N., and Hanein, D. (2006). Three-dimensional structure of vinculin bound to actin filaments. Mol. Cell *21,* 271-281.

Julien, O., Mercier, P., Crane, M.L., and Sykes, B.D. (2009). The effect of the cosolvent trifluoroethanol on a tryptophan side chain orientation in the hydrophobic core of troponin C. Protein Sci. *18,* 1165-1174.

Julien, O., Sun, Y., Wang, X., Lindhout, D.A., Thiessen, A., Irving, M., and Sykes, B.D. (2008). Tryptophan mutants of cardiac troponin C: 3D structure, troponin I affinity, and in situ activity. Biochemistry *47,* 597-606.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22,* 2577-2637.

Kimura, A., Harada, H., Park, J.E., Nishi, H., Satoh, M., Takahashi, M., Hiroi, S., Sasaoka, T., Ohbuchi, N., Nakamura, T*., et al.* (1997). Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. Nat. Genet. *16,* 379-382.

Kretsinger, R.H., and Nockolds, C.E. (1973). Carp muscle calcium-binding protein. II. Structure determination and general description. J. Biol. Chem. *248,* 3313-3326.

Kulikovskaya, I., McClellan, G., Flavigny, J., Carrier, L., and Winegrad, S. (2003). Effect of MyBP-C binding to actin on contractility in heart muscle. J. Gen. Physiol. *122,* 761-774.

Kumar, V., Abbas, A.K., and Aster, J.C. (2013). Robbins Basic Pathology, Ninth Edition (Philadelphia, PA: Elsevier Saunders).

Kunst, G., Kress, K.R., Gruen, M., Uttenweiler, D., Gautel, M., and Fink, R.H. (2000). Myosin binding protein C, a phosphorylation-dependent force regulator in muscle that controls the attachment of myosin heads by its interaction with myosin S2. Circ. Res. *86,* 51-58.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R*., et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23,* 2947-2948.

Lee, H., Bellin, R.M., Walker, D.L., Patel, B., Powers, P., Liu, H., Garcia-Alvarez, B., de Pereda, J.,M., Liddington, R.C., Volkmann, N.*, et al.* (2004). Characterization of an actin-binding site within the talin FERM domain. J. Mol. Biol. *343,* 771-784.

Lesk, A.M., and Chothia, C. (1982). Evolution of proteins formed by beta-sheets. II. The core of the immunoglobulin domains. J. Mol. Biol. *160,* 325-342.

Li, A.Y., Stevens, C.M., Liang, B., Rayani, K., Little, S., Davis, J., and Tibbits, G.F. (2013). Familial hypertrophic cardiomyopathy related cardiac troponin C L29Q mutation alters length-dependent activation and functional effects of phosphomimetic troponin I*. PLoS One *8,* e79363-e79363.

Li, A.Y., Lee, J., Borek, D., Otwinowski, Z., Tibbits, G.F., and Paetzel, M. (2011). Crystal structure of cardiac troponin C regulatory domain in complex with cadmium and deoxycholic acid reveals novel conformation. J. Mol. Biol. *413,* 699-711.

Li, M.X., Spyracopoulos, L., and Sykes, B.D. (1999). Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. Biochemistry *38,* 8289-8298.

Li, Y., Love, M.L., Putkey, J.A., and Cohen, C. (2000). Bepridil opens the regulatory N-terminal lobe of cardiac troponin C. Proc. Natl. Acad. Sci. U. S. A. *97,* 5140-5145.

Liang, B., Chung, F., Qu, Y., Pavlov, D., Gillis, T.E., Tikunova, S.B., Davis, J.P., and Tibbits, G.F. (2008). Familial hypertrophic cardiomyopathy-related cardiac troponin C mutation L29Q affects Ca2+ binding and myofilament contractility. Physiol Genomics *33,* 257-266.

Lindhout, D.A., and Sykes, B.D. (2003). Structure and dynamics of the C-domain of human cardiac troponin C in complex with the inhibitory region of human cardiac troponin I. J. Biol. Chem. *278,* 27024-27034.

Lipari, G., and Szabo, A. (1982). Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. J Am Chem Soc *104,* 4546-4559.

Liu, W., Hanson, M.A., Stevens, R.C., and Cherezov, V. (2010). LCP-Tm: an assay to measure and understand stability of membrane proteins in a membrane environment. Biophys. J. *98,* 1539-1548.

Maron, B.J., Gardin, J.M., Flack, J.M., Gidding, S.S., Kurosaki, T.T., and Bild, D.E. (1995). Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults. Circulation *92,* 785-789.

McKay, R.T., Tripet, B.P., Hodges, R.S., and Sykes, B.D. (1997). Interaction of the second binding region of troponin I with the regulatory domain of skeletal muscle troponin C as determined by NMR spectroscopy. J. Biol. Chem. *272,* 28494-28500.

Mogensen, J., Klausen, I.C., Pedersen, A.K., Egeblad, H., Bross, P., Kruse, T.A., Gregersen, N., Hansen, P.S., Baandrup, U., and Borglum, A.D. (1999). Alpha-cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. J. Clin. Invest. *103,* R39-R43.

Mun, J.Y., Gulick, J., Robbins, J., Woodhead, J., Lehman, W., and Craig, R. (2011). Electron microscopy and 3D reconstruction of F-actin decorated with cardiac myosin-binding protein C (cMyBP-C). J. Mol. Biol. *410,* 214-225.

Niesen, F.H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. Nat Protoc *2,* 2212-2221.

Oakley, C.E., Hambly, B.D., Curmi, P.M.G., and Brown, L.J. (2004). Myosin binding protein C: structural abnormalities in familial hypertrophic cardiomyopathy. Cell Res. *14,* 95-110.

Offer, G., Moos, C., and Starr, R. (1973). A new protein of the thick filaments of vertebrate skeletal myofibrils. Extractions, purification and characterization. J. Mol. Biol. *74,* 653-676.

Okagaki, T., Weber, F.E., Fischman, D.A., Vaughan, K.T., Mikawa, T., and Reinach, F.C. (1993). The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif. J. Cell Biol. *123,* 619-626.

Oleszczuk, M., Robertson, I.M., Li, M.X., and Sykes, B.D. (2010). Solution structure of the regulatory domain of human cardiac troponin C in complex with the switch region of cardiac troponin I and W7: the basis of W7 as an inhibitor of cardiac muscle contraction. J. Mol. Cell. Cardiol. *48,* 925-933.

Orwig, S.D., and Lieberman, R.L. (2011). Biophysical characterization of the olfactomedin domain of myocilin, an extracellular matrix protein implicated in inherited forms of glaucoma. PLoS One *6,* e16347-e16347.

Otsuki, I., Masaki, T., Nonomura, Y., and Ebashi, S. (1967). Periodic distribution of troponin along the thin filament. J Biochem *61,* 817-819.

Page, S.P., Kounas, S., Syrris, P., Christiansen, M., Frank-Hansen, R., Andersen, P.S., Elliott, P.M., and McKenna, W.J. (2012). Cardiac myosin binding protein-C mutations in families with hypertrophic cardiomyopathy: disease expression in relation to age, gender, and long term outcome. Circ Cardiovasc Genet *5,* 156-166.

Perczel, A., Hollósi, M., Tusnády, G., and Fasman, G.D. (1991). Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins. Protein Eng. *4,* 669-679.

Pineda-Sanabria, S., Robertson, I.M., and Sykes, B.D. (2011). Structure of trans-resveratrol in complex with the cardiac regulatory protein troponin C. Biochemistry *50,* 1309-1320.

Poetter, K., Jiang, H., Hassanzadeh, S., Master, S.R., Chang, A., Dalakas, M.C., Rayment, I., Sellers, J.R., Fananapazir, L., and Epstein, N.D. (1996). Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. Nat. Genet. *13,* 63-69.

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., Benaiche, A., Isnard, R., Dubourg, O., Burban, M.*, et al.* (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation *107,* 2227-2232.

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., Benaiche, A., Isnard, R., Dubourg, O., Burban, M.*, et al.* (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation *107,* 2227-2232.

Robertson, I.M., Li, M.X., and Sykes, B.D. (2009). Solution structure of human cardiac troponin C in complex with the green tea polyphenol, (-)-epigallocatechin 3-gallate. J. Biol. Chem. *284,* 23012-23023.

Robertson, I.M., Sun, Y., Li, M.X., and Sykes, B.D. (2010). A structural and functional perspective into the mechanism of Ca2+-sensitizers that target the cardiac troponin complex. J. Mol. Cell. Cardiol. *49,* 1031-1041.

Robertson, S.P., Johnson, J.D., Holroyde, M.J., Kranias, E.G., Potter, J.D., and Solaro, R.J. (1982). The effect of troponin I phosphorylation on the Ca2+-binding properties of the Ca2+-regulatory site of bovine cardiac troponin. J. Biol. Chem. *257,* 260-263.

Saltzman, A.J., Mancini-DiNardo, D., Li, C., Chung, W.K., Ho, C.Y., Hurst, S., Wynn, J., Care, M., Hamilton, R.M., Seidman, G.W.*, et al.* (2010). The cardiac myosin binding protein C Arg502Trp mutation. A common cause of hypertrophic cardiomyopathy. Circ. Res. *106,* 1549-1552.

Satoh, M., Takahashi, M., Sakamoto, T., Hiroe, M., Marumo, F., and Kimura, A. (1999). Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. Biochem. Biophys. Res. Commun. *262,* 411-417.

Sattler, M., Schleucher, J., and Griesinger, C. (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Prog. Nucl. Mag. Res. Spec. *34,* 93-158.

Schaub, M.C., and Perry, S.V. (1969). Resolution of the troponin complex into inhibitory and calcium sensitizing factors and their relationship to tropomyosin. Biochem. J. *115,* 993-1004.

Sharma, D., and Rajarathnam, K. (2000). 13C NMR chemical shifts can predict disulfide bond formation. Journal of Biomolecular NMR *18(2),* 165-171.

Shen, Y., and Bax, A. (2012). Identification of helix capping and b-turn motifs from NMR chemical shifts. J. Biomol. NMR *52,* 211-232.

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biomol. NMR *44,* 213-223.

Sia, S.K., Li, M.X., Spyracopoulos, L., Gagné,S.M., Liu, W., Putkey, J.A., and Sykes, B.D. (1997). Structure of cardiac muscle troponin C unexpectedly reveals a closed regulatory domain. J. Biol. Chem. *272,* 18216-18221.

Sjöström, M., and Squire, J.M. (1977). Fine structure of the A-band in cryo-sections. The structure of the A-band of human skeletal muscle fibres from ultra-thin cryo-sections negatively stained. J. Mol. Biol. *109,* 49-68.

Slupsky, C.M., and Sykes, B.D. (1995). NMR solution structure of calcium-saturated skeletal muscle troponin C. Biochemistry *34,* 15953-15964.

Spyracopoulos, L., Li, M.X., Sia, S.K., Gagné,S.M., Chandra, M., Solaro, R.J., and Sykes, B.D. (1997). Calcium-induced structural transition in the regulatory domain of human cardiac troponin C. Biochemistry *36,* 12138-12146.

Squire, J.M., Luther, P.K., and Knupp, C. (2003). Structural evidence for the interaction of C-protein (MyBP-C) with actin and sequence identification of a possible actin-binding domain. J. Mol. Biol. *331,* 713-724.

Stacklies, W., Vega, M.C., Wilmanns, M., and Gräter, F. (2009). Mechanical network in titin immunoglobulin from force distribution analysis. PLoS Comput Biol *5,* e1000306-e1000306.

Strynadka, N.C., and James, M.N. (1989). Crystal structures of the helix-loop-helix calcium-binding proteins. Annu. Rev. Biochem. *58,* 951-998.

Takeda, S., Yamashita, A., Maeda, K., and Maéda, Y. (2003). Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. Nature *424,* 35-41.

Tanaka, K., Tamura, T., Yoshimura, T., and Ichihara, A. (1992). Proteasomes: protein and gene structures. New Biol. *4,* 173-187.

The UniProt Consortium. (2014). Activities at the Universal Protein Resource (UniProt). Nucleic Acids Res. *42,* D191-D198.

Thierfelder, L., Watkins, H., MacRae, C., Lamas, R., McKenna, W., Vosberg, H.P., Seidman, J.G., and Seidman, C.E. (1994). Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. Cell *77,* 701-712.

Tollinger, M., Skrynnikov, N.R., Mulder, F.A., Forman-Kay, J., and Kay, L.E. (2001). Slow dynamics in folded and unfolded states of an SH3 domain. J. Am. Chem. Soc. *123,* 11341-11352.

Wang, J., Cieplak, P., and Kollman, P.A. (2000). How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? Journal of Computational Chemistry *21 (12),* 1049-1074.

Wang, X., Li, M.X., Spyracopoulos, L., Beier, N., Chandra, M., Solaro, R.J., and Sykes, B.D. (2001). Structure of the C-domain of human cardiac troponin C in complex with the Ca2+ sensitizing drug EMD 57033. J. Biol. Chem. *276,* 25456-25466.

Wang, X., Li, M.X., and Sykes, B.D. (2002). Structure of the regulatory N-domain of human cardiac troponin C in complex with human cardiac troponin I147-163 and bepridil. J. Biol. Chem. *277,* 31124-31133.

Wang, X., Mercier, P., Letourneau, P., and Sykes, B.D. (2005). Effects of Phe-to-Trp mutation and fluorotryptophan incorporation on the solution structure of cardiac troponin C, and analysis of its suitability as a potential probe for in situ NMR studies. Protein Sci. *14,* 2447-2460.

Watkins, H., Conner, D., Thierfelder, L., Jarcho, J.A., MacRae, C., McKenna, W.J., Maron, B.J., Seidman, J.G., and Seidman, C.E. (1995). Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. Nat. Genet. *11,* 434-437.

Watkins, H., McKenna, W.J., Thierfelder, L., Suk, H.J., Anan, R., O'Donoghue, A., Spirito, P., Matsumori, A., Moravec, C.S., Seidman, J.G., and et. al. (1995). Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy. N. Engl. J. Med. *332,* 1058-1064.

Watkins, H., Rosenzweig, A., Hwang, D.S., Levi, T., McKenna, W., Seidman, C.E., and Seidman, J.G. (1992). Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy. N. Engl. J. Med. *326,* 1108-1114.

Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., and Hochstrasser, D.F. (1999). Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. *112,* 531-552.

Witt, C.C., Gerull, B., Davies, M.J., Centner, T., Linke, W.A., and Thierfelder, L. (2001). Hypercontractile properties of cardiac muscle fibers in a knock-in mouse model of cardiac myosin-binding protein-C. J. Biol. Chem. *276,* 5353-5359.

Yamazaki, T., Forman-Kay, J.D., and Kay, L.E. (1993). Two dimensional NMR experiment for correlating ^{13}Cβ and 1Hδ/ε chemical shifts of aromatic residues in 13C labeled proteins via scalar couplings. J Am Chem Soc *115,* 11054-11055.

Yarmola, E.G., Edison, A.S., Lenox, R.H., and Bubb, M.R. (2001). Actin filament cross-linking by MARCKS: characterization of two actin-binding sites within the phosphorylation site domain. J. Biol. Chem. *276,* 22351-22358.

Ye, Y., and Godzik, A. (2003). Flexible structure alignment by chaining aligned fragment pairs allowing twists. Bioinformatics *19 Suppl 2,* ii246-ii255.

Zwahlen, C., Gardner, K.H., Sarma, S.P., Horita,D.A., Byrd, R.A., and Kay, L.E. (1998). An NMR Experiment for Measuring Methyl−Methyl NOEs in 13C-Labeled Proteins with High Resolution. Journal of the American Chemical Society *120 (30),* 7617-7625.

Appendix A.
List of constructs

A list of the constructs used in the thesis project.

Construct	Residues	Vector	MWP#	Uniprot ID	pI*	Extinction Coefficient ($M^{-1}cm^{-1}$)*
Human WT cNTnC	1-89	pET-21	750	P63316	3.99	1490
Human L29Q cNTnC	1-89	pET-21	751		3.99	1490
Human NIQD cNTnC	1-89	pET-21	752		3.99	1490
Mouse WT cTnC	1-161	pET-21	754	P19123	4.02	4470
Mouse L29Q cTnC	1-161	pET-21	755		4.02	4470
Mouse C35S cNTnC	1-89	pET-21	851		4.15	1490
Mouse C84S cTnC	1-161	pET-21	758		4.02	4470
Mouse C35S/C84S cTnC	1-161	pET-21	759		4.02	4470
Mouse WT cCTnC	90-161	pET-21	875		3.89	2980
Mouse WT cTnI	1-211	psBETa	756	P48787	9.57	11460
Mouse WT cTnT	1-301	pET-24a	757	P50752	4.98	20970
Human WT cMyBP-C C3	453-543	pET-28	824	Q14896	5.00	8480
Human R502W cMyBP-C C3	453-543	pET-28	917		5.00	8480

* Calculated by ProtParam (Gasteiger et al., 2005)

Appendix B.
Co-crystallization of the mouse cTnC/cTnI complex and troponin complexes

Efforts have been made to co-crystallize the mouse cTnC/cTnI and troponin complexes, however, no crystals have formed to date. This appendix will document the over-expression, purification and crystallization methods used in the attempt to co-crystallize the cTnC/cTnI and the troponin complexes.

The cTnC/cTnI-peptide protein preparation and crystallization:

Full-length cTnC, cNTnC and cCTnC were over-expressed and purified as described in Section 2.2.1. The cTnI peptides were synthesized commercially and combined with cTnC in 5:1 and 10:1 molar ratios before being used for crystallization.

List of combinations used for cTnC/cTnI co-crystallization.

cTnC/cTnI	Full-length cTnC / cTnI (128-147)
	Full-length cTnC / cTnI (147-163)
	cNTnC / cTnI (128-147)
	cNTnC / cTnI (147-163)
	cCTnC / cTnI (128-147)

The cardiac troponin complex protein preparation and crystallization:

Mouse WT and L29Q cTnC over-expression and purification

Mouse WT and L29Q cTnC were over-expressed and purified using the same procedure as described for the human cNTnC (Section 2.2.1).

Mouse WT cTnI and S23D/S24D cTnI over-expression and purification

The BL21(DE3) glycerol stocks containing the psBETa plasmid with the mouse WT and with the mouse mutant cTnI insert were generously contributed by Dr. Glen Tibbits' lab. Cells containing the plasmids were grown at 37°C in LB media (50 µg/mL kanamycin) until the OD_{600} reached 0.6. The cells were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and were incubated for an additional 4 hours at 37°C. Cell pellets were collected by centrifugation at 6000 g for 7 minutes at 4 °C and resuspended in extraction buffer [25 mM Triethanolamine Hydrochloride (TEA-HCl) pH 7.5, 5 mM EDTA, 6M urea, 15mM β-mercaptoethanol] (20 mL buffer per liter of culture). Cells were sonicated for 45 seconds at 30% amplitude using a Model 500 Dismembrator (Fisher Scientific) and lysed using an Avestin EmulsiFlex-C3 high-pressure homogenizer set at 15,000 - 20,000 psi for 5 minutes. The

lysate was centrifuged at 28,964 g for 35 minutes at 4°C and then the supernatant was collected. Twenty-five grams of ammonium sulphate were added to each 100 mL of supernatant, and the supernatant was stirred at 4°C for 30 minutes. The sample was centrifuged again at 28,964 g for 35 minutes at 4°C and then the supernatant was dialyzed with column buffer A [25 mM TEA-HCl, pH7.5, 1mM EDTA, 6M urea, 15mM β-mercaptoethanol] at 4°C. The dialyzed sample was filtered through a 0.45 µm filter before being loaded onto a CM sepharose column. The proteins was eluted with Column Buffer B [25 mM TEA-HCl, pH7.5, 1mM EDTA, 6M urea, 15mM β-mercaptoethanol, 0.5M NaCl] on a gradient. Samples from the elution fractions were run on a 12% SDS-PAGE gel. Fractions containing cTnI were combined and dialyzed with DEAE Column Buffer A [20 mM Tris-HCl, 1mM EDTA, 6M urea, 0.5 mM DTT]. The protein was then further purified using a DEAE FF column and eluted on a gradient using DEAE Column Buffer B [20 mM Tris-HCl, 1mM EDTA, 6M urea, 0.5 mM DTT, 0.5M NaCl]. All fractions containing cTnI were identified using 12% SDS-PAGE and dialyzed extensively with 1 mM HCl. The protein was lyophilized and stored at -80°C.

WT cTnT over-expression and purification

The BL21(DE3) glycerol stock containing the pET-24a plasmid with the mouse WT cTnT insert was generously contributed by Dr. Glen Tibbits' lab. Cells containing the plasmid were grown at 37°C in LB media (100 µg/mL kanamycin) until the OD_{600} reached 0.6. The cells were then induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and were incubated for an additional 4 hours at 37°C. Cell pellets were collected by centrifugation at 6000 g for 7 minutes at 4°C and resuspended in resuspension buffer [25 mM TEA-HCl pH 7.5, 1 mM DTT, 6M urea, 5mM EDTA, 1 tablet of protease cocktail inhibitor (Roche Diagnostics)] (10mL buffer per 1 liter of culture). Cells were sonicated for 45 seconds at 30% amplitude using a Model 500 Dismembrator (Fisher Scientific) and lysed using an Avestin EmulsiFlex-C3 high-pressure homogenizer set at 15,000 - 20,000 psi for 5 minutes. The lysate was centrifuged at 28,964 g for 35 minutes at 4°C and then the supernatant was dialyzed against CM Column Buffer A [50mM sodium citrate, pH6.0, 6M urea, 1mM EDTA, 0.1mM DTT] at 4°C. The dialyzed sample was filtered through a 0.45 µm filter before being loaded onto a CM sepharose column and eluted on a gradient with CM Column Buffer B [50mM sodium citrate, pH6.0, 6M urea, 1mM EDTA, 0.1mM DTT, 0.5M NaCl].

Samples from the elution fractions were run on a 12% SDS-PAGE gel. Fractions containing cTnT were combined and dialyzed against DEAE Column Buffer A [50mM Tris-HCl, pH8.0, 1mM EDTA, 6M urea, 0.1mM DTT] at 4°C. The dialyzed protein was then further purified using a DEAE column and eluted on a gradient with DEAE Column Buffer B [50mM Tris-HCl, pH8.0, 1mM EDTA, 6M urea, 0.1mM DTT, 0.5M NaCl]. All fractions containing cTnT were identified using 12% SDS-PAGE and dialyzed extensively against 5mM ammonium hydrogen carbonate. The dialyzed sample was then lyophilized and stored at -80°C.

Generating the mouse troponin complex using serial buffer dialysis

To make the cardiac troponin complex, purified mouse cTnC, cTnI and cTnT proteins were dialyzed against 10mM MOPS, pH7.0, 4.6M urea, 0.01% NaN_3, 0.5 mM DTT at 4°C. The dialyzed proteins were combined at a 1 cTnC : 1.5 cTnI: 1.5 cTnT ratio. The mixture was incubated at room temperature for 20 minutes after gentle mixing and then dialyzed against a series of buffers (Table B.1) to make the troponin complex (no stirring was used during dialysis). The dialyzed sample was centrifuged at 90K rpm for 20 minutes at 4°C using a benchtop centrifuge. The concentration of the complex was determined by absorbance at 280nm with an extinction coefficient of 31600 and a molecular weight of 77.20 kDa.

Table B.1 Serial dialysis buffers used to generate the mouse cardiac troponin complex.

Buffer 1	10mM MOPS, pH 7.0, 4M urea, 1M KCl, 0.01% NaN_3, 0.5mM DTT, 3mM $MgCl_2$
Buffer 2	10mM MOPS, pH 7.0, 2M urea, 750mM KCl, 0.01% NaN_3, 0.5mM DTT, 3mM $MgCl_2$
Buffer 3	10mM MOPS, pH 7.0, 500mM KCl, 0.01% NaN_3, 0.5mM DTT, 3mM $MgCl_2$
Buffer 4	10mM MOPS, pH 7.0, 150mM KCl, 0.01% NaN_3, 0.5mM DTT, 3mM $MgCl_2$

Four different complexes listed below were prepared and crystallized with various crystal screens using a range of crystallization techniques and strategies.

List of combinations used for troponin complex co-crystallization.

Full-length Troponin Complex	WT cTnC / WT cTnI / WT cTnT
	WT cTnC / SD cTnI / WT cTnT
	L29Q cTnC / WT cTnI / WT cTnT
	L29Q cTnC / SD cTnI / WT cTnT

Appendix C.
List of metals and screens used for cTnC crystallization

Metal	Compound	Concentration Used
Europium	Europium Acetate, $Eu(CH_3COO)_3$	10 mM in water
Gadolinium	Gadolinium(III) chloride hexahydrate, $GdCl_3.6H_2O$	10 mM in water
Gold	Gold potassium cyanide, $KAu(CN)_4$	10 mM in water
Lead	Trimethyl Lead Acetate, $(CH_3)_3Pb(CH_3COO)$	10 mM in water
Mercury	Mercury chloride, $HgCl_2$	10 mM in water
	Potassium tetraiodomercurate(II), K_2HgI_4	10 mM in water
Palladium	Potassium chloropalladite, K_2PdCl_4	10 mM in water
Platinum	Potassium tetrachloroplatinate(II), K_2PtCl_4	10 mM in water
	Potassium hexachloroplatinate(IV), K_2PtCl_6	10 mM in water
	Platinum potassium thiocyanate, $K_2Pt(CNS)_6$	10 mM in water
	Potassium thiocyanoplatinite, $K_2Pt(CNS)_4$	10 mM in water
	Potassium platinum iodide, K_2PtI_6	10 mM in water
	Platinum chloride diammine, $PtCl_2(NH_3)_2$	10 mM in water
Uranium	Uranyl acetate, $UO_2(CH_3COO)_2.2H_2O$	5 mM in ethanol
	Uranyl formate, $UO_2(CHO_2)_2.H_2O$	5 mM in ethanol
	Uranyl nitrate, $UO_2(NO_3)_2$	10 mM in water

Crystal Screens
Crystal Screen 1 (Hampton Research)
Crystal Screen 2 (Hampton Research)
PEG Ion Screen (Hampton Research)
Wizard Screen 1 (Rigaku)
Wizard Screen 2 (Rigaku)
PACT Premier Screen (Molecular Dimensions)
JCSG+ Screen (Molecular Dimensions)
Helsinki Random Screen 1 (Cudney et al., 1994)