


**University of Alberta**

**“Cardiomyopathy mutations in cardiac troponin C: functional and structural consequences”**

by

**Olga Konstantinovna Baryshnikova** 

A thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

Department of Biochemistry

Edmonton, Alberta

Fall 2008



Library and  
Archives Canada

Bibliothèque et  
Archives Canada

Published Heritage  
Branch

Direction du  
Patrimoine de l'édition

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file    Votre référence*

*ISBN: 978-0-494-46280-5*

*Our file    Notre référence*

*ISBN: 978-0-494-46280-5*

#### NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

#### AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

*моему отцу*

## ABSTRACT

Recent genetic analysis revealed a number of human mutations in cardiac troponin C (cTnC), which were linked to cardiomyopathy, the most common genetic disorder of the heart. This thesis explores structural and functional consequences of these mutations obtained using solution NMR spectroscopy. The L29Q and E59D/D75Y mutations did not affect the  $\text{Ca}^{2+}$  affinity of the N-domain of cTnC (cNTnC). Instead the affinity towards the switch region of cardiac troponin I (cTnI) was perturbed, especially in the presence of the cTnI N-terminus, responsible for the transferring of PKA signaling. PKA signal is transferred via modulating the affinity of cNTnC towards the switch region of cTnI, which was altered in the L29Q and E59D/D75Y mutants possibly resulting in the development of malfunctions. The other mutant studied, G159D, had a weaker affinity towards the anchoring region of cTnI, which likely results in a weaker anchoring of the entire troponin complex to the thin filament, thus leading to the heart dysfunction. Differences in chemical shifts and NOE connectivities support the notion of the weaker binding for the anchoring region of cTnI, however, the structure and dynamics of the G159D mutant remained unchanged. Since the N-terminus of cTnI plays an important role in the function of cNTnC and its mutants, the attempt has been made to determine individual pKa's for side chain carboxyl groups possibly involved in the formation of the complex using the modified pulse sequence HBHG(CBCG)CO. It was also determined in this work that the formation of the cNTnC•cTnI<sub>1-29</sub> complex is electrostatically driven. These results elaborate on the current model of  $\text{Ca}^{2+}$  regulated cardiac contraction and provide an insight on the nature of perturbed interactions in cardiomyopathy mutants of cTnC, which can be beneficial for the development of therapeutic strategies in cardiomyopathy.



A portion of this thesis (Appendices C and D) addresses structural and dynamic aspects of SDF-1 $\alpha$  (a chemokine involved in cancer metastasis and HIV propagation) and the peptide agonists developed on the basis of SDF-1 $\alpha$ . NMR solution structures of the peptide agonists, RSVM and ASLW, and the dimerization of SDF-1 $\alpha$  studied by NMR relaxation methods are described.

## ACKNOWLEDGMENTS

This document is a testament of scientific growth that I watched myself going through while being in Dr. Brian Sykes's lab. Comparing my graduate career to the ski hill, most of the time, the black diamonds I was taking barely matched my skills, and the expectations for easy pleasure were hardly ever met. However, finishing off my run and finding my legs intact is an exhilarating feeling, that gives me the courage to climb another chair lift and glimpse onto another mountain top. I would like to thank people who shared my run: Marion Benedict, who set me on this trip; Dr. Howard Young and Dr. Mark Glover, who held the hand on the pulse of my progress all the way down; Sue Smith and the secretaries, who helped me to navigate through mountains of paperwork; technicians, Angela Thiessen and Dave Corson, who taught me a few of the most important steps; Dr. Monica Li, who taught me how to write; computer specialists, Robert Boyko, Dr. Pascal Mercier, and Olivier Julien for not giving up on me; and many others, who lent me a hand or a piece of advice or a word of encouragement.

This has been also the time of personal growth, shared by the extended family of Brian's lab and our department. I would like to thank Drs. Brian Sykes, Monica Li, Leo Spyropoulos, Howard Young, Mark Glover, Ross Edwards, and Andrew MacMillan for sharing their experiences and thoughts on philosophy and nature of science, life, politics, and music. I also thank my peers, who ended up with the same set of skis on the same hill, Ryan Hoffman, Xu Wang, Ian Robertson, Olivier Julien, and others, with whom we shared our days, questioned ourselves and the world around us, and at times admired the view.

The skier would have given up, was it not for the guardian angel, Dr. Sykes, who with paternal devotion followed me from the top to the bottom through the pits of despair, taught me all the strategies, and whose faith in the skier's ability has never failed. Apart from Brian's awe-inspiring expertise and love of NMR, his professional and personal qualities set an unsurpassable example to follow. For me, his love of science and love of people, scientific depth and creativity, his energy and wisdom have been (and always will be) inspirational.

Thank you all.

## TABLE OF CONTENTS

	Page number
<b>CHAPTER I</b>	<b>1</b>
<b>INTRODUCTION</b>	
References	13
<b>CHAPTER II</b>	<b>18</b>
<b>The effect of L29Q and E59D/D75Y cardiomyopathy mutations on the function of cNTnC</b>	
Outline	18
Introduction	18
Results	19
Discussion	27
Conclusions	33
Materials and Methods	34
References	47
<b>CHAPTER III</b>	<b>51</b>
<b>Consequences of the DCM mutation G159D on the structure and dynamics of cardiac troponin C and its interaction with troponin I</b>	
Outline	51
Introduction	51
Results	53
Discussion	59
Materials and Methods	61
References	72
<b>CHAPTER IV</b>	<b>76</b>
<b>Exploring the interaction between cNTnC and the N-terminus of cTnI</b>	
Outline	76
Introduction	76

Results and Discussion	78
Conclusions	80
Materials and Methods	80
References	86
<b>CHAPTER V</b>	<b>88</b>
<b>Towards the determination of the pKa's of side chain carboxyl groups involved in electrostatic interactions on the interface between cNTnC and the N-terminus of cTnI</b>	
Outline	88
Introduction	88
Results and Discussion	89
Conclusions	91
Materials and Methods	91
References	95
<b>CHAPTER VI</b>	<b>97</b>
<b>Conclusions</b>	
References	102
<b>APPENDIX A</b>	<b>103</b>
<b>The code for the HBHG(CBCG)CO pulse sequence</b>	
<b>APPENDIX B</b>	<b>108</b>
<b>Buffer components as internal pH indicators for biomolecular NMR</b>	
Outline	108
Introduction	108
Results and Discussion	109
Conclusions	111
References	113
<b>APPENDIX C</b>	<b>114</b>
<b>Structures of peptide agonists derived on the basis of SDF-1<math>\alpha</math> reveal a</b>	

## **common motif possibly implicated in the binding to the CXCR4 receptor**

Outline	114
Introduction	114
Results	116
Discussion	120
Materials and Methods	122
References	129

## **APPENDIX D**

### **Dimerization of SDF-1 $\alpha$ studied by NMR relaxation methods**

Outline	132
Introduction	132
Results	134
Discussion	139
Materials and Methods	142
References	152

## LIST OF TABLES

	Page number
Table I-1	8
Table II-1	23
Table II-2	25
Table II-3	26
Table II-4	31
Table III-1	55
Table III-2	57
Table III-3	63
Table B-1	110
Table C-1	117
Table C-2	118
Table C-3	119
Table C-4	124
Table C-5	125
Table D-1	135
Table D-2	137

## LIST OF FIGURES

	Page number
Figure I-1	1
Figure I-2	2
Figure I-3	3
Figure I-4	10
Figure I-5	11
Figure II-1	37
Figure II-2	38
Figure II-3	39
Figure II-4	40
Figure II-5	41
Figure II-6	42
Figure II-7	45
Figure II-8	46
Figure III-1	64
Figure III-2	65
Figure III-3	66
Figure III-4	67
Figure III-5	68
Figure III-6	69
Figure III-7	70
Figure III-8	71
Figure IV-1	82
Figure IV-2	83
Figure IV-3	84
Figure IV-4	85
Figure V-1	92

Figure V-2	93
Figure V-3	94
Figure VI-1	100
Figure VI-2	101
Figure B-1	112
Figure C-1	126
Figure C-2	126
Figure C-3	127
Figure C-4	128
Figure D-1	144
Figure D-2	145
Figure D-3	146
Figure D-4	147
Figure D-5	147
Figure D-6	148
Figure D-7	149
Figure D-8	150
Figure D-9	151



## LIST OF ABBREVIATIONS

AEDANS	5-((2-((iodoacetyl)amino)ethyl)amino)naphthalene-1-sulfonic acid
AIDS	acquired immunodeficiency syndrome
BAPTA	1,2-bis(o-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid
cCTnC	the C-domain of cardiac troponin C
CD	circular dichroism
cNTnC	the N-domain of cardiac troponin C
CSD	chemical shift deviation
DCM	dilated cardiomyopathy
E59D/D75Y	the E59D/D75Y mutant of the N-domain of cardiac troponin C
FHC	familial hypertrophic cardiomyopathy
FID	Free induction decay
G159D	the G159D mutant of the C-domain of cardiac troponin C
HSQC	heteronuclear single quantum coherence
IAANS	2-(4'-(2''-iodoacetamido)phenyl)aminonaphthalene-6-sulfonic acid
INEPT	insensitive nuclei enhancement by polarization transfer
ITC	isothermal titration calorimetry
L29Q	the L29Q mutant of the N-domain of cardiac troponin C
MyBPC	myosin binding protein C
NCX	Na <sup>+</sup> /Ca <sup>2+</sup> exchanger
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	Nuclear Overhauser Enhancement Spectroscopy
PENCE	Protein Engineering Network of Centres of Excellence
PKA	protein kinase A
PKC	protein kinase C
PKD	protein kinase D
PKG	protein kinase G
PLB	Phospholamban
RCM	restricted cardiomyopathy
RyR	ryanodine receptor
SDF-1 $\alpha$	stromal derived factor

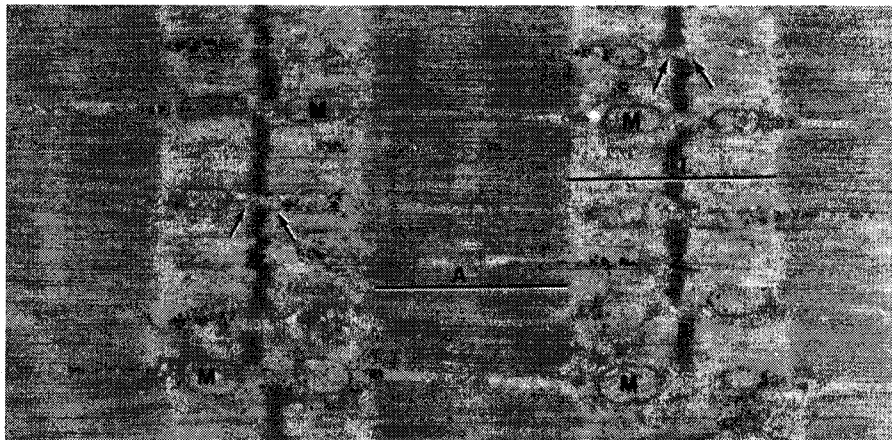
SERCA	sarco/endoplasmic reticulum $\text{Ca}^{2+}$ -ATPase
ssTnI (sTnI)	slow skeletal isoform of troponin I
Tn (cTn)	troponin complex (cardiac troponin complex)
TnC (cTnC)	troponin C (cardiac troponin C)
TnI (cTnI)	troponin I (cardiac troponin I)
TnT (cTnT)	troponin T (cardiac troponin T)
TOCSY	total correlation spectroscopy

## CHAPTER I INTRODUCTION

### *Cardiac muscle contraction: overall picture*

The contractile unit in striated muscle consists of thick and thin filaments sliding along each other during contraction, as first proposed in 1954 (1, 2). The main proteins involved in this process are actin and tropomyosin, composing the thin filament, and myosin, composing the thick filament. The troponin complex was identified in 1963 as a 'protein factor' required to reproduce 'the natural actomyosin' from its parts (3). According to the current knowledge, a dozen of other proteins, such as titin, desmin, dystrophin, and nebulin are also important participants (4, 5). Muscle is organized as repeating dark (A-band) and light (I-band) zones, with the Z-discs located in the middle of the I-band (Fig. I-1). A sarcomere, the unit of contraction, has one dark A-band, the half of the light I-band at each end, and is limited by Z-disks.

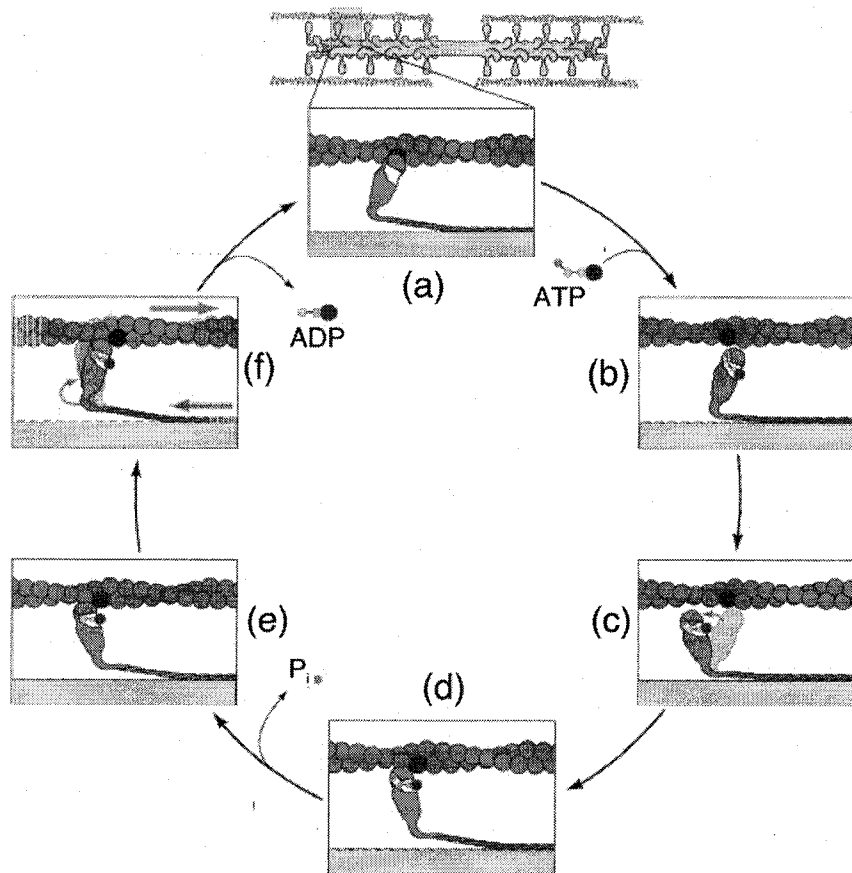
**Figure I-1.** Sarcomere appearance in electron microscope, where the A-, I-, and Z-zones are indicated (taken from <http://www.apsu.edu/thompsonj/Physiology/>). M is mitochondria.



Thick filaments consist of myosin, which has a long rod-like domain connected to the head. Myosin heads bind to actin of the thin filament and hydrolyze ATP to promote the power stroke, leading to the pulling of actin ~10 nm toward the center of a sarcomere. According to the current understanding (6), at the beginning of the cycle, the myosin head is tightly bound to actin (Fig. I-2a), which dissociates from its actin site after binding ATP (Fig. I-2b). Upon hydrolysis of ATP, myosin undergoes a large conformational change, altering the angle between the myosin head and the thick

filament (Fig. I-2c). This allows myosin to bind weakly to another actin binding site (Fig. I-2d), which strengthens after the release of inorganic phosphate (Fig. I-2e), promoting the pulling of the thin filament (Fig. I-2f). Release of ADP results in a power stroke, upon completion of which, the myosin head is tightly bound to actin as at the beginning of the cycle (Fig. I-2a).

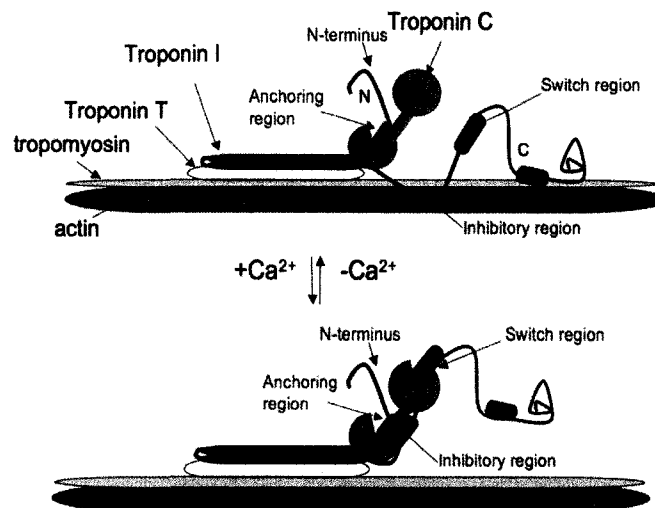
**Figure I-2.** Mechanism of force generation in muscle, with myosin heads interacting with actin of the thin filament (adopted from (7)). Different stages are shown as explained in the text.



Myosin heads bind actin and perform the work in the open state of the thin filament. In the blocked state, myosin sites on actin are inhibited. The inhibition of actin is relieved upon  $Ca^{2+}$  stimulation, which is sensed by the cardiac troponin complex (cTn). cTn consists of troponin T (cTnT) that anchors cTn to the thin filament via tropomyosin; troponin I (cTnI) that inhibits the interaction between myosin and actin; and troponin C (cTnC) that binds  $Ca^{2+}$  and triggers the switch from the relaxed to the active state (Fig. I-

3). cTnC is a dumbbell shaped molecule with the structural C-domain (cCTnC) linked to the regulatory N-domain (cNTnC) through a stretch of 11 residues. The C-domain has a higher affinity for  $\text{Ca}^{2+}$  and stays saturated with  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$  during the contracting-relaxing cycle. It is thought to be bound at all times to the anchoring region of cTnI (~ residues 34-71), which tethers cTnC to the thin filament. The N-domain has a lower affinity for  $\text{Ca}^{2+}$  and its  $\text{Ca}^{2+}$  binding site is occupied only when the concentration of  $\text{Ca}^{2+}$  is relatively high. This happens upon the release of  $\text{Ca}^{2+}$  from the sarcoplasmic reticulum after the stimulation. Upon  $\text{Ca}^{2+}$  binding, cNTnC binds to the switch region of cTnI (~ residues 147-163), which pulls the inhibitory region of cTnI (~ residues 128-147) away from its binding site on actin resulting in the movement of tropomyosin and the exposure of myosin binding sites on actin (Fig. I-3). This model of  $\text{Ca}^{2+}$  regulation was formulated based on numerous biochemical, biophysical, and physiological studies (8-12) including high resolution structures of cardiac (13) and skeletal troponin complexes (14). New reports have emerged that tackle other questions clarifying the position of the troponin complex on the actin fiber (15), the role of globular C-domain of cTnI (16), and the mechanism of signal transduction from troponin to tropomyosin (16), developing a more complete picture of muscle contraction.

**Figure I-3.** Model of  $\text{Ca}^{2+}$ -dependent regulation of cardiac muscle contraction. Figure adapted from (11, 13). cTnC is depicted in green, cTnT in white, cTnI in red, and actin and tropomyosin are in grey.



### *Ca<sup>2+</sup> cycling in cardiomyocytes*

The stimulation initiating cardiac contraction starts with the action potential generated by depolarizing of the sarcolemma by the voltage gated L-type Ca<sup>2+</sup> channel (17-20). Ca<sup>2+</sup> entry is sensed by the sarcoplasmic Ca<sup>2+</sup> receptor, called ryanodine receptor (RyR). RyR mediates a process of Ca<sup>2+</sup>-induced Ca<sup>2+</sup> release from the sarcoplasm. Upon the release from the sarcoplasm, Ca<sup>2+</sup> binds to cTnC of the thin filament promoting the contraction of muscle. Removal of Ca<sup>2+</sup> from the cytosol is achieved by several mechanisms. The first mechanism is the sarcoplasmic uptake by the Ca<sup>2+</sup> transporter (SERCA), which is regulated by phospholamban (PLB). The second mechanism of Ca<sup>2+</sup> removal is the extrusion through the sarcolemmal Na<sup>+</sup>/Ca<sup>2+</sup> exchanger (NCX). During a Ca<sup>2+</sup> cycle, the intracellular concentration of Ca<sup>2+</sup> rises from 0.1 to > 1  $\mu$ M, while the local concentration of Ca<sup>2+</sup> in the proximity to a cluster of RyR's can reach > 100  $\mu$ M (20-22). Ca<sup>2+</sup> cycles are typically completed within ~ 400-600 ms (21, 23). Interestingly, different kinds of myocytes have different Ca<sup>2+</sup> cycling patterns. For example, in ventricular myocytes local Ca<sup>2+</sup> sparks give rise to a synchronized global Ca<sup>2+</sup> signal due to the extensive network of T-tubules and sarcoplasmic invaginations delivering Ca<sup>2+</sup> to the sarcomeres. In atrial myocytes, Ca<sup>2+</sup> concentration sharply declines with the distance away from the sarcolemma due to the lack of inward spreading (18).

### *Regulation of Ca<sup>2+</sup> regulated contraction by phosphorylation*

In order to meet various needs of an organism, contractile machinery is fine tuned by the phosphorylation of its composing parts. The stimulation of  $\beta$ -adrenergic receptor by norepinephrine results in the generation of cAMP leading to the activation of downstream protein kinase A (PKA). Activated PKA phosphorylates several proteins central to the regulation of contractility. These proteins include the sarcolemmal Ca<sup>2+</sup> channel, RyR, PLB, myosin binding protein C (MyBPC) and cTnI (24, 25). As a result of phosphorylation, the current through the sarcolemmal Ca<sup>2+</sup> channel, the rate of Ca<sup>2+</sup> reuptake by SERCA, and the rate of Ca<sup>2+</sup> cycling are increased, attributed to the phosphorylation of the Ca<sup>2+</sup> cycling machinery. The relaxation rate of the muscle is enhanced and the Ca<sup>2+</sup> sensitivity of actomyosin ATPase is decreased, attributed to the phosphorylation of cTnI (26). The return to the baseline contractility is associated with the phosphorylation of PLB (27). The increase in cross-bridge cycling rate and the change in stretch-activation was shown to result from the phosphorylation of MyBPC (28). The overall consequence of  $\beta$ -adrenergic stimulation is the ~20% increase in power

output, due to the speeding up of the steps in  $\text{Ca}^{2+}$  and cross-bridge cycling (29). In spite of many efforts, the relative contribution of individual phosphorylated proteins to the increase in power output remains a matter of controversy.

The other mechanism for regulation of cardiac contraction is through the action of PKC, which phosphorylates cTnI and cTnT. At least 12 different PKC isoforms have been identified (25). The resultant contractility depends on many factors and has a wide range of effects on muscle function. Although the precise effects of PKC-dependent phosphorylation remain to be fully understood, there is an agreement that PKC phosphorylation of cTnI results in the decrease of  $\text{Ca}^{2+}$  activated tension and in sliding velocity. Other kinases, such as recently found PKG and PKD (25, 30), and their action on the sarcomeric proteins are beyond the scope of this introduction.

#### *Phosphorylation of cTnI*

PKA phosphorylates cTnI at residues Ser<sup>22</sup> and Ser<sup>23</sup> located in the cardiac specific N-terminus of cTnI (~ residues 1-29) adjacent to the anchoring region of cTnI (Fig. I-3). This results in the decrease of  $\text{Ca}^{2+}$  sensitivity of ATPase (31-33) and the increase of  $\text{Ca}^{2+}$  off-rates for the Tn complex (34). The N-terminus of cTnI has been cross-linked to cTnT (35); however, this interaction has not been visualized in the crystal structures (13) or in NMR studies (36-38), except for the computational models (39, 40). Attempts have been made to obtain the binding constant using ITC (35) and fluorescent titrations (37, 41, 42); however, experiments were performed with long constructs of cTnI and therefore represent the complex function of several binding events. Available experimental data provides little insight on the molecular details of cTnT binding to the N-terminus of cTnI, but suggests that the interaction of cTnT with the N-terminus of cTnI (~ residues 1-29) affects the function of cTnT and thus confer the signaling by PKA.

PKC phosphorylates cTnI at residues Ser<sup>43</sup>, Ser<sup>45</sup>, and Thr<sup>144</sup>. First two residues are located at the anchoring region of cTnI (Fig. I-3), the affinity of which to cTnT is expected to be altered by the presence of phosphorylation, but the binding was shown to be indistinguishable in NMR studies (43). The phosphorylation of Ser<sup>43</sup> and Ser<sup>45</sup> was attributed to the decrease in maximal  $\text{Ca}^{2+}$  activated tension (44). Thr<sup>144</sup> is located at the inhibitory region of cTnI, and its phosphorylation was shown to decrease the affinity of cTnT towards the cTnI<sub>128-147</sub> (45), which is in agreement with the increased inhibition of acto-myosin and the decrease in sliding velocity (44). PKC can also phosphorylate Ser<sup>22</sup>

and Ser<sup>23</sup>, suggesting the interdependence between PKA- and PKC- mediated effects and the complexity of the cTnI regulation by phosphorylation (25).

#### *Mutations in sarcomeric proteins that cause cardiomyopathies: overview*

Since the contractile machinery is composed of an array of proteins, with each of them playing a specific role, the number of vulnerable spots that can impair contractility is also significant. One of the leading genetic disorders of the heart, cardiomyopathy, can stem from mutations in proteins constructing the sarcomere, in proteins that align filaments inside myocytes (cytoskeletal proteins), and in the Ca<sup>2+</sup> cycling machinery (46-49). Cardiomyopathies are divided into *dilated cardiomyopathy*, DCM, characterized by the enlarged ventricular chamber, weakened heart wall, and impaired systolic function; *hypertrophic cardiomyopathy*, FHC (an abbreviation for Familial Hypertrophic Cardiomyopathy), characterized by the thickened ventricular muscle, reduced volume of the ventricle, and impaired diastolic function (48, 50); and RCM, *restricted cardiomyopathy*, the least common type, characterized by the restrictive diastolic dysfunction. The prevalence of cardiomyopathy mutations in the general population is high (~0.2% for FHC for young adults 18 to 30 years of age (51)). Both, FHC and DCM, can result in sudden cardiac death especially in younger FHC patients (mean age, 45 ± 20 years) (52). At present, the following sarcomeric genes have been linked to cardiomyopathies: α-actin (53), α- and β-tropomyosin (54), troponin (55, 56), α- and β-myosin heavy chain (57, 58), regulatory and essential myosin light chains (59), and myosin binding protein C (60), with the total number of mutations exceeding 450 (61). Although unifying hypotheses are present, such as the inefficient utilization of ATP in FHC (62) or the reduced force development in DCM (47, 63), the majority of data suggests that cardiomyopathy is represented by a group of genetically distinct disorders that should be approached individually.

#### *Cardiomyopathy mutations in cTn*

Cardiomyopathy mutations were found in all three subunits of cTn (cTnC, cTnI, and cTnT) (55, 56), with the majority of mutations found in cTnI and cTnT (Table I-1). The structures and functions of cTnC and cTnI have been studied the most thoroughly, which allows for the assessment of a mutation in terms of specific molecular interactions, possibly perturbed by the mutation (Table I-1). In many cases, especially for cTnT, the



functions are overlapping. For example, a long stretch of cTnT is thought to interact with cTnC, actin, cTnI, and tropomyosin (64-66). In cTnI and cTnC, mutations are found throughout the sequence including all functional regions, and sometimes are in the regions, functions of which are unknown, suggesting the deficiencies in the current model of cardiac contraction. Physiological data, including motility assays, maximal force generated, and the  $\text{Ca}^{2+}$  sensitivity of force development and ATPase efficiency are available for a number of mutations (for reviews (55, 56, 66, 67)); however, the structural information pertaining to a particular perturbed interaction is scarce. Structural studies are complicated due to the location of many mutations in intrinsically unstructured regions; for example, the inhibitory region of cTnI, and the N- and C-termini of cTnI.

Often the linkage analysis is needed to establish the connection of a mutation with the development of the disease. For mutations in cTnC studied in this work, only one mutation, G159D, was subjected to the analysis of inheritance (63). Two other mutations, L29Q and E59D/D75Y, were found in single patients and suspected to cause cardiomyopathy, however the possibility of polymorphism remains (68, 69).

#### *Biophysical techniques available for studying cardiomyopathies*

To elucidate the mechanisms by which the mutations in sarcomeric proteins bring about the symptoms of cardiomyopathy, several biophysical techniques have proved to be useful. The study of motility was one of the first methods used to address this issue (59). In this type of assay, myosin is deposited onto a nitrocellulose platform and the translocation rate of fluorescently labeled actin filaments is recorded. In another method, the ATPase activity of myosin is measured in an assay performed on the mixture of F-actin, tropomyosin, myosin, and troponin complex (70, 71). After incubation with ATP the reaction is terminated and the released inorganic phosphate is measured. The efficiency of ATP hydrolysis is a measure of the myosin ATPase activity. Force development is another powerful technique available for probing the functional consequences of mutations in sarcomeric proteins (72, 73). In this assay, strips of cardiac muscle are mounted on a force transducer and immersed alternatively in relaxation (low  $\text{Ca}^{2+}$ ) and contraction (high  $\text{Ca}^{2+}$ ) solutions. The force generated during contraction is then measured. The advantage of this method is that soluble muscle proteins such as troponin can be stripped off the muscle and substituted by their recombinant analogues containing mutations or other modifications.

**Table I-1.** Cardiomyopathy mutations in human cTnT, cTnI, and cTnC, their location, and physiological consequences (55, 56, 66, 67, 74).

Gene	Mutation	Location	Binding partners of the particular region	Phenotype
cTnT	A28V, S69F, F70L, P77L, I79N, E83K, D86A, R92W, R92L, R92Q, R94C, R94L, K97N, A104V, F110I, F110L, F110V, R130C	T1 domain	Tropomyosin	HCM
	DE160, E163K, E163R, S179F, DK210, E244D, K247E, N271I, K273E, R278C, R278P, R286C, R286H	T2 domain	cTnI, tropomyosin, cTnC, and actin	
	R141W	T1 domain	Tropomyosin	DCM
	A172S	T2 domain	cTnI, tropomyosin, cTnC, and actin	
cTnI	R21C	N-terminus	cNTnC	HCM
	P82S			
	Q130R, R141Q, R145G, R145Q	inhibitory region	cCTnC and actin	
	A157V, R162Q, R162P, R162W, S166F	switch region	cNTnC	
	DK177, K183E, DK183, R186Q, I195M, D196N, S199N, G203R, G203S, R204C, K206Q	C-domain	Actin	
	A2V	N-terminus	cNTnC	DCM
	L144Q, R145W	inhibitory region	cCTnC and actin	RCM
	A171T	switch region	cNTnC	
	K178E, D190G, R192H	C-domain	Actin	
cTnC	A8V, L29Q, C84Y	N-domain	the switch region and the N-terminus of cTnI	HCM
	E134D, D145E	C-domain	the anchoring and inhibitory regions of cTnI, and cTnT	
	E59D/D75Y	N-domain	the switch region and the N-terminus of cTnI	DCM
	G159D	C-domain	the anchoring and inhibitory regions of cTnI, and cTnT	

Structural techniques such as X-ray crystallography and NMR spectroscopy, when applicable, have proved to be very useful in tackling various aspects of protein biochemistry; however, in the quest for understanding cardiomyopathies, structural studies are lagging behind. In several reports, FHC and DCM mutations were mapped on the available structures of myosin (75) and cTn (11, 63). Studies of the mutations in MyBPC using 1D NMR spectroscopy were also reported (76). In our laboratory 2D NMR spectroscopy was used to elucidate the effects of FHC mutations of cTnI (R144G,

S149P, R162W) on the binding affinity of cTnI towards cTnC (77). However, no high-resolution structure of the sarcomeric protein with cardiomyopathy mutations is currently available and no atomic details of perturbed protein-protein interactions are known.

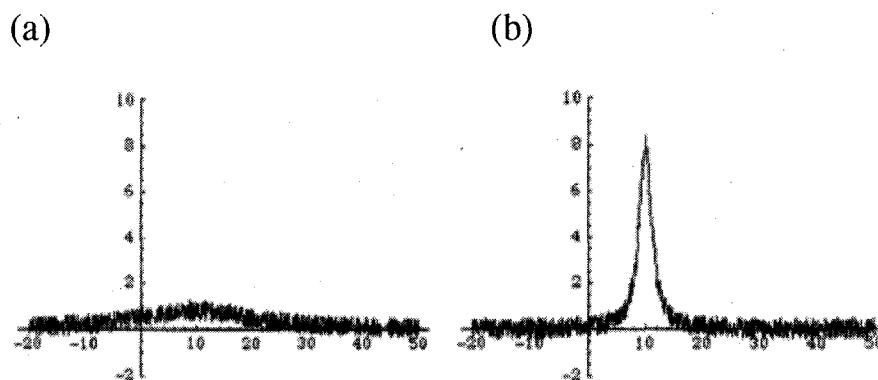
#### *Methodology employed in present work*

We have attempted to investigate the effect of cardiomyopathy mutations on the structure and function of cTnC using solution NMR spectroscopy. cTnC is a well behaved protein in NMR studies and the structure of cTnC has been detailed during years of research in Dr. Sykes' laboratory. We did expect that the cTnC mutants are also folded and well behaved proteins, so that their structures as well as other structural information could be obtained using NMR. In contrast, many mutations were found in cTnI, yet cTnI is a difficult target for traditional solution NMR due to its largely intrinsically disordered structure.

NMR is an extremely versatile technique that can provide a wealth of information about the system studied (78). Proteins (unlabeled or commonly labeled with  $^{15}\text{N}$  and/or  $^{13}\text{C}$  isotopes) are examined in solution at  $\sim$  mM concentrations, often together with their binding partners. The first very important parameter obtained by NMR is the chemical shift of individual nuclei (typically  $^1\text{H}$ ,  $^{15}\text{N}$ , or  $^{13}\text{C}$ ), which is a parameter related to the chemical structure of the molecule and its chemical environment. NMR chemical shifts are measured relative to a reference compound and are field dependant. Differences in chemical shifts between the wild type protein and the mutant or the protein complex indicate the change in chemical environment and can be used to monitor changes during the titrations. They can be also mapped on the 3D structures, showing the surface affected by the mutation or complex formation. Another parameter is NOE (nuclear Overhauser effect), which is a change in the intensity of NMR signal due to through-space interactions between the nuclei. This parameter is used to obtain the constraints for structure calculations from 2D/3D NOESY NMR spectra. In relaxation experiments, NOE provides the information on flexibility of nuclei. Other parameters are  $T_1$  and  $T_2$  relaxation rates, describing the rate of decay for the NMR signal observed during pulsed NMR experiments.  $T_1$  and  $T_2$  depend on the tumbling rate of the molecule, exchange processes, and dynamic behavior of the nuclei. The line width of the NMR peak at the half height is related to  $T_2$  ( $\Delta\nu_{1/2} = 1/2\pi T_2$ ), so that the line broadening in NMR experiment (sharpness of the peaks) often indicative of the formation of complexes or chemical exchange processes (Fig. I-4) as explained further. Other parameters

obtained by NMR, such as dipolar couplings, spin-spin couplings, etc, are useful in many experiments, however, were not employed in this work.

**Figure I-4.** Differences in line broadening for the sample containing the same amount of protein, where the line is broadened due to the aggregation (a), and where the line is narrow, in agreement with the protein molecular weight (b). Traces were taken from  $\{^{15}\text{N}, ^1\text{H}\}$  HSQC experiments, with the frequency in Hz on the x-axis and the intensity in arbitrary units on the y-axis.



The dependence of  $T_2$  on the molecular weight of a protein or its complex was often used in this work to indicate the increase in the molecular weight of protein species studied. When two predominant mechanisms for relaxation are dipole-dipole interaction between  $^{15}\text{N}$  and  $^1\text{H}$  nuclei and the chemical anisotropy of the  $^{15}\text{N}$  nucleus, the relaxation parameters  $T_2$ ,  $T_1$ , and NOE are given as following:

$$\frac{1}{T_1} = d^2 [J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N),$$

$$\frac{1}{T_2} = \frac{d^2}{8} [4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{6} [4J(0) + 3J(\omega_N)],$$

$$\text{NOE} = 1 + \frac{d^2}{4R_1} \frac{\gamma_H}{\gamma_N} [6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)].$$

where  $d = (\mu_o \hbar \gamma_H \gamma_N / 4\pi^2) r_{NH}^{-3}$ ,  $c = \omega_N \Delta\sigma / \sqrt{3}$ ,  $\mu_o$  is the permeability of free space,  $\hbar$  is Planck's constant,  $\gamma_H$  and  $\gamma_N$  are magnetogyric ratios of the  $^{15}\text{N}$  and  $^1\text{H}$  nuclei,  $r_{NH}$  is the  $^{15}\text{N} - ^1\text{H}$  bond length,  $\Delta\sigma$  is the chemical shift anisotropy of the  $^{15}\text{N}$  nucleus, and  $\omega_N$  is the Larmor frequency of the  $^{15}\text{N}$  nucleus (79). The spectral density function is described in Model Free approach as following:

$$J(\omega) = \frac{2}{5} \left( \frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} + \frac{(1 - S^2) \tau}{1 + \omega^2 \tau^2} \right),$$

where  $S^2$  is the generalized order parameter, and  $\frac{1}{\tau} = \left( \frac{1}{\tau_m} + \frac{1}{\tau_e} \right)$ , with  $\tau_e$  being the internal correlation time for molecular motions, and  $\tau_m$  is the overall correlation time of a protein (80). The spectral density function for a typical protein depends on the frequency as shown in Figure I-V, where  $J(0)$  (at zero) being an order of magnitude larger than  $J(\omega_N)$  (at the frequency of  $^{15}\text{N}$ ), and similarly larger than  $J(\omega_N + \omega_H)$ , and  $J(\omega_H - \omega_N)$ . This will allow for the simplification of the equation for  $T_2$ , but not for  $T_1$  and NOE.

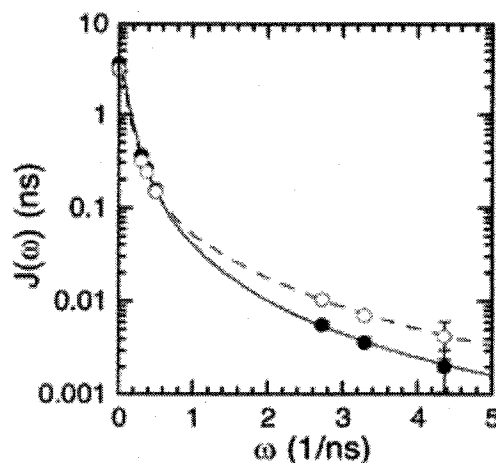
$$\frac{1}{T_2} \approx \frac{d^2}{8} [4J(0)] + \frac{c^2}{6} [4J(0)]$$

Assuming negligible internal motions the spectral density will be denoted as following:

$$J(\omega) = \frac{2}{5} \frac{S^2 \tau_m}{(1 + \omega^2 \tau_m^2)}.$$

At frequency zero, spectral density function  $J(0)$  has a linear dependence on the overall correlation time. Thus the relaxation rate  $T_2$  can be approximated to have a linear relationship with the overall correlation time as well, and consequently, with the molecular weight of the species.

**Figure I-V.** Spectral density function for backbone amide  $^{15}\text{N}$  spins in *E.coli* ribonuclease. Figure is taken from (79), where the data shows the values of spectral density at several frequencies, including those acquired at different magnetic fields.



### *Objectives and organization of this thesis*

The goal of this study was to address the structure and function of cardiomyopathy mutations in cTnC using solution state NMR spectroscopy. Chapter II describes the structural and functional aspects of the two mutants found in the N-domain of cTnC, L29Q and E59D/D75Y. We determined how these mutations affect the function of cTnC and proposed the mechanism explaining the hypercontractility in L29Q and hypocontractility in E59D/D75Y. Chapter III presents the structure and function of the mutation found in the C-domain of cTnC, G159D. Based on the results obtained we have developed the hypothesis of the mechanism of G159D and the development of disease. Effects of the L29Q and E59D/D75Y mutations described in Chapter II were observed in the presence of the N-terminus of cTnI (~ residue 1-29), and the effect of the G159D mutation described in Chapter III was suggested to be mediated through the N-terminus of cTnI. The interaction between the N-terminus of cTnI and cTnC as a function of ionic strength was explored in Chapter IV. Chapter V presents the pulse sequence, HBHG(CBCG)CO, developed to determine the individual side chain ionization constants for residues involved in the electrostatic interactions such as between cTnC and the N-terminus of cTnI. Chapter VI concludes this thesis with the future outlook for the field of cardiac muscle contraction. The following additional material is gathered in the Appendices: the HBHG(CBCG)CO pulse sequence code (Appendix A); the method for pH determination of the NMR samples using buffer components (Appendix B); the NMR structures of peptide agonists developed on the base of SDF-1 $\alpha$  (Appendix C); and the dimerization of SDF-1 $\alpha$  studied by the NMR relaxation methods (Appendix D). The latter two projects were accomplished during the first three years of my PhD career in Dr. Sykes' laboratory and were a part of a collaborative PENCE project, discontinued due to the demise of our key collaborator Dr. Ian Clark-Lewis (UBC). In order to communicate a streamlined and focused discussion we decided to build our narrative around the cardiomyopathy mutations in cTnC leaving the story of SDF-1 $\alpha$  in Appendices.

## References

- (1) Huxley, H., and Hanson, J. (1954) Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature* 173, 973-6.
- (2) Huxley, A. F., and Niedergerke, R. (1954) Structural changes in muscle during contraction; interference microscopy of living muscle fibres. *Nature* 173, 971-3.
- (3) Ebashi, S. (1963) Third Component Participating In The Superprecipitation Of 'Natural Actomyosin'. *Nature* 200, 1010.
- (4) Geeves, M. A., and Holmes, K. C. (1999) Structural mechanism of muscle contraction. *Annu Rev Biochem* 68, 687-728.
- (5) Clark, K. A., McElhinny, A. S., Beckerle, M. C., and Gregorio, C. C. (2002) Striated muscle cytoarchitecture: an intricate web of form and function. *Annu Rev Cell Dev Biol* 18, 637-706.
- (6) Rayment, I., Holden, H. M., Whittaker, M., Yohn, C. B., Lorenz, M., Holmes, K. C., and Milligan, R. A. (1993) Structure of the actin-myosin complex and its implications for muscle contraction. *Science* 261, 58-65.
- (7) Murphy, C. T., Rock, R. S., and Spudich, J. A. (2001) A myosin II mutation uncouples ATPase activity from motility and shortens step size. *Nat Cell Biol* 3, 311-5.
- (8) Sykes, B. D. (2003) Pulling the calcium trigger. *Nat Struct Biol* 10, 588-9.
- (9) Gordon, A. M., Homsher, E., and Regnier, M. (2000) Regulation of contraction in striated muscle. *Physiol Rev* 80, 853-924.
- (10) Slupsky, C. M., and Sykes, B. D. (1996) *Muscle Proteins*, John Wiley & Sons, Chichester, UK.
- (11) Li, M. X., Wang, X., and Sykes, B. D. (2004) Structural based insights into the role of troponin in cardiac muscle pathophysiology. *J Muscle Res Cell Motil* 25, 559-79.
- (12) Solaro, R. J., and Rarick, H. M. (1998) Troponin and tropomyosin: proteins that switch on and tune in the activity of cardiac myofilaments. *Circ Res* 83, 471-80.
- (13) Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. *Nature* 424, 35-41.
- (14) Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) Ca(2+)-regulated structural changes in troponin. *Proc Natl Acad Sci U S A* 102, 5038-43.
- (15) Ferguson, R. E., Sun, Y. B., Mercier, P., Brack, A. S., Sykes, B. D., Corrie, J. E., Trentham, D. R., and Irving, M. (2003) In situ orientations of protein domains: troponin C in skeletal muscle fibers. *Mol Cell* 11, 865-74.
- (16) Murakami, K., Yumoto, F., Ohki, S. Y., Yasunaga, T., Tanokura, M., and Wakabayashi, T. (2005) Structural basis for Ca<sup>2+</sup>-regulated muscle relaxation at interaction sites of troponin with actin and tropomyosin. *J Mol Biol* 352, 178-201.
- (17) Woodcock, E. A., and Matkovich, S. J. (2005) Cardiomyocytes structure, function and associated pathologies. *Int J Biochem Cell Biol* 37, 1746-51.
- (18) Bootman, M. D., Higazi, D. R., Coombes, S., and Roderick, H. L. (2006) Calcium signalling during excitation-contraction coupling in mammalian atrial myocytes. *J Cell Sci* 119, 3915-25.
- (19) Bers, D. M. (2008) Calcium cycling and signaling in cardiac myocytes. *Annu Rev Physiol* 70, 23-49.

- (20) Bers, D. M., and Guo, T. (2005) Calcium signaling in cardiac ventricular myocytes. *Ann N Y Acad Sci* 1047, 86-98.
- (21) Cheng, H., Lederer, M. R., Lederer, W. J., and Cannell, M. B. (1996) Calcium sparks and  $[Ca^{2+}]_i$  waves in cardiac myocytes. *Am J Physiol* 270, C148-59.
- (22) Takamatsu, T., and Wier, W. G. (1990) Calcium waves in mammalian heart: quantification of origin, magnitude, waveform, and velocity. *Faseb J* 4, 1519-25.
- (23) Williams, D. A., Delbridge, L. M., Cody, S. H., Harris, P. J., and Morgan, T. O. (1992) Spontaneous and propagated calcium release in isolated cardiac myocytes viewed by confocal microscopy. *Am J Physiol* 262, C731-42.
- (24) Metzger, J. M., and Westfall, M. V. (2004) Covalent and noncovalent modification of thin filament action: the essential role of troponin in cardiac muscle regulation. *Circ Res* 94, 146-58.
- (25) Layland, J., Solaro, R. J., and Shah, A. M. (2005) Regulation of cardiac contractile function by troponin I phosphorylation. *Cardiovasc Res* 66, 12-21.
- (26) Wolska, B. M., Arteaga, G. M., Pena, J. R., Nowak, G., Phillips, R. M., Sahai, S., de Tombe, P. P., Martin, A. F., Kranias, E. G., and Solaro, R. J. (2002) Expression of slow skeletal troponin I in hearts of phospholamban knockout mice alters the relaxant effect of beta-adrenergic stimulation. *Circ Res* 90, 882-8.
- (27) Garvey, J. L., Kranias, E. G., and Solaro, R. J. (1988) Phosphorylation of C-protein, troponin I and phospholamban in isolated rabbit hearts. *Biochem J* 249, 709-14.
- (28) Stelzer, J. E., Patel, J. R., Walker, J. W., and Moss, R. L. (2007) Differential roles of cardiac myosin-binding protein C and cardiac troponin I in the myofibrillar force responses to protein kinase A phosphorylation. *Circ Res* 101, 503-11.
- (29) Herron, T. J., Korte, F. S., and McDonald, K. S. (2001) Power output is increased after phosphorylation of myofibrillar proteins in rat skinned cardiac myocytes. *Circ Res* 89, 1184-90.
- (30) Avkiran, M., Rowland, A. J., Cuello, F., and Haworth, R. S. (2008) Protein kinase d in the cardiovascular system: emerging roles in health and disease. *Circ Res* 102, 157-63.
- (31) Holroyde, M. J., Howe, E., and Solaro, R. J. (1979) Modification of Calcium requirements for activation of cardiac myofibrillar ATPase by cyclic AMP dependent phosphorylation. *Biochim. Biophys. Acta* 586, 63-69.
- (32) Wattanapermpool, J., Guo, X., and Solaro, R. J. (1995) The unique amino-terminal peptide of cardiac troponin I regulates myofibrillar activity only when it is phosphorylated. *J Mol Cell Cardiol* 27, 1383-91.
- (33) Zhang, R., Zhao, J., and Potter, J. D. (1995) Phosphorylation of both serine residues in cardiac troponin I is required to decrease the  $Ca^{2+}$  affinity of cardiac troponin C. *J Biol Chem* 270, 30773-80.
- (34) 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  $Ca^{2+}$ -binding properties of the  $Ca^{2+}$ -regulatory site of bovine cardiac troponin. *J Biol Chem* 257, 260-3.
- (35) Ward, D. G., Brewer, S. M., Cornes, M. P., and Trayer, I. P. (2003) A cross-linking study of the N-terminal extension of human cardiac troponin I. *Biochemistry* 42, 10324-32.
- (36) Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) NMR analysis of cardiac troponin C-troponin I complexes: effects of phosphorylation. *FEBS Lett* 453, 107-12.



- (37) Abbott, M. B., Dong, W. J., Dvoretzky, A., DaGue, B., Caprioli, R. M., Cheung, H. C., and Rosevear, P. R. (2001) Modulation of cardiac troponin C-cardiac troponin I regulatory interactions by the amino-terminus of cardiac troponin I. *Biochemistry* 40, 5992-6001.
- (38) Ward, D. G., Brewer, S. M., Gallon, C. E., Gao, Y., Levine, B. A., and Trayer, I. P. (2004) NMR and mutagenesis studies on the phosphorylation region of human cardiac troponin I. *Biochemistry* 43, 5772-81.
- (39) Howarth, J. W., Meller, J., Solaro, R. J., Trewhella, J., and Rosevear, P. R. (2007) Phosphorylation-dependent conformational transition of the cardiac specific N-extension of troponin I in cardiac troponin. *J Mol Biol* 373, 706-22.
- (40) Solaro, R. J., Rosevear, P., and Kobayashi, T. (2008) The unique functions of cardiac troponin I in the control of cardiac muscle contraction and relaxation. *Biochem Biophys Res Commun* 369, 82-7.
- (41) Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) Effects of protein kinase A phosphorylation on signaling between cardiac troponin I and the N-terminal domain of cardiac troponin C. *Biochemistry* 36, 13305-11.
- (42) Dong, W. J., Chandra, M., Xing, J., She, M., Solaro, R. J., and Cheung, H. C. (1997) Phosphorylation-induced distance change in a cardiac muscle troponin I mutant. *Biochemistry* 36, 6754-61.
- (43) Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) Phosphorylation and mutation of human cardiac troponin I differentially destabilize the interaction of the functional regions of troponin I with troponin C. *Biochemistry* 42, 14460-8.
- (44) Burkart, E. M., Sumandea, M. P., Kobayashi, T., Nili, M., Martin, A. F., Homsher, E., and Solaro, R. J. (2003) Phosphorylation or glutamic acid substitution at protein kinase C sites on cardiac troponin I differentially depress myofilament tension and shortening velocity. *J Biol Chem* 278, 11265-72.
- (45) Lindhout, D. A., Li, M. X., Schieve, D., and Sykes, B. D. (2002) Effects of T142 phosphorylation and mutation R145G on the interaction of the inhibitory region of human cardiac troponin I with the C-domain of human cardiac troponin C. *Biochemistry* 41, 7267-74.
- (46) Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc Res* 44, 20-36.
- (47) Morita, H., Seidman, J., and Seidman, C. E. (2005) Genetic causes of human heart failure. *J Clin Invest* 115, 518-26.
- (48) Seidman, J. G., and Seidman, C. (2001) The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell* 104, 557-67.
- (49) Karkkainen, S., and Peuhkurinen, K. (2007) Genetics of dilated cardiomyopathy. *Ann Med* 39, 91-107.
- (50) Bruce, J. (2005) Getting to the heart of cardiomyopathies. *Nursing* 35, 44-7.
- (51) 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, 785-9.
- (52) Maron, B. J., Olivotto, I., Spirito, P., Casey, S. A., Bellone, P., Gohman, T. E., Graham, K. J., Burton, D. A., and Cecchi, F. (2000) Epidemiology of hypertrophic cardiomyopathy-related death: revisited in a large non-referral-based patient population. *Circulation* 102, 858-64.

- (53) 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, R39-43.
- (54) 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, 701-12.
- (55) Harada, K., and Morimoto, S. (2004) Inherited cardiomyopathies as a troponin disease. *Jpn J Physiol* 54, 307-18.
- (56) Gomes, A. V., and Potter, J. D. (2004) Molecular and cellular aspects of troponin cardiomyopathies. *Ann N Y Acad Sci* 1015, 214-24.
- (57) Tanigawa, G., Jarcho, J. A., Kass, S., Solomon, S. D., Vosberg, H. P., Seidman, J. G., and Seidman, C. E. (1990) A molecular basis for familial hypertrophic cardiomyopathy: an alpha/beta cardiac myosin heavy chain hybrid gene. *Cell* 62, 991-8.
- (58) Geisterfer-Lowrance, A. A., Christe, M., Conner, D. A., Ingwall, J. S., Schoen, F. J., Seidman, C. E., and Seidman, J. G. (1996) A mouse model of familial hypertrophic cardiomyopathy. *Science* 272, 731-4.
- (59) 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, 63-9.
- (60) Oakley, C. E., Hambly, B. D., Curmi, P. M., and Brown, L. J. (2004) Myosin binding protein C: structural abnormalities in familial hypertrophic cardiomyopathy. *Cell Res* 14, 95-110.
- (61) Alcalai, R., Seidman, J. G., and Seidman, C. E. (2008) Genetic basis of hypertrophic cardiomyopathy: from bench to the clinics. *J Cardiovasc Electrophysiol* 19, 104-10.
- (62) Ashrafian, H., Redwood, C., Blair, E., and Watkins, H. (2003) Hypertrophic cardiomyopathy: a paradigm for myocardial energy depletion. *Trends Genet* 19, 263-8.
- (63) Mirza, M., Marston, S., Willott, R., Ashley, C., Mogensen, J., McKenna, W., Robinson, P., Redwood, C., and Watkins, H. (2005) Dilated cardiomyopathy mutations in three thin filament regulatory proteins result in a common functional phenotype. *J Biol Chem* 280, 28498-506.
- (64) Perry, S. V. (1998) Troponin T: genetics, properties and function. *J Muscle Res Cell Motil* 19, 575-602.
- (65) Blumenschein, T. M., Tripet, B. P., Hodges, R. S., and Sykes, B. D. (2001) Mapping the interacting regions between troponins T and C. Binding of TnT and TnI peptides to TnC and NMR mapping of the TnT-binding site on TnC. *J Biol Chem* 276, 36606-12.
- (66) Knollmann, B. C., and Potter, J. D. (2001) Altered regulation of cardiac muscle contraction by troponin T mutations that cause familial hypertrophic cardiomyopathy. *Trends Cardiovasc Med* 11, 206-12.
- (67) Gomes, A. V., and Potter, J. D. (2004) Cellular and molecular aspects of familial hypertrophic cardiomyopathy caused by mutations in the cardiac troponin I gene. *Mol Cell Biochem* 263, 99-114.
- (68) 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.

- (69) Lim, C. C., Yang, H., Yang, M., Wang, C. K., Shi, J., Berg, E. A., Pimentel, D. R., Gwathmey, J. K., Hajjar, R. J., Helmes, M., Costello, C. E., Huo, S., and Liao, R. (2008) A Novel Mutant Cardiac Troponin C Disrupts Molecular Motions Critical For Calcium Binding Affinity And Cardiomyocyte Contractility. *Biophys J*.
- (70) Elliott, K., Watkins, H., and Redwood, C. S. (2000) Altered regulatory properties of human cardiac troponin I mutants that cause hypertrophic cardiomyopathy. *J Biol Chem* 275, 22069-74.
- (71) Szczesna, D., Zhang, R., Zhao, J., Jones, M., Guzman, G., and Potter, J. D. (2000) Altered regulation of cardiac muscle contraction by troponin T mutations that cause familial hypertrophic cardiomyopathy. *J Biol Chem* 275, 624-30.
- (72) Gomes, A. V., Venkatraman, G., Davis, J. P., Tikunova, S. B., Engel, P., Solaro, R. J., and Potter, J. D. (2004) Cardiac troponin T isoforms affect the Ca(2+) sensitivity of force development in the presence of slow skeletal troponin I: insights into the role of troponin T isoforms in the fetal heart. *J Biol Chem* 279, 49579-87.
- (73) Hernandez, O., Szczesna-Cordary, D., Knollmann, B. C., Miller, T., Bell, M., Zhao, J., Sirenko, S. G., Diaz, Z., Guzman, G., Xu, Y., Wang, Y., Kerrick, W. G., and Potter, J. D. (2005) F110I and R278C troponin T mutations that cause familial hypertrophic cardiomyopathy affect muscle contraction in transgenic mice and reconstituted human cardiac fibers. *J Biol Chem*.
- (74) Parvatiyar, M., Pinto, J. R., Jones, M. A., Ackerman, M. J., and Potter, J. D. (2008) in *Biophysical Society, Annual Meeting*, Long Beach, CA.
- (75) Rayment, I., Holden, H. M., Sellers, J. R., Fananapazir, L., and Epstein, N. D. (1995) Structural interpretation of the mutations in the beta-cardiac myosin that have been implicated in familial hypertrophic cardiomyopathy. *Proc Natl Acad Sci U S A* 92, 3864-8.
- (76) 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, 745-61.
- (77) Lindhout, D. A., Boyko, R. F., Corson, D. C., Li, M. X., and Sykes, B. D. (2005) The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C. *Biochemistry* 44, 14750-9.
- (78) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
- (79) Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 104, 3623-40.
- (80) 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.

## CHAPTER II

### The effect of L29Q and E59D/D75Y cardiomyopathy mutations on the function of cNTnC\*

#### Outline

The cardiac specific N-terminus of cTnI is known to modulate the activity of troponin upon phosphorylation with PKA by decreasing its  $\text{Ca}^{2+}$  affinity and increasing the relaxation rate of the thin filament. The molecular details of this modulation have not been elaborated to date. We have established that the N-terminus (cTnI<sub>1-29</sub>/cTnI<sub>1-29</sub>PP) and the switch region of cTnI (cTnI<sub>147-163</sub>) bind to cNTnC simultaneously, and that the PKA signal is transferred via the cTnI N-terminus modulating the cNTnC affinity towards cTnI<sub>147-163</sub> but not towards  $\text{Ca}^{2+}$ . The  $K_D$  of cNTnC for cTnI<sub>147-163</sub> was found to be 600  $\mu\text{M}$  in the presence of cTnI<sub>1-29</sub> and 370  $\mu\text{M}$  in the presence of cTnI<sub>1-29</sub>PP, which can explain the difference in muscle relaxation rates upon the phosphorylation with PKA in experiments with cardiac fibers. In the light of newly found mutations in cNTnC, associated with cardiomyopathies, the role of the cTnI N-terminus is emerging also to be important in the development of heart disorders. The mutants studied, L29Q and E59D/D75Y, demonstrated unchanged  $\text{Ca}^{2+}$  affinity per se and in complex with the cTnI N-terminus (cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP). The affinity of L29Q and E59D/D75Y towards cTnI<sub>147-163</sub> was significantly perturbed, both alone and in complex with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP, which is likely to be responsible for the development of malfunctions.

#### Introduction

Recent studies suggested the involvement of the N-terminal cTnI region (Fig. I-3) in the development of cardiomyopathies, in which the mutations in cTnC perturb interactions with the N-terminus of cTnI, giving rise to the malfunction of the heart (1, 2). One of these cNTnC mutations, L29Q, has been found in a patient with hypertrophy of the heart (3) suggesting it to be a cause for FHC. L29Q is located on silent  $\text{Ca}^{2+}$  binding site I of cNTnC and is potentially capable of changing the  $\text{Ca}^{2+}$  affinity by reviving defunct site I. This hypothesis is especially attractive since the L29Q substitution was

---

\* The version of this chapter has been published. Baryshnikova O.K., Li M.X., and Sykes B.D. (2008) *J Mol Biol* 375, 735-751.

found in cold water fish where it was proposed to increase the  $\text{Ca}^{2+}$  sensitivity of the cardiac muscle at low temperatures (4). It is also possible that L29Q affects the interaction with the switch region of cTnI (~residues 147-163) or, as suggested in literature, the interaction with the N-terminal region of cTnI (~residues 1-29) (Fig. I-3) (1). The latter will in turn affect the way cTnI<sub>1-29</sub> modulates the  $\text{Ca}^{2+}$  affinity of cTn, which might be responsible for the development of FHC.

The other cNTnC mutation mentioned in literature, E59D/D75Y, was connected to a dilated cardiomyopathy, DCM (5, 6). For this mutant, the change in  $\text{Ca}^{2+}$  affinity can be expected since residue 75 is mapped on the loop corresponding to functional  $\text{Ca}^{2+}$  binding site II. Residue 59 is located in close proximity to the bound cTnI<sub>147-163</sub> peptide (7), which if mutated might also affect the affinity towards the switch region of cTnI (~residues 147-163). In addition, the E59D/D75Y mutation might affect the binding of the cTnI N-terminus (~residues 1-29) and perturb its modulation of cNTnC interactions (Fig. I-3), leading to the development of DCM.

Since the molecular details of the effect of the cTnI N-terminus on the interactions of cNTnC have not been fully elaborated to date (8), we have addressed these questions. The affinities of cNTnC towards  $\text{Ca}^{2+}$  and cTnI<sub>147-163</sub> were measured in the presence of the phosphorylated and unphosphorylated cTnI peptides (cTnI<sub>1-29</sub>PP and cTnI<sub>1-29</sub>), where the minimal length of the peptides required to confer the modulation by PKA was determined previously (9). The nature of the binding site for cTnI<sub>1-29</sub> on cNTnC, the simultaneous binding of the cTnI N-terminus and the switch region, and the possible mechanisms by which the phosphorylation signal is transferred within cTn are discussed. Second, we have examined how the presence of mutations in cNTnC, L29Q and E59D/D75Y, affects the  $\text{Ca}^{2+}$  and cTnI<sub>147-163</sub> binding affinities per se and in the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP.

## Results

### *Titration of L29Q and E59D/D75Y with $\text{Ca}^{2+}$ using NMR spectroscopy*

In a typical 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectrum, the observed cross-peaks correlate the chemical shifts of backbone amide  $^{15}\text{N}$  to the chemical shifts of backbone amide  $^1\text{H}$ . Backbone amide chemical shifts are exquisitely sensitive to the changes in structure and dynamics of a protein. Therefore, a 2D NMR spectrum provides a high resolution residue specific probe to follow changes in a protein; for example, upon binding a ligand. This method has been called chemical shift mapping. The 2D  $\{^1\text{H}, ^{15}\text{N}\}$

HSQC NMR spectra of L29Q and E59D/D75Y were very similar to the spectrum of cNTnC, implying that both mutant proteins have a similar structure to the native fold. This is also supported by the  $^{15}\text{N}$  backbone amide NMR relaxation measurements discussed below.

As it can be seen from the overlay of 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC spectra obtained during  $\text{Ca}^{2+}$  titrations,  $\text{Ca}^{2+}$  binding to L29Q (Fig. II-1a) and E59D/D75Y (Fig. II-1b) causes gradual chemical shift changes that were very similar to each other and to the changes induced upon the  $\text{Ca}^{2+}$  binding to cNTnC (7), indicating binding to the same site and in a similar fashion. Several cross-peaks in Figure II-1 are labeled based on the assignments for cNTnC. In both cases, proportional movement of cross-peaks implies that the  $\text{Ca}^{2+}$  binding kinetics is in fast exchange on the NMR time scale. Linear movement of cross-peaks in titrations suggests that a single binding event is involved. Multiple  $\text{Ca}^{2+}$  binding sites would cause non-linear chemical shift changes. The movements of cross-peaks in Figure II-1 are primarily linear. The possibility for the weak second binding site, suggested for L29Q on the basis of comparison with cold water fish cTnC (4), will be considered further below.

Binding curves for L29Q and for E59D/D75Y are shown in Figures II-1a and II-1b (insets) respectively. When the curves were fitted with the single binding site model, the dissociation constants,  $K_D$ 's, were found to be 1.5  $\mu\text{M}$  for L29Q and 2  $\mu\text{M}$  for E59D/D75Y (Table II-1). The upper boundary errors for the fitted  $K_D$ 's were 2  $\mu\text{M}$ , whereas the lower boundary errors could not be established since  $K_D$ 's tighter than 1  $\mu\text{M}$  cannot be discriminated within the current concentration conditions. When a second binding site was taken into account in the case of L29Q, the quality of fit was not noticeably improved as the chemical shifts almost completely stopped changing at the  $\text{Ca}^{2+}$  : protein ratio equal to 1. Within the limitations of this method, we assume that the  $K_D$ 's determined ( $\sim 2$   $\mu\text{M}$  for E59D/D75Y and  $\sim 1.5$   $\mu\text{M}$  for L29Q) were essentially the same as the  $K_D$  of 2.6  $\mu\text{M}$  for cNTnC (10), with no indication of a weak second  $\text{Ca}^{2+}$  binding site in case of L29Q.

*$\text{Ca}^{2+}$  affinity towards cNTnC, L29Q, and E59D/D75Y determined using competition assays with the chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA*

The NMR method for determining affinities is most accurate in 0.1-1.0 mM range, where  $K_D$ 's and the concentrations of protein are of the same order of magnitude. To measure high affinity  $K_D$ 's more accurately, we used competition assays with the

chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA. In this type of assays, the concentrations of proteins were ~20-30  $\mu$ M. The chelator, 5,5'-Br<sub>2</sub>-BAPTA, was chosen because its  $K_D$  for Ca<sup>2+</sup> is close to the expected  $K_D$ 's for the proteins of interest (11). Titration data for cNTnC, L29Q, and E59D/D75Y obtained in competitive assays are presented in Figure II-2. Using a single site binding model for curve fitting, determined  $K_D$ 's for cNTnC, L29Q, and E59D/D75Y were  $5 \pm 4$   $\mu$ M,  $4 \pm 2.5$   $\mu$ M, and  $6 \pm 3$   $\mu$ M respectively (Table II-1). The errors in fitted  $K_D$ 's reflect mostly the error in the  $K_D$  of Ca<sup>2+</sup> binding to 5,5'-Br<sub>2</sub>-BAPTA. Data obtained are in agreement with the results of NMR titrations, where the L29Q and E59D/D75Y mutants showed essentially the same Ca<sup>2+</sup> affinity as cNTnC within the limitation of the technique.

Since the L29Q substitution occurs naturally in trout cTnC, which possesses the second weak Ca<sup>2+</sup> binding site I (12), we attempted to fit the NMR data for Ca<sup>2+</sup> titration of L29Q by a two site binding model, using  $K_{D1}$  for site II (determined in competitive assays) as a fixed parameter. Three curves corresponding to  $K_{D2}$  of 5 mM, 0.5 mM, and 0.05 mM shown in Figure II-3 imply that  $K_{D2} \geq 5$  mM. Thus, if L29Q makes defunct site I slightly active, its affinity would be exceptionally low, on the order of 5 mM or higher.

It has been suggested that the Ca<sup>2+</sup> affinity of cTn is dependent on the presence of the cardiac specific N-terminus of cTnI in phosphorylated state (13). Since the affinities of cNTnC towards Ca<sup>2+</sup> in the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP have not been determined previously, we have performed Ca<sup>2+</sup> titrations of cNTnC in the complex with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP using competitive assays with 5,5'-Br<sub>2</sub>-BAPTA. Complexes formed with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP had the ratio of protein : peptide equal to 1:3. Figure II-4a demonstrates that the  $K_D$ 's towards Ca<sup>2+</sup> for cNTnC, cNTnC•cTnI<sub>1-29</sub>, and cNTnC•cTnI<sub>1-29</sub>PP were indistinguishable and equal to  $5 \pm 3$   $\mu$ M (Table II-1).

Similar experiments were performed with the L29Q and E59D/D75Y mutants in the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP. Figures II-4b and II-4c indicate that these mutations do not affect the Ca<sup>2+</sup> affinity of L29Q and E59D/D75Y in 1:3 complex with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP within the limitations of the method. Determined  $K_D$ 's for L29Q, L29Q•cTnI<sub>1-29</sub>, and L29Q•cTnI<sub>1-29</sub>PP were  $4 \pm 2$   $\mu$ M and  $K_D$ 's for E59D/D75Y, E59D/D75Y•cTnI<sub>1-29</sub>, and E59D/D75Y•cTnI<sub>1-29</sub>PP were  $6 \pm 3$   $\mu$ M (Table II-1).

#### *cTnI<sub>147-163</sub> titrations of L29Q•Ca<sup>2+</sup> and E59D/D75Y•Ca<sup>2+</sup>*

The switch region of cTnI (~residues 147-163) is an important binding partner of cNTnC, that binds in a Ca<sup>2+</sup> dependent manner to cNTnC, relays the Ca<sup>2+</sup> signal, and

relieves the inhibition of the actomyosin ATPase. To address the question whether the cardiomyopathy mutations confer the changes in the cTnI<sub>147-163</sub> binding, L29Q•Ca<sup>2+</sup> and E59D/D75Y•Ca<sup>2+</sup> were titrated with the cTnI<sub>147-163</sub> peptide using 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy. The overlay of 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectra obtained during titrations (Fig. II-5) demonstrated that induced chemical shift changes were remarkably similar to those observed for cTnC•Ca<sup>2+</sup> (7) indicating that cTnI<sub>147-163</sub> binds in a similar fashion to the same binding site. The per residue chemical shift changes caused by the cTnI<sub>147-163</sub> binding to L29Q•Ca<sup>2+</sup> and E59D/D75Y•Ca<sup>2+</sup> were averaged for several of the most perturbed residues and plotted against the peptide/protein ratio (Figs. II-5a and II-5b). The value of K<sub>D</sub> obtained for L29Q•Ca<sup>2+</sup> was 200 ± 20 μM, which is ~1.3-fold weaker than the value of 154 μM determined for cTnC•Ca<sup>2+</sup> earlier (7) (Table I-1). The value of K<sub>D</sub> obtained for E59D/D75Y•Ca<sup>2+</sup> was 60 ± 40 μM, which is ~2.6-fold tighter than that of for cTnC•Ca<sup>2+</sup> (Table I-1).

*cTnI<sub>147-163</sub> titrations of cTnC•Ca<sup>2+</sup>, L29Q•Ca<sup>2+</sup>, and E59D/D75Y•Ca<sup>2+</sup> in complex with cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP*

The question whether the cardiac specific extension of cTnI in its phosphorylated or unphosphorylated state affects the affinity of cTnC•Ca<sup>2+</sup> towards cTnI<sub>147-163</sub> was addressed using 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy. Residues that underwent chemical shift changes during the binding of cTnI<sub>147-163</sub> to cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> and to cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP were essentially the same as that of for cTnC•Ca<sup>2+</sup> (Fig. II-6). This suggests that cTnI<sub>147-163</sub> binds in the same manner to the same binding site on cTnC•Ca<sup>2+</sup> alone and in complex with cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP. However, binding affinities determined for cTnC•Ca<sup>2+</sup> in the presence of cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP were substantially different from the K<sub>D</sub> of 154 μM determined for cTnC•Ca<sup>2+</sup> (7). The K<sub>D</sub> obtained for the cTnI<sub>147-163</sub> binding towards cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (Fig. II-7a and Table II-1) was equal to 600 ± 100 μM (~4 fold increase compared to cTnC•Ca<sup>2+</sup> alone) and the K<sub>D</sub> obtained for the cTnI<sub>147-163</sub> binding towards cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP was 370 ± 30 μM (~2.4 fold increase compared to cTnC•Ca<sup>2+</sup> alone).

It has been proposed that the L29Q mutation of troponin C perturbs the interaction with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP, and thus is responsible for the development of FHC (1). We have measured whether the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP affects the affinity towards cTnI<sub>147-163</sub> in case of L29Q and E59D/D75Y. 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR



spectroscopy was employed to monitor the changes in chemical shifts of L29Q•Ca<sup>2+</sup> and E59D/D75Y•Ca<sup>2+</sup> in complex with unlabeled cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP during titrations with the cTnI<sub>147-163</sub> peptide. The overlay of 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectra shows the similarity between cTnI<sub>147-163</sub> binding sites in case of L29Q•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>, L29Q•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP, E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>, and E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP (Figs. II-6c, II-6d, II-6e, and II-6f). In spite of these similarities, determined affinities were noticeably different (Table II-1). The K<sub>D</sub> obtained for cTnI<sub>147-163</sub> towards L29Q•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> was equal to 410 ± 30 μM and the K<sub>D</sub> obtained for cTnI<sub>147-163</sub> towards L29Q•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP was 450 ± 130 μM (Fig. II-7b). The K<sub>D</sub> obtained for cTnI<sub>147-163</sub> towards E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> was equal to 290 ± 100 μM and the K<sub>D</sub> obtained for cTnI<sub>147-163</sub> towards E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP was 830 ± 150 μM (Fig. II-7c).

**Table II-1.** Binding affinities towards Ca<sup>2+</sup> and cTnI<sub>147-163</sub> for cTnTnC, L29Q, and E59D/D75Y determined by solution NMR and in competitive assays with the chelator 5,5'-Br<sub>2</sub>-BAPTA.

	Ca <sup>2+</sup> affinity determined by solution NMR (μM)	Ca <sup>2+</sup> affinity determined using competitive assays with the chelator, 5,5'-Br <sub>2</sub> -BAPTA (μM)	Ca <sup>2+</sup> affinity determined using competitive assays with the chelator, 5,5'-Br <sub>2</sub> -BAPTA (μM) in a complex with	
			cTnI <sub>1-29</sub>	cTnI <sub>1-29</sub> PP
cTnTnC	2.6 (10)	5 ± 4	5 ± 3	5 ± 3
L29Q	1.5 ± 2	4 ± 2.5	4 ± 2	4 ± 2
E59D/D75Y	2 ± 2	6 ± 3	6 ± 3	6 ± 3
	cTnI <sub>147-163</sub> affinity determined using solution NMR (μM)		cTnI <sub>147-163</sub> affinity determined using solution NMR (μM) in a complex with	
			cTnI <sub>1-29</sub>	cTnI <sub>1-29</sub> PP
cTnTnC•Ca <sup>2+</sup>	154 (7)		600 ± 100	370 ± 30
L29Q•Ca <sup>2+</sup>	200 ± 20		410 ± 30	450 ± 130
E59D/D75Y•Ca <sup>2+</sup>	60 ± 40		290 ± 100	830 ± 150

#### <sup>15</sup>N backbone relaxation data

The <sup>15</sup>N backbone NMR relaxation parameters T<sub>1</sub>, T<sub>2</sub>, and NOE depend on the motions of a protein, including the tumbling of a protein as a whole and the motions of its

independent parts.  $T_1$ ,  $T_2$ , and NOE are the complex functions of protein motions; however, the interpretation of  $T_2$  is more straightforward since the major contribution to  $T_2$  is the tumbling of a protein as a whole, which makes  $T_2$  directly proportional to the overall correlation time,  $\tau_m$ , and therefore to the molecular weight of the species for a wide variety of conditions (14). The other significant contribution to  $T_2$  arises from the chemical exchange processes including self-association.  $T_2$  can be used to indicate the presence of chemical exchange or self-association if it is noticeably lower than the theoretical  $T_2$  values, calculated using Model Free approach (15), where  $S^2 \approx 0.85$  (the approximation valid for many proteins (16)) and the contribution of internal motions are considered negligible ( $\tau_f < 100 \mu s$ ) (17).  $T_1$ ,  $T_2$ , and NOE were determined for apo cNTnC, cNTnC•Ca<sup>2+</sup>, L29Q•Ca<sup>2+</sup>, and E59D/D75Y•Ca<sup>2+</sup> alone, and in the presence of cTnI<sub>147-163</sub>, cTnI<sub>1-29</sub>, and cTnI<sub>1-29</sub>PP (Table II-2, data for  $T_1$  and NOE are not shown). Experimental  $T_2$  values are mostly in the same range as the theoretical  $T_2$  values that are directly proportional to the molecular weight of the species (Table II-2), which indicates the absence of significant exchange processes or self-association. Care should be taken for E59D/D75Y, where  $T_2$  is ~20 ms lower in the case of the Ca<sup>2+</sup> and cTnI<sub>147-163</sub> bound forms. Line broadening observed during the cTnI<sub>147-163</sub> titration of E59D/D75Y•Ca<sup>2+</sup> (for ex. residues 34, 73, 66, Fig. II-5a) argues that the change in the cTnI<sub>147-163</sub> binding kinetics from fast to slow on NMR time scale is likely to be responsible for the decrease in  $T_2$ , which brings the contribution from the chemical exchange to the  $T_2$  relaxation rate. In case of the Ca<sup>2+</sup> bound form, the decrease in  $T_2$  can also result from the changes in Ca<sup>2+</sup> binding kinetics but more likely it stems from the slight dimerization, supported also by the increase in the overall correlation time,  $\tau_m$ , as discussed below.

The overall correlation time,  $\tau_m$ , obtained from relaxation parameters  $T_1$ ,  $T_2$ , and NOE (17, 18) often serves as a useful indicator of molecular weight of the species in solution (19). The increase in  $\tau_m$ , as compared to the theoretical estimate, equal to  $MW(kDa)/2$ , can sometimes result from the presence of mobile domains that slows down the tumbling of a protein or from the dimerization (20), in which case the Model Free analysis is not strictly suitable (21) unless modified for the presence of two species. Based on  $T_2$  data, the dimerization was absent for almost all species except for E59D/D75Y•Ca<sup>2+</sup>. Based on the similarity between {<sup>1</sup>H, <sup>15</sup>N} HSQC spectra (Figs. II-1, II-5, and II-6) for cNTnC, L29Q, and E59D/D75Y, we conclude that the proteins have a similar fold to that of cNTnC, for which no mobile domains or significantly long tails were found (7, 22). Thus, obtained  $\tau_m$ 's (Table II-3, Fig. II-8) serve as a direct indicator

of the increase in the molecular weight of species due to the formation of complexes with the cTnI peptides.

**Table II-2.**  $^{15}\text{N}$  backbone amide NMR relaxation parameter,  $T_2$ , for cTnTnC, L29Q, and E59D/D75Y in various complexes obtained at 500 MHz.  $T_2$ 's were averaged over all residues and presented together with their standard deviations. The errors in  $T_2$ 's represent the quality of the fit. Theoretical  $T_2$ 's values were calculated for a protein of the same molecular weight using Model Free formalism, where  $S^2$  is equal to 0.85 and internal motions are considered negligible (15).

	Average $T_2$ (ms)	$T_2$ error (ms)	Theoretical $T_2$ (ms)
apo cTnTnC	$134 \pm 31$	4	149
apo cTnTnC•cTnI <sub>1-29</sub>	$113 \pm 19$	3	122
apo cTnTnC•cTnI <sub>1-29</sub> PP	$137 \pm 48$	3	121
cTnTnC•Ca <sup>2+</sup>	$142 \pm 52$	5	149
L29Q•Ca <sup>2+</sup>	$139 \pm 15$	4	149
E59D/D75Y•Ca <sup>2+</sup>	$130 \pm 46$	4	149
cTnTnC•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	$138 \pm 27$	8	132
L29Q•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	$123 \pm 49$	4	132
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	$111 \pm 47$	2	132
cTnTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	$118 \pm 39$	5	122
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	$128 \pm 51$	4	122
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	$114 \pm 47$	4	122
cTnTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	$139 \pm 57$	5	121
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	$148 \pm 55$	4	121
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	$120 \pm 45$	5	121
cTnTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	$102 \pm 47$	3	110
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	$111 \pm 49$	5	110
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	$101 \pm 51$	4	110
cTnTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	$115 \pm 51$	4	109
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	$121 \pm 56$	4	109
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	$111 \pm 47$	4	109

**Table II-3.** Theoretical versus experimental overall correlation time  $\tau_m$  for cNTnC, L29Q, and E59D/D75Y in various complexes. Theoretical  $\tau_m$  values were calculated as MW(kDa)/2, similarly to the described previously (19).

	Overall correlation time, $\tau_m$ (ns)	Theoretical overall correlation time, $\tau_m$ (ns)
apo cNTnC	4.99	5.08
apo cNTnC•cTnI <sub>1-29</sub>	6.66	6.66
apo cNTnC•cTnI <sub>1-29</sub> PP	5.14	6.74
cNTnC•Ca <sup>2+</sup>	4.95	5.08
cNTnC•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	5.79	5.98
cNTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	6.05	6.66
cNTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	6.89	7.56
cNTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	5.16	6.74
cNTnC•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	6.09	7.64
L29Q•Ca <sup>2+</sup>	5.07	5.09
L29Q•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	5.87	5.99
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	5.81	6.67
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	6.44	7.57
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	5.12	6.75
L29Q•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	5.91	7.65
E59D/D75Y•Ca <sup>2+</sup>	5.49	5.1
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>147-163</sub>	6.39	6.00
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub>	6.33	6.68
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> •cTnI <sub>147-163</sub>	6.84	7.58
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP	5.82	6.76
E59D/D75Y•Ca <sup>2+</sup> •cTnI <sub>1-29</sub> PP•cTnI <sub>147-163</sub>	6.37	7.67

Based on the  $\tau_m$  data (Table II-3, Fig. II-8a), cNTnC•Ca<sup>2+</sup> binds all three peptides cTnI<sub>1-29</sub>, cTnI<sub>1-29</sub>PP, and cTnI<sub>147-163</sub> independently. Apo cNTnC binds both cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP as well, with similar avidity. The interaction of cNTnC•Ca<sup>2+</sup> with cTnI<sub>1-29</sub> is more efficient than with cTnI<sub>1-29</sub>PP since  $\tau_m$  for cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> is larger than  $\tau_m$  for cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP and  $\tau_m$  for apo cNTnC•cTnI<sub>1-29</sub> is larger than  $\tau_m$  for apo cNTnC•cTnI<sub>1-29</sub>PP (Table II-3, Fig. II-8a). Complexes with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP are formed with the stoichiometry less than 1:1 ( $\tau_m$ 's are smaller than the theoretical values in Table II-3 and below the diagonal in Figure II-8a), which indicates incomplete binding. cNTnC•Ca<sup>2+</sup> in complex with cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP binds cTnI<sub>147-163</sub> forming a triple complex ( $\tau_m$ 's for triple complexes are higher than  $\tau_m$ 's for the corresponding duplexes). The addition of cTnI<sub>147-163</sub> increases the weight of the cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>

complex to a smaller degree than that of  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$  (Table II-3) supporting the titration data where  $\text{cTnI}_{147-163}$  binds tighter to  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$ .

Similar qualitative conclusions can be drawn for L29Q and E59D/D75Y (Table II-3, Figs. II-8b and II-8c). Both mutants bind three peptides,  $\text{cTnI}_{1-29}$ ,  $\text{cTnI}_{1-29}\text{PP}$ , and  $\text{cTnI}_{147-163}$ , independently, and form triplexes with  $\text{cTnI}_{1-29}$  or  $\text{cTnI}_{1-29}\text{PP}$  and  $\text{cTnI}_{147-163}$ .  $\text{L29Q} \cdot \text{Ca}^{2+}$  binds  $\text{cTnI}_{1-29}$  and  $\text{cTnI}_{1-29}\text{PP}$  less efficiently than  $\text{cTnC} \cdot \text{Ca}^{2+}$ . The  $\text{L29Q} \cdot \text{Ca}^{2+}$  binding to  $\text{cTnI}_{1-29}\text{PP}$  is especially small ( $\tau_m$ 's for  $\text{L29Q} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$  and  $\text{L29Q} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP} \cdot \text{cTnI}_{147-163}$  are almost the same as the  $\tau_m$ 's for  $\text{L29Q} \cdot \text{Ca}^{2+}$  and  $\text{L29Q} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{147-163}$ ). The binding of  $\text{cTnI}_{147-163}$  to  $\text{L29Q} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$  is almost as efficient as binding to  $\text{L29Q} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$ .

Since several  $\tau_m$ 's in Figure II-8c are above the theoretical values and possibly slightly distorted due to the dimerization, conclusions about the E59D/D75Y behavior should be drawn with care. It is likely that  $\text{E59D/D75Y} \cdot \text{Ca}^{2+}$  binds  $\text{cTnI}_{1-29}$  weaker than  $\text{cTnC} \cdot \text{Ca}^{2+}$  and binds  $\text{cTnI}_{1-29}\text{PP}$  stronger than  $\text{cTnC} \cdot \text{Ca}^{2+}$ . The binding of  $\text{cTnI}_{147-163}$  to  $\text{E59D/D75Y} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$  might be as efficient as the binding of  $\text{cTnI}_{147-163}$  to  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$  not contradicting our titration data.

## Discussion

### *$\text{Ca}^{2+}$ and $\text{cTnI}_{147-163}$ binding affinity for $\text{cTnC}$ in the presence of $\text{cTnI}_{1-29}$ and $\text{cTnI}_{1-29}\text{PP}$*

It has been established that the phosphorylation of  $\text{cTnI}$  at position 22 and 23 results in the decreased  $\text{Ca}^{2+}$  sensitivity of actomyosin ATPase (23, 24) and increased rates of muscle relaxation and cross-bridge cycling (25, 26).  $\text{Ca}^{2+}$  off-rates in phosphorylated Tn complex are increased as well (27), which might stem from the alterations in either equilibria involved in Scheme 1. Within the errors of the method our data demonstrates that the  $\text{Ca}^{2+}$  binding affinity of  $\text{cTnC}$  is not affected by the presence of  $\text{cTnI}_{1-29}$  or  $\text{cTnI}_{1-29}\text{PP}$  (Fig. II-4, Table II-1), suggesting that the interaction with  $\text{cTnI}_{1-29}$  or  $\text{cTnI}_{1-29}\text{PP}$  does not interrupt  $\text{Ca}^{2+}$  ligation and does not significantly affect the structure and dynamics of  $\text{cTnC}$ , which might also potentially perturb the  $\text{Ca}^{2+}$  binding energetics. The affinities towards  $\text{cTnI}_{147-163}$  are instead decreased substantially:  $\sim 4$  times for  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$  and  $\sim 2$  times for  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP}$ . The  $\text{cTnI}_{147-163}$  binding affinity towards  $\text{cTnC} \cdot \text{Ca}^{2+}$  is decreasing in the following order:  $\text{cTnC} \cdot \text{Ca}^{2+} > \text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP} > \text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$ , which corresponds to the  $\text{cTnI}$  inhibition activity increasing in the similar order:  $\text{cTnC} \cdot \text{Ca}^{2+} < \text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}\text{PP} < \text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$ . This is in excellent agreement with the physiological data

on muscle myofibrillar relaxation (25), where relaxation rates followed the similar trend: a muscle preparation from transgenic mice expressing ssTnI (where ssTnI is a slow skeletal version of cTnI lacking the first 32 residues) relaxed faster (~70 ms) than the phosphorylated preparation from wt mice muscle (~90 ms), which in turn relaxed faster than the unphosphorylated version (~150 ms). This serves as an evidence of the importance of the cTnI<sub>147-163</sub> binding kinetics in determining the rate of muscle relaxation. The more efficiently the switch region of cTnI (~residues 147-163) binds to cTnC•Ca<sup>2+</sup>, the more efficiently the inhibitory region of cTnI (~residues 128-147) is sequestered away from its binding site on actin, which is likely to accelerate the formation of cross-bridges, their cycling rate, and the relaxation rate of muscle, thus transferring the signal of phosphorylation by PKA.

The binding interactions between cTnC•Ca<sup>2+</sup> and cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP are of enormous interest, but notoriously difficult to detect. The quenching of fluorescent Tyr<sup>25</sup> and Tyr<sup>28</sup> on cTnI and the changes in the CD spectra during titration with cTnC•Ca<sup>2+</sup> were negligible (data not shown). The signal in ITC titrations was remarkably small (28) and the majority of binding studies presented in literature had to utilize a label, AEDANS (8) or IAANS (29, 30), possibly affecting the charge distribution on the cTnC surface and the binding equilibrium. Chemical shifts changes in the NMR spectra of the N-domain of cTnC•Ca<sup>2+</sup> bound to cTnI<sub>1-80</sub> were exceptionally small (31), so that a binding curve plot was not possible to construct (our data not shown). In this work the following insights about the cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP binding to cTnC•Ca<sup>2+</sup> have been obtained: first, based on the  $\tau_m$  values, both, cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP bind to cTnC•Ca<sup>2+</sup> and apo cTnC. In agreement with literature data, cTnI<sub>1-29</sub> binds more efficiently than cTnI<sub>1-29</sub>PP (28, 30). Second, the cTnI<sub>1-29</sub> (cTnI<sub>1-29</sub>PP) binding to cTnC•Ca<sup>2+</sup> (apo cTnC) does not induce significant conformational changes and most probably is driven by electrostatics since the changes in cTnC•Ca<sup>2+</sup> and apo cTnC chemical shifts upon binding to cTnI<sub>1-29</sub> (cTnI<sub>1-29</sub>PP) are very small, which is characteristic of a typical electrostatic interaction. In comparison, the cTnI<sub>147-163</sub> binding to cTnC•Ca<sup>2+</sup> is accompanied by the large chemical shift changes in NMR spectra, and induces the substantial conformation change in cTnC from a closed to an opened conformation. Third, both complexes, cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> and cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP, bind cTnI<sub>147-163</sub> into its original binding site and form triplexes, which implies that cTnI<sub>147-163</sub> and cTnI<sub>1-29</sub> (cTnI<sub>1-29</sub>PP) have independent binding sites and can bind simultaneously. Fourth, the presence of cTnI<sub>1-29</sub> (cTnI<sub>1-29</sub>PP) affects the affinity of cTnC•Ca<sup>2+</sup> towards cTnI<sub>147-163</sub>, but not the

affinity of cTnC towards  $\text{Ca}^{2+}$ . One of the hypotheses, explaining the change in the cTnI<sub>147-163</sub> affinity, proposes that the cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP binding sites on cTnC• $\text{Ca}^{2+}$  partially overlap with the binding site for cTnI<sub>147-163</sub>. This overlap is small since the cTnI<sub>147-163</sub> peptide still binds fairly tightly and the cTnI<sub>1-29</sub> binding alone does not induce the open conformation of cTnC• $\text{Ca}^{2+}$ . This hypothesis would be in agreement with the cross-linking data, showing that cTnI residues 5 (or 8) and 154 are in close proximity (32). Another hypothesis proposes that cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP sites on cTnC• $\text{Ca}^{2+}$  do not overlap with the cTnI<sub>147-163</sub> binding site, however, cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP stabilize cTnC• $\text{Ca}^{2+}$  in a closed conformation, albeit with different strengths, thus requiring more energy for cTnI<sub>147-163</sub> to bind to the cleft on cTnC• $\text{Ca}^{2+}$ . The affinity of cTnI<sub>147-163</sub> will be lessened the most when cTnI<sub>1-29</sub> is bound unphosphorylated, since more cTnI<sub>1-29</sub> is bound to cTnC• $\text{Ca}^{2+}$ , locking cTnC• $\text{Ca}^{2+}$  in the closed conformation. This hypothesis is in agreement with the reported NMR data showing the line broadening in areas of  $\beta$ -sheets on cTnC• $\text{Ca}^{2+}$  bound to cTnI<sub>1-80</sub> (31, 33). The area of  $\beta$ -sheets on cTnC• $\text{Ca}^{2+}$  is located on the opposite side from the cTnI<sub>147-163</sub> binding site (7). We favor the model where cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP bind on cTnC• $\text{Ca}^{2+}$  without overlapping with cTnI<sub>147-163</sub>; however, the distance between both cTnI regions when bound to cTnC• $\text{Ca}^{2+}$  should be short enough to afford the cross-linking between them (32). More detailed information on binding of cTnI<sub>1-29</sub> (cTnI<sub>1-29</sub>PP) to cTnC• $\text{Ca}^{2+}$ , e.g. the nature of the forces driving the formation of the cTnC• $\text{Ca}^{2+}$ •cTnI<sub>1-29</sub> complex will be discussed in Chapter IV.

It is desirable to compare other binding data available with the  $K_D$ 's obtained in this work. For instance, data on the binding of whole cTnI to cTnC has been reported (8, 29), however, upon phosphorylation one segment of cTnI should bind weaker (~residues 1-29) and another segment of cTnI should bind stronger (~residues 147-163), thus reported  $K_D$ 's represent a complex function of several independent events that are difficult to dissect. This emphasizes the need for binding studies with independent peptides, where individual  $K_D$ 's can be obtained and compared. In addition, the introduction of a charged label such as AEDANS, might perturb the charge distribution on the cTnC surface, which similarly to E59D/D75Y might reverse the effect of the cTnI N-terminus (8), strengthening the binding towards cTnI<sub>147-163</sub> in the presence of TnI<sub>1-29</sub> and weakening in the presence of TnI<sub>1-29</sub>PP.

It is also highly attractive to explain the increase in  $\text{Ca}^{2+}$  off rates and the decrease in  $\text{Ca}^{2+}$  sensitivity upon phosphorylation. The  $\text{Ca}^{2+}$  sensitivity of force development decreases in the following order ( $\text{Ca}^{2+}$  required for 50% activation): ssTnI

(0.74  $\mu\text{M}$ ) < cTnI (1.52  $\mu\text{M}$ ) < cTnI-PP (2.15  $\mu\text{M}$ ). The  $\text{Ca}^{2+}$  off rates increase in the following order: cTnI (21 /s) < cTnI-PP (48 /s) (27). As it is known from literature,  $\text{Ca}^{2+}$  sensitivity and  $\text{Ca}^{2+}$  off-rates depends on the length of a sarcomere (34), on the presence of ATP and  $\text{Mg}^{2+}$  (35), and on the activation of cross-bridges (36, 37). Theoretical considerations advocate that “neither activation of force nor that of ATPase is proportional to  $\text{Ca}^{2+}$  absorption to the activating sites”, which underlines the complexity of  $\text{Ca}^{2+}$  regulation in muscle (38). At this point, we are unable to provide a mechanistic explanation for the differences in  $\text{Ca}^{2+}$  off-rates and filament  $\text{Ca}^{2+}$  sensitivities in phosphorylated, unphosphorylated states, and without the N-terminus of cTnI due to the complexity of the system.

#### *$\text{Ca}^{2+}$ binding affinity for L29Q and E59D/D75Y*

A consensus EF hand motif in  $\text{Ca}^{2+}$  binding proteins is composed of 12 residues, 6 of which are directly involved in the coordination of  $\text{Ca}^{2+}$  (Table II-4) (39). In cTnNC silent  $\text{Ca}^{2+}$  binding site I, 4 residues remain available for the coordination of  $\text{Ca}^{2+}$  out of 6 conserved residues, providing 5 out of 7 ligating atoms. Several consensus residues that do not directly interact with  $\text{Ca}^{2+}$  but provide the required rigidity and charge balance for a  $\text{Ca}^{2+}$  binding site are also substituted in cTnNC site I, and one residue preceding binding site I is inserted (Table II-4). The L29Q mutation in site I is potentially capable of providing the side chain for  $\text{Ca}^{2+}$  coordination, however, the amide group of Gln<sup>29</sup> is a less potent ligand than the carboxylic group of consensus Asp<sup>29</sup>, and the chain length for a ligand is slightly longer than that of Asp<sup>29</sup>. Site II remained intact and the affinity of site II in L29Q should not be changed. The question of whether the L29Q substitution increases the overall  $\text{Ca}^{2+}$  affinity of cTnNC was addressed using two independent methods, NMR titrations and UV/VIS titrations with a chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA. Both methods yielded consistent results with the L29Q affinity towards  $\text{Ca}^{2+}$  essentially the same as that of for cTnNC. In fish cTnNC, the engineered Q29L mutation was shown to decrease the originally high  $\text{Ca}^{2+}$  sensitivity to the level comparable with human cTnNC, thus proving that the L29Q substitution in fish is required for the higher sensitivity towards  $\text{Ca}^{2+}$ , however, L29Q is not the only substitution necessary for a higher  $\text{Ca}^{2+}$  sensitivity (40). While the increase in  $\text{Ca}^{2+}$  sensitivity in fish cTnNC is the most noticeable at cold temperatures, where it allows the fish to overcome the desensitization of the heart, the increase in  $\text{Ca}^{2+}$  affinity for human L29Q at 30°C was not significant enough to be recorded. Little evidence for a weak second  $\text{Ca}^{2+}$  binding site for



L29Q was observed and the estimated  $K_{D2}$  for this binding is ~5 mM or weaker, which limits its possible physiological consequences.

In case of E59D/D75Y, the  $\text{Ca}^{2+}$  affinity also did not noticeably change, although the slight decrease in  $\text{Ca}^{2+}$  affinity of site II might have been expected. The substitution of conserved Asp<sup>75</sup> to Tyr<sup>75</sup> should not affect  $\text{Ca}^{2+}$  coordination directly, however, it can disrupt electrostatic balance and the packing within the  $\text{Ca}^{2+}$  binding loop (Table II-4). In competition assays with a chelator, 5,5'-Br<sub>2</sub>-BAPTA, and in NMR titrations E59D/D75Y showed no change in  $\text{Ca}^{2+}$  affinity within the errors of the method. It is still possible that the changes in affinities smaller than the resolution of the techniques used in this work could be physiologically significant.

**Table II-4.** Residues involved in the coordination of  $\text{Ca}^{2+}$  in the EF hand motifs of various TnC's.

Position	X		Y		Z		-Y	-X		-Z				
Consensus sequence	D	Basic	D/N	G	D/N	G	Any	hydro-phobic	D/E/T/S	Hydro-phobic	Acidic	E		
Coordinating group	COO <sup>-</sup> side chain		C=O COO <sup>-</sup> side chain		C=O COO <sup>-</sup> side chain		C=O back-bone		COO <sup>-</sup> CONH <sub>2</sub> OH Water		COO <sup>-</sup>			
cNTnC (site I)	F <sup>27</sup>	V <sup>28</sup>	<b>L<sup>29</sup></b>	G <sup>30</sup>	A <sup>31</sup>	E <sup>32</sup>	D <sup>33</sup>	G <sup>34</sup>	S <sup>35</sup>	I <sup>36</sup>	S <sup>37</sup>	T <sup>38</sup>	K <sup>39</sup>	E <sup>40</sup>
L29Q	F <sup>27</sup>	V <sup>28</sup>	<b>Q<sup>29</sup></b>	G <sup>30</sup>	A <sup>31</sup>	E <sup>32</sup>	D <sup>33</sup>	G <sup>34</sup>	S <sup>35</sup>	I <sup>36</sup>	S <sup>37</sup>	T <sup>38</sup>	K <sup>39</sup>	E <sup>40</sup>
Trout cNTnC	F <sup>27</sup>	I <sup>28</sup>	<b>Q<sup>29</sup></b>	D <sup>30</sup>	A <sup>31</sup>	E <sup>32</sup>	D <sup>33</sup>	G <sup>34</sup>	S <sup>35</sup>	I <sup>36</sup>	S <sup>37</sup>	T <sup>38</sup>	K <sup>39</sup>	E <sup>40</sup>
sNTnC	F <sup>28</sup>	-	<b>D<sup>29</sup></b>	A <sup>30</sup>	D <sup>31</sup>	G <sup>32</sup>	D <sup>33</sup>	G <sup>34</sup>	D <sup>35</sup>	I <sup>36</sup>	S <sup>37</sup>	V <sup>38</sup>	K <sup>39</sup>	E <sup>40</sup>
cNTnC (site II)	E <sup>63</sup>	V <sup>64</sup>	D <sup>65</sup>	E <sup>66</sup>	D <sup>67</sup>	G <sup>68</sup>	S <sup>69</sup>	G <sup>70</sup>	T <sup>71</sup>	V <sup>72</sup>	D <sup>73</sup>	F <sup>74</sup>	<b>D<sup>75</sup></b>	E <sup>76</sup>
E59D/D75Y	E <sup>63</sup>	V <sup>64</sup>	D <sup>65</sup>	E <sup>66</sup>	D <sup>67</sup>	G <sup>68</sup>	S <sup>69</sup>	G <sup>70</sup>	T <sup>71</sup>	V <sup>72</sup>	D <sup>73</sup>	F <sup>74</sup>	<b>Y<sup>75</sup></b>	E <sup>76</sup>
Trout cNTnC	E <sup>63</sup>	V <sup>64</sup>	D <sup>65</sup>	E <sup>66</sup>	D <sup>67</sup>	G <sup>68</sup>	S <sup>69</sup>	G <sup>70</sup>	T <sup>71</sup>	V <sup>72</sup>	D <sup>73</sup>	F <sup>74</sup>	<b>D<sup>75</sup></b>	E <sup>76</sup>
sNTnC	E <sup>63</sup>	V <sup>64</sup>	D <sup>65</sup>	E <sup>66</sup>	D <sup>67</sup>	G <sup>68</sup>	S <sup>69</sup>	G <sup>70</sup>	T <sup>71</sup>	I <sup>71</sup>	D <sup>73</sup>	F <sup>74</sup>	<b>D<sup>75</sup></b>	E <sup>76</sup>

*cTnI<sub>147-163</sub> binding affinity for L29Q•Ca<sup>2+</sup> and E59D/D75Y•Ca<sup>2+</sup>*

The slight decrease in affinity for L29Q•Ca<sup>2+</sup> towards cTnI<sub>147-163</sub> was anticipated since the L29Q mutation is likely to perturb the binding interface with the cTnI<sub>147-163</sub> peptide. On the structure reported by Li et al. (7), the  $\delta$ -hydrogen of Leu<sup>29</sup> on cNTnC is positioned 6.3 Å away from the  $\delta$ -proton and 6.6 Å away from the  $\beta$ -proton of Leu<sup>158</sup> on cTnI<sub>147-163</sub>, and 7.1 Å away from the  $\delta$ -proton of Leu<sup>159</sup> on cTnI<sub>147-163</sub>. Although these residues do not directly interact with Leu<sup>29</sup> on cNTnC, the decrease in the affinity of L29Q•Ca<sup>2+</sup> towards cTnI<sub>147-163</sub> can stem from the perturbed bonding interactions within the hydrophobic core of the interface, for example, through Val<sup>44</sup> and Ile<sup>26</sup>.

The increase in the affinity of E59D/D75Y•Ca<sup>2+</sup> towards cTnI<sub>147-163</sub> can be explained by the strengthening of the interaction between Asp<sup>59</sup> on E59D/D75Y and Arg<sup>147</sup> of cTnI<sub>147-163</sub> due to the shortening of the side chain in the position of residue 59. In the structure of cNTnC•Ca<sup>2+</sup> in complex with cTnI<sub>147-163</sub> (7) the guanidine group of Arg<sup>147</sup> on cTnI<sub>147-163</sub> is ~ 6 Å away from the carboxylic oxygen and ~ 4 Å away from the γ-protons of Glu<sup>59</sup> on cNTnC. Thus, the shortening of the side chain of residue 59 is likely to result in a stronger electrostatic contact between E59D/D75Y•Ca<sup>2+</sup> and cTnI<sub>147-163</sub>.

*Ca<sup>2+</sup> and cTnI<sub>147-163</sub> binding affinity for L29Q, and E59D/D75Y in the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP*

Similarly to cNTnC, within the experimental error the affinities of L29Q and E59D/D75Y towards Ca<sup>2+</sup> were not affected by the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP. The affinity towards cTnI<sub>147-163</sub> was distinct for every mutant and different from that of cNTnC•Ca<sup>2+</sup>. L29Q•Ca<sup>2+</sup> affinity towards cTnI<sub>147-163</sub> was essentially the same in the presence of cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP, which should confer the same cTnI inhibition of actin regardless of cTnI phosphorylation state. This is in accordance with the data showing that the impact of phosphorylation in L29Q was abolished, and both, ATPase activity and the sliding velocity of the thin filament, remained unaffected by cTnI phosphorylation (1). Based on our τ<sub>m</sub> data, L29Q•Ca<sup>2+</sup> binds both cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP less efficiently than cNTnC•Ca<sup>2+</sup>, which might explain the absence of bound species on the binding plate in Schmidtman et al. (1). A slight decrease in the Ca<sup>2+</sup> sensitivity of force development (41) is in agreement with a decreased affinity towards cTnI<sub>147-163</sub> (in the absence of phosphorylation by PKA) that might increase the inhibition of actin and presumably decelerate the rate of cross-bridge formation. When phosphorylated, the L29Q containing fibers might demonstrate a decrease in inhibition of actin and increase in the rates of cross-bridge formation and muscle relaxation, for which physiological data is not available. The introduction of the L29Q mutation has brought the change in the cTnI<sub>147-163</sub> binding affinity per se, and the change in the interactions of L29Q•Ca<sup>2+</sup> with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP, which affects their modulation of the cTnI<sub>147-163</sub> affinity with and without PKA phosphorylation. The interplay between these events is likely to be responsible for the development of FHC.

The E59D/D75Y mutation brings a drastic change in the cTnI<sub>147-163</sub> binding equilibrium to cNTnC•Ca<sup>2+</sup> alone and in complex with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP. The strength of cTnI<sub>147-163</sub> binding becomes reversed: E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> binds cTnI<sub>147-</sub>

<sub>163</sub> with higher affinity than E59D/D75Y•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>PP. This data can be explained by the charge perturbation on the cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP binding interface. In E59D/D75Y, the surface charge is changed to the less negative as compared to cTnI, decreasing the electrostatic repulsion of cTnI<sub>1-29</sub>PP. This leads to a slightly tighter binding towards cTnI<sub>1-29</sub>PP as compared to cTnI (qualitatively supported by the  $\tau_m$  data) and a weaker binding towards cTnI<sub>147-163</sub> (supported by our titration data). When phosphorylated, the E59D/D75Y containing filaments are likely to exhibit the increased inhibition of actin, which is in agreement with the reduced ability to activate myosin ATPase activity (42). The surface charge perturbation of E59D/D75Y•Ca<sup>2+</sup> should also lead to a weaker binding of cTnI<sub>1-29</sub> and consequently a tighter binding of cTnI<sub>147-163</sub> (supported by our titration data), which in muscle relaxation experiments is likely to result in the decreased muscle inhibition and accelerated relaxation rates. The combined effects of increased affinity towards cTnI<sub>147-163</sub> for uncomplexed E59D/D75Y•Ca<sup>2+</sup> and perturbed modulation of the cTnI<sub>147-163</sub> affinity by cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP are likely to result in the development of a heart disorder such as DCM.

## Conclusions

The cTnI N-terminus is a crucial participant of the Ca<sup>2+</sup> regulated cardiac muscle contraction. It binds to cTnT simultaneously with the switch region of cTnI (~residues 147-163) and upon phosphorylation modulates the cTnT affinity towards cTnI<sub>147-163</sub>, thus transferring the PKA phosphorylation signal further along the thin filament. In case of the cTnT mutants associated with cardiomyopathy, L29Q and E59D/D75Y, the interactions with the N-terminus and the switch region of cTnI are significantly perturbed, which underlines the importance of these regions in the development of malfunctions and advances our understanding of cardiac muscle regulation.

## Materials and Methods

### *Protein Expression and Purification*

The DNA encoding cNTnC (1-89) with the mutations C35S and C84S was subcloned previously into pET-3a expression vector as described in Pearlstone et al. (43). Using cNTnC(C35S, C84S) as a template, the L29Q and E59D/D75Y mutations were engineered using a site-directed mutagenesis kit (QuikChange purchased from Stratagene). *E. coli* strain BL21(DE3)pLysS was transformed with the expression vector and incubated at 37°C to the OD<sub>600</sub> of 0.6-0.9. Cell cultures were induced with IPTG and harvested after incubation for 3 hours. Uniformly <sup>15</sup>N-labeled cNTnC, L29Q, and E59D/D75Y were expressed in a minimal media enriched with (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (44). Cell pellet was lysed using French press and applied to a DEAE-sephadex (A25) column (50 mM Tris, 0.1 M NaCl, pH 8.0). Proteins were further purified using a Superdex-75 size exclusion column (50 mM Tris, 0.15 M NaCl, pH 8.0), desalted using Sephadex G25 column (10 mM NH<sub>4</sub>HCO<sub>3</sub>), decalcified as described previously (45), and lyophilized. Molecular masses for unlabeled proteins, cNTnC, L29Q, and E59D/D75Y, as determined by MALDI mass spectrometry were equal to the expected values for cNTnC (10,030 ± 1 Da), L29Q (10,045 ± 1 Da), and E59D/D75Y (10,064 ± 1 Da). Amino acid composition was confirmed independently by amino acid analysis.

### *Peptides*

Synthetic peptides, cTnI<sub>147-163</sub> (acetyl-RISADAMMQALLGARAK-amide), cTnI<sub>1-29</sub> (acetyl-ADGSSDAAREPRPAPAPIRRRSSNYRAYA-amide), and cTnI<sub>1-29</sub>PP (acetyl-ADGSSDAAREPRPAPAPIRRRS(P)S(P)NYRAYA-amide) were prepared using standard methodology for a typical TnI peptide (46). The sequences were confirmed by amino acid analysis and the masses were verified by electrospray mass spectrometry.

### *Ca<sup>2+</sup> titrations of <sup>15</sup>N-L29Q and <sup>15</sup>N-E59D/D75Y monitored by 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy*

Decalcified <sup>15</sup>N-L29Q and <sup>15</sup>N-E59D/D75Y, respectively, were dissolved in 500 µl of NMR buffer containing 100 mM KCl, 10 mM imidazole, 5 mM NaN<sub>3</sub>, and 0.2 mM DSS in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. The pH was adjusted to 6.8 for both samples. Solutions of 50 mM and 100 mM CaCl<sub>2</sub> in water used in titrations of both samples, were prepared from 1.0 M CaCl<sub>2</sub> solution purchased from Sigma-Aldrich. The concentration of L29Q (0.9 mM) and E59D/D75Y (1 mM) were determined by amino acid analysis. Changes in protein concentration due to dilution were taken into account during data analysis. To plot the binding curve, per residue chemical shift changes,  $\Delta_{total}$ , were obtained using the following equation:  $\Delta_{total} = \sqrt{(0.2\Delta\delta_{15N})^2 + (\Delta\delta_{1H})^2}$ , where  $\Delta\delta_{15N}$  is the chemical shift change in ppm in <sup>15</sup>N dimension and  $\Delta\delta_{1H}$  is the chemical shift change in ppm in <sup>1</sup>H dimension. The coefficient of 0.2 accounts for the difference between sensitivities of <sup>15</sup>N and <sup>1</sup>H nuclei. The per residue chemical shift changes for the most perturbed residues (with large chemical shift changes,  $\Delta_{total}$ ) were averaged and analyzed using the program xcrvfit (www.bionmr.ualberta.ca).

### *Ca<sup>2+</sup> titrations of cNTnC, L29Q, and E59D/D75Y monitored by UV/VIS spectroscopy using competition assays with the chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA*

A buffer containing 150 mM KCl and 10 mM imidazole (pH 6.8) was decalcified with Chelex-100 resin (BioRad). The affinity of a chelator towards Ca<sup>2+</sup> was determined by titrating of 40-50 µM of tetra potassium salt of 5,5'-Br<sub>2</sub>-BAPTA (CalBiochem) with

CaCl<sub>2</sub>. 3  $\mu$ L aliquots of 3 mM CaCl<sub>2</sub> were added at every titration point. The UV absorbance was recorded at  $\lambda=263$  nm. The initial concentration of 5,5'-Br<sub>2</sub>-BAPTA (Ca<sup>2+</sup> free) was calculated using the extinction coefficient of  $2 \cdot 10^4$  L $\cdot$ mol<sup>-1</sup> $\cdot$ cm<sup>-1</sup> at  $\lambda=263$  nm (47). The K<sub>D</sub> for 5,5'-Br<sub>2</sub>-BAPTA under experimental conditions was determined to be  $0.5 \pm 0.3$   $\mu$ M from the average of five independent titrations, which is comparable to the literature data (11, 47).

In competition assays, the mixture of a chelator and a protein (or its complex with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP) in a total volume of 1 mL was titrated with Ca<sup>2+</sup>. 3  $\mu$ L aliquots of 3 mM CaCl<sub>2</sub> were added at every titration point. In order to assure the equal concentration of proteins during titrations alone and in complexes with cTnI<sub>1-29</sub> and cTnI<sub>1-29</sub>PP, protein stock solutions were divided in three portions, into which cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP were subsequently added. The concentrations of protein stock solutions were determined by amino acid analysis. The K<sub>D</sub> of a protein was deconvoluted from obtained binding curves using Microsoft Excel and xcrvfit software using the K<sub>D</sub> of the chelator and the concentrations of individual components as known parameters during the fitting (11).

*cTnI<sub>147-163</sub> titration of L29Q $\cdot$ Ca<sup>2+</sup> and E59D/D75Y $\cdot$ Ca<sup>2+</sup> monitored by 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy*

The synthetic peptide, cTnI<sub>147-163</sub>, was added in the form of powder during titration of the solution of 0.9 mM L29Q $\cdot$ Ca<sup>2+</sup> and 0.6 mM E59D/D75Y $\cdot$ Ca<sup>2+</sup> respectively. The buffer was the same as the buffer used for the NMR Ca<sup>2+</sup> titrations. CaCl<sub>2</sub> was added to a concentration of 35  $\mu$ M. pH of solutions was adjusted to  $6.75 \pm 0.1$ . The concentration of proteins and peptides was determined by amino acid analysis at every titration point.

*cTnI<sub>147-163</sub> titration of cTnTnC $\cdot$ Ca<sup>2+</sup>, L29Q $\cdot$ Ca<sup>2+</sup> and E59D/D75Y $\cdot$ Ca<sup>2+</sup> in complex with cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP monitored by 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy*

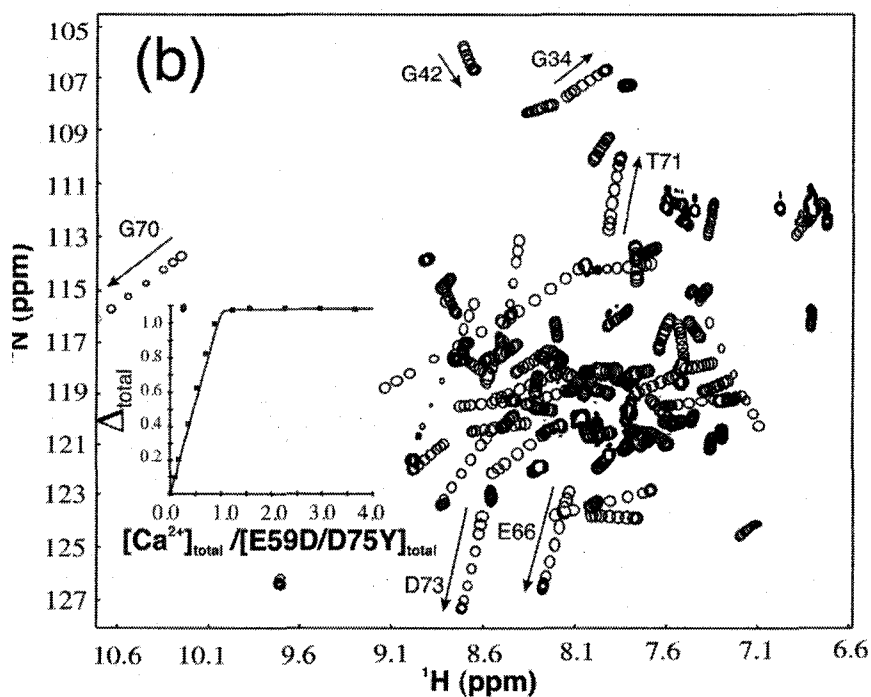
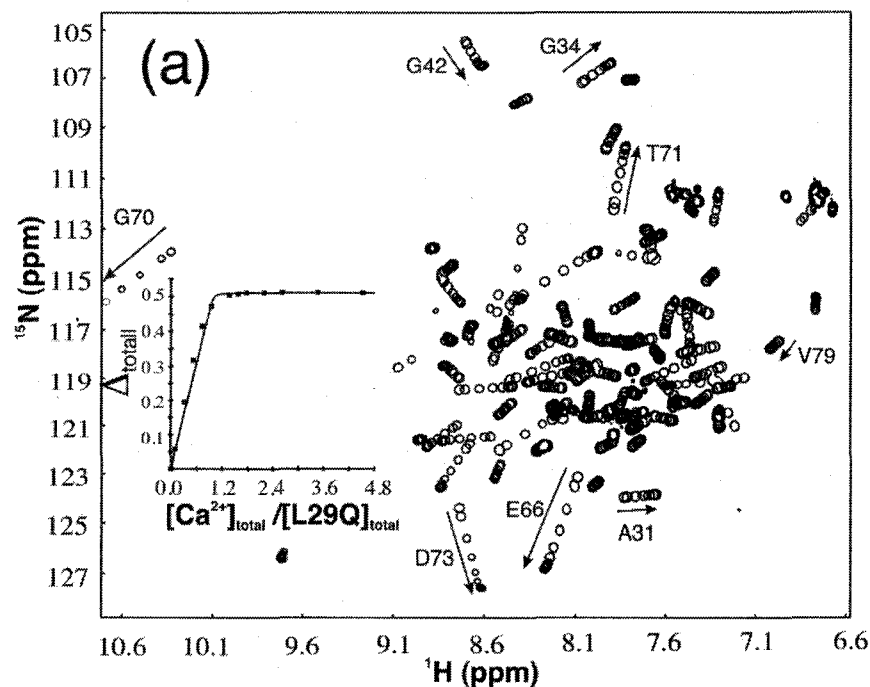
The complexes of cTnTnC $\cdot$ Ca<sup>2+</sup>, L29Q $\cdot$ Ca<sup>2+</sup> and E59D/D75Y $\cdot$ Ca<sup>2+</sup> with cTnI<sub>1-29</sub> or cTnI<sub>1-29</sub>PP were prepared to the ratio 1 : 2.3. The concentration of proteins was equal to 0.6 mM in the same NMR buffer described above, containing 35  $\mu$ M of CaCl<sub>2</sub>. The concentration of peptides was equal to  $1.4 \pm 0.05$  mM in all samples. The formation of complexes was verified using the measurements of <sup>15</sup>N backbone amide NMR relaxation rates. During titrations, cTnI<sub>147-163</sub> was added in the form of powder until the ratio of cTnTnC $\cdot$ Ca<sup>2+</sup> (L29Q $\cdot$ Ca<sup>2+</sup> or E59D/D75Y $\cdot$ Ca<sup>2+</sup>) : cTnI<sub>147-163</sub> was equal to 1 : ~3.5. The pH of solutions was kept at  $6.75 \pm 0.1$ . The results of amino acid analysis for the mixture of three components gave rise to a less accurate estimation of concentrations as compared to the concentrations determined by weight. Thus, in these experiments, the weight of added peptide powder was used to calculate the concentration of cTnI<sub>147-163</sub>.

*<sup>15</sup>N backbone amide NMR relaxation data*

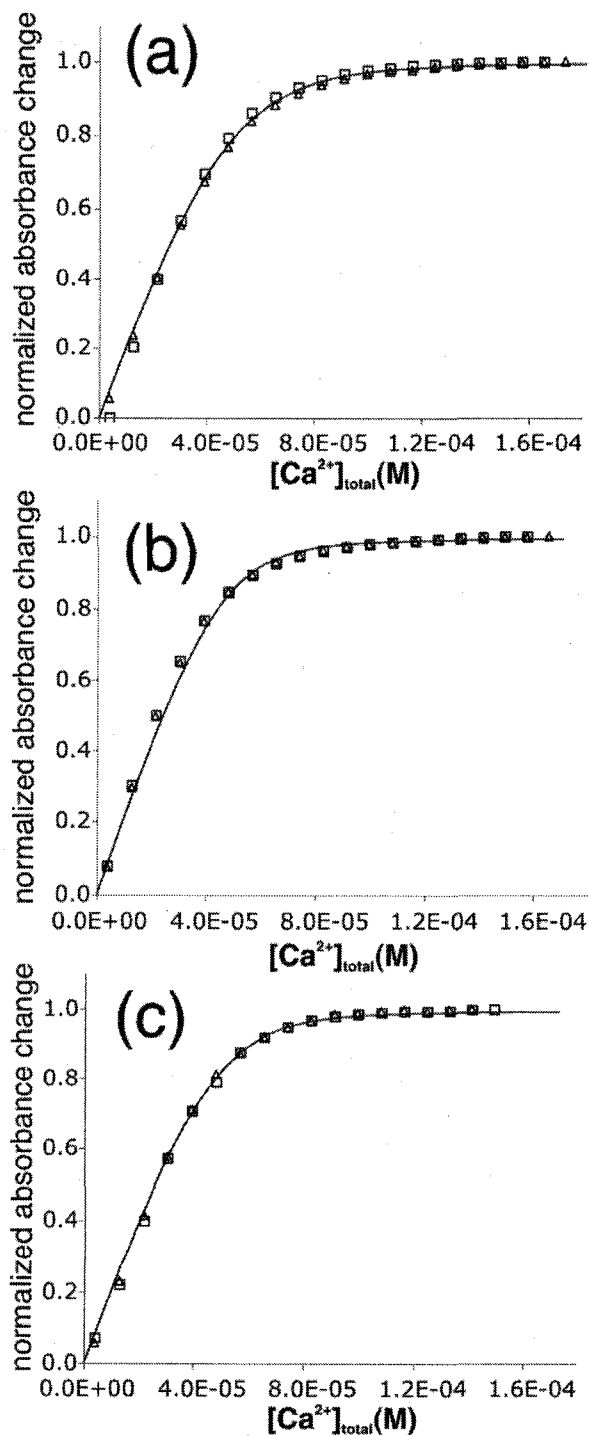
Relaxation data were acquired from <sup>15</sup>N-T<sub>1</sub>, <sup>15</sup>N-T<sub>2</sub>, and {<sup>1</sup>H-<sup>15</sup>N} NOE experiments (Biopack, Varian Associates) for 500  $\mu$ L samples containing in all experiments 0.6 mM of cTnTnC, L29Q, and E59D/D75Y. T<sub>1</sub> data were acquired using relaxation delays of 10, 50, 100, 200, 300, and 400 ms, and T<sub>2</sub> data were acquired using relaxation delays of 10, 30, 50, 70, 90, and 110 ms. The delays between transients in <sup>15</sup>N-T<sub>2</sub> and <sup>15</sup>N-T<sub>1</sub> experiments were set to 3 s. {<sup>1</sup>H-<sup>15</sup>N} NOE experiments were performed using delays of 5 s for spectra recorded without proton saturation and delays of 2 s for spectra recorded with proton saturation. The proton saturation was set to 3 s so that the total time between transients was equal to 5 s. NMR experiments were conducted at 30

°C on a Varian Inova 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance probe. Acquired spectra were processed using NMRPipe (48) and visualized using NMRView (49). Relaxation parameters,  $T_1$ ,  $T_2$ , and NOE, were extracted using the rate analysis module built in NMRView. The overall rotational tumbling  $\tau_m$  was averaged from the per residue fitting of the relaxation data to the Lipari-Szabo  $S^2$ - $\tau_m$ - $\tau_f$  model (15). In this method of  $\tau_m$  calculation, the relaxation data for residues with significant internal motions, as judged by  $\text{NOE} < 0.65$ , and the data for residues with possible slow motions, as judged by significantly decreased  $T_2$ , were excluded. Mathematica scripts used for the analysis of relaxation data were provided by Dr. Leo Spyracopoulos ([www.bionmr.ualberta.ca/~lspy/](http://www.bionmr.ualberta.ca/~lspy/)) (50).

**Figure II-1.** The overlay of 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectra obtained during the  $\text{Ca}^{2+}$  titrations of L29Q (a) and E59D/D75Y (b). Residue assignments are shown for resonances with significant chemical shift changes. Binding curves (insets) represent the average of chemical shift changes for indicated residues versus the ratio of the total calcium to protein concentrations. The results of fitting are shown for one binding site model.

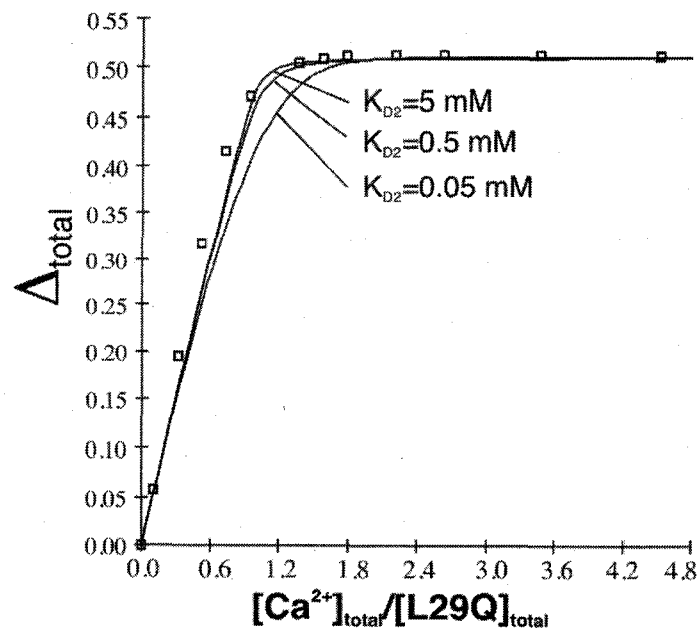


**Figure II-2.** Normalized absorbance changes during  $\text{Ca}^{2+}$  titrations for cNTnC (a), L29Q (b), and E59D/D75Y (c) obtained in competitive assays with the chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA. Two independent binding experiments (triangles and squares) for each protein are shown to demonstrate reproducibility of the method.

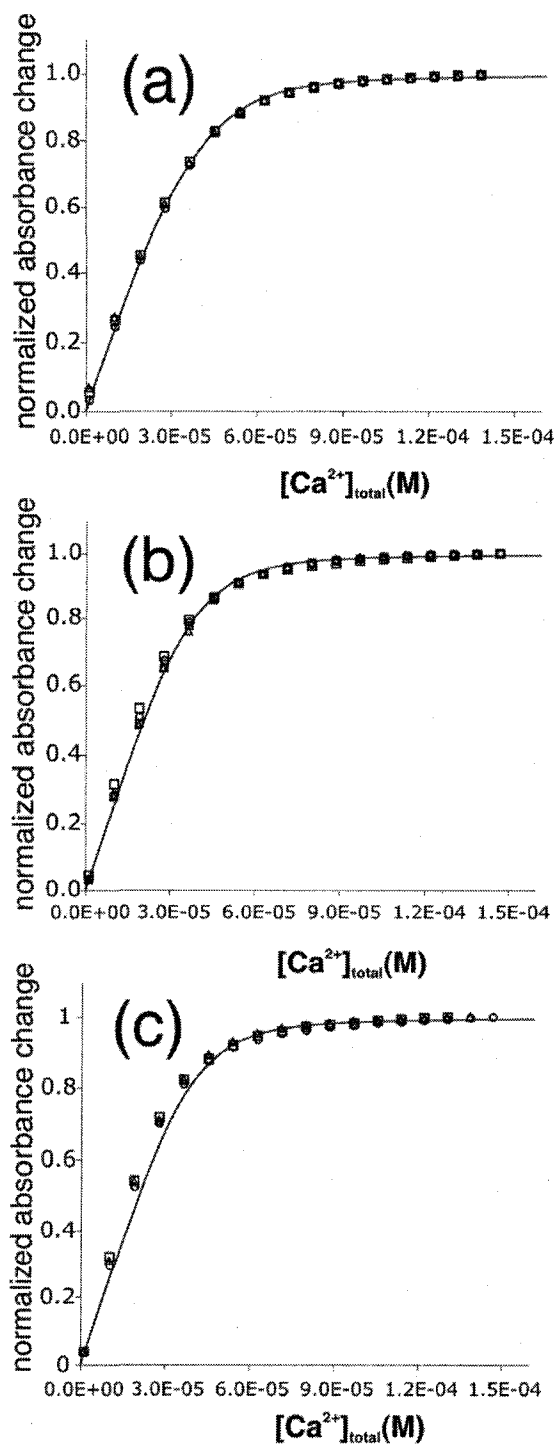




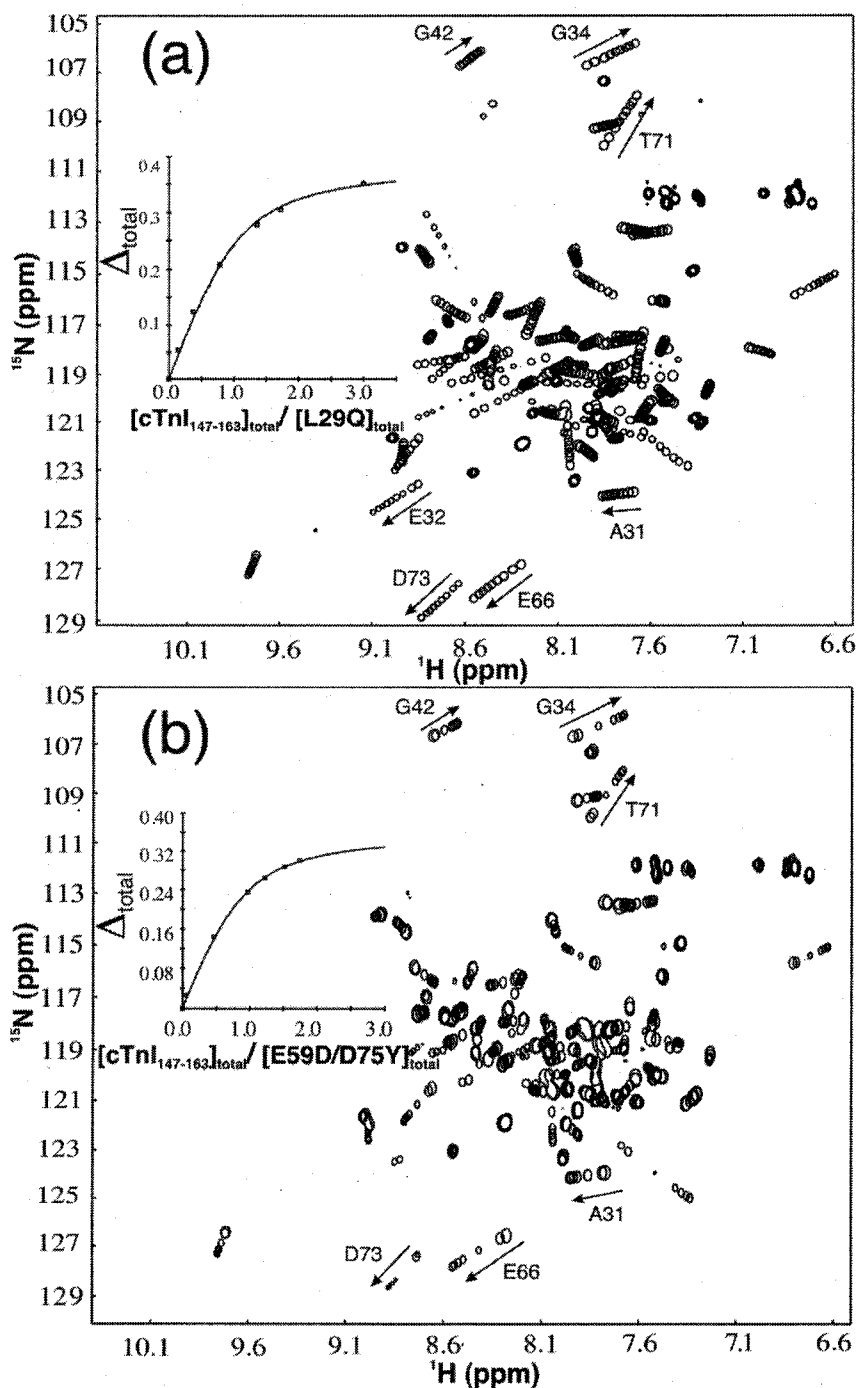
**Figure II-3.**  $\text{Ca}^{2+}$  titration of L29Q determined by NMR and analyzed using the two binding site model. Affinity for the first binding site ( $K_{D1} = 5 \mu\text{M}$ ) was taken as a fixed parameter during the fitting. Examples of the curves obtained with different  $K_{D2}$ 's are shown for comparison.



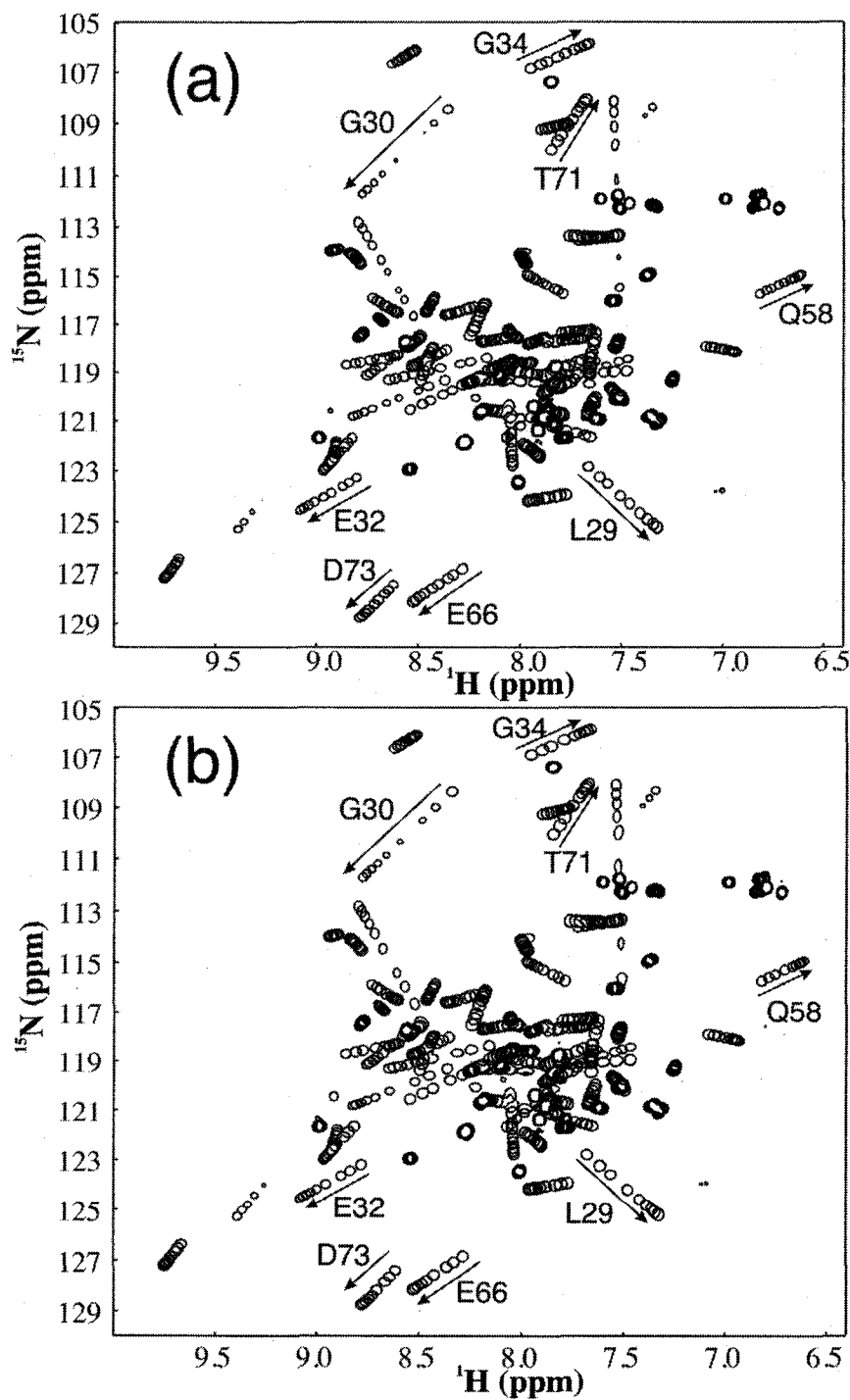
**Figure II-4.** Normalized absorbance changes during  $\text{Ca}^{2+}$  titrations for cTnC (triangles), cTnC•cTnI<sub>1-29</sub> (squares), and cTnC•cTnI<sub>1-29</sub>PP (circles) (a), L29Q (triangles), L29Q•cTnI<sub>1-29</sub> (squares), and L29Q•cTnI<sub>1-29</sub>PP (circles) (b), and E59D/D75Y (triangles), E59D/D75Y•cTnI<sub>1-29</sub> (squares), and E59D/D75Y•cTnI<sub>1-29</sub>PP (circles) (c) obtained in competitive assays with the chromophoric chelator, 5,5'-Br<sub>2</sub>-BAPTA.

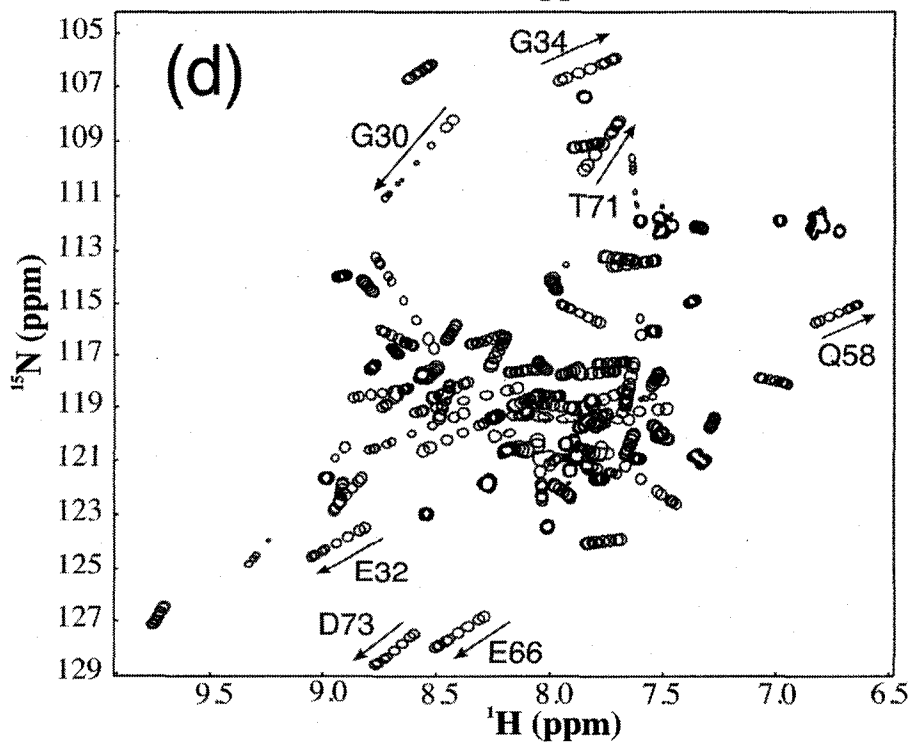
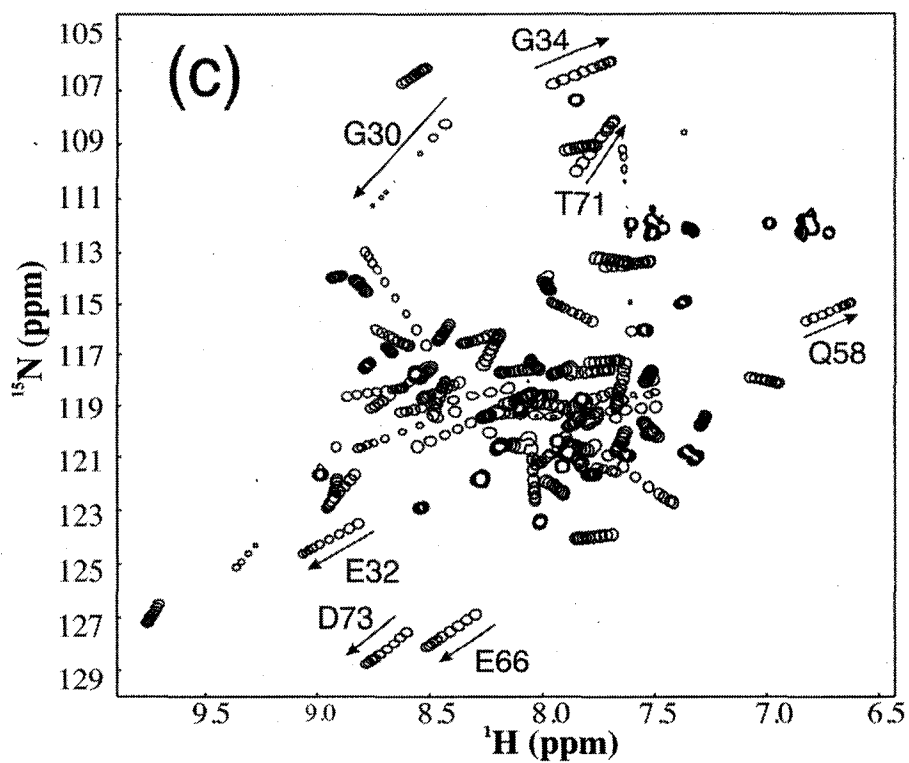


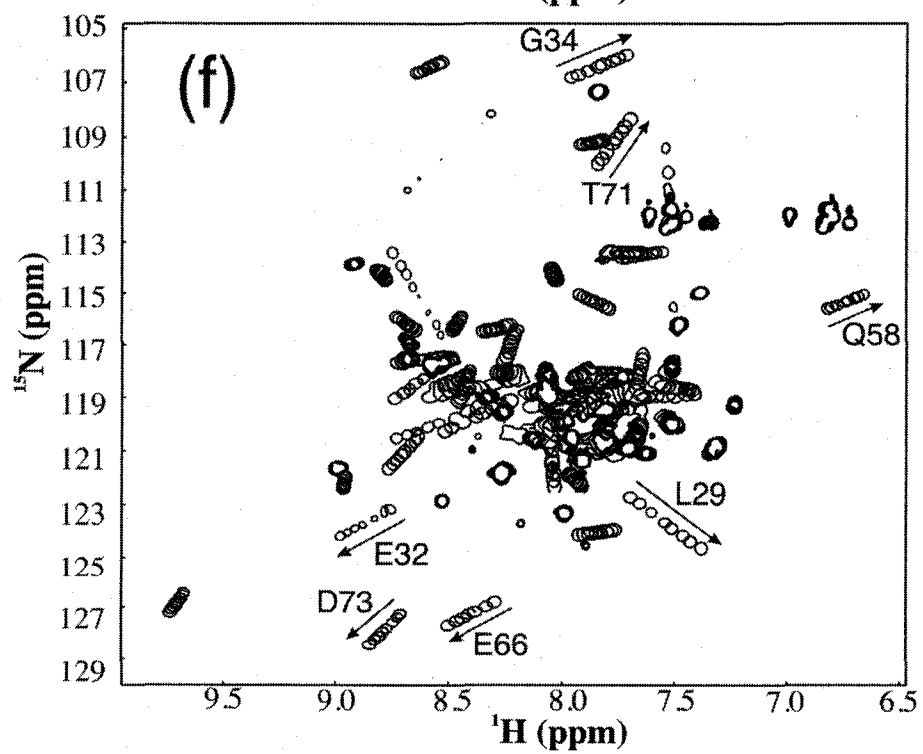
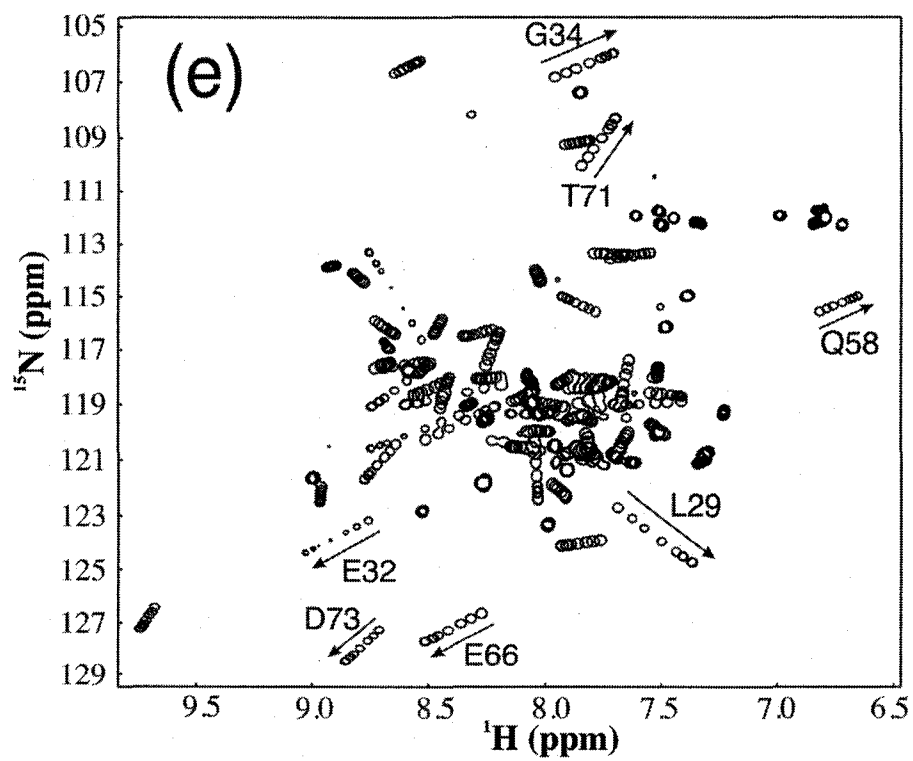
**Figure II-5.** The overlay of 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectra acquired during titration of L29Q•Ca<sup>2+</sup> (a) and E59D/D75Y•Ca<sup>2+</sup> (b) with cTnI<sub>147-163</sub>. The contour level for spectra was chosen to demonstrate the fact of line broadening in the middle of titration in case of E59D/D75Y. Binding curves (insets) are averaged over the several residues with substantial chemical shift changes, for which the residue assignments are indicated.



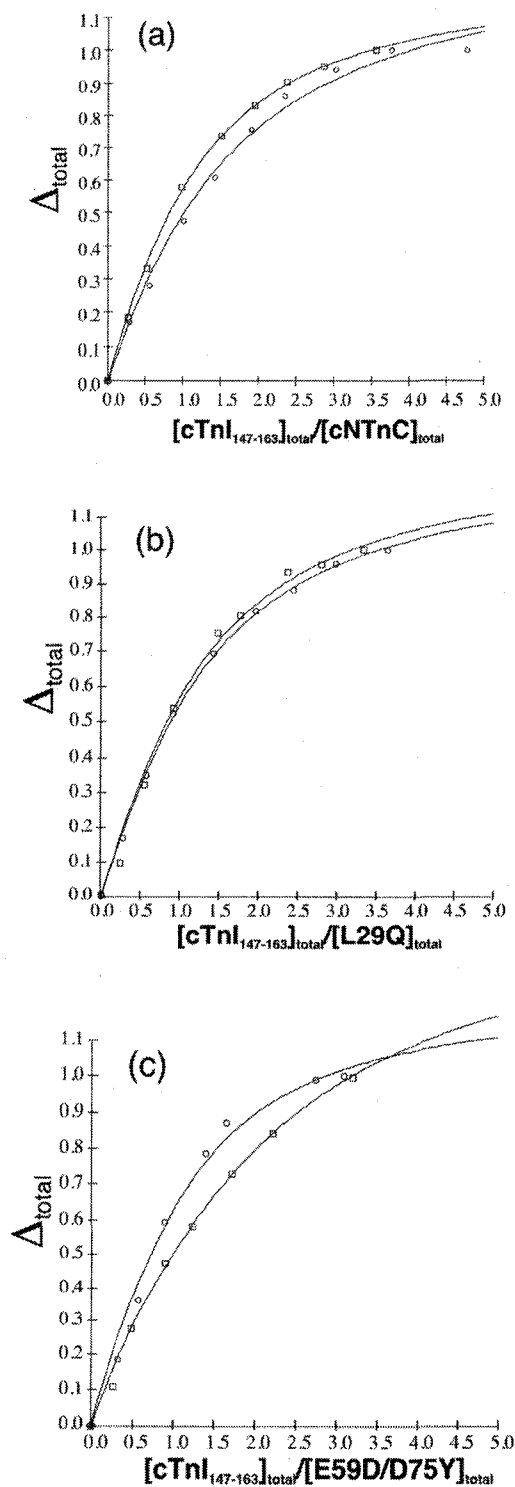
**Figure II-6.** The overlay of 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectra acquired during the cTnI<sub>147-163</sub> titration of cTnTc•Ca<sup>2+</sup> in the presence of cTnI<sub>1-29</sub> (a) and cTnI<sub>1-29</sub>PP (b), L29Q•Ca<sup>2+</sup> in the presence of cTnI<sub>1-29</sub> (c) and cTnI<sub>1-29</sub>PP (d), and E59D/D75Y•Ca<sup>2+</sup> in the presence of cTnI<sub>1-29</sub> (e) and cTnI<sub>1-29</sub>PP (f). Residue assignments are indicated for the same resonances in each panel.



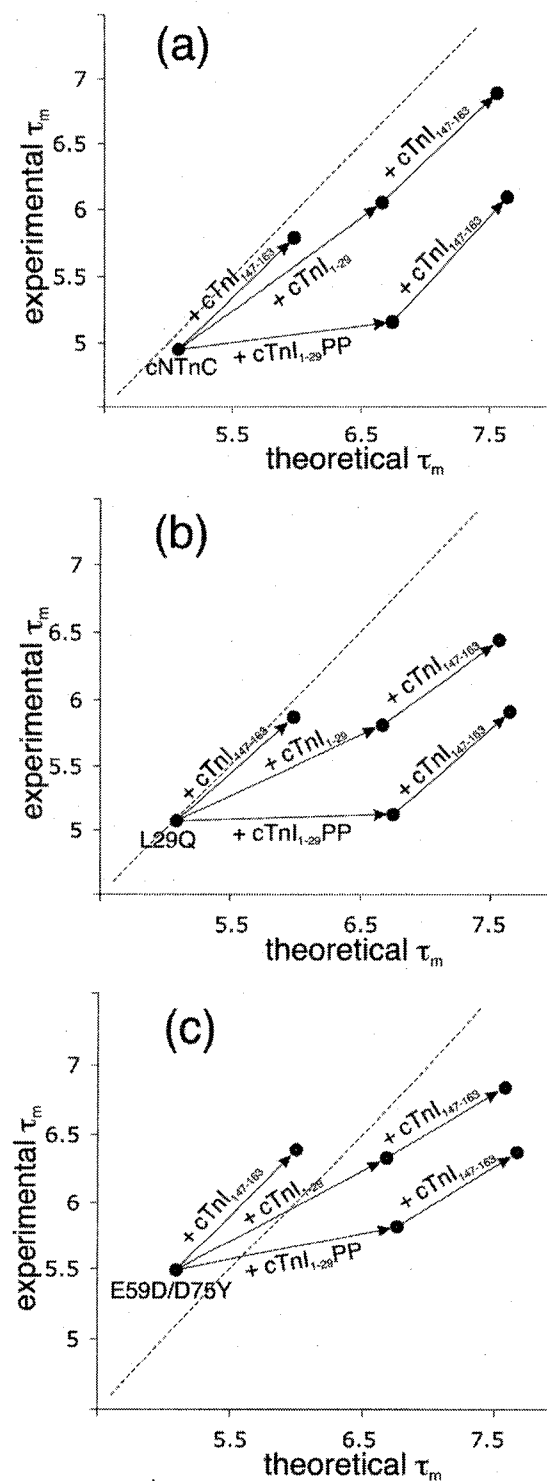




**Figure II-7.** Binding curves obtained during the  $cTnI_{147-163}$  titrations of  $cNTnC \cdot Ca^{2+}$  in the complex with  $cTnI_{1-29}$  (circles) and  $cTnI_{1-29}PP$  (squares) (a),  $L29Q \cdot Ca^{2+}$  in complex with  $cTnI_{1-29}$  (circles) and  $cTnI_{1-29}PP$  (squares) (b), and  $E59D/D75Y \cdot Ca^{2+}$  in complex with  $cTnI_{1-29}$  (circles) and  $cTnI_{1-29}PP$  (squares) (c).



**Figure II-8.** Experimental NMR rotational correlation times plotted versus theoretical estimates calculated based on the expected molecular weights for cNTnC (a), L29Q (b), and E59D/D75Y (d) in various complexes. Complexes below diagonal line are formed with the stoichiometry less than 1:1. Complexes above diagonal line most likely undergo dimerization. In the case of E59D/D75Y, the suspected dimerization is mild.





## References

- (1) Schmidtman, A., Lindow, C., Villard, S., Heuser, A., Mugge, A., Gessner, R., Granier, C., and Jaquet, K. (2005) Cardiac troponin C-L29Q, related to hypertrophic cardiomyopathy, hinders the transduction of the protein kinase A dependent phosphorylation signal from cardiac troponin I to C. *Febs J* 272, 6087-97.
- (2) Biesiadecki, B. J., Kobayashi, T., Walker, J. S., John Solaro, R., and de Tombe, P. P. (2007) The Troponin C G159D Mutation Blunts Myofilament Desensitization Induced by Troponin I Ser23/24 Phosphorylation. *Circ Res* 100, 1486-1493.
- (3) 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.
- (4) Churcott, C. S., Moyes, C. D., Bressler, B. H., Baldwin, K. M., and Tibbits, G. F. (1994) Temperature and pH effects on Ca<sup>2+</sup> sensitivity of cardiac myofibrils: a comparison of trout with mammals. *Am J Physiol* 267, R62-70.
- (5) Gomes, A. V., and Potter, J. D. (2004) Molecular and cellular aspects of troponin cardiomyopathies. *Ann N Y Acad Sci* 1015, 214-24.
- (6) Lim, C. C., Yang, H., Yang, M., Wang, C. K., Shi, J., Berg, E. A., Pimentel, D. R., Gwathmey, J. K., Hajjar, R. J., Helmes, M., Costello, C. E., Huo, S., and Liao, R. (2008) A Novel Mutant Cardiac Troponin C Disrupts Molecular Motions Critical For Calcium Binding Affinity And Cardiomyocyte Contractility. *Biophys J* 94, 3577-89.
- (7) Li, M. X., Spyropoulos, L., and Sykes, B. D. (1999) Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. *Biochemistry* 38, 8289-98.
- (8) Abbott, M. B., Dong, W. J., Dvoretzky, A., DaGue, B., Caprioli, R. M., Cheung, H. C., and Rosevear, P. R. (2001) Modulation of cardiac troponin C-cardiac troponin I regulatory interactions by the amino-terminus of cardiac troponin I. *Biochemistry* 40, 5992-6001.
- (9) Ward, D. G., Cornes, M. P., and Trayer, I. P. (2002) Structural consequences of cardiac troponin I phosphorylation. *J Biol Chem* 277, 41795-801.
- (10) Li, M. X., Gagne, S. M., Spyropoulos, L., Klocks, C. P., Audette, G., Chandra, M., Solaro, R. J., Smillie, L. B., and Sykes, B. D. (1997) NMR studies of Ca<sup>2+</sup> binding to the regulatory domains of cardiac and E41A skeletal muscle troponin C reveal the importance of site I to energetics of the induced structural changes. *Biochemistry* 36, 12519-25.
- (11) Linse, S. (2002) in *Methods in Molecular Biology* (Vogel, H. J., Ed.), Humana Press, Totowa, New Jersey.
- (12) Blumenschein, T. M., Gillis, T. E., Tibbits, G. F., and Sykes, B. D. (2004) Effect of temperature on the structure of trout troponin C. *Biochemistry* 43, 4955-63.
- (13) Solaro, R. J., Moir, A. J., and Perry, S. V. (1976) Phosphorylation of troponin I and the inotropic effect of adrenaline in the perfused rabbit heart. *Nature* 262, 615-7.
- (14) Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 104, 3623-40.
- (15) 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.

- (16) Goodman, J. L., Pagel, M. D., and Stone, M. J. (2000) Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J Mol Biol* 295, 963-78.
- (17) Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by <sup>15</sup>N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28, 8972-9.
- (18) Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by <sup>15</sup>N NMR relaxation. *Biochemistry* 33, 5984-6003.
- (19) Lindhout, D. A., Boyko, R. F., Corson, D. C., Li, M. X., and Sykes, B. D. (2005) The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C. *Biochemistry* 44, 14750-9.
- (20) Baryshnikova, O. K., and Sykes, B. D. (2006) Backbone dynamics of SDF-1 $\alpha$  determined by NMR: interpretation in the presence of monomer-dimer equilibrium. *Protein Sci* 15, 2568-78.
- (21) Schurr, J. M., Babcock, H. P., and Fujimoto, B. S. (1994) A test of the model-free formulas. Effects of anisotropic rotational diffusion and dimerization. *J Magn Reson B* 105, 211-24.
- (22) Sia, S. K., Li, M. X., Spyropoulos, L., Gagne, 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, 18216-21.
- (23) Ray, K. P., and England, P. J. (1976) Phosphorylation of the inhibitory subunit of troponin and its effect on the calcium dependence of cardiac myofibril adenosine triphosphatase. *FEBS Lett* 70, 11-6.
- (24) Holroyde, M. J., Howe, E., and Solaro, R. J. (1979) Modification of Calcium requirements for activation of cardiac myofibrillar ATPase by cyclic AMP dependent phosphorylation. *Biochim Biophys Acta* 586, 63-69.
- (25) Kentish, J. C., McCloskey, D. T., Layland, J., Palmer, S., Leiden, J. M., Martin, A. F., and Solaro, R. J. (2001) Phosphorylation of troponin I by protein kinase A accelerates relaxation and crossbridge cycle kinetics in mouse ventricular muscle. *Circ Res* 88, 1059-65.
- (26) Herron, T. J., Korte, F. S., and McDonald, K. S. (2001) Power output is increased after phosphorylation of myofibrillar proteins in rat skinned cardiac myocytes. *Circ Res* 89, 1184-90.
- (27) 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 Ca<sup>2+</sup>-binding properties of the Ca<sup>2+</sup>-regulatory site of bovine cardiac troponin. *J Biol Chem* 257, 260-3.
- (28) Ward, D. G., Brewer, S. M., Cornes, M. P., and Trayer, I. P. (2003) A cross-linking study of the N-terminal extension of human cardiac troponin I. *Biochemistry* 42, 10324-32.
- (29) Dong, W. J., Chandra, M., Xing, J., She, M., Solaro, R. J., and Cheung, H. C. (1997) Phosphorylation-induced distance change in a cardiac muscle troponin I mutant. *Biochemistry* 36, 6754-61.
- (30) Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) Effects of protein kinase A phosphorylation on signaling between cardiac troponin I and the N-terminal domain of cardiac troponin C. *Biochemistry* 36, 13305-11.
- (31) Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and

- Rosevear, P. R. (1999) NMR analysis of cardiac troponin C-troponin I complexes: effects of phosphorylation. *FEBS Lett* 453, 107-12.
- (32) Warren, C. M., Kobayashi, T., and Solaro, R. J. (2007) in *Biophysical Society, Annual Meeting* pp 473a, Baltimore, Maryland.
  - (33) Gaponenko, V., Abusamhadneh, E., Abbott, M. B., Finley, N., Gasmi-Seabrook, G., Solaro, R. J., Rance, M., and Rosevear, P. R. (1999) Effects of troponin I phosphorylation on conformational exchange in the regulatory domain of cardiac troponin C. *J Biol Chem* 274, 16681-4.
  - (34) Wang, Y. P., and Fuchs, F. (1995) Osmotic compression of skinned cardiac and skeletal muscle bundles: effects on force generation, Ca<sup>2+</sup> sensitivity and Ca<sup>2+</sup> binding. *J Mol Cell Cardiol* 27, 1235-44.
  - (35) Pan, B. S., and Solaro, R. J. (1987) Calcium-binding properties of troponin C in detergent-skinned heart muscle fibers. *J Biol Chem* 262, 7839-49.
  - (36) Engel, P. L., Kobayashi, T., Biesiadecki, B., Davis, J., Tikunova, S., Wu, S., and Solaro, R. J. (2007) Identification of a region of troponin I important in signaling cross-bridge-dependent activation of cardiac myofilaments. *J Biol Chem* 282, 183-93.
  - (37) Solzin, J., Iorga, B., Sierakowski, E., Gomez Alcazar, D. P., Ruess, D. F., Kubacki, T., Zittrich, S., Blaudeck, N., Pfitzer, G., and Stehle, R. (2007) Kinetic mechanism of the Ca<sup>2+</sup>-dependent switch-on and off of cardiac troponin in myofibrils. *Biophys J* 93, 3917-31.
  - (38) Shiner, J. S., and Solaro, R. J. (1982) Activation of thin-filament-regulated muscle by calcium ion: considerations based on nearest-neighbor lattice statistics. *Proc Natl Acad Sci U S A* 79, 4637-41.
  - (39) Strynadka, N. C., and James, M. N. (1989) Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu Rev Biochem* 58, 951-98.
  - (40) Gillis, T. E., Moyes, C. D., and Tibbits, G. F. (2003) Sequence mutations in teleost cardiac troponin C that are permissive of high Ca<sup>2+</sup> affinity of site II. *Am J Physiol Cell Physiol* 284, C1176-84.
  - (41) Muir, L. A., Tschirgi, M. L., Rajapakse, I., and Chandra, M. (2006) in *Biophysical Society, Annual Meeting* pp 529, Salt Lake City, Utah.
  - (42) Dweck, D., Gomes, A. V., and Potter, J. D. (2005) in *Biophysical Society, Annual Meeting* pp 317a, Long Beach, CA.
  - (43) Pearlstone, J. R., Chandra, M., Sorenson, M. M., and Smillie, L. B. (2000) Biological function and site II Ca<sup>2+</sup>-induced opening of the regulatory domain of skeletal troponin C are impaired by invariant site I or II Glu mutations. *J Biol Chem* 275, 35106-15.
  - (44) Li, M. X., Corson, D. C., and Sykes, B. D. (2002) Structure determination by NMR. Isotope labeling. *Methods Mol Biol* 173, 255-65.
  - (45) Li, M. X., Gagne, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) Calcium binding to the regulatory N-domain of skeletal muscle troponin C occurs in a stepwise manner. *Biochemistry* 34, 8330-40.
  - (46) Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) Mapping of a second actin-tropomyosin and a second troponin C binding site within the C terminus of troponin I, and their importance in the Ca<sup>2+</sup>-dependent regulation of muscle contraction. *J Mol Biol* 271, 728-50.
  - (47) Tsien, R. Y. (1980) New calcium indicators and buffers with high selectivity against magnesium and protons: design, synthesis, and properties of prototype structures. *Biochemistry* 19, 2396-404.

- (48) 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, 277-93.
- (49) Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603-614.
- (50) Spyropoulos, L. (2006) A suite of Mathematica notebooks for the analysis of protein main chain <sup>15</sup>N NMR relaxation data. *J Biomol NMR* 36, 215-24.

## CHAPTER III

### Consequences of the DCM mutation G159D on the structure and dynamics of cardiac troponin C and its interaction with troponin I\*

#### Outline

NMR spectroscopy has been employed to elucidate the molecular consequences of the DCM mutation G159D on structure and dynamics of troponin C, and its interaction with troponin I. Since the effects of this human mutation are expected to be subtle, all NMR experiments were run as direct side-by-side comparisons of the wild type and G159D mutant proteins. The affinity of G159D towards the anchoring region of cTnI (cTnI<sub>34-71</sub>) was reduced ( $K_D = 6 \pm 3 \mu\text{M}$ ) as compared to that for cCTnC ( $K_D < 1 \mu\text{M}$ ). The overall structure and dynamics of the G159D•cTnI<sub>34-71</sub> complex were very similar to those of cCTnC•cTnI<sub>34-71</sub>. There were, however, significant changes in the <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N NMR chemical shifts of troponin C in the complex, especially for the residues involved in the direct contact with cTnI<sub>34-71</sub>, and changes in the NOE connectivity patterns between G159D and cTnI<sub>34-71</sub>. Thus, one possibility is that the development of disease results from the poor anchoring of the cTnI to cCTnC, with the resulting increase in acto-myosin inhibition in agreement with physiological data. Another possibility is that long range electrostatic interactions affect the binding of the switch region of cTnI (cTnI<sub>147-163</sub>) and/or the cardiac specific N-terminus of cTnI (cTnI<sub>1-29</sub>) to the N-domain of cTnC. These important interactions are all spatially close in the X-ray structure of the core of cardiac TnC (1).

#### Introduction

Ca<sup>2+</sup> regulation of contraction in cardiac muscle is a fine tuned process that depends on the reliable function of its protein parts. The development of cardiac malfunctions can be caused by the presence of mutations in contractile proteins. Indeed, as revealed by genetic analysis, a large number of mutations in  $\beta$ -myosin,  $\alpha$ -tropomyosin, cardiac troponin I (cTnI), and cardiac troponin T (cTnT) causes various forms of cardiomyopathy (2, 3). Cardiac troponin C (cTnC), the protein directly responsible for Ca<sup>2+</sup> sensitivity of the thin filament, is highly conserved among mammals and thought to have fewer mutations until recently (4-7).

cTnC consists of two domains: the structural domain, cCTnC, and the regulatory domain, cNTnC. At all times, cCTnC, stays saturated with  $\text{Ca}^{2+}$  and bound to the anchoring region of cTnI (~ residues 34-71), which tethers cTn to the thin filament. The regulatory domain, cNTnC, binds  $\text{Ca}^{2+}$  only at high  $\text{Ca}^{2+}$  concentration upon muscle stimulation. In  $\text{Ca}^{2+}$  form, cNTnC binds to the switch region of cTnI (~ residues 147-163), which pulls the inhibitory region of cTnI (~ residues 128-147) away from its binding site on actin, triggering contraction (for reviews see (8, 9)).

Two mutations found in cNTnC, L29Q and E59D/D75Y, were linked to hypertrophic cardiomyopathy (FHC) and dilated cardiomyopathy (DCM), respectively (4, 5). In Chapter II, the physiological consequences of the L29Q and E59D/D75Y mutations were linked to the impaired thermodynamics of the cNTnC interactions with its binding partners, cTnI<sub>1-29</sub> and cTnI<sub>147-163</sub>. Another mutation, G159D, associated with DCM, was found in cCTnC (7). In reconstituted rat (10) and rabbit (11) cardiac trabeculae, the presence of G159D did not significantly alter myofilament  $\text{Ca}^{2+}$  sensitivity. However, the presence of G159D in guinea pig trabeculae resulted in the reduced  $\text{Ca}^{2+}$  sensitivity (12). Available physiological data on G159D reported the reduced activity of ATPase (12); however, the  $\text{Ca}^{2+}$  sensitivity of ATPase was not affected (10). It might be difficult to interpret physiological results on the molecular level, especially since  $\text{Ca}^{2+}$  sensitivity is a complex function of many consequent events, such as the binding of  $\text{Ca}^{2+}$ , binding of cTnI<sub>147-163</sub> and attachment of cross-bridges, and since the reproducibility of results was mentioned to be dependent on the muscle fiber type (11). We have attempted to address this question using the G159D mutant of cCTnC and the cTnI peptide (cTnI<sub>34-71</sub>), so that the interactions between troponin parts can be unambiguously traced. The knowledge of these interactions can be afterwards compiled into the overall picture assisting the interpretation of the physiological data.

There are several functions attributed to cCTnC, which could be affected by the presence of G159D: first, in its  $\text{Ca}^{2+}$  bound form, cCTnC binds to the anchoring region of cTnI (~ residues 34-71) and tethers the entire troponin complex to the thin filament; second, cCTnC binds the inhibitory region of cTnI (~ residues 128-147), the location of which has not been visualized in available crystal structures (1) and is suspected to be driven by electrostatics (13, 14). The binding of cTnI<sub>128-147</sub> occurs without replacement of cTnI<sub>34-71</sub> and is hypothesized to be localized next to the highly charged E-D linker connecting the two domains of cTnC. Third, cCTnC was also shown to interact with cTnT (15). Fourth, G159D might affect the interaction with the cardiac specific N-

terminus of cTnI (~ residues 1-29), as was suspected in experiments with skinned rat trabeculae (10) and with reconstituted troponin complex (16). In the crystal structure, Gly<sup>159</sup> is located at the very end of the H-helix, 5 Å away from Leu<sup>47</sup> of the anchoring region of cTnI (~residues 34-71) (1). This suggests that the presence of mutation might first of all affect the interaction with cTnI<sub>34-71</sub>, which has not been addressed previously. Using NMR spectroscopy, we have determined the affinity of G159D towards cTnI<sub>34-71</sub> and examined the consequences of this mutation on structure and dynamics of the G159D•cTnI<sub>34-71</sub> complex. Based on NOE connectivity patterns, NMR chemical shifts changes, dynamics, and determined 3D structures, we conclude that, compared to cCTnC, G159D binds to cTnI<sub>34-71</sub> in a similar fashion, albeit with a lower affinity. This interaction is a key function of cCTnC, and when impaired, can perturb the function of the entire Tn complex and lead to severe physiological consequences, such as the development of cardiomyopathy.

## Results

### *cTnI<sub>34-71</sub> titrations of cCTnC and G159D*

In a 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectrum, cross peaks correlate the <sup>15</sup>N and <sup>1</sup>H chemical shifts of backbone amide NH. Chemical shifts are exceptionally sensitive to changes in the structure and dynamics of a protein. 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC spectra obtained for G159D (Fig. III-1a) and cCTnC (Fig. III-1b), with and without their binding partner, cTnI<sub>34-71</sub>, demonstrate the difference in the positions of several {<sup>1</sup>H, <sup>15</sup>N} HSQC cross peaks, corresponding to the residues adjacent to the site of mutation. The rest of the amide chemical shifts, and the changes they undergo during the titration with cTnI<sub>34-71</sub>, are similar, suggesting resemblance in the overall structural folds for G159D and cCTnC, in agreement with the structures determined herein (see below). The kinetics of the binding of cTnI<sub>34-71</sub> to G159D (Fig. III-1a) and cCTnC (Fig. III-1b) is on the NMR slow exchange time scale, in which the intensities of cross-peaks for complexed and uncomplexed species vary depending on the degree of binding. Binding curves were obtained by measuring the intensity decrease for several representative cross-peaks corresponding to the unliganded protein, which were averaged and fit to a single binding site model (Fig. III-1, insets).  $K_D$ 's were found to be  $6 \pm 3 \mu\text{M}$  in the case of G159D and  $< 1 \mu\text{M}$  in the case of cCTnC. The later is in agreement with literature data (17).  $K_D$ 's tighter than  $1 \mu\text{M}$  cannot be determined precisely within the concentration limits of NMR.

### <sup>15</sup>N backbone relaxation data

<sup>15</sup>N backbone NMR relaxation parameters,  $T_1$ ,  $T_2$ , and NOE, are the functions of the protein motions, including the tumbling of a protein as a whole, the motions of its independent parts, and the chemical exchange processes. Per residue  $T_1$ ,  $T_2$ , and NOE were determined for G159D and cCTnC alone, and in complex with cTnI<sub>34-71</sub> (Fig. III-2).  $T_1$ ,  $T_2$ , and NOE for cCTnC•cTnI<sub>34-71</sub> and G159D•cTnI<sub>34-71</sub> were virtually identical within the errors of the experiments (Fig. III-2), implying that the presence of the mutation did not affect the dynamic properties, as well as the conformational flexibility, of the complex. Importantly, <sup>15</sup>N backbone  $T_1$ ,  $T_2$ , and NOE for cCTnC•cTnI<sub>34-71</sub> and G159D•cTnI<sub>34-71</sub> were also similar in the area adjacent to the mutation site. Relaxation data for cCTnC was in agreement with qualitative data presented earlier (18, 19).

$T_1$ ,  $T_2$ , and NOE are per residue parameters, whereas the overall correlation time,  $\tau_m$ , calculated from  $T_1$ ,  $T_2$ , and NOE (Table III-1), is the function of the hydrated radius of the entire protein,  $r_h$ , and consequently its molecular weight. According to the Stokes-Einstein-Debye theory (20),

$$\tau_m = \frac{4}{3} \pi \eta r_h^3 / k_B T,$$

where  $\eta$  is the solvent viscosity,  $k_B$  is the Boltzmann constant, and  $T$  is temperature. As a rule of thumb,  $\tau_m$  (ns) is approximated to the half of the molecular weight of a protein (kDa), provided that the protein tumbling is mostly isotropic. This can serve as a good estimate of the molecular weight of a protein or its complexes (13). An increase in  $\tau_m$ , as compared to the calculated molecular weight can result from issues such as dimerization (21), which is the case for skeletal protein, sCTnC (22). We have determined  $\tau_m$ 's for G159D and cCTnC alone, and in complex with cTnI<sub>34-71</sub> (Table III-1). For cCTnC and G159D,  $\tau_m$ 's are comparable to each other and smaller than the rule of thumb estimate, which could be explained by a lesser degree of bound water, for example (23, 24). The values of  $\tau_m$ 's determined for G159D•cTnI<sub>34-71</sub> and cCTnC•cTnI<sub>34-71</sub> were comparable to each other and to the theoretical estimate, which suggests the formation of 1:1 complexes in both cases, and the absence of dimerization and significant conformational exchange.



**Table III-1.** NMR relaxation parameters  $T_1$ ,  $T_2$ , and  $\tau_m$  determined for G159D and cCTnC alone and in complexes with cTnI<sub>34-71</sub>. Theoretical estimate for correlation time  $\tau_m$  is equal to the half of the molecular weight in kDa.

	Average of per residue $T_1$ (ms)	Average of per residue $T_2$ (ms)	Overall correlation time, $\tau_m$ (ns)	Theoretical estimate for correlation time, $\tau_m$ (ns)
G159D	356 ± 47	193 ± 31	3.3	4.2
G159D•cTnI <sub>34-71</sub>	436 ± 46	112 ± 9	6.6	6.4
cCTnC	348 ± 42	185 ± 25	3.4	4.2
cCTnC•cTnI <sub>34-71</sub>	438 ± 53	109 ± 13	6.5	6.4

*Differences in the chemical shifts of cCTnC and G159D in the complexes with cTnI<sub>34-71</sub>*

Chemical shifts changes are uniquely indicative of changes in chemical environment of nuclei and hence, in protein structure and dynamics. We have calculated the chemical shift changes of  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and  $^1\text{H}$  for G159D and cCTnC in complexes with cTnI<sub>34-71</sub> (Figs. III-3 and III-4). Backbone amide  $^{15}\text{N}$  chemical shifts underwent the most significant perturbations at the end of the H-helix near to the mutation site (Figs. III-3a and III-4a).  $^{15}\text{N}$  chemical shifts in the region where the H-helix is contacting the E-helix (residues 93-97) are also affected by the presence of mutation but to a much lesser degree. As demonstrated by  $^{15}\text{N}$  chemical shifts changes, the overall fold of G159D and cCTnC remains similar with the only significant perturbation being near to the mutation site. Changes in  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts, especially those for the side chains, are sensitive not only to the changes in overall fold but also to the changes in the configurational patterns, e.g. the number, energetics, and the geometry of connections that stabilize protein core or protein complex. As demonstrated by  $^{13}\text{C}$  chemical shifts changes (Figs. III-3b and III-4b), the major perturbations occurred at the several hydrophobic residues (for example Leu<sup>100</sup>, Met<sup>120</sup>, Leu<sup>121</sup>, and Ile<sup>128</sup>) known to be responsible for the interaction with cTnI<sub>34-71</sub> (I).  $^1\text{H}$  atoms directly responsible for the majority of van der Waals and hydrophobic contacts in proteins, are the most sensitive indicators of changes in protein core and binding interfaces. As demonstrated by  $^1\text{H}$  chemical shifts differences (Figs. III-3c and III-4c), significant changes occurred with the residues forming the interface between cCTnC and cTnI<sub>34-71</sub> (for example Leu<sup>100</sup>, Met<sup>103</sup>, Met<sup>120</sup>, Leu<sup>121</sup>, Ile<sup>128</sup>, Leu<sup>136</sup>), implying that the binding of cTnI<sub>34-71</sub> to cCTnC is perturbed in the presence of the mutation (I). Interestingly, the chemical shifts of two interior residues, Ile<sup>148</sup> and Ile<sup>112</sup>, which form the hydrophobic core of cCTnC, have also been affected.

#### *Differences in the NOE connectivities of cCTnC and G159D in the complex with cTnI<sub>34-71</sub>*

NOE connectivities obtained from 3D  $^{13}\text{C}$  NOESY HSQC and  $^{15}\text{N}$  NOESY HSQC NMR spectra are typically used as distance restraints during structure calculations. It is often informative therefore to make pair-wise comparisons of NOE patterns. NOE, however, is a less sensitive parameter to long range perturbations as compared to a chemical shift since the intensity of NOE decreases rapidly with the distance between two nuclei:  $\text{NOE} \propto \frac{1}{r^6}$  (25). We have compared strips from 3D  $^{13}\text{C}$  NOESY HSQC and  $^{15}\text{N}$  NOESY HSQC NMR spectra for G159D and cCTnC for residues with the significant chemical shift changes. Spectra were acquired using identical parameters, such as mixing time, which is crucial for such a comparison. The vast majority of contacts in the cCTnC spectra were accurately reproduced in the G159D spectra; however, there were some differences (Fig. III-5). For example, in the  $^{13}\text{C}$  NOESY HSQC spectrum, the  $\gamma$ -proton of Met<sup>103</sup> has more contacts in cCTnC compared to G159D, with some of these contacts being tentatively assigned to the peptide resonances (Fig. III-5a). The  $\gamma$ -proton of Ile<sup>112</sup> contacted the  $\alpha$ -proton of Asp<sup>149</sup> in G159D, and the  $\alpha$ -proton of Asp<sup>105</sup> and the  $\delta$ -proton of Tyr<sup>150</sup> in cCTnC (Fig. III-5b). The  $\alpha$ -proton of Gly<sup>140</sup> contacted the  $\gamma$ -proton of Glu<sup>152</sup> in cCTnC but not in G159D (Fig. III-5c). The  $\gamma$ -proton of Val<sup>160</sup> contacted the  $\epsilon$ -proton of Phe<sup>156</sup> in G159D but not in cCTnC (Fig. III-5d). In the  $^{15}\text{N}$  NOESY HSQC NMR spectrum, the differences in NOE patterns were less significant, with the noticeable change observed in the amide proton of Ala<sup>108</sup>, which had contacts with the  $\beta$ -protons of Lys<sup>106</sup> and Asn<sup>107</sup> in G159D but not in cCTnC. These results demonstrate that the residues with the large chemical shift differences (Fig. III-5) had also distinct NOE connectivities in  $^{13}\text{C}$  NOESY HSQC and  $^{15}\text{N}$  NOESY HSQC spectra.

#### *Structures of cCTnC and G159D in the complexes with cTnI<sub>34-71</sub>*

The ensemble containing 20 best NMR structures out of 50 calculated was analyzed for G159D and cCTnC. All structures had a good geometry after water refinement with > 92 % in most favorable and > 7 % in additional allowed regions of the Ramachandran plot for G159D and cCTnC (Table III-2). Structures of both proteins converged with rmsd < 1.4 Å for all atoms for residues 95-155, including loops. The overall topology and the secondary structure for G159D and cCTnC were similar to the cCTnC structure obtained by X-ray crystallography, PDB code 1J1D (1) and previously by NMR, PDB code 1FI5 (19). Rmsd's for backbone atoms between cCTnC and 1J1D

was 1.2 Å for all residues, and between cCTnC and 1FI5 1.4 Å (Fig. III-6a). Similar rmsd of 1.7 Å was between reported 1FI5 and 1J1D. The G159D structure, although similar overall, overlapped with 1J1D with rmsd of 1.7 Å for backbone atoms for all residues (Fig. III-6b). G159D overlapped with cCTnC with rmsd of 1.4 Å and with 1FI5 with 1.3 Å (Fig. III-6b). The resolution of X-ray diffraction data for 1J1D was 2.6 Å (1), within which, structures determined in this work and reported previously have similar folds.

**Table III-2.** Statistics for 20 NMR structures of G159D and cCTnC in complex with cTnI<sub>34-71</sub>. Structures were calculated using CYANA (26), refined using XPLOR-NIH (27, 28), and analyzed using PROCHECK (29) and WHATCHECK (30).

	G159D	cCTnC
NOE restraints	1075	957
Short range ( $ i-j =1$ )	624	561
Medium range ( $1< i-j <5$ )	231	215
Long range ( $ i-j \geq 5$ )	220	181
Ca <sup>2+</sup> binding restraints	16	16
Dihedral restraints ( $\phi/\psi$ ) <sup>a</sup>	84	86
NOE violations		
> 0.5 Å	0.15 ± 0.37	0.0 ± 0.0
> 0.3 Å	0.25 ± 0.55	1.8 ± 1.20
> 0.1 Å	6.3 ± 1.95	10.45 ± 2.28
Dihedral Violations (°)	0.0 ± 0.0	0.0 ± 0.0
Ramachandran plot statistics <sup>b</sup>		
$\phi/\psi$ in most favorable regions (%)	93 ± 4	92 ± 4
$\phi/\psi$ in additionally allowed regions (%)	7 ± 3	7 ± 3
$\phi/\psi$ in generously allowed regions (%)	0 ± 1	1 ± 1
$\phi/\psi$ in disallowed regions (%)	0 ± 1	1 ± 1
Pairwise RMSD (Å) <sup>c</sup>		
Before water refinement	0.71 ± 0.13	0.59 ± 0.12
After water refinement	1.29 ± 0.28	1.31 ± 0.24
WHAT CHECK structure Z-scores after water refinement		
Second-generation packing quality	-0.9 ± 0.4	-0.9 ± 0.3
Ramachandran plot appearance	0.9 ± 0.7	1.8 ± 0.8

<sup>a</sup>Predicted from chemical shifts using TALOS

<sup>b</sup>PROCHECK was used over all the residues (90-161)

<sup>c</sup>Main chain nuclei over the residues 95-155

The packing of the cCTnC core occurs between the hydrophobic residues of  $\alpha$ -helical regions (Leu<sup>97</sup>, Phe<sup>101</sup>, Leu<sup>117</sup>, Met<sup>120</sup>, Leu<sup>121</sup>, Ile<sup>133</sup>, Leu<sup>136</sup>, Phe<sup>153</sup>, and Met<sup>157</sup>) and the two residues of the  $\beta$ -sheet region (Ile<sup>112</sup> and Ile<sup>148</sup>). Orientation of these key residues did not change in the cCTnC and G159D structures calculated in this work as compared

to 1J1D and 1FI5. Several residues such as Phe<sup>104</sup> and Phe<sup>156</sup> in cCTnC and G159D are lacking the restraints with the peptide and thus their positions are more flexible, but are in the good agreement with the NMR structure, 1FI5. The packing interactions between helices and the protein cores of cCTnC have not been affected by the presence of G159D as demonstrated by the resultant structures. In spite of some differences in NOE contacts for Ile<sup>112</sup> and Ile<sup>148</sup>, the calculated structures were similar within the limits of the method. This demonstrates that among the restraints used (Table III-2), many were identical, which determined the similarity between the structures.

*Differences in the NOE connectivities between cCTnC (and G159D) and cTnI<sub>34-71</sub> in <sup>13</sup>C-edited, filtered NOESY spectra*

The NOE connectivities between <sup>15</sup>N, <sup>13</sup>C-labeled cCTnC and the unlabeled peptide cTnI<sub>34-71</sub> were obtained using the <sup>13</sup>C-edited, filtered NOESY pulse sequence (31) optimized in our laboratory for this particular complex (32). In this experiment, each cross peak corresponds to a <sup>1</sup>H protein resonance on the vertical axis and to a <sup>1</sup>H peptide resonance on the horizontal axis. <sup>13</sup>C resonance in the third dimension connects to the corresponding <sup>1</sup>H resonances on the protein, which allows for the assignment. Peptide <sup>1</sup>H assignments can be typically obtained with the 2D <sup>15</sup>N, <sup>13</sup>C-edited TOCSY and 2D <sup>15</sup>N, <sup>13</sup>C-edited NOESY (33-35); however, in our case, few resonances can be unambiguously assigned due to the poor dispersion of peptide chemical shifts, which precluded the structure calculation for the entire complex. Similarly to previous NMR studies (19, 22, 36), structure determination of the entire complex was difficult to achieve without the additional restraints between a protein and a peptide, which might be found in crystal structures but are not always desirable to use due to possible introduction of artifacts. The peptide in complex is predominantly  $\alpha$ -helical, which is supported by the values of peptide chemical shifts in 2D <sup>15</sup>N, <sup>13</sup>C-edited TOCSY and 2D <sup>15</sup>N, <sup>13</sup>C-edited NOESY and their poor dispersion. The superposition of <sup>13</sup>C-edited, filtered NOESY spectra for G159D and cCTnC in complexes with cTnI<sub>34-71</sub> shows a similarity between contacts formed (Fig. III-7). Some contacts are virtually identical, for example, the contacts between Ala<sup>57</sup> of cTnI<sub>34-71</sub> and Leu<sup>100</sup> and Met<sup>120</sup> of cCTnC and G159D (Fig. III-7a) or the contact between Leu<sup>62</sup> of cTnI<sub>34-71</sub> and Met<sup>103</sup> of cCTnC and G159D (Fig. III-7c). Some contacts were slightly changed, for example, the contacts between Ile<sup>56</sup> of cTnI<sub>34-71</sub> and Thr<sup>124</sup> of cCTnC and G159D (Fig. III-7c). One contact, between Leu<sup>54</sup> of cTnI<sub>34-71</sub> and Val<sup>160</sup> of cCTnC, was clearly missing in G159D (Fig. III-7b). No contacts were observed

in the spectrum of G159D that were missing in case of cTnC. Also, the intensities of existing connectivities were not stronger in the case of G159D. This demonstrates that G159D interacts with cTnI<sub>34-71</sub> similarly to cTnC, without making new connections between the protein and the peptide and with the existing connections being qualitatively not stronger.

## Discussion

The task of defining the effect of mutations on protein behavior represents a substantial challenge, especially for human mutations where the effect on the protein has to be mild to allow an individual to function even in disease. Structural approaches, including the determination of structure using NMR or crystallography might not bring conclusive answers when the differences in structures of mutant and wild type proteins are not significant enough to be resolved within the resolution of the chosen technique. Often the affinities of mutated and wild type proteins for their targets and/or ligands are unchanged. It remains unknown, therefore, whether the mutant phenotype has been brought about by the changes undetectable within the limits of methods. However, the necessity to study individual human mutations is evident by the growing interest in genetically tailored nutrition and medicine, and overall will be beneficial to our understanding of current biochemical models, in our example, cardiac muscle contraction.

We have attempted to dissect the effect of the G159D mutation on the structure, dynamics, and interactions of troponin C using solution state NMR. This technique allows not only for the structure determination but also provides a wealth of information pertaining to the chemical environment and dynamics of individual residues. Since the effects are expected to be subtle, all NMR experiments were run as direct side-by-side comparisons of the wild type and G159D mutant proteins. We have found that the secondary structure, tertiary structure, and the packing of the protein core did not changed for G159D•cTnI<sub>34-71</sub> in comparison to cTnC•cTnI<sub>34-71</sub>. Dynamic behavior for G159D was also unaffected, even in the vicinity of the mutation. However, the affinity of G159D towards cTnI<sub>34-71</sub> was reduced, and the NMR chemical shifts of protein residues were perturbed. The differences in NMR chemical shifts were well beyond the resolution in <sup>1</sup>H dimension (0.03 ppm), <sup>13</sup>C dimension (0.3 ppm), and <sup>15</sup>N dimension (0.2 ppm) in our experiments, with the most significant changes observed for residues directly involved in binding to cTnI<sub>34-71</sub>. The K<sub>D</sub> of G159D towards cTnI<sub>34-71</sub> has been increased >

6 fold. We have also observed the changes in NOE connectivities between the protein and the peptide, with some connectivities changed or missing. All of the above evidence indicates the weakened binding and perturbed interaction between G159D and cTnI<sub>34-71</sub>. The structure and dynamics of the protein core remain the same, however, in agreement with the G159D mutation locating at the very end of the H-helix, away from the protein core.

The binding of cCTnC to cTnI<sub>34-71</sub> is a key function of the C-domain of cTnC. In skeletal TnC, the binding of the inhibitory region of sTnI is dependent upon the presence of the anchoring region of sTnI, with the anchoring region modulating the binding of the inhibitory region (37-39) which directly influences contractility. In insect flight muscle, the regulation of contraction is achieved entirely through the C-domain of TnC without the involvement of the N-domain and its binding to the region analogous to the switch region of cTnI (40). Our data supports the possibility that G159D might lead to the impaired contractility and the disease through the impaired anchoring of cTnC and enhanced inhibition of acto-myosin. When this interaction is impaired in G159D, the weaker binding to cTnI<sub>34-71</sub> might lead to a weaker anchoring of troponin C to the thin filament, lower  $k_{on}$  for the switch region of cTnI (~residues 147-163) due to the decrease in proximity effect, and a more pronounced inhibition of contraction. This is in agreement with a slower activation kinetics in rabbit psoas fibers (11), which might be a more important parameter than the steady state force development (41, 42). Our conclusions are also in agreement with the decrease in actin-tropomyosin activated ATPase rate and the decrease in sliding velocity (12).

The second possibility is that the presence of G159D impaired the function of cTnC via the perturbations in the long-range electrostatic interactions. Electrostatic potential drops off as a function of  $\frac{1}{r}$ , where  $r$  is the distance from the charge, so that it can be active 10-20 Å away from the site of mutation. In crystal structure (1), the N-domain of cTnC is folded over the C-domain of cTnC (Fig. III-8), so that Gly<sup>159</sup> is in proximity to the binding site for the switch region of cTnI (~residues 147-163) (43), to the possible binding site for the cardiac specific N-terminus of cTnI (~ residues 1-29) (44, 45), and to the binding site of the inhibitory region of cTnI (~residues 128-147). The binding of the cardiac specific N-terminus and the inhibitory regions of TnI are suspected to be electrostatically driven (14,18), as well as influenced by phosphorylation (46), and the change in electrostatic potential of cCTnC might affect these interactions. A

perturbed interaction of G159D with the N-terminus of cTnI has been reported (10, 16), which resulted in blunting the effect of PKA phosphorylation. Interestingly, the L29Q mutation has also blunted the effect of PKA phosphorylation (16, 47, 48) but lead to hypertrophy instead (4). The physical parameter crucial for the development of hyper- vs hypocontractility may be the resultant affinity towards the switch region of cTnI (~residues 147-163) (47) and the kinetics of this binding interaction (16).

In this work, we have determined that G159D binds to cTnI<sub>34-71</sub> with weaker affinity, supported by structural data, such as changes in chemical shifts and NOE connectivities. This weakened interaction may impair the anchoring of troponin C, with the resulting increase of acto-myosin inhibition. The G159D mutation may also perturb the binding of the adjacent N-terminus of cTnI (~ residues 1-29) to cTnNC through the long range electrostatic interactions, which will affect the cTnNC binding of the switch region of cTnI (~ residues 147-163) and overall contractility.

## Materials and methods

### *Protein Expression and Purification*

The DNA encoding cCTnC (91-161) was subcloned previously into pET-3a expression vector similarly to Pearlstone et al. (49). Using cCTnC as a template, the G159D mutation was engineered using a site-directed mutagenesis kit (QuikChange purchased from Stratagene). *E. coli* strain BL21(DE3)pLysS was transformed with the expression vector and incubated at 37°C to the OD<sub>600</sub> of 0.6-0.9. Cell cultures were induced with IPTG and harvested after incubation for 3 hours. Uniformly <sup>15</sup>N-labeled cCTnC and G159D were expressed in a minimal media enriched with (<sup>15</sup>NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (50). Cell pellet was lysed using French press and applied to a DEAE column (50 mM Tris, 0.1 M NaCl, pH 8.0). Proteins were further purified using a Superdex-75 size exclusion column (50 mM Tris, 0.15 M NaCl, pH 8.0), desalted using Sephadex G25 column (10 mM NH<sub>4</sub>HCO<sub>3</sub>), and lyophilized. Molecular masses for unlabeled proteins, cCTnC and G159D, as determined by MALDI mass spectrometry were equal to the expected values. Amino acid composition was confirmed independently by amino acid analysis.

### *cTnI<sub>34-71</sub> titration of cCTnC and G159D monitored by 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy*

The peptide, cTnI<sub>34-71</sub>, (acetyl-AKKKSKISASRKLQLKTLTLLQIAKQELEREA-EERRGEK-amide) was synthesized using standard methodology for a typical TnI peptide (51). The sequence was confirmed by amino acid analysis and the mass was verified by electrospray mass spectrometry. During titration, the peptide was added in the form of powder to the 500 µL solution of 0.57 mM cCTnC and 0.56 mM G159D. The buffer consisted of 100 mM KCl, 10 mM imidazole, 2 mM CaCl<sub>2</sub>, 5 mM NaN<sub>3</sub>, and 0.2 mM DSS in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. The pH was adjusted to 6.7 ± 0.05 at every titration point. Concentrations of proteins and peptides were determined by weight and confirmed by amino-acid analysis. Since the reaction proceeds within the slow exchange limits for both cCTnC and G159D, the chemical shifts for two species (cCTnC or G159D alone and in

complex with cTnI<sub>34-71</sub>) were observed during titrations. The decreasing intensities of uncomplexed cCTnC or G159D were used to calculate  $K_D$ .

#### *<sup>15</sup>N backbone amide NMR relaxation data*

Relaxation data were acquired at 30 °C on a Varian Inova 500 MHz spectrometer from <sup>15</sup>N-T<sub>1</sub>, <sup>15</sup>N-T<sub>2</sub>, and {<sup>1</sup>H-<sup>15</sup>N} NOE experiments (Biopack, Varian Associates) for 500 µL samples containing 0.57 mM of cCTnC or 0.56 mM G159D and 1.2 mM of cTnI<sub>34-71</sub>. T<sub>1</sub> data were acquired using relaxation delays of 10, 50, 100, 200, 300, and 400 ms, and T<sub>2</sub> data were acquired using relaxation delays of 10, 30, 50, 70, 90, and 110 ms. Delays between transients in <sup>15</sup>N-T<sub>2</sub> and <sup>15</sup>N-T<sub>1</sub> experiments were set to 3 s. {<sup>1</sup>H-<sup>15</sup>N} NOE experiments were performed using delays of 5 s for spectra recorded without proton saturation and delays of 2 s for spectra recorded with proton saturation. Proton saturation was set to 3 s so that the total time between transients was equal to 5 s. Relaxation parameters, T<sub>1</sub>, T<sub>2</sub>, and NOE, were extracted using the rate analysis module built in NMRView (52) and analyzed using Mathematica scripts provided by Dr. Leo Spyropoulos ([www.bionmr.ualberta.ca/~lspy/](http://www.bionmr.ualberta.ca/~lspy/)) (53). Overall rotational tumbling time,  $\tau_m$ , was averaged from the per residue fitting of relaxation data to the Lipari-Szabo  $S^2$ - $\tau_m$ - $\tau_j$  model (54). Relaxation data for residues with significant internal motions, as judged by NOE<0.65, and data for residues with possible slow motions, as judged by significantly decreased T<sub>2</sub>, were excluded in this method of  $\tau_m$  calculation.

#### *Assignment and structure calculation for cCTnC•cTnI<sub>34-71</sub> and G159D•cTnI<sub>34-71</sub>*

Samples used for NMR data acquisition contained ~0.5 mM of G159D or cCTnC, ~1.2 mM of cTnI<sub>34-71</sub> in the buffer consisted of 100 mM KCl, 10 mM imidazole, 2 mM CaCl<sub>2</sub>, 5 mM NaN<sub>3</sub>, and 0.2 mM DSS in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. Buffer used for data acquisition in D<sub>2</sub>O contained 100 mM KCl, 0.5 mM imidazole, 9.5 mM deuterated imidazole, 2 mM CaCl<sub>2</sub>, 5 mM NaN<sub>3</sub>, and 0.2 mM DSS. NMR spectra were acquired at 30°C on Varian INOVA 500, Unity 600, or INOVA 800 spectrometers equipped with 5 mm triple resonance probes and z axis pulsed field gradients (Table III-3). NMR data was processed using NMRPipe (55) and analyzed with NMRView (52). The backbone resonances for cCTnC and G159D in the [<sup>13</sup>C,<sup>15</sup>N] cCTnC•cTnI<sub>34-71</sub> and [<sup>13</sup>C,<sup>15</sup>N] G159D•cTnI<sub>34-71</sub> complexes were assigned using SmartNotebook v5.1.3 (56). Side chain assignment was carried out using HCCH-TOCSY, HCCONH, CCONH, HNHA, HNHB, and <sup>15</sup>N TOCSY HSQC. The majority of side chain resonances were unambiguously assigned. Resonances for aromatic residues were assigned using 2D homonuclear DQF-COSY and NOESY experiments in D<sub>2</sub>O. Resonances for the cTnI<sub>34-71</sub> peptide were not dispersed well enough in 2D <sup>13</sup>C/<sup>15</sup>N-filtered TOCSY and NOESY experiments to allow for the assignment of many residues except for amide protons, HN, for residues 41-64, H $\alpha$  for residues 41-59, H $\beta$  for residues 41-43, 50, 53, 56, and H $\delta/\gamma$  for residues 46-53, 55, and 58-60.

Structure calculations were performed with CYANA (26) using the 'noeassign' automatic assignment procedure (57) and distance restraints from <sup>13</sup>C NOESY HSQC and <sup>15</sup>N NOESY HSQC experiments. Unambiguous restraints were assigned manually and were forced to keep their assignments during the first four runs of CYANA calculations, after which they were open for automatic assignment by CYANA. Distance restraints were calibrated with CYANA standard procedure using upper limits of 6 Å. TALOS dihedral restraints for helical regions for G159D and cCTnC (58), and Ca<sup>2+</sup> restraints for Ca<sup>2+</sup> binding loops were obtained from X-ray crystallographic data and added into the calculation (Table III-3). TALOS restraints did not affect the overall fold of the structures (data not shown), however, in case of G159D, in final structure calculations the restraints



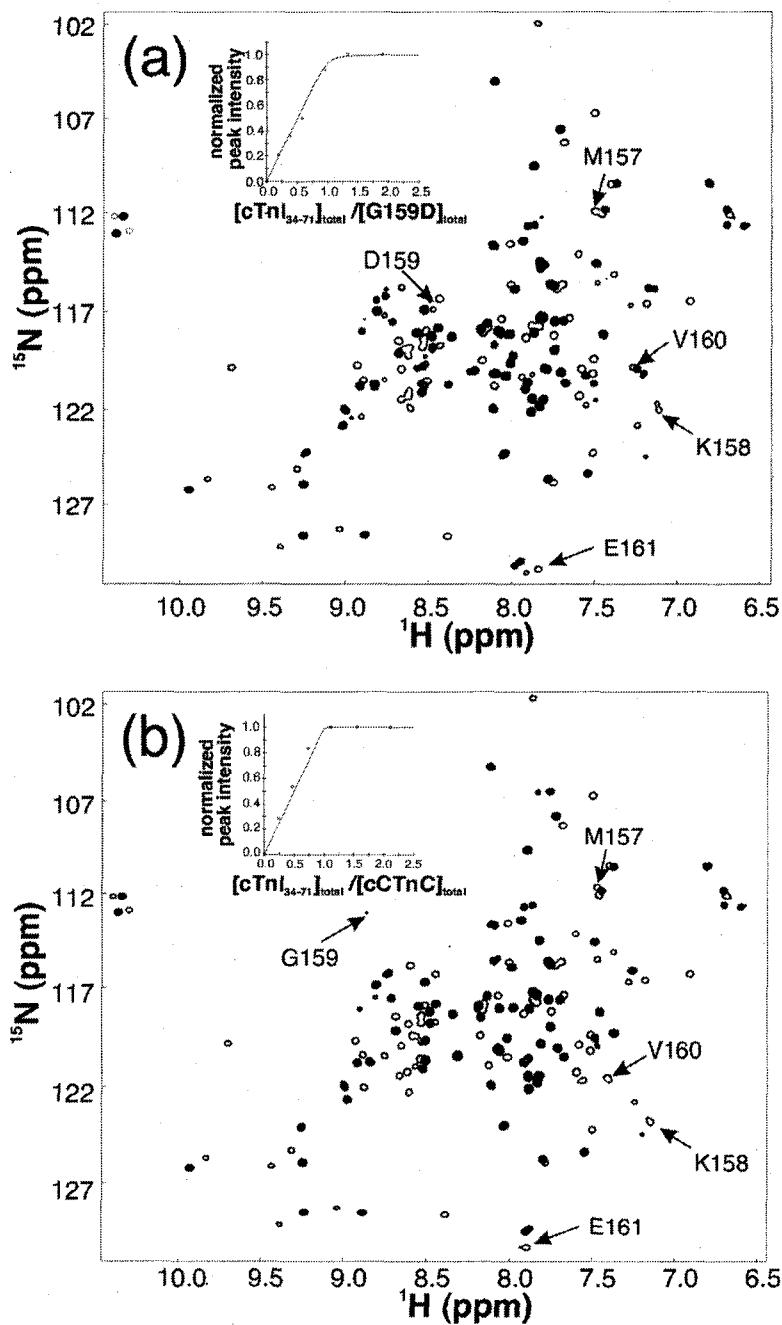
were truncated up to residue 151, to avoid artifactual secondary structure. CYANA was used to calculate 50 structures, and the 20 conformers with the lowest target function were refined in explicit solvent by XPLOR-NIH (27, 28) with a water box of 18.856 Å. The final ensembles were averaged and refined using the same water refinement protocol as done previously, and are the structures discussed herein and deposited (PDBs). Structures were validated using PROCHECK (29) and WHATCHECK (30).

**Table III-3.** NMR experiments acquired for structure calculations and chemical shift assignments.

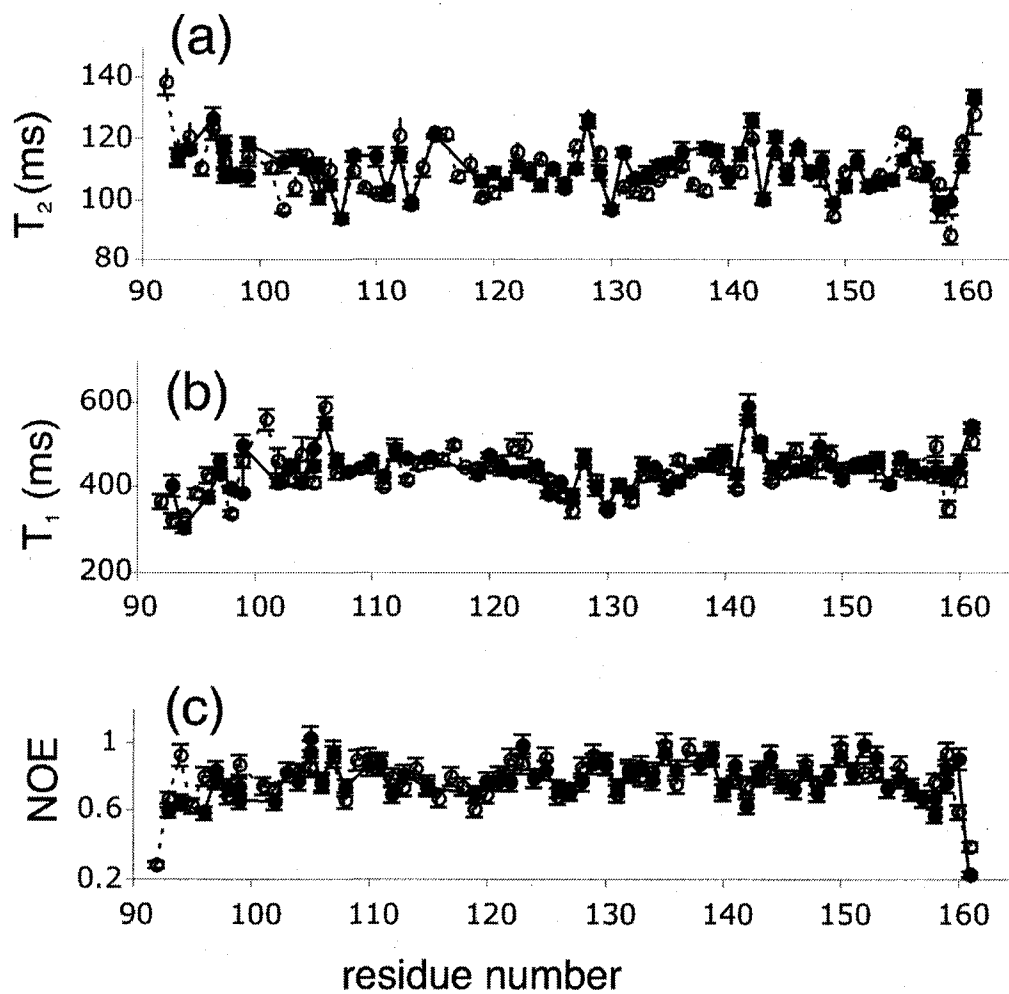
Experiment	Nuclei <sup>a</sup>	<sup>1</sup> H	x-pts <sup>b</sup>	y-pts	z-pts	x-sw	y-sw	z-sw	Mix (ms)
<sup>15</sup> N-HSQC	<sup>1</sup> H, <sup>15</sup> N	500	1074	192		8384.9		2026.3	–
HCCH-TOCSY	<sup>1</sup> H, <sup>13</sup> C, <sup>1</sup> H	500	1024	128	104	8384.9	5998.5	10057.8	–
CBCA(CO)NNH	<sup>1</sup> H, <sup>13</sup> C, <sup>15</sup> N	500	1024	118	64	8384.9	10057.2	2026.3	–
HNCACB	<sup>1</sup> H, <sup>13</sup> C, <sup>15</sup> N	500	1024	118	64	8384.9	10057.2	2026.3	–
HCCONH	<sup>1</sup> H, <sup>13</sup> C, <sup>15</sup> N	600	1024	160	64	8398.1	5998.5	2431.6	–
CCONH	<sup>1</sup> H, <sup>13</sup> C, <sup>15</sup> N	600	1024	184	64	8398.1	12067.2	2431.6	–
<sup>13</sup> C NOESY HSQC	<sup>1</sup> H, <sup>1</sup> H, <sup>13</sup> C	600	1024	160	80	8398.1	5998.7	6033.7	100
<sup>15</sup> N NOESY HSQC	<sup>1</sup> H, <sup>1</sup> H, <sup>15</sup> N	600	1024	160	80	8398.1	5998.7	2431.6	150
<sup>15</sup> N TOCSY HSQC	<sup>1</sup> H, <sup>1</sup> H, <sup>15</sup> N	600	1024	160	64	8398.1	5998.7	2431.6	50
HNHA	<sup>1</sup> H, <sup>1</sup> H, <sup>15</sup> N	600	1024	160	64	8398.1	5998.7	2431.6	–
HNHB	<sup>1</sup> H, <sup>1</sup> H, <sup>15</sup> N	500	1024	160	64	8384.9	4998.8	2026.3	–
DQF-COSY (D <sub>2</sub> O)	<sup>1</sup> H, <sup>1</sup> H	800							
NOESY (D <sub>2</sub> O)	<sup>1</sup> H, <sup>1</sup> H	800							
<sup>13</sup> C/ <sup>15</sup> N-filt. TOCSY	<sup>1</sup> H, <sup>1</sup> H	600	4096	512	–	8384.9	8385.7	–	0.06
<sup>13</sup> C/ <sup>15</sup> N-filt. NOESY	<sup>1</sup> H, <sup>1</sup> H	600	4096	512	–	8384.9	8385.7	–	0.25
<sup>13</sup> C-filt., ed. NOESY	<sup>1</sup> H, <sup>1</sup> H, <sup>13</sup> C	600	1024	240	64	3595.9	3595.9	3014.0	0.15

<sup>a</sup> The nucleus acquired in each dimension (e.g., <sup>1</sup>H, <sup>15</sup>N indicates proton x, nitrogen y). <sup>b</sup> x-, y-, and z-pts and -sw are the number of complex points and the sweep width in each respective dimension (x is the directly detected dimension).

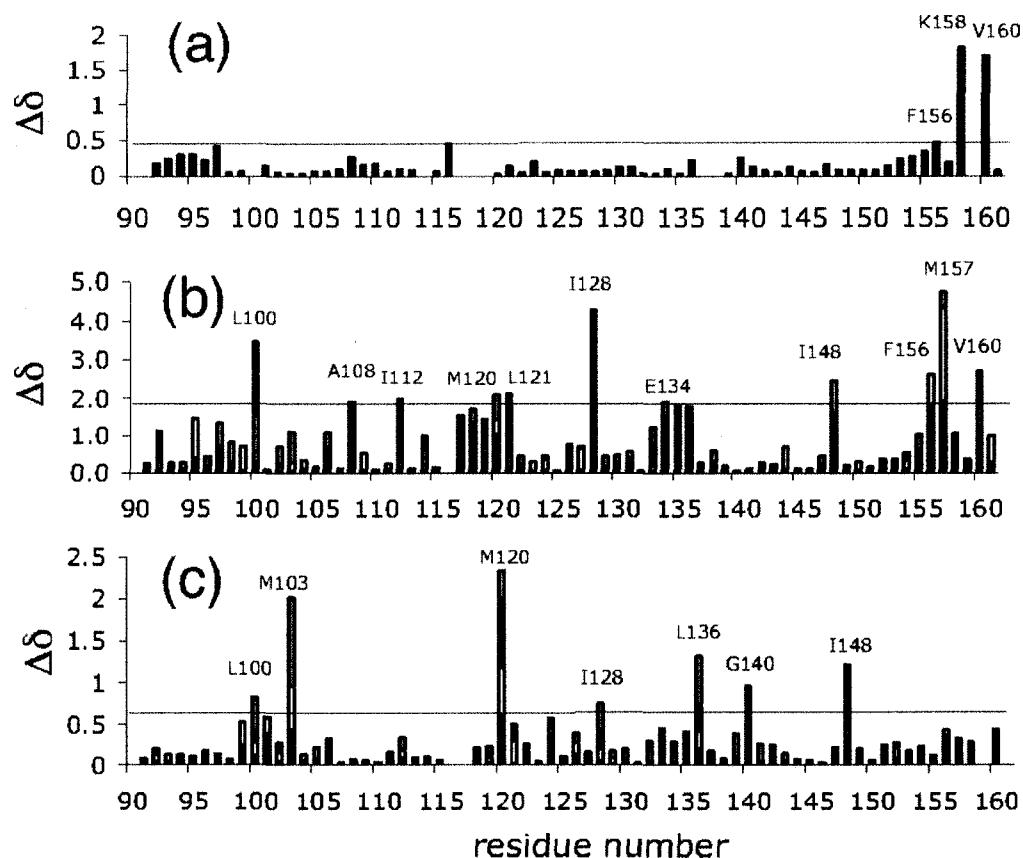
**Figure III-1.** Titration of G159D (a) and cTnC (b) with cTnI<sub>34-71</sub> using 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectroscopy. Closed cross-peaks correspond to uncomplexed cTnC or G159D and open cross-peaks correspond to the complex of cTnC or G159D with cTnI<sub>34-71</sub>. Corresponding binding curves are shown in the insets. Annotated residues have the largest  $^{15}\text{N}$  backbone chemical shift differences and are adjacent to the mutation site.



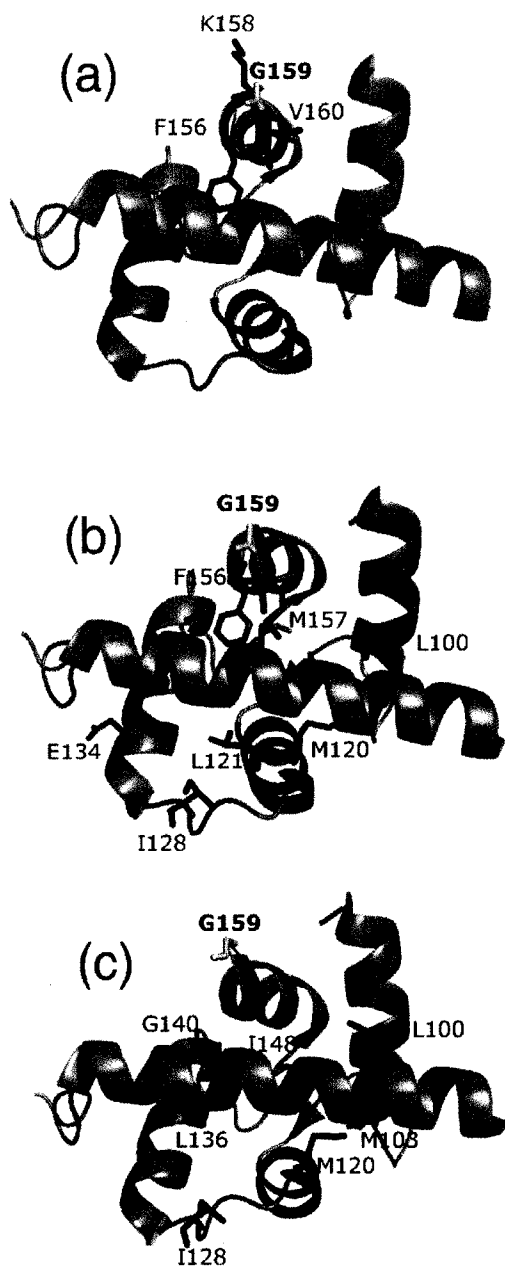
**Figure III-2.** Per residue NMR relaxation parameters  $T_2$  (a),  $T_1$  (b), NOE (c) for G159D (closed circles) and cCTnC (open circles) in complexes with cTnI<sub>34-71</sub>.



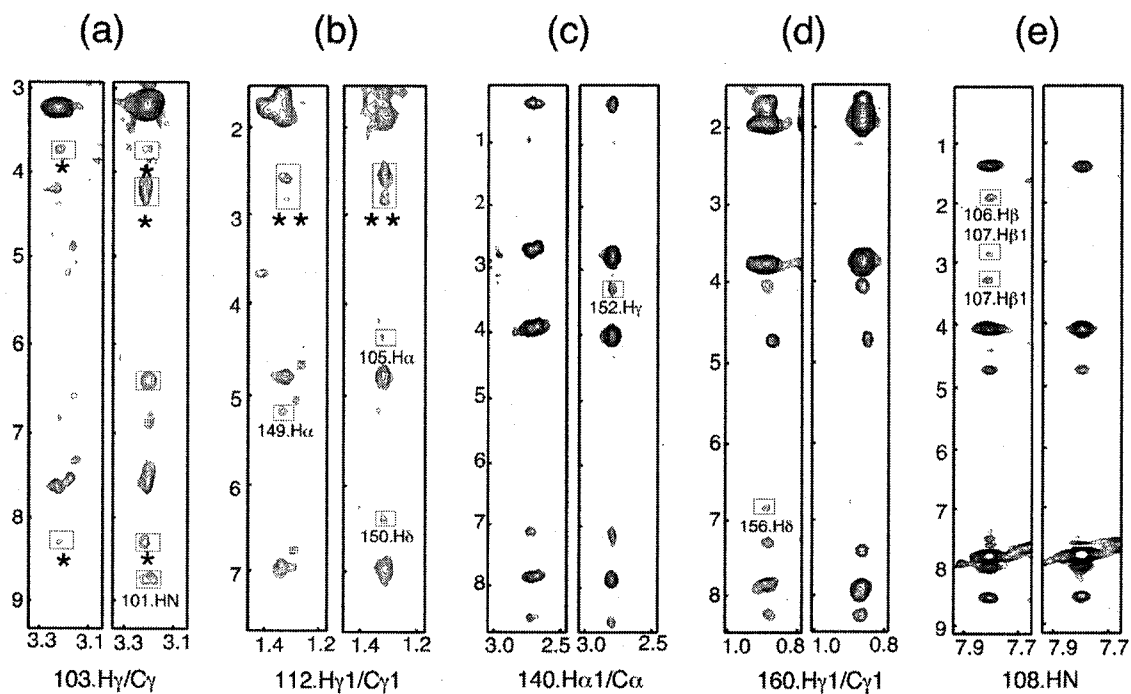
**Figure III-3.** Chemical shifts differences between G159D and cCTnC in complexes with cTnI<sub>34-71</sub> for backbone amide <sup>15</sup>N (a), side chain <sup>13</sup>C (b), and side chain <sup>1</sup>H (c). Side chain chemical shift changes for <sup>13</sup>C and <sup>1</sup>H are added together for every residue, with C $\alpha$  and H $\alpha$  in black, C $\beta$  and H $\beta$  in white, C $\delta$  and H $\delta$  in red, C $\epsilon$  and H $\epsilon$  in dark blue, and C $\delta$  and H $\delta$  in green. Horizontal lines correspond to the average value plus one standard deviation.



**Figure III-4.** Residues with chemical shift changes visualized on the structure of cCTnC determined by X-ray crystallography (1J1D). Residues with the significant  $^{15}\text{N}$  chemical shift changes are shown in black (a). Residues with the significant side chain  $^{13}\text{C}$  chemical shift changes are shown in green (b). Residues with the significant side chain  $^1\text{H}$  chemical shift changes are shown in red (c). Gly<sup>159</sup> is shown in yellow in all three panels.



**Figure III-5.** NOE strips for G159D (left) and cCTnC (right) in complexes with cTnI<sub>34-71</sub> obtained in 3D <sup>13</sup>C NOESY HSQC NMR experiments (a-d) and in <sup>15</sup>N NOESY HSQC NMR experiments (e), where stars represent the resonances tentatively assigned to the peptide, and double stars represent resonances tentatively assigned to the protein. Assignments are indicated below each pair of plots.



**Figure III-6.** Secondary structures of cCTnC (a) and G159D (b) in complexes with the anchoring region of cTnI (or cTnI<sub>34-71</sub>) in comparison with the previously determined structures 1J1D (grey) and 1FI5 (green). The structure of cCTnC determined in this work is shown on both panels in red and the structure of G159D is in yellow.

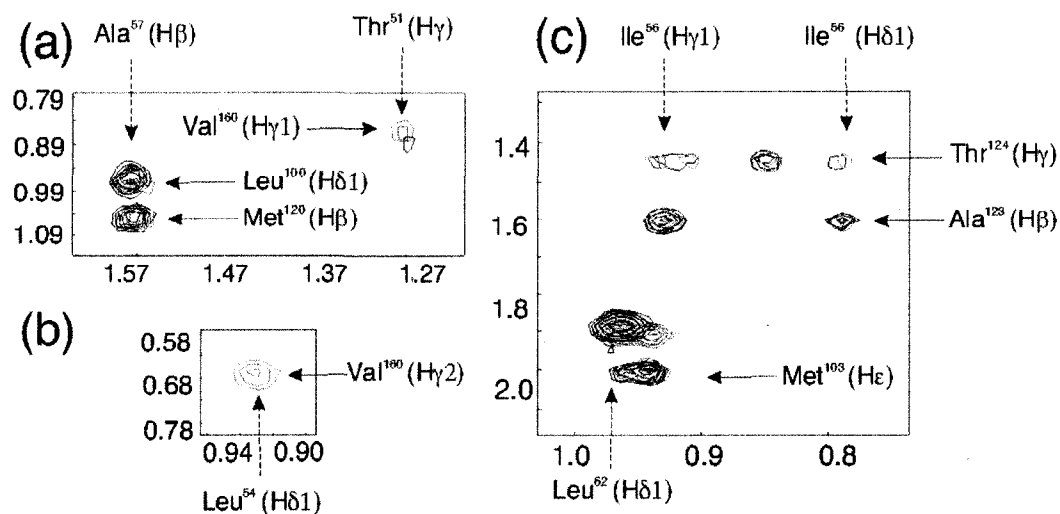
(a)



(b)

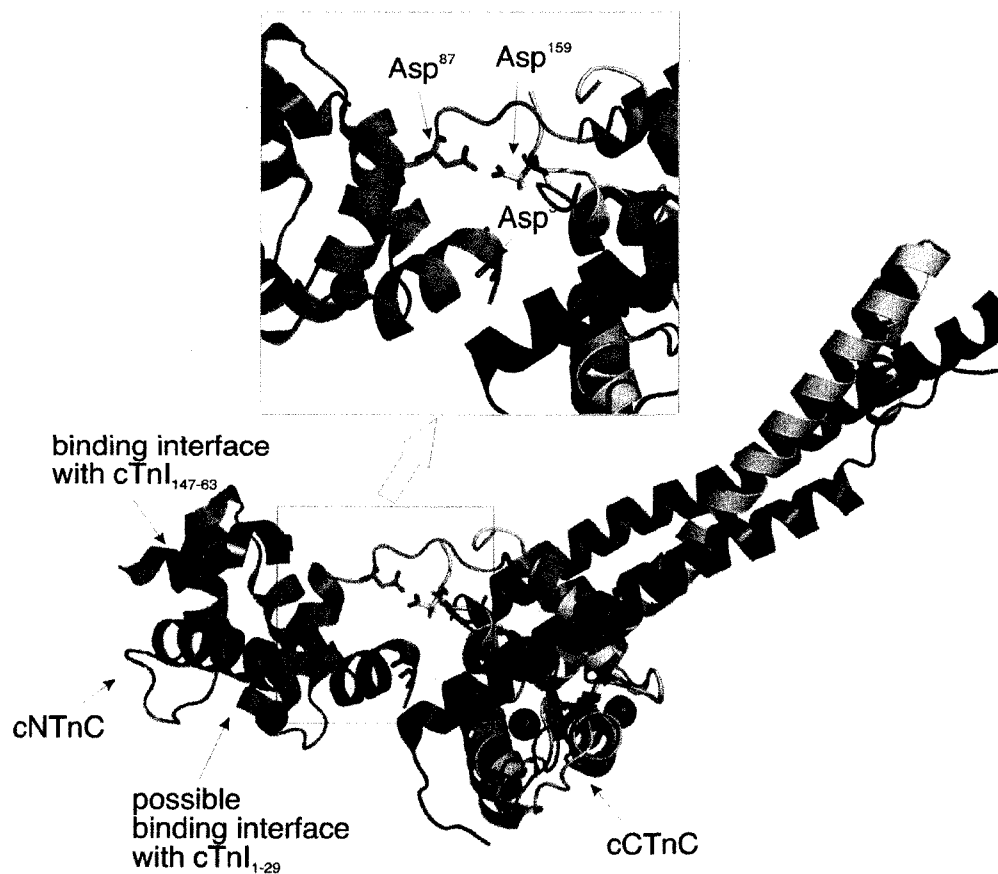


**Figure III-7.** Overlay of the regions of  $^{13}\text{C}$ -filtered, edited NOESY spectra showing the direct contacts between G159D (black contours) and cCTnC (red contours) with cTnI<sub>34-71</sub>. The resonances of proteins are shown with the horizontal solid arrows, the resonances for the peptide are shown with the vertical dashed arrows. The contact for G159D in the panel B is missing.





**Figure III-8.** The X-ray structure of the core of cardiac troponin complex (1JID) overlaid with the G159D structure to demonstrate the proximity of Asp<sup>159</sup> to cNTnC. Ribbon diagram is shown for cTnC (green), cTnI (red), cTnT (grey), and G159D (yellow). The cNTnC binding site for cTnI<sub>147-163</sub> and the possible cNTnC binding site for cTnI<sub>1-29</sub> are indicated. Electrostatic residues of cNTnC in close proximity to Asp<sup>159</sup> are shown in sticks (Asp<sup>87</sup> and Asp<sup>3</sup>).



## References

- (1) Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. *Nature* 424, 35-41.
- (2) Redwood, C. S., Moolman-Smook, J. C., and Watkins, H. (1999) Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc Res* 44, 20-36.
- (3) Seidman, J. G., and Seidman, C. (2001) The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell* 104, 557-67.
- (4) 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.
- (5) Lim, C. C., Yang, H., Yang, M., Wang, C. K., Shi, J., Berg, E. A., Pimentel, D. R., Gwathmey, J. K., Hajjar, R. J., Helmes, M., Costello, C. E., Huo, S., and Liao, R. (2008) A Novel Mutant Cardiac Troponin C Disrupts Molecular Motions Critical For Calcium Binding Affinity And Cardiomyocyte Contractility. *Biophys J* 94, 3577-89.
- (6) Parvatiyar, M., Pinto, J. R., Jones, M. A., Ackerman, M. J., and Potter, J. D. (2008) in *Biophysical Society, Annual Meeting*, Long Beach, CA.
- (7) Mogensen, J., Murphy, R. T., Shaw, T., Bahl, A., Redwood, C., Watkins, H., Burke, M., Elliott, P. M., and McKenna, W. J. (2004) Severe disease expression of cardiac troponin C and T mutations in patients with idiopathic dilated cardiomyopathy. *J Am Coll Cardiol* 44, 2033-40.
- (8) Li, M. X., Wang, X., and Sykes, B. D. (2004) Structural based insights into the role of troponin in cardiac muscle pathophysiology. *J Muscle Res Cell Motil* 25, 559-79.
- (9) Sykes, B. D. (2003) Pulling the calcium trigger. *Nat Struct Biol* 10, 588-9.
- (10) Biesiadecki, B. J., Kobayashi, T., Walker, J. S., John Solaro, R., and de Tombe, P. P. (2007) The Troponin C G159D Mutation Blunts Myofilament Desensitization Induced by Troponin I Ser23/24 Phosphorylation. *Circ Res* 100, 1486-1493.
- (11) Preston, L. C., Ashley, C. C., and Redwood, C. S. L. (2007) DCM troponin C mutant Gly159Asp blunts the response to troponin phosphorylation. *Biochem Biophys Res Commun* 360, 27-32.
- (12) Mirza, M., Marston, S., Willott, R., Ashley, C., Mogensen, J., McKenna, W., Robinson, P., Redwood, C., and Watkins, H. (2005) Dilated cardiomyopathy mutations in three thin filament regulatory proteins result in a common functional phenotype. *J Biol Chem* 280, 28498-506.
- (13) Lindhout, D. A., Boyko, R. F., Corson, D. C., Li, M. X., and Sykes, B. D. (2005) The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C. *Biochemistry* 44, 14750-9.
- (14) Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) Ca(2+)-regulated structural changes in troponin. *Proc Natl Acad Sci U S A* 102, 5038-43.
- (15) Blumenschein, T. M., Tripet, B. P., Hodges, R. S., and Sykes, B. D. (2001) Mapping the interacting regions between troponins T and C. Binding of TnT and TnI peptides to TnC and NMR mapping of the TnT-binding site on TnC. *J Biol Chem* 276, 36606-12.

- (16) Dong, W. J., Xing, J., Ouyang, Y., An, J., and Cheung, H. C. (2008) Structural kinetics of cardiac troponin C mutants linked to familial hypertrophic and dilated cardiomyopathy in troponin complexes. *J Biol Chem* 283, 3424-32.
- (17) Li, M. X., Wang, X., Lindhout, D. A., Buscemi, N., Van Eyk, J. E., and Sykes, B. D. (2003) Phosphorylation and mutation of human cardiac troponin I differentially destabilize the interaction of the functional regions of troponin I with troponin C. *Biochemistry* 42, 14460-8.
- (18) 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, 27024-34.
- (19) Gasmi-Seabrook, G. M., 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, 8313-22.
- (20) Bloembergen, N., Purcell, E. M., and Pound, R. V. (1948) Relaxation Effects in Nuclear Magnetic Resonance Absorption. *Physical Review* 73, 679-712.
- (21) Baryshnikova, O. K., and Sykes, B. D. (2006) Backbone dynamics of SDF-1 $\alpha$  determined by NMR: interpretation in the presence of monomer-dimer equilibrium. *Protein Sci* 15, 2568-78.
- (22) Mercier, P., Spyropoulos, L., and Sykes, B. D. (2001) Structure, dynamics, and thermodynamics of the structural domain of troponin C in complex with the regulatory peptide 1-40 of troponin I. *Biochemistry* 40, 10063-77.
- (23) Cantor, C. R., and Schimmel, P. R. (1980) *Biophysical Chemistry, Part II, Techniques for the study of biological structure and function*, W.H. Freeman, New York.
- (24) Tanford, C. (1961) *Physical chemistry of macromolecules*, John Wiley & Sons, Inc., New York.
- (25) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
- (26) Guntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353-78.
- (27) Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* 50, 496-506.
- (28) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* 160, 65-73.
- (29) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26P, 283-291.
- (30) Hooft, R. W., Vriend, G., Sander, C., and Abola, E. E. (1996) Errors in protein structures. *Nature* 381, 272.
- (31) Lee, W., Revington, M. J., Arrowsmith, C., and Kay, L. E. (1994) A pulsed field gradient isotope-filtered 3D <sup>13</sup>C HMQC-NOESY experiment for extracting intermolecular NOE contacts in molecular complexes. *FEBS Lett* 350, 87-90.
- (32) Robertson, I. M., Spyropoulos, L., and Sykes, B. D. (2007) in *Proceedings for the international School of Biological Magnetic Resonance, 8th Course on Biophysics and the Challenges of Emerging Threats* (Puglisi, J. D., Ed.), NATO Science Series: Life and Behavioural Sciences, IOS press, The Netherlands.
- (33) Gemmecker, G., Olejniczak, E. T., and Fesik, S. W. (1992) An Improved Method for Selectively Observing Protons Attached to C-12 in the Presence of H-1-C-13 spin Pairs. *J Mag Res* 96, 199-204.

- (34) Ikura, M., and Bax, A. (1992) Isotope-Filtered 2d NMR of a Protein Peptide Complex - Study of a Skeletal-Muscle Myosin Light Chain Kinase Fragment Bound to Calmodulin. *J Am Chem Soc* 114, 2433-2440.
- (35) Ogura, K., Terasawa, H., and Inagaki, F. (1996) An improved double-tuned and isotope-filtered pulse scheme based on a pulsed field gradient and a wide-band inversion shaped pulse. *J Biomol NMR* 8, 492-498.
- (36) Finley, N. L., Howarth, J. W., and Rosevear, P. R. (2004) Structure of the Mg<sup>2+</sup>-loaded C-lobe of cardiac troponin C bound to the N-domain of cardiac troponin I: comparison with the Ca<sup>2+</sup>-loaded structure. *Biochemistry* 43, 11371-9.
- (37) Sheng, Z., Pan, B. S., Miller, T. E., and Potter, J. D. (1992) Isolation, expression, and mutation of a rabbit skeletal muscle cDNA clone for troponin I. The role of the NH<sub>2</sub> terminus of fast skeletal muscle troponin I in its biological activity. *J Biol Chem* 267, 25407-13.
- (38) Potter, J. D., Sheng, Z., Pan, B. S., and Zhao, J. (1995) A direct regulatory role for troponin T and a dual role for troponin C in the Ca<sup>2+</sup> regulation of muscle contraction. *J Biol Chem* 270, 2557-62.
- (39) Tripet, B., De Crescenzo, G., Grothe, S., O'Connor-McCourt, M., and Hodges, R. S. (2003) Kinetic analysis of the interactions between troponin C (TnC) and troponin I (TnI) binding peptides: evidence for separate binding sites for the 'structural' N-terminus and the 'regulatory' C-terminus of TnI on TnC. *J Mol Recognit* 16, 37-53.
- (40) De Nicola, G., Burkart, C., Qiu, F., Agianian, B., Labeit, S., Martin, S., Bullard, B., and Pastore, A. (2007) The structure of Lethocerus troponin C: insights into the mechanism of stretch activation in muscles. *Structure* 15, 813-24.
- (41) Cheng, H., Lederer, M. R., Lederer, W. J., and Cannell, M. B. (1996) Calcium sparks and [Ca<sup>2+</sup>]<sub>i</sub> waves in cardiac myocytes. *Am J Physiol* 270, C148-59.
- (42) Takamatsu, T., and Wier, W. G. (1990) Calcium waves in mammalian heart: quantification of origin, magnitude, waveform, and velocity. *Faseb J* 4, 1519-25.
- (43) Li, M. X., Spyropoulos, L., and Sykes, B. D. (1999) Binding of cardiac troponin-1147-163 induces a structural opening in human cardiac troponin-C. *Biochemistry* 38, 8289-98.
- (44) Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) NMR analysis of cardiac troponin C-troponin I complexes: effects of phosphorylation. *FEBS Lett* 453, 107-12.
- (45) Gaponenko, V., Abusamhadneh, E., Abbott, M. B., Finley, N., Gasmi-Seabrook, G., Solaro, R. J., Rance, M., and Rosevear, P. R. (1999) Effects of troponin I phosphorylation on conformational exchange in the regulatory domain of cardiac troponin C. *J Biol Chem* 274, 16681-4.
- (46) Keane, N. E., Quirk, P. G., Gao, Y., Patchell, V. B., Perry, S. V., and Levine, B. A. (1997) The ordered phosphorylation of cardiac troponin I by the cAMP-dependent protein kinase--structural consequences and functional implications. *Eur J Biochem* 248, 329-37.
- (47) Baryshnikova, O. K., Li, M. X., and Sykes, B. D. (2008) Modulation of cardiac troponin C function by the cardiac-specific N-terminus of troponin I: influence of PKA phosphorylation and involvement in cardiomyopathies. *J Mol Biol* 375, 735-51.
- (48) Schmidtmann, A., Lindow, C., Villard, S., Heuser, A., Mugge, A., Gessner, R., Granier, C., and Jaquet, K. (2005) Cardiac troponin C-L29Q, related to hypertrophic cardiomyopathy, hinders the transduction of the protein kinase A

- dependent phosphorylation signal from cardiac troponin I to C. *Febs J* 272, 6087-97.
- (49) Pearlstone, J. R., Chandra, M., Sorenson, M. M., and Smillie, L. B. (2000) Biological function and site II Ca<sup>2+</sup>-induced opening of the regulatory domain of skeletal troponin C are impaired by invariant site I or II Glu mutations. *J Biol Chem* 275, 35106-15.
  - (50) Li, M. X., Corson, D. C., and Sykes, B. D. (2002) Structure determination by NMR. Isotope labeling. *Methods Mol Biol* 173, 255-65.
  - (51) Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) Mapping of a second actin-tropomyosin and a second troponin C binding site within the C terminus of troponin I, and their importance in the Ca<sup>2+</sup>-dependent regulation of muscle contraction. *J Mol Biol* 271, 728-50.
  - (52) Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603-614.
  - (53) Spyropoulos, L. (2006) A suite of Mathematica notebooks for the analysis of protein main chain <sup>15</sup>N NMR relaxation data. *J Biomol NMR* 36, 215-24.
  - (54) 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.
  - (55) 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, 277-93.
  - (56) Slupsky, C. M., Boyko, R. F., Booth, V. K., and Sykes, B. D. (2003) Smartnotebook: a semi-automated approach to protein sequential NMR resonance assignments. *J Biomol NMR* 27, 313-21.
  - (57) Jee, J., and Guntert, P. (2003) Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment. *J Struct Funct Genomics* 4, 179-89.
  - (58) Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13, 289-302.

## CHAPTER IV

### Exploring the interaction between cTnNC and the N-terminus of cTnI

#### Outline

In Chapter II we have investigated the interaction between cTnI<sub>1-29</sub> and cTnNC and demonstrated that it is important in modulation of cTnC function. This modulation is perturbed in the presence of cardiomyopathy mutations studied, and therefore suggests that it may be involved in the development of cardiomyopathies. Various lines of evidence suggest that electrostatic interactions play a major role in the binding of cTnI<sub>1-29</sub> to cTnNC. Using NMR relaxation data as a function of ionic strength, we have obtained experimental evidence supporting the hypothesized electrostatic nature of the interaction between cTnNC•Ca<sup>2+</sup> and cTnI<sub>1-29</sub>.

#### Introduction

There is an increasing amount of evidence for interaction between the N-terminus of cTnI (~residues 1-29) and cTnNC (with and without Ca<sup>2+</sup>) (1-5), which was not visualized in crystal structures (6) nor in NMR studies (1, 7). This may be due to the fairly weak binding and/or flexibility of this region of cTnI. Estimation of binding affinity for cTnI<sub>1-29</sub> has proven to be difficult. This interaction, however, plays a paramount role in the modulation of the cTnC function in response to  $\beta$ -adrenergic stimulation (8-10), and in disease, for example in the presence of FHC and DCM mutations (11-13). Visualization of this interaction, as well as others that occur with intrinsically disordered and/or charged protein fragments, remains both technically challenging and of the outmost biological interest (14, 15).

As mentioned in Chapter II, a major contribution to the interaction between cTnNC and the N-terminus of cTnI (cTnI<sub>1-29</sub>) is thought to come from electrostatics based on following facts: first, the backbone chemical shift changes induced by the binding of cTnI<sub>1-29</sub>/ cTnI<sub>1-29</sub>PP to cTnNC are very small upon the formation of complex (1), similar to the interaction of the inhibitory region of cTnI (~residues 128-147) with cTnNC•Ca<sup>2+</sup> (16). The lack of the large backbone chemical shift changes is indicative of electrostatic interactions between sidechains, in which the core of the protein is not significantly affected and the backbone chemical shifts remain unchanged. Second, line broadening in the NMR study of backbone dynamics for the cTnNC•cTnI<sub>1-80</sub> complex indicated that the

possible site of binding was located in the region of  $\beta$ -sheets, opposite to the hydrophobic cleft, which is the binding site for the switch region of cTnI (~residues 147-163) (17). Formation of the triple complex was confirmed by the increase in the overall correlation time, supporting the notion of the independent binding site for cTnI<sub>1-29</sub>, different from the binding site for cTnI<sub>147-163</sub> (5). Third, the N-terminus of cTnI is positively charged, with the first 29 residues being ADGSSDAAREPRPAPAPIRRRSSNYRAYA, where the three consequent arginines (underlined) are suspected to interact with the negatively charged residues on cTnNC (3, 7). Upon phosphorylation, as shown by the NMR studies of phosphorylated peptides, the conformation of the cTnI N-terminus changes so that two phosphogroups screen the charges on the arginines, possibly precluding strong electrostatic contacts to be formed and promoting the dissociation of the cTnI N-terminus from cTnNC (18).

In Debye-Huckel formalism, the presence of electrolyte has a screening effect on the electric field of individual ions, which varies with the square root of the ionic strength (19). Coulomb potential has the following form:

$$\varphi_C(r) = \frac{z}{4\pi\epsilon} \cdot \frac{1}{r},$$

where the elementary charge is equal to 1,  $z$  is a charge of the ion,  $\epsilon$  is the medium permittivity, and  $r$  is the distance from the center of ion. The effect of charge shielding of the central ion is denoted as following:

$$\varphi_C(r) = \frac{ze^{-\frac{r}{r_D}}}{4\pi\epsilon} \cdot \frac{1}{r},$$

where  $r_D$  is the Debye radius with the following dependence on ionic strength:

$$r_D = \sqrt{\frac{\epsilon_0 \epsilon_r kT}{2NI}},$$

where  $\epsilon_0$  is the permittivity of free space,  $\epsilon_r$  is the dielectric constant,  $k$  is Boltzmann's constant,  $T$  is the absolute temperature,  $N$  is Avogadro's number, and  $I$  is the ionic strength. Ionic strength is expressed as a function of concentration of all ions in solution:

$$I = \frac{1}{2} \sum_A c_A z_A^2$$

where  $c_A$  is the concentration of ion A and  $z_A$  is the charge of ion A in solution. Thus, the higher the concentration of salt and ionic strength are, the greater the shielding, and the weaker electrostatic potential.

One of the methods to probe electrostatic interactions is to vary the ionic strength of the solution. While this method gives straightforward results for some systems (20, 21), electrostatic interactions in proteins can contribute not only to the formation of interfaces but in some cases to the stabilization of a protein core complicating the interpretation of results (22, 23). We have determined the dependence of the  $^{15}\text{N}$  NMR relaxation parameter  $T_2$  as a function of ionic strength for different  $^{15}\text{N}$  labeled cTnC complexes, apo cTnC, cTnC•Ca $^{2+}$ , cTnC•cTnI $_{1-29}$ , and cTnC•Ca $^{2+}$ •cTnI $_{1-29}$ , in order to probe electrostatic interactions between cTnI $_{1-29}$  and cTnC. Average  $T_2$  values (obtained by averaging the per residue  $T_2$ 's) as a function of ionic strength provided evidence for the decrease in binding of cTnI $_{1-29}$  to apo cTnC and cTnC•Ca $^{2+}$  in high salt, supporting the hypothesis of electrostatic interaction. Per residue  $T_2$  trends, together with chemical shift changes, demonstrate that increased flexibility and possibly K $^{+}$  binding to Ca $^{2+}$  binding site of apo cTnC take place at low salt. In case of cTnC•Ca $^{2+}$ , per residue  $T_2$  trends and chemical shift changes show that cTnC•Ca $^{2+}$  remains essentially unchanged (up to ionic strength of ~400 mM) and therefore serve as a fairly good control in our experiments. Decrease in  $T_2$  upon the addition of cTnI $_{1-29}$  in low salt substantiated the hypothesis of electrostatic interactions between cTnI $_{1-29}$  and cTnC•Ca $^{2+}$ .

## Results and Discussion

*Dependence of per residue and average  $T_2$  on ionic strength for apo cTnC, cTnC•Ca $^{2+}$ , cTnC•cTnI $_{1-29}$ , and cTnC•Ca $^{2+}$ •cTnI $_{1-29}$*

Relaxation parameter  $T_2$  is a complex function of protein motions (24, 25), but can be used to correlate with the increase in molecular weight of the protein (16, 26). Average  $T_2$  data is presented in Figure IV-1 as a function of ionic strength. For cTnC•Ca $^{2+}$  the dependence on ionic strength was minimal (open red squares) as compared to cTnC•Ca $^{2+}$ •cTnI $_{1-29}$  (closed red squares), which underwent a decrease in  $T_2$  values in lower salt. At high salt, the average  $T_2$  for cTnC•Ca $^{2+}$ •cTnI $_{1-29}$  was almost equal to the average  $T_2$  for cTnC•Ca $^{2+}$ , which implies that complex formation is very weak at high salt and therefore provides direct evidence for electrostatic interactions playing the major role in stabilizing the complex between cTnC•Ca $^{2+}$  and cTnI $_{1-29}$ .

Similar results were observed for apo cTnC, with average  $T_2$  values remained approximately independent of salt concentration (open black circles), whereas for cTnC•cTnI $_{1-29}$  (closed black circles), the dependence on ionic strength is clearly



noticeable. This data similarly supports the notion of electrostatic interactions playing the major role in the complex between cTnI<sub>1-29</sub> and cNTnC, since contacts are stronger in lower salt and therefore, the complex is heavier as demonstrated by lower values for T<sub>2</sub>. The average T<sub>2</sub> data for cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> and cNTnC•cTnI<sub>1-29</sub> are in agreement at ionic strength >100 mM, supporting the notion of formation similar complexes.

Per residue T<sub>2</sub> trends followed similar dependences to the average T<sub>2</sub>'s for cNTnC•Ca<sup>2+</sup>, cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub>, and cNTnC•cTnI<sub>1-29</sub> (Figs. IV-2a, IV-2b, and IV-2d). The control, cNTnC•Ca<sup>2+</sup>, could be considered roughly unperturbed by ionic strength based on per residue and average T<sub>2</sub> data. Per residue T<sub>2</sub> trends for apo cNTnC followed patterns different from the average T<sub>2</sub> (Fig. IV-2c), demonstrating that the addition of KCl affected the conformational flexibility of cNTnC. Some residues of apo cNTnC in low salt became more flexible (higher T<sub>2</sub> values) and some residues were likely to be involved in conformational exchange (lower T<sub>2</sub> values). The possibility of aggregation for apo cNTnC can be ruled out since the average T<sub>2</sub> values remain high. It is possible that electrostatic repulsion within the protein has been increased leading to the conformational flexibility. It is also possible that at high concentration of KCl, K<sup>+</sup> might occupy the Ca<sup>2+</sup> binding site on cNTnC. This complicates the interpretation of data pertaining to the apo cNTnC complex with cTnI<sub>1-29</sub>. At low salt, the T<sub>2</sub> values for cNTnC•cTnI<sub>1-29</sub> are very low suggesting aggregation, possibly triggered by the increased flexibility of apo cNTnC.

*2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy of apo cNTnC, cNTnC•Ca<sup>2+</sup>, cNTnC•cTnI<sub>1-29</sub>, and cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> in varying ionic strength*

Chemical shift changes are sensitive parameters that can be also used to monitor protein changes upon changing the ionic strength of solution. 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectra for apo cNTnC, cNTnC•Ca<sup>2+</sup>, cNTnC•cTnI<sub>1-29</sub>, and cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> were acquired in varying ionic strength and are shown in Figures IV-3 and IV-4. The spectrum of cNTnC•Ca<sup>2+</sup> underwent no or very slight perturbation as a function of ionic strength (Fig. IV-3b), with only a few charged residues moving slightly, as it is shown with arrows. Overall chemical shift changes were small, barely above the resolution in the {<sup>1</sup>H, <sup>15</sup>N} HSQC spectrum, which was ~0.02 ppm (Fig. IV-4). Small changes in chemical shifts accompanied also the addition of salt to apo cNTnC, cNTnC•cTnI<sub>1-29</sub>, and cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (Figs. IV-3a, IV-3c, and IV-3d). For cNTnC•cTnI<sub>1-29</sub>, the changes in chemical shift were the largest (Fig. IV-4). Interestingly, in case of apo cNTnC, some

chemical shift changes followed curved trajectory, supporting the existence of several processes. Based on chemical shift changes, the structure of  $\text{cTnC} \cdot \text{Ca}^{2+}$  was essentially unaffected by the increase in ionic strength, and thus,  $\text{cTnC} \cdot \text{Ca}^{2+}$  can serve as a good control in our experiments. Changes observed for  $\text{cTnC} \cdot \text{cTnI}_{1-29}$  and  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$  were similar, demonstrating the similarity between formed complexes, however, since the behavior of apo cTnC as a function of ionic strength is more complex, the data related to  $\text{cTnC} \cdot \text{cTnI}_{1-29}$  should be treated with caution.

## Conclusions

Varying ionic strength is a common method to probe electrostatic interactions within the complexes. However, protein folding may also be dependent to a some degree on particular ionic conditions. Based on our data, the structure of  $\text{cTnC} \cdot \text{Ca}^{2+}$  was largely unaffected by the differences in ionic strength (up to ~400 mM), and thus, the changes observed with  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$  can be directly attributed to the changes in interactions between  $\text{cTnC} \cdot \text{Ca}^{2+}$  and  $\text{cTnI}_{1-29}$ . For  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$  we have observed a decrease in  $T_2$  values at low salt, and an increase in  $T_2$  at high salt. At high ionic strength  $T_2$  for the complex is almost equal to  $T_2$  for  $\text{cTnC} \cdot \text{Ca}^{2+}$ , suggesting that the complex is not formed under these conditions and providing the evidence for the hypothesis that electrostatic interactions play a major role in the interaction between  $\text{cTnC} \cdot \text{Ca}^{2+}$  and  $\text{cTnI}_{1-29}$ .

In case of apo cTnC, the dependence of average  $T_2$  and chemical shifts on ionic strength for  $\text{cTnC} \cdot \text{cTnI}_{1-29}$  shows generally similar trends to  $\text{cTnC} \cdot \text{Ca}^{2+} \cdot \text{cTnI}_{1-29}$ ; however, these results should be treated with care since the per residue  $T_2$  for apo cTnC demonstrated a dependence on ionic strength in the absence of  $\text{cTnI}_{1-29}$ .

The binding of  $\text{cTnI}_{1-29}$  to cTnC is of substantial interest due to its involvement in the regulation of cardiac contraction and the development of cardiac malfunctions. The data presented here provides experimental support for the electrostatic nature of the interaction between  $\text{cTnI}_{1-29}$  and  $\text{cTnC} \cdot \text{Ca}^{2+}$ .

## Materials and Methods

### *Protein Expression and Purification*

The DNA encoding cTnC (1-89) with the mutations C35S and C84S was used in pET-3a expression vector. *E. coli* strain BL21(DE3)pLysS was transformed and incubated at 37°C to the  $\text{OD}_{600}$  of 0.6-0.9. Cell cultures were induced with IPTG and harvested after incubation for 3 hours. Uniformly  $^{15}\text{N}$ -labeled cTnC was expressed in a

minimal media enriched with  $(^{15}\text{NH}_4)_2\text{SO}_4$  (27). Cell pellet was lysed using French press and applied to a DEAE-sephadex (A25) column (50 mM Tris, 0.1 M NaCl, pH 8.0). Protein was further purified using a Superdex-75 size exclusion column (50 mM Tris, 0.15 M NaCl, pH 8.0), decalcified, and desalted using Sephadex G25 column (10 mM  $\text{NH}_4\text{HCO}_3$ ), decalcified as described previously (28), and lyophilized. Molecular mass for the unlabeled protein was confirmed with mass spectrometry.

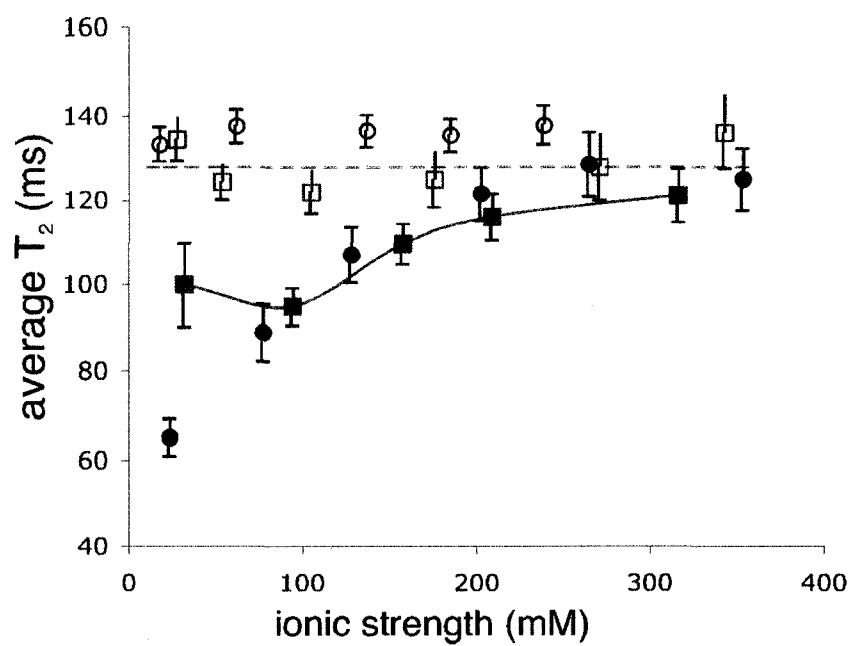
### *Peptides*

Synthetic peptide, cTnI<sub>1-29</sub> (acetyl-ADGSSDAAREPRPAPAPIRRRSSNYRAYA-amide), was synthesized using standard methodology (29). The sequence was confirmed by amino acid analysis and the masses were verified by electrospray mass spectrometry.

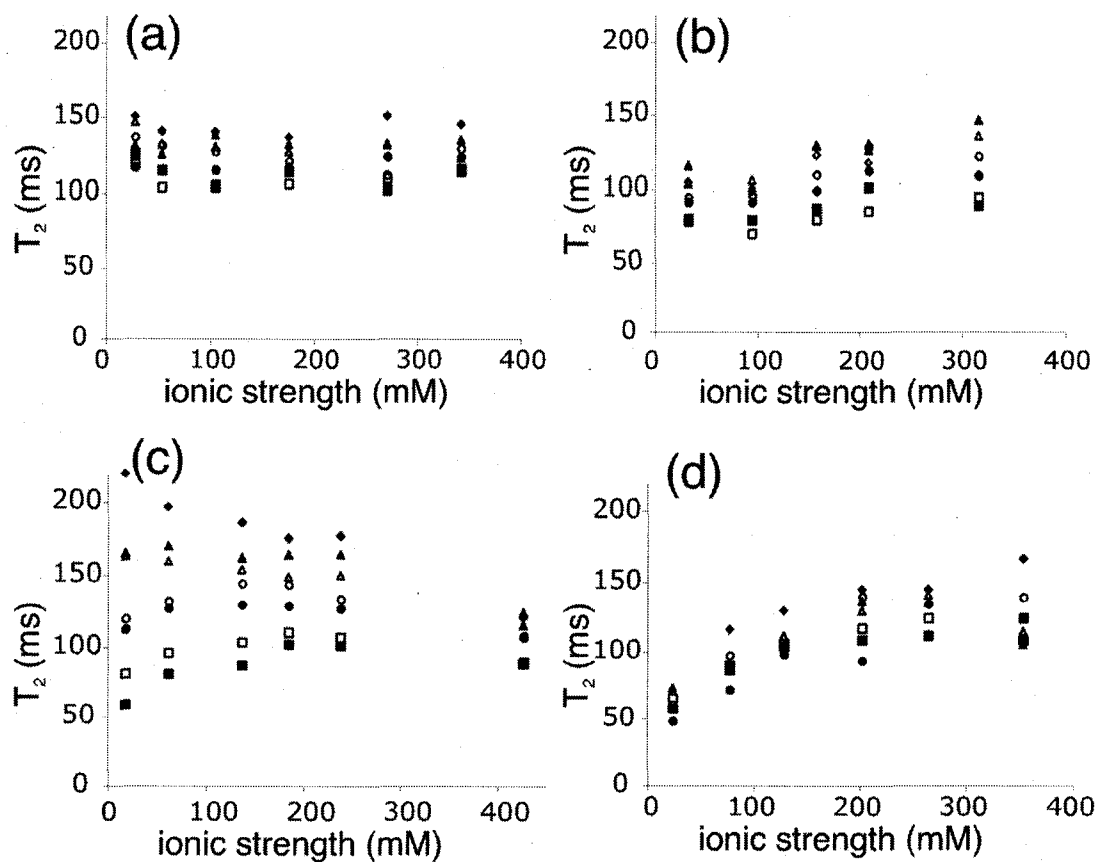
### *NMR spectroscopy*

NMR samples were prepared in 500  $\mu\text{l}$  of NMR buffer containing 10 mM imidazole, 5 mM  $\text{NaN}_3$ , and 0.2 mM DSS in 90%  $\text{H}_2\text{O}$ /10%  $\text{D}_2\text{O}$ . In case of cTnC• $\text{Ca}^{2+}$ , 1 ml of 1M  $\text{CaCl}_2$  was added. Concentration of cTnC in all experiments was  $\sim 0.6$  mM and the concentration of cTnI<sub>1-29</sub> was  $\sim 1.8$  mM. The increase in ionic strength was achieved by adding solid KCl. Whenever needed, pH was adjusted to 6.75. At every salt concentration the following experiments were acquired:  $^1\text{H}$  1D, 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC, and the  $T_2$  series using relaxation delays of 10, 30, 50, 70, 90, and 110 ms. NMR experiments were conducted at 30 °C on a Varian Inova 500 MHz spectrometer equipped with a z-axis pulsed field gradient triple resonance probe. Acquired spectra were processed using NMRPipe (30) and visualized using NMRView (31). Per residue relaxation parameters,  $T_2$ , were extracted using the rate analysis module built in NMRView and presented along with the  $T_2$  averaged over a dataset. Per residue chemical shift changes,  $\Delta\delta$ , were obtained using the following equation:  $\Delta\delta = \sqrt{(0.2\Delta\delta_{^{15}\text{N}})^2 + (\Delta\delta_{^1\text{H}})^2}$ , where  $\Delta\delta_{^{15}\text{N}}$  is the chemical shift change in ppm in  $^{15}\text{N}$  dimension and  $\Delta\delta_{^1\text{H}}$  is the chemical shift change in ppm in  $^1\text{H}$  dimension. The coefficient of 0.2 accounts for the difference between sensitivities of  $^{15}\text{N}$  and  $^1\text{H}$  nuclei. Easily traceable changes in chemical shifts for several residues were analyzed and presented as an averaged value.

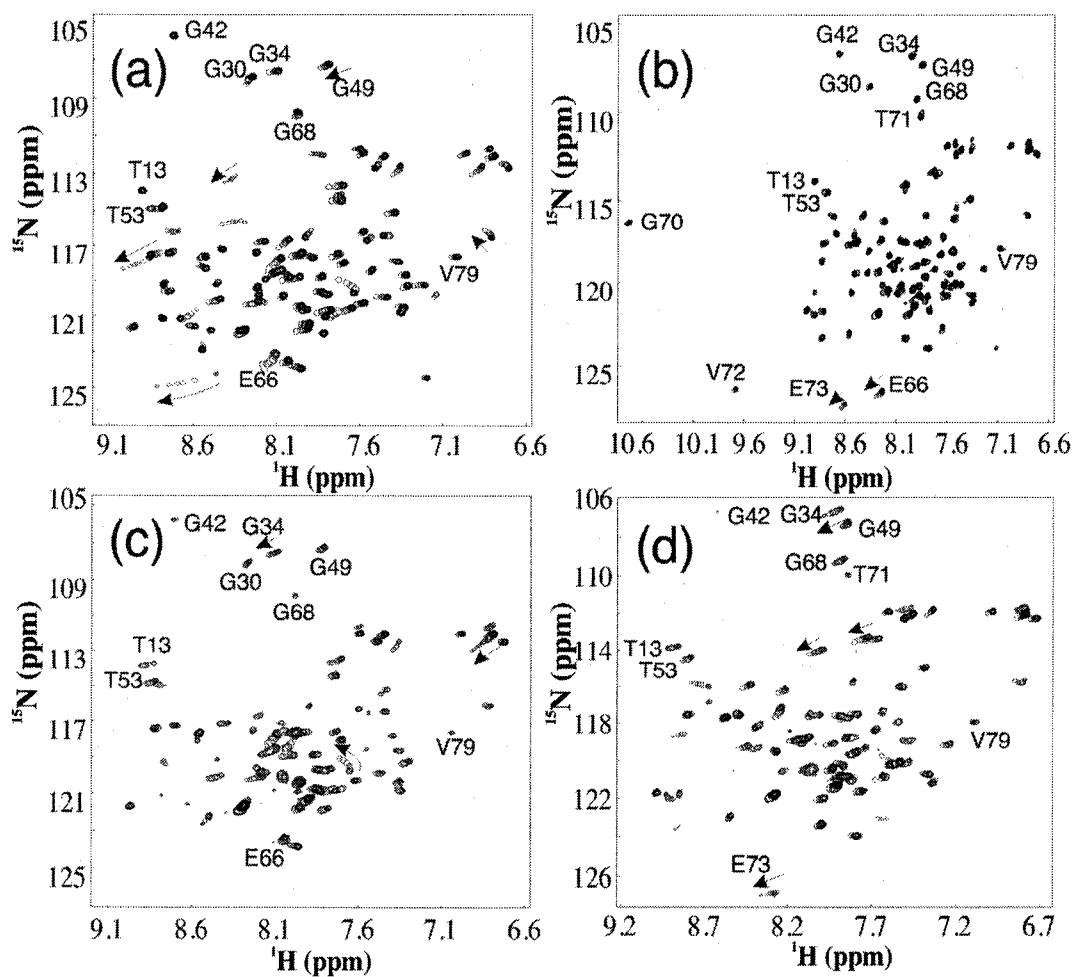
**Figure IV-1.** Average  $T_2$  values for apo cTnC (open black circles), cTnC•Ca<sup>2+</sup> (open red squares), cTnC•cTnI<sub>1-29</sub> (closed black circles), and cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (closed red squares).



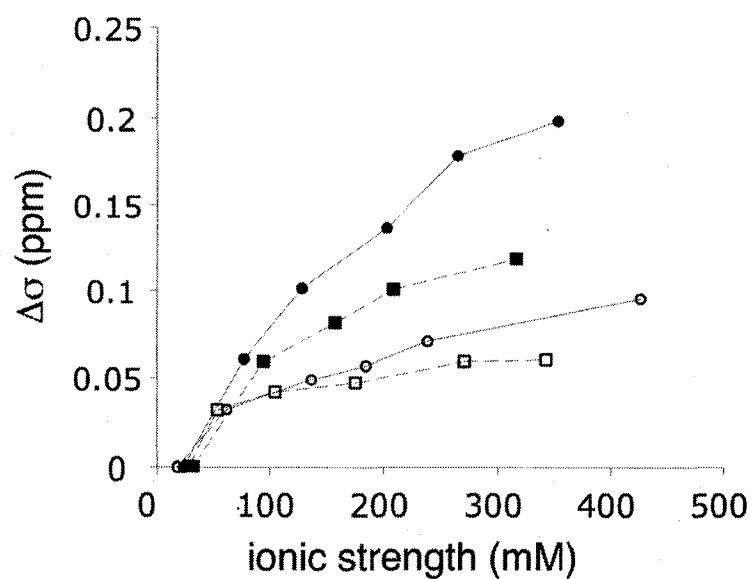
**Figure IV-2.** Dependence of per residue  $T_2$ 's on ionic strength. Data is shown for several residues to demonstrate the range of  $T_2$  values for cNTnC•Ca<sup>2+</sup> (a), cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (b), apo cNTnC (c), and cNTnC•cTnI<sub>1-29</sub> (d). Notice the spread in individual  $T_2$ 's for apo cNTnC in low salt (c).



**Figure IV-3.** Overlays of 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectra obtained for apo cNTnC (a), cNTnC•Ca<sup>2+</sup> (b), cNTnC•cTnI<sub>1-29</sub> (c), and cNTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (d) as a function of ionic strength. Assignments are taken from (32).



**Figure IV-4.** Chemical shift difference as a function of ionic strength for several easily traceable residues for apo cTnC (open circles), cTnC•Ca<sup>2+</sup> (open squares), cTnC•cTnI<sub>1-29</sub> (closed circles), and cTnC•Ca<sup>2+</sup>•cTnI<sub>1-29</sub> (closed squares).



## References

- (1) Finley, N., Abbott, M. B., Abusamhadneh, E., Gaponenko, V., Dong, W., Gasmi-Seabrook, G., Howarth, J. W., Rance, M., Solaro, R. J., Cheung, H. C., and Rosevear, P. R. (1999) NMR analysis of cardiac troponin C-troponin I complexes: effects of phosphorylation. *FEBS Lett* 453, 107-12.
- (2) Abbott, M. B., Dong, W. J., Dvoretzky, A., DaGue, B., Caprioli, R. M., Cheung, H. C., and Rosevear, P. R. (2001) Modulation of cardiac troponin C-cardiac troponin I regulatory interactions by the amino-terminus of cardiac troponin I. *Biochemistry* 40, 5992-6001.
- (3) Ward, D. G., Cornes, M. P., and Trayer, I. P. (2002) Structural consequences of cardiac troponin I phosphorylation. *J Biol Chem* 277, 41795-801.
- (4) Ward, D. G., Brewer, S. M., Cornes, M. P., and Trayer, I. P. (2003) A cross-linking study of the N-terminal extension of human cardiac troponin I. *Biochemistry* 42, 10324-32.
- (5) Baryshnikova, O. K., Li, M. X., and Sykes, B. D. (2008) Modulation of cardiac troponin C function by the cardiac-specific N-terminus of troponin I: influence of PKA phosphorylation and involvement in cardiomyopathies. *J Mol Biol* 375, 735-51.
- (6) Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. *Nature* 424, 35-41.
- (7) Ward, D. G., Brewer, S. M., Gallon, C. E., Gao, Y., Levine, B. A., and Trayer, I. P. (2004) NMR and mutagenesis studies on the phosphorylation region of human cardiac troponin I. *Biochemistry* 43, 5772-81.
- (8) 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 Ca<sup>2+</sup>-binding properties of the Ca<sup>2+</sup>-regulatory site of bovine cardiac troponin. *J Biol Chem* 257, 260-3.
- (9) Holroyde, M. J., Howe, E., and Solaro, R. J. (1979) Modification of Calcium requirements for activation of cardiac myofibrillar ATPase by cyclic AMP dependent phosphorylation. *Biochim Biophys Acta* 586, 63-69.
- (10) Chandra, M., Dong, W. J., Pan, B. S., Cheung, H. C., and Solaro, R. J. (1997) Effects of protein kinase A phosphorylation on signaling between cardiac troponin I and the N-terminal domain of cardiac troponin C. *Biochemistry* 36, 13305-11.
- (11) Schmidtman, A., Lindow, C., Villard, S., Heuser, A., Mugge, A., Gessner, R., Granier, C., and Jaquet, K. (2005) Cardiac troponin C-L29Q, related to hypertrophic cardiomyopathy, hinders the transduction of the protein kinase A dependent phosphorylation signal from cardiac troponin I to C. *Febs J* 272, 6087-97.
- (12) Dong, W. J., Xing, J., Ouyang, Y., An, J., and Cheung, H. C. (2008) Structural kinetics of cardiac troponin C mutants linked to familial hypertrophic and dilated cardiomyopathy in troponin complexes. *J Biol Chem* 283, 3424-32.
- (13) Biesiadecki, B. J., Kobayashi, T., Walker, J. S., John Solaro, R., and de Tombe, P. P. (2007) The Troponin C G159D Mutation Blunts Myofilament Desensitization Induced by Troponin I Ser23/24 Phosphorylation. *Circ Res* 100, 1486-1493.
- (14) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6, 197-208.
- (15) Fink, A. L. (2005) Natively unfolded proteins. *Curr Opin Struct Biol* 15, 35-41.



- (16) Lindhout, D. A., Boyko, R. F., Corson, D. C., Li, M. X., and Sykes, B. D. (2005) The role of electrostatics in the interaction of the inhibitory region of troponin I with troponin C. *Biochemistry* 44, 14750-9.
- (17) Gaponenko, V., Abusamhadneh, E., Abbott, M. B., Finley, N., Gasmi-Seabrook, G., Solaro, R. J., Rance, M., and Rosevear, P. R. (1999) Effects of troponin I phosphorylation on conformational exchange in the regulatory domain of cardiac troponin C. *J Biol Chem* 274, 16681-4.
- (18) Keane, N. E., Quirk, P. G., Gao, Y., Patchell, V. B., Perry, S. V., and Levine, B. A. (1997) The ordered phosphorylation of cardiac troponin I by the cAMP-dependent protein kinase--structural consequences and functional implications. *Eur J Biochem* 248, 329-37.
- (19) Debye, P., and Huckel, E. (1923) The theory of electrolytes. I. Lowering of freezing point and related phenomena. *Physikalische Zeitschrift* 24, 185-206.
- (20) Suh, J. Y., Tang, C., and Clore, G. M. (2007) Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J Am Chem Soc* 129, 12954-5.
- (21) Davis, S. J., Davies, E. A., Tucknott, M. G., Jones, E. Y., and van der Merwe, P. A. (1998) The role of charged residues mediating low affinity protein-protein recognition at the cell surface by CD2. *Proc Natl Acad Sci U S A* 95, 5490-4.
- (22) Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry* 29, 7133-55.
- (23) Stigter, D., Alonso, D. O., and Dill, K. A. (1991) Protein stability: electrostatics and compact denatured states. *Proc Natl Acad Sci U S A* 88, 4176-80.
- (24) Palmer, A. G., 3rd. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem Rev* 104, 3623-40.
- (25) Palmer, A. G., 3rd. (1997) Probing molecular motion by NMR. *Curr Opin Struct Biol* 7, 732-7.
- (26) Blumenschein, T. M., Stone, D. B., Fletterick, R. J., Mendelson, R. A., and Sykes, B. D. (2005) Calcium-dependent changes in the flexibility of the regulatory domain of troponin C in the troponin complex. *J Biol Chem* 280, 21924-32.
- (27) Li, M. X., Corson, D. C., and Sykes, B. D. (2002) Structure determination by NMR. Isotope labeling. *Methods Mol Biol* 173, 255-65.
- (28) Li, M. X., Gagne, S. M., Tsuda, S., Kay, C. M., Smillie, L. B., and Sykes, B. D. (1995) Calcium binding to the regulatory N-domain of skeletal muscle troponin C occurs in a stepwise manner. *Biochemistry* 34, 8330-40.
- (29) Tripet, B., Van Eyk, J. E., and Hodges, R. S. (1997) Mapping of a second actin-tropomyosin and a second troponin C binding site within the C terminus of troponin I, and their importance in the Ca<sup>2+</sup>-dependent regulation of muscle contraction. *J Mol Biol* 271, 728-50.
- (30) 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, 277-93.
- (31) Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603-614.
- (32) Li, M. X., Spyropoulos, L., and Sykes, B. D. (1999) Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. *Biochemistry* 38, 8289-98.

## CHAPTER V

### Towards the determination of the pKa's of side chain carboxyl groups involved in electrostatic interactions on the interface between cNTnC and the N-terminus of cTnI

#### Outline

Side chain pKa's for residues forming an interface with a ligand are often shifted from their typical values in solution. This information can be used to delineate residues participating in the interface, and to understand the interactions that stabilize the protein complex. We attempted to determine the pKa's of side chain carboxyl groups of cNTnC using a 2D HBHG(CBCG)CO NMR pulse sequence, which we developed on the basis of 3D  $^{13}\text{C}'-(^{13}\text{C})-^1\text{H}$  HSQC/HSQC (*I*). Although the pulse sequence was successful, the pH titrations of cNTnC had to be discontinued due to the precipitation of cNTnC below pH ~5. The approach should be useful, however, for other studies.

#### Introduction

One approach to probe electrostatic interactions in a complex involves the determination of side chain pKa's for residues suspected to participate in the formation of the interface. It is known that catalytically important residues (2-7), as well as residues that determine protein folding (8-10), often have side chain pKa's different from those of the individual amino acids in solution. Residues forming interfaces were shown to have different pKa's in the absence and presence of a binding partner (*II*). The electrostatic nature of the interaction between the N-terminus of cTnI (cTnI<sub>1-29</sub>) and cNTnC was investigated in Chapter IV. Possible contacts between cTnI<sub>1-29</sub> and cNTnC can exist between Arg<sup>19</sup>, Arg<sup>20</sup>, and Arg<sup>21</sup> of cTnI<sub>1-29</sub> and Glu<sup>32</sup>, Glu<sup>40</sup>, Glu<sup>55</sup>, Glu<sup>63</sup>, Glu<sup>66</sup>, Asp<sup>33</sup>, Asp<sup>62</sup>, and Asp<sup>75</sup> of cNTnC. Determination of pKa's of side chain carboxyl groups in cNTnC should provide direct evidence for their involvement in the formation of the interface with cTnI<sub>1-29</sub>. Similarly, the identification of ionization constants for guanidinium groups in arginines of cTnI<sub>1-29</sub> might provide the evidence for their participation in the formation of the interface with cNTnC.

A common and straightforward method for the determination of side chain pKa's involves the pH titration of a protein monitored by the solution state NMR spectroscopy. In one of the earliest applications of biomacromolecular NMR spectroscopy, 1D  $^1\text{H}$  NMR

experiments were employed to study the pKa's of histidines of ribonuclease (2). With the development of multidimensional multinuclear NMR methods and protein isotope labeling techniques, other experiments became available for determining the pKa of side chain carboxyl groups; for example, 2D  $\{^{13}\text{C}, ^1\text{H}\}$  H(C)CO TOCSY (11), 2D homonuclear TOCSY (4, 9), 1D  $^{13}\text{C}$  (12-14), and  $^{13}\text{C}'-(^{13}\text{C})-^1\text{H}$  HSQC/HSQC (1, 3, 5). New experiments are emerging for detecting arginine side chain resonances, such as the Arg-H $^{\epsilon}$ (N $^{\epsilon}\text{C}^{\delta}$ )H $^{\delta}$ , Arg-H $^{\epsilon}$ (N $^{\epsilon}\text{C}^{\zeta}$ )H $^{\eta}$ , and Arg-H $^{\eta}$ (N $^{\eta}\text{C}^{\zeta}$ )H $^{\epsilon}$  pulse sequences (15, 16). Lysine side chains have been recently studied using an HISQC pulse sequence (17).

We attempted to determine the pKa's of side chain carboxyl groups for cNTnC•Ca $^{2+}$  using the HBHG(CBCG)CO pulse sequence, developed on the basis of 3D  $^{13}\text{C}'-(^{13}\text{C})-^1\text{H}$  HSQC/HSQC (1). The modified and tested pulse sequence worked well. However, the lower limit of pH where cNTnC•Ca $^{2+}$  remain a folded protein with the known structure was ~5, below which cNTnC•Ca $^{2+}$  precipitated. Thus, the NMR method for identification of pKa's of side chain carboxyl groups was not suitable for studying cNTnC•Ca $^{2+}$ . However, the HBHG(CBCG)CO pulse sequence, might find applications in further studies in Dr. Sykes' laboratory.

## Results and Discussion

### *HBHG(CBCG)CO pulse sequence*

To follow the side chain carboxyl groups of cNTnC, we have modified the 3D pulse sequence  $^{13}\text{C}'-(^{13}\text{C})-^1\text{H}$  HSQC/HSQC from Yamazaki et al. (1) to create a 2D pulse sequence, referred herein as HBHG(CBCG)CO, in which H $^{\beta}$ /C $^{\beta}$  (or H $^{\gamma}$ /C $^{\gamma}$ ) and carboxyl resonances (C $^{\gamma}$ ) are correlated.  $^{13}\text{C}'-(^{13}\text{C})-^1\text{H}$  HSQC/HSQC was developed similarly to the HCACO pulse sequence, in which H $^{\alpha}$ , C $^{\alpha}$ , and C $^{\gamma}$  are correlated (18). At time point *a* (Fig. V-1), the single quantum coherence C $^{\beta}_{\gamma}$ H $^{\beta}_{\gamma}$  (or C $^{\gamma}_{\gamma}$ H $^{\gamma}_{\gamma}$ ) is generated as a result of INEPT. During evolution for a total time period of  $1/2J_{\text{C}^{\beta}\text{C}^{\gamma}}$ , the C $^{\beta}$  (C $^{\gamma}$ ) coherences evolve with the C $^{\beta}$ C $^{\gamma}$  (C $^{\gamma}$ C $^{\gamma}$ ) J-couplings active and the carbon chemical shifts inverted and refocused by the composite Shaka pulse (19, 20). The magnetization is thus transferred on the carboxyl carbon C $^{\gamma}$  (point *b*). C $^{\gamma}$  coherences are labeled with the chemical shift during evolution time  $t_1$ , while the chemical shifts for H and C $^{\beta}$  are refocused (point *c*). After this, magnetization from C $^{\gamma}$  is transferred to C $^{\beta}$  (C $^{\gamma}$ ) during the evolution under the active C $^{\beta}$ -C $^{\gamma}$  (C $^{\gamma}$ -C $^{\gamma}$ ) J-couplings, with chemical shifts inverted and refocused by the composite Shaka pulse (point *d*). Reversed INEPT is used to transfer the magnetization from C $^{\beta}$  (C $^{\gamma}$ ) to the observable magnetization on H $^{\beta}$  (H $^{\gamma}$ ), point *e*, followed

by the acquisition of FID. Pulse sequence was coded and compiled for VnmrJ, with its code presented in Appendix A. HBHG(CBCG)CO was tested on the previously assigned complex of cTnC with EMD57033 (21). The carboxyl  $^{13}\text{C}$  NMR chemical shifts were in agreement with those previously assigned (data not shown).

*pH titration of cTnC•Ca<sup>2+</sup> monitored by HBHG(CBCG)CO spectroscopy*

The lower pH limit for these experiments was determined by the precipitation of cTnC•Ca<sup>2+</sup>, which became pronounced at pH ~5. Ionization constants for glutamates and aspartates are typically found in the range of pKa  $\approx$  2–5. The majority of cross peaks correlating the H <sup>$\beta$</sup>  (H <sup>$\gamma$</sup> ) resonance on the x-axis and the carboxyl  $^{13}\text{C}$  resonance on the y-axis did not shift significantly within the pH range used (Fig. V-2). Interestingly, a few cross peaks changed their chemical shifts starting from pH ~7–8. These residues are glutamates (left panel) and aspartates (right panel) with pKa's higher than typical, which might suggest their biological significance. However, protein precipitation precluded further pH titrations, and pKa's were not determined.

*pH titration of cTnC•Ca<sup>2+</sup> monitored by {<sup>1</sup>H, <sup>15</sup>N} NMR spectroscopy*

2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy provides a useful insight on protein integrity due to the unique sensitivity of backbone amide chemical shifts. In acidic and basic conditions the exchange of labile amide protons is often facilitated, which renders the observation of labile protons difficult (22). The general trend of pH dependences (Fig. V-3) supported the previous assignment and correlates with the known secondary structure of the protein (23). Residues involved in hydrogen bonding interactions, such as elements of  $\alpha$ -helices, are protected against <sup>1</sup>H exchange the most; for example, Leu<sup>12</sup>, Ile<sup>26</sup>, Met<sup>47</sup>, Glu<sup>56</sup>, and Phe<sup>77</sup> (Figs. V-3a and V-3b), whose intensities have not significantly decreased up to pH ~9. Residues belonging to the loops, for example, Glu<sup>32</sup>, Gln<sup>50</sup>, and Gly<sup>68</sup>, disappeared from the 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC spectra earlier in titrations (Fig. V-3c). At high pH, the chemical shifts for remaining observable amides remained the same (data not shown), which indicates that the overall structure of cTnC•Ca<sup>2+</sup> remained unchanged up to ~ pH 9.5. In spite of the fact that many amides were not observable (50% cross-peaks were missing at pH 9.8),  $^{13}\text{C}$  and <sup>1</sup>H NMR of side chain atoms can be still used to explore the behavior of cTnC•Ca<sup>2+</sup> at these pH conditions. At pH ~5 and lower, cTnC•Ca<sup>2+</sup> underwent conformational changes, as demonstrated by

the shifting of some cross-peaks and appearance of new cross-peaks in 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC spectra (data not shown).

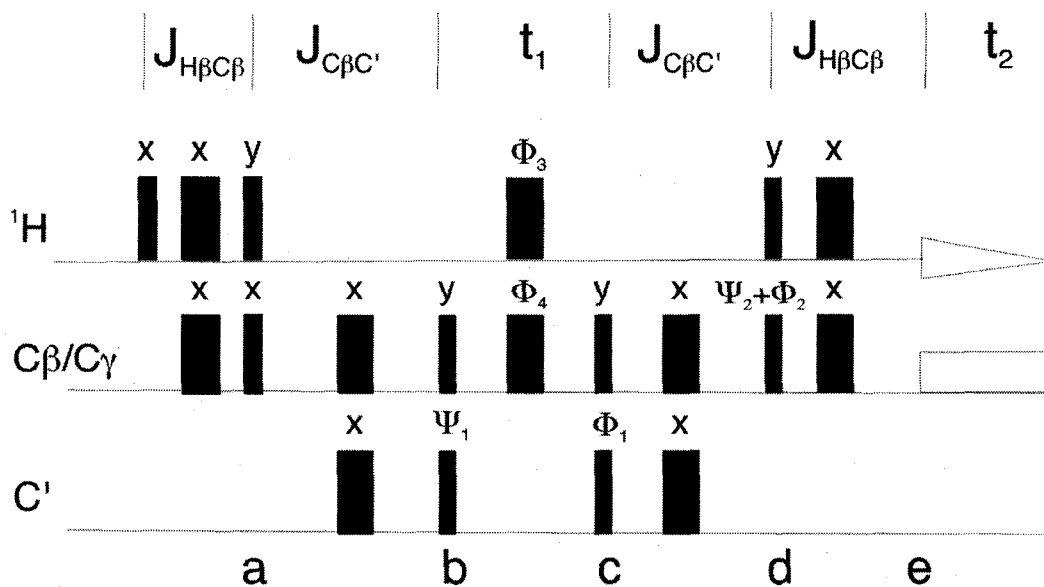
## Conclusions

Typically, studies of troponin C has been carried out within the physiological range of pH from ~6.7 to ~7 (23-26). In experiments aiming to determine the pKa of side chain carboxyl groups and their role in electrostatic interactions, the limits of pH in which cNTnC•Ca<sup>2+</sup> is studied have to be expanded. As it was determined here, the range of pH within which cNTnC•Ca<sup>2+</sup> exists as a folded protein with the known structure was limited by pH ~5, thus making the chosen method for determination of pKa's unsuitable. The modified and tested pulse sequence HBHG(CBCG)CO might find its application in the studies of different protein complexes.

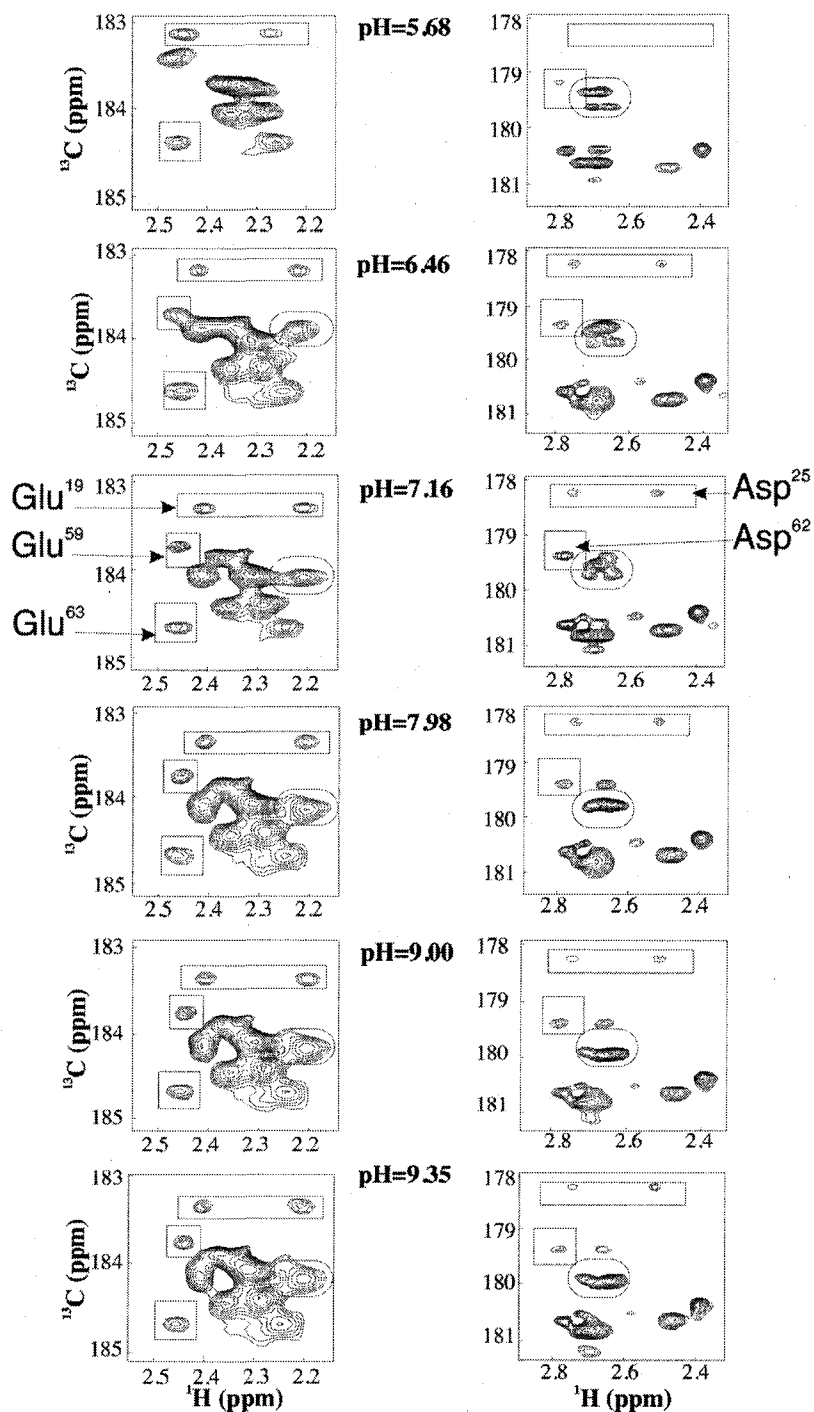
## Materials and Methods

Titration were performed on the sample containing 0.6 mM of <sup>15</sup>N labeled cNTnC, in a buffer consisted of 100 mM KCl, 5% of D<sub>2</sub>O, and 0.2 mM DSS and 2 mM of each, formate, imidazole, tris, and piperazine, that were used as internal pH indicators. The protein, cNTnC•Ca<sup>2+</sup>, was expressed and purified as it is described earlier (27). Aliquots of 1N NaOH or HCl were added in order to adjust the pH of a sample. 1D <sup>1</sup>H, 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR and 2D HBHG(CBCG)CO spectra were acquired at 30 °C on a Varian Unity 600 MHz spectrometer. <sup>1</sup>H chemical shifts in 1D <sup>1</sup>H NMR spectra corresponding to pH indicators (formate, tris, imidazole and piperazine (Appendix B)) were used to determine pH.

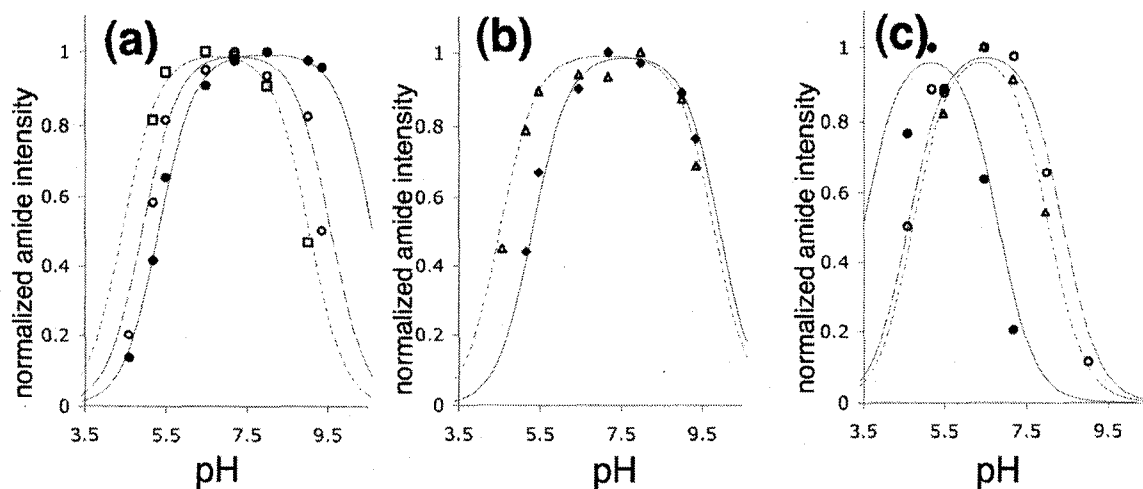
**Figure V-1.** Pulse scheme for the experiment employed to correlate the  $C^\beta/H^\beta$  (or  $C'/H'$ ) and carboxyl carbon coherences ( $C'$ ). Narrow and wide pulses correspond to  $90^\circ$  and  $180^\circ$  flip angles. Phases are as indicated. The carrier is set to 35 ppm for the  $C^\beta$  (or  $C'$ ) pulses and to 180 ppm for the carboxyl pulses. Phase cycling is the following:  $\Phi_1=(x, -x)$ ,  $\Phi_2=(2x, 2(-x))$ ,  $\Phi_3=\Phi_4=(4x, 4(-x))$ ,  $\Psi_1=(x,y)$ ,  $\Psi_1=(x,-y)$ .



**Figure V-2.** Expanded regions of the 2D HBHG(CBCG)CO NMR spectra of cNTnC•Ca<sup>2+</sup> obtained at different pH values. Cross peaks that demonstrate little or no movement are shown inside the boxes, cross peaks that undergo some movement are shown inside the ovals. First signs of possible precipitation were observed at pH of 5.68. The assignment of some side chain carboxyl groups is indicated.



**Figure V-3.** The pH dependence of backbone amide cross peak intensities for residues Leu<sup>12</sup> (open square), Ile<sup>26</sup> (closed circle), and Met<sup>47</sup> (open circle) located on the helices N, A, and B respectively (a); for residues Glu<sup>56</sup> (open triangle) and Phe<sup>77</sup> (closed diamond) located on the helices C and D respectively (b), and for residues Glu<sup>32</sup> (closed circle), Gln<sup>50</sup> (open triangle), and Gly<sup>68</sup> (open circle) located on the loops, including the Ca<sup>2+</sup> binding loop for residue Gly<sup>68</sup> (c). Solid and dashed lines represent the fits of the data using equations from (28). Resonances were assigned previously (23).





## References

- (1) Yamazaki, T., Yoshida, M., and Nagayama, K. (1993) Complete assignments of magnetic resonances of ribonuclease H from *Escherichia coli* by double- and triple-resonance 2D and 3D NMR spectroscopies. *Biochemistry* 32, 5656-69.
- (2) Bradbury, J. H., and Scheraga, H. A. (1966) Structural studies of ribonuclease. XXIV. The application of Nuclear Magnetic Resonance spectroscopy to distinguish between the histidine residues of ribonuclease. *J Am Chem Soc* 88, 4240.
- (3) Jeng, M. F., and Dyson, H. J. (1996) Direct measurement of the aspartic acid 26 pKa for reduced *Escherichia coli* thioredoxin by <sup>13</sup>C NMR. *Biochemistry* 35, 1-6.
- (4) Perez-Canadillas, J. M., Campos-Olivas, R., Lacadena, J., Martinez del Pozo, A., Gavilanes, J. G., Santoro, J., Rico, M., and Bruix, M. (1998) Characterization of pKa values and titration shifts in the cytotoxic ribonuclease alpha-sarcin by NMR. Relationship between electrostatic interactions, structure, and catalytic function. *Biochemistry* 37, 15865-76.
- (5) Oda, Y., Yamazaki, T., Nagayama, K., Kanaya, S., Kuroda, Y., and Nakamura, H. (1994) Individual ionization constants of all the carboxyl groups in ribonuclease HI from *Escherichia coli* determined by NMR. *Biochemistry* 33, 5275-84.
- (6) Forman-Kay, J. D., Clore, G. M., and Gronenborn, A. M. (1992) Relationship between electrostatics and redox function in human thioredoxin: characterization of pH titration shifts using two-dimensional homo- and heteronuclear NMR. *Biochemistry* 31, 3442-52.
- (7) Joshi, M. D., Hedberg, A., and McIntosh, L. P. (1997) Complete measurement of the pKa values of the carboxyl and imidazole groups in *Bacillus circulans* xylanase. *Protein Sci* 6, 2667-70.
- (8) Kohda, D., Sawada, T., and Inagaki, F. (1991) Characterization of pH titration shifts for all the nonlabile proton resonances a protein by two-dimensional NMR: the case of mouse epidermal growth factor. *Biochemistry* 30, 4896-900.
- (9) Kuhlman, B., Luisi, D. L., Young, P., and Raleigh, D. P. (1999) pKa values and the pH dependent stability of the N-terminal domain of L9 as probes of electrostatic interactions in the denatured state. Differentiation between local and nonlocal interactions. *Biochemistry* 38, 4896-903.
- (10) Khare, D., Alexander, P., Antosiewicz, J., Bryan, P., Gilson, M., and Orban, J. (1997) pKa measurements from nuclear magnetic resonance for the B1 and B2 immunoglobulin G-binding domains of protein G: comparison with calculated values for nuclear magnetic resonance and X-ray structures. *Biochemistry* 36, 3580-9.
- (11) Chen, H. A., Pfuhl, M., McAlister, M. S., and Driscoll, P. C. (2000) Determination of pK(a) values of carboxyl groups in the N-terminal domain of rat CD2: anomalous pK(a) of a glutamate on the ligand-binding surface. *Biochemistry* 39, 6814-24.
- (12) McIntosh, L. P., Hand, G., Johnson, P. E., Joshi, M. D., Korner, M., Plesniak, L. A., Ziser, L., Wakarchuk, W. W., and Withers, S. G. (1996) The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a <sup>13</sup>C-NMR study of *bacillus circulans* xylanase. *Biochemistry* 35, 9958-66.

- (13) Pascal, S. M., Singer, A. U., Yamazaki, T., Kay, L. E., and Forman-Kay, J. D. (1995) Structural and dynamic characterization of an SH2 domain-phosphopeptide complex by NMR approaches. *Biochem Soc Trans* 23, 729-33.
- (14) Joshi, M. D., Sidhu, G., Nielsen, J. E., Brayer, G. D., Withers, S. G., and McIntosh, L. P. (2001) Dissecting the electrostatic interactions and pH-dependent activity of a family 11 glycosidase. *Biochemistry* 40, 10115-39.
- (15) Pascal, S. M., Yamazaki, T., Singer, A. U., Kay, L. E., and Forman-Kay, J. D. (1995) Structural and dynamic characterization of the phosphotyrosine binding region of a Src homology 2 domain--phosphopeptide complex by NMR relaxation, proton exchange, and chemical shift approaches. *Biochemistry* 34, 11353-62.
- (16) Yamazaki, T., Pascal, S. M., Singer, A. U., Forman-Kay, J. D., and Kay, L. E. (1995) NMR Pulse Schemes for the Sequence-Specific Assignment of Arginine Guanidino  $^{15}\text{N}$  and  $^1\text{H}$  Chemical Shifts in Proteins. *J Am Chem Soc* 117, 3556-3564.
- (17) Iwahara, J., Jung, Y. S., and Clore, G. M. (2007) Heteronuclear NMR spectroscopy for lysine NH(3) groups in proteins: unique effect of water exchange on  $(^{15}\text{N})$  transverse relaxation. *J Am Chem Soc* 129, 2971-80.
- (18) Grzesiek, S., and Bax, A. (1993) The origin and removal of artifacts in 3D HACO spectra of proteins uniformly enriched with  $^{13}\text{C}$ . *J Mag Res B* 102, 103-106.
- (19) Shaka, A. J. (1985) Composite pulses for ultra-broadband spin inversion. *Chem Phys Lett* 120, 201-205.
- (20) Mohebbi, A., and Shaka, A. J. (1991) Improvements in carbon-13 broadband homonuclear cross-polarization for 2D and 3D NMR. *Chem Phys Lett* 178, 374-378.
- (21) Wang, X., Li, M. X., Spyropoulos, 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  $\text{Ca}^{2+}$  sensitizing drug EMD 57033. *J Biol Chem* 276, 25456-66.
- (22) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
- (23) Sia, S. K., Li, M. X., Spyropoulos, L., Gagne, 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, 18216-21.
- (24) Li, M. X., Spyropoulos, L., and Sykes, B. D. (1999) Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. *Biochemistry* 38, 8289-98.
- (25) 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, 27024-34.
- (26) Gagne, S. M., Tsuda, S., Li, M. X., Smillie, L. B., and Sykes, B. D. (1995) Structures of the troponin C regulatory domains in the apo and calcium-saturated states. *Nat Struct Biol* 2, 784-9.
- (27) Li, M. X., Corson, D. C., and Sykes, B. D. (2002) Structure determination by NMR. Isotope labeling. *Methods Mol Biol* 173, 255-65.
- (28) O'Neil, J. D., and Sykes, B. D. (1989) NMR studies of the influence of dodecyl sulfate on the amide hydrogen exchange kinetics of a micelle-solubilized hydrophobic tripeptide. *Biochemistry* 28, 699-707.

## CHAPTER VI

### Conclusions

The aim of this thesis was to provide an insight on the mechanism by which cardiomyopathy mutations in troponin C disrupt its normal function and may lead to the development of the disease. We have established, first, that the L29Q and E59D/D75Y mutants have unperturbed affinity to  $\text{Ca}^{2+}$ . Instead their affinity towards the switch region of cTnI (cTnI<sub>147-163</sub>) has been significantly changed especially in the presence of the N-terminus of cTnI (cTnI<sub>1-29</sub>). This points to the complex nature of the interactions on the biochemical scale where  $\text{Ca}^{+2}$  binding leads to the conformational changes in cTnC, which modulates and being modulated by the interactions with cTnI. Second, we have determined that the G159D mutation results in the weakened interaction with the anchoring region of cTnI (cTnI<sub>34-71</sub>), which might lead to a weaker anchoring of cTn to the thin filament. Third, we have established the important role played by the N-terminus of cTnI (cTnI<sub>1-29</sub>/cTnI<sub>1-29</sub>PP) in modulating the binding of the switch region of cTnI (cTnI<sub>147-163</sub>). Fourth, we have determined that the interaction between the N-terminus of cTnI and the N-domain of cTnC is  $\text{Ca}^{2+}$  independent and electrostatically driven.

Results pertaining to the modulation of the cNTnC binding to the switch region of cTnI can be summarized as the following (Fig. VI-1): the N-terminus of cTnI in unphosphorylated form is bound to cNTnC (Fig. VI-1a) which lessens the cNTnC affinity towards the switch region of cTnI. In original model, this interaction was not present. Upon PKA phosphorylation, the N-terminus of cTnI partially or completely dissociates from cNTnC, resulting in the increased affinity of cNTnC towards the switch region of cTnI (Fig. VI-1b). This would likely result in the increased muscle contractility induced by PKA phosphorylation as a response to  $\beta$ -adrenergic stimulation. Thus, as a result of PKA phosphorylation, the cNTnC affinity towards the switch region of cTnI is modulated, whereas the affinity towards  $\text{Ca}^{2+}$  remains unperturbed. These findings compliment the existing model proposed to explain the mechanism of PKA signaling in cTn (1, 2).

Within the cTn complex, some of the interactions between cTnC and cTnI remain to be visualized at the molecular level. For example, the binding of cTnC to the N-terminus or the inhibitory region of cTnI, which occur within the poorly structured regions of cTnI, are not easily amenable to traditional structural techniques. A major advance in technology might be needed in order to visualize the complete interactions

between the contractile proteins in the Tn complex. Without understanding of the molecular details of these interactions, the picture of cardiac contractility, as well as the effect of mutations, will remain incomplete.

As it is becoming more evident, the dynamics and conformational flexibility of proteins or their domains play a major role in the functioning of biological molecules. In order for cTnC to reach the N-terminus of cTnI, the N-domain of cTnC must be folded onto the C-domain of cTnC as was observed in the crystal structure of the cTn core (3). This 3D quaternary arrangement was not emphasized previously, and our data (Chapters II and IV) provide evidence for such an arrangement. However, in order to interact with other regions of cTnI, the central linker of cTnC might need to be straighter similar to sTnC (4). This suggests the possibility of the large conformational change for cTnC during the relaxation-contraction cycle (Fig. VI-2) and the need for the additional states in the overall model of  $\text{Ca}^{2+}$  regulated cardiac contraction. The attempt to visualize individual conformations along the relaxation-contraction cycle might increase the level of complexity of the  $\text{Ca}^{2+}$  regulated cardiac contraction mechanism, but might help explain special features of cardiac muscle such as the effect of mechanical activation (5).

The organization of contractile proteins within the sarcomere, although studied in much detail, is not yet completely understood. For example, the interaction of cTnT with tropomyosin is not well studied (6). The mechanism of stretch activation is just beginning to be understood (5, 7). The effects of PKA, PKC, PKG, and PKD phosphorylation on the molecular interactions within the contractile machinery remain in rudimentary stage (8). The understanding of the interplay between the expression patterns of particular cTnI and cTnT isoforms during the development (and in disease) and the molecular details of cardiac muscle contraction during the development will require much future effort (9). In addition, cardiomyocytes vary in appearance and function depending on their location in the heart (10), the molecular details of which the current model of cardiac contraction does not explain. In spite of tremendous amount of research in this field and many keystone discoveries, we still have a partially illuminated picture, with many pieces of the puzzle missing and poorly connected.

The mathematical reconstruction of the individual events of the relaxation-contraction cycle, from  $\text{Ca}^{2+}$  release by the sarcolemmal  $\text{Ca}^{2+}$  channel to the binding of cross-bridges, with the understanding of thermodynamics and kinetics of every individual steps, would allow us to answer many questions regarding the perturbed contractility in a quantitative manner. There is a large body of thermodynamic and kinetic data presented

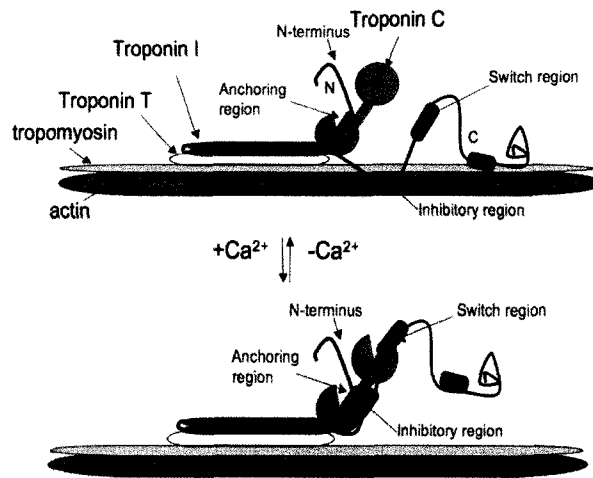
in literature, which is often obtained with different techniques providing partially or completely incompatible results. Due to the differences in solution compositions, pH, ionic strength, viscosity, etc, thermodynamic and kinetic data obtained *in vitro* might be very different from the *in vivo* values. Such a mathematical model will also require the consideration of non-equilibrium conditions and appears to be a major endeavor, not easily feasible. Recent computational efforts used global parameters such as the  $\text{Ca}^{2+}$  sensitivity of the myofilament in order to simulate the effect of FHC mutations on the development of force (11), without considering the fine tuned machinery of the  $\text{Ca}^{2+}$  regulation.

Studies of individual cardiomyopathy mutations greatly promote the expansion of our knowledge pertaining to the  $\text{Ca}^{2+}$  regulated contraction. As it was demonstrated in Chapter II, the effects of L29Q and E59D/D75Y mutations needed to be considered in the context of the N-terminus of cTnI. However, there are following challenges in dissecting the effects of mutations on contractility. First, traditional methods have their limitations and associated errors, within which physiologically important differences might remain undetected. It is noteworthy, that human mutations with severe effects would be lethal, and only mild mutations or polymorphisms are expected. Second, the model of  $\text{Ca}^{2+}$  contraction is incomplete, and therefore the differences found in protein functions might not be important in the overall picture of contraction or compensated by other mechanisms. Third, the number of human mutations is increasingly large and is expected to increase even more, which indicates the need for technological advances, where these mutations can be studied in throughput manner, side-by-side, and with careful controls.

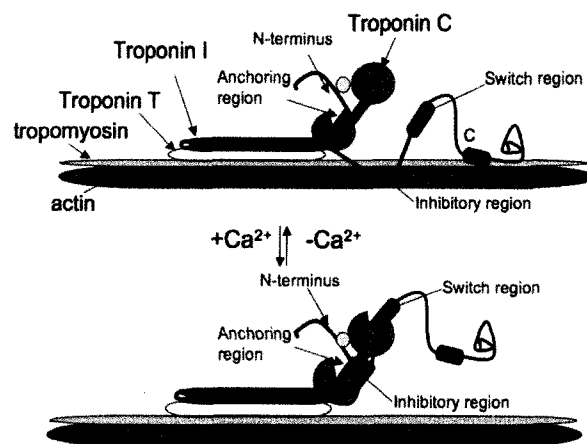
Understanding of human mutations is an exciting quest, which may help us to unravel the basis of individual differences in human pathology and physiology. With the knowledge of mutation effects and their interplay with the environment, we might approach in every individual case the understanding of the role of nature vs nurture. In future, it might be possible to predict and treat health issues, control dieting and exercising, deal with psychological problems, etc, in a genetically tailored manner. As a field of science, biochemistry might eventually exhaust the range of problems traditionally addressed by biochemists, but the biochemical approach should spread into other fields of science: medicine, nutrition, psychology, etc. Using biochemistry as a platform in other fields of sciences, human curiosity will find even more unanswered questions. This thesis is a part of the ongoing process undertaken by humankind aimed towards the understanding of ourselves and the world around us.

**Figure VI-1.** The model of  $\text{Ca}^{2+}$  regulated cardiac muscle contraction with the results of our work added. Two states of contractility are considered, with the first state in the absence of PKA phosphorylation (a), and the second in the presence of PKA phosphorylation (b). Phosphorylation is denoted as a yellow dot. Notice that in both states, the switch region of cTnI binds to cTnT, however, with different affinities.

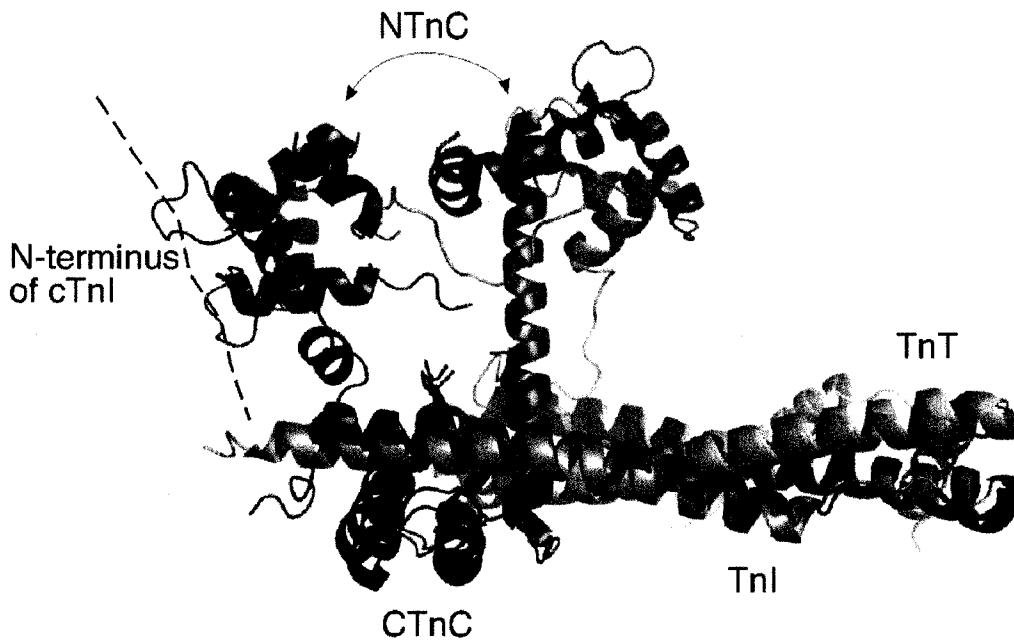
(a)



(b)



**Figure VI-2.** Possible conformational change of cTnC during the contraction-relaxing cycle. Crystallographic structures of sTn and cTn (3, 4) are aligned over their cTnI subunits (grey). The C-domain of skeletal TnC (green) and cardiac TnC (red) are well overlapped, however the N-domains are oriented differently. The conformational change needed to bring cNTnC to overlap with sNTnC is shown with the arrow. In cardiac structure, the N-terminus of cTnI is positioned next to cNTnC, making possible the interaction with cNTnC (shown as a dashed line). If cNTnC moved similarly to sNTnC, such an interaction would be impossible.



## References

- (1) Howarth, J. W., Meller, J., Solaro, R. J., Trewhella, J., and Rosevear, P. R. (2007) Phosphorylation-dependent conformational transition of the cardiac specific N-extension of troponin I in cardiac troponin. *J Mol Biol* 373, 706-22.
- (2) Solaro, R. J., Rosevear, P., and Kobayashi, T. (2008) The unique functions of cardiac troponin I in the control of cardiac muscle contraction and relaxation. *Biochem Biophys Res Commun* 369, 82-7.
- (3) Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form. *Nature* 424, 35-41.
- (4) Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) Ca(2+)-regulated structural changes in troponin. *Proc Natl Acad Sci U S A* 102, 5038-43.
- (5) De Nicola, G., Burkart, C., Qiu, F., Agianian, B., Labeit, S., Martin, S., Bullard, B., and Pastore, A. (2007) The structure of Lethocerus troponin C: insights into the mechanism of stretch activation in muscles. *Structure* 15, 813-24.
- (6) Perry, S. V. (1998) Troponin T: genetics, properties and function. *J Muscle Res Cell Motil* 19, 575-602.
- (7) Fukuda, N., and Granzier, H. L. (2005) Titin/connectin-based modulation of the Frank-Starling mechanism of the heart. *J Muscle Res Cell Motil* 26, 319-23.
- (8) Metzger, J. M., and Westfall, M. V. (2004) Covalent and noncovalent modification of thin filament action: the essential role of troponin in cardiac muscle regulation. *Circ Res* 94, 146-58.
- (9) Gomes, A. V., Venkatraman, G., Davis, J. P., Tikunova, S. B., Engel, P., Solaro, R. J., and Potter, J. D. (2004) Cardiac troponin T isoforms affect the Ca(2+) sensitivity of force development in the presence of slow skeletal troponin I: insights into the role of troponin T isoforms in the fetal heart. *J Biol Chem* 279, 49579-87.
- (10) Bootman, M. D., Higazi, D. R., Coombes, S., and Roderick, H. L. (2006) Calcium signalling during excitation-contraction coupling in mammalian atrial myocytes. *J Cell Sci* 119, 3915-25.
- (11) Kataoka, A., Hemmer, C., and Chase, P. B. (2007) Computational simulation of hypertrophic cardiomyopathy mutations in troponin I: influence of increased myofilament calcium sensitivity on isometric force, ATPase and [Ca2+]i. *J Biomech* 40, 2044-52.



## APPENIDIX A

### The code for the HBHG(CBCG)CO pulse sequence

```
/* gChsqc_Chsgc.c
Sequence for CB/CG CO HB/HG experiment for Asp and Glu
      F2  F1  F3 (Presently F1F3 2D experiment)
Modified from Yamazaki et al. Biochem 32 5656 (1994)
Original pulse sequence is called 3D Y-X-1H HSQC/HSQC

set dof for F2 to 35.0 ppm (centered on CB/CG)
F1 is set to 175.0 ppm (centered on CO)

Set:  tauHC = 1/(4JHC)
      tauCC = 1/(4 JCC)
      pwC90 for selective CB/CG or CO pulse depending upon dof
      pwC180 for selective SLP CB/CG 180 null at CO, decpwr same as pwC90 pulse
above

satmode = 'ynn' for presat
satdly = length of presaturation period
satpwr = power level for presat `*/

#include <standard.h>
#include "Pbox_bio.h" /* Pbox Bio Pack Pulse Shaping Utilities */

shaka6(pwC90)
double pwC90;
{
    /* Shaka composite (10.9333*pwC90) */
    decrgpulse(pwC90*158.0/90.0, zero, 0.0, 0.0);
    decrgpulse(pwC90*171.2/90.0, two, 0.0, 0.0);
    decrgpulse(pwC90*342.8/90.0, zero, 0.0, 0.0);
    decrgpulse(pwC90*145.5/90.0, two, 0.0, 0.0);
    decrgpulse(pwC90*81.2/90.0, zero, 0.0, 0.0);
    decrgpulse(pwC90*85.3/90.0, two, 0.0, 0.0);
}

static double d2_init=0.0;

static int
phi1[2] = {0,2}, /* phi 123&5 as per Yamazaki */
phi2[4] = {0,0,2,2},
phi3[8] = {0,0,0,0,2,2,2,2},
phi5[16] = {0,0,0,0,0,0,0,0,2,2,2,2,2,2,2,2},
phi11[1] = {0}, /* psi 1 as per Yamazaki */
rec[4] = {0,2,2,0};

pulsesequenece()
{
```

```

/* Declare variables */
char
    satmode[MAXSTR];

getstr("satmode",satmode);

int
    t1_counter = 0,
    phase = getval("phase");

double
    d0 = getval("d0"),
    satdly = getval("satdly"),
    satpwr = getval("satpwr"),
    satfrq = getval("satfrq"),
    tauHC = getval("tauHC"),
    tauCC = getval("tauCC"),
    pwC = getval("pwC"),
    pwClvl = getval("pwClvl"),
    compC = getval("compC"),
    pwC90 = getval("pwC90"), /* 90 deg CB/CG or CO pulse at dpwrf rf2*/ pwC180,
    pwN = getval("pwN"),
    pwNlvl = getval("pwNlvl"),
    gt0 = getval("gt0"),
    gt1 = getval("gt1"),
    gt3 = getval("gt3"),
    gt5 = getval("gt5"),
    gzlvl0 = getval("gzlvl0"),
    gzlvl1 = getval("gzlvl1"),
    gzlvl3 = getval("gzlvl3"),
    gzlvl5 = getval("gzlvl5"),
    rf0,
    rf2; /* fine dec power for CB/CG region selective pulse*/
    /* also used for fine dec power for CO 90 deg pulse at CB/CG */
    pwC180=2*pwC90;

/* Parameter checks */
if ( pw > 20e-6 )
{
    printf( "your pw seems too long !!! ");
    printf( " pw must be <= 20 usec \n");
    psg_abort(1);
}
if ( pwC > 30e-6 )
{
    printf( "your pwC seems too long !!! ");
    printf( " pwC must be <= 30 usec \n");
    psg_abort(1);
}
if ( pwN > 50e-6 )

```

```

    {
        printf( "your pwN seems too long !!! ");
        printf( " pwN must be <= 50 usec \n");
        psg_abort(1);
    }
    if ( dm[C] == 'y')
    {
        if ( at > 0.11 )
        {
            printf( " acquisition time is too long for decoupling !!!\n");
            psg_abort(1);
        }
        if ( dpwr > 46 )
        {
            printf( "incorrect power for dpwr !!! ");
            printf( " must limit dpwr <= 46 \n");
            psg_abort(1);
        }
    }
    if ( ( dm[A] == 'y') || ( dm[B] == 'y'))
    {
        printf( "no decoupling should be done during status periods A or B !!!\n");
        psg_abort(1);
    }
    if ( gt1 > 1.5e-3 ) { printf( "gt1 is too long !!!\n");
        psg_abort(1);
    }
}

/* Set variables */
settable(t1, 2, phi1);
settable(t2, 4, phi2);
settable(t3, 8, phi3);
settable(t5, 16, phi5);
settable(t11, 1, phi11);
settable(t60, 4, rec);

/* Phase incrementation for hypercomplex data */
if ( phase == 2 )
{
    tsadd(t11,1,4);
}

if( ix == 1) d2_init = d2;
t1_counter = (int) ( (d2-d2_init)*sw1 + 0.5 );
if(t1_counter % 2)
{ tsadd(t11,2,4); tsadd(t60,2,4); }

/* 90 degree pulse at C0 (175ppm) null at CB/CG (35ppm) */
/* 90 degree regional selective pulse at CB/CG (35ppm) */
rf0 = 4095.0;
rf2 = (compC*4095.0*pwC)/pwC90; /* power level for CB/CG pulses */

```

```

rf2 = (int) (rf2 + 0.5);

/* Begin the actual pulse sequence */
status(A);
delay(d0);
txphase(zero);
obspower(tpwr);
decphase(zero);
decpower(pwClvl);
dec2phase(zero);
dec2power(pwNlvl);

if ( satmode[A] == 'y' )
{
obspower(satpwr);
obsoffset(satfrq);
rgpulse(satdly,zero,rof1,rof2);
obspower(tpwr);
obsoffset(tof);
delay(1.0e-3);
}
else
{
delay(d1);
}
rcvroff();
status(B);

if(gt1>0.2e-6)
{
delay(10.0e-6);
decrpulse(pwC,zero,0.0,0.0);
delay(0.2e-6);
zgradpulse(gzlvl1,gt1);
}
delay(1.0e-3);

rgpulse(pw,zero,0.0,0.0); /* INEPT section HA->CA, HB->CB, HG->CG */
delay(2.0e-6);
zgradpulse(gzlvl0,gt0);
delay(tauHC-gt0-2.0e-6);
simpulse(2.0*pw,2*pwC,zero,zero,0.0,0.0);
delay(tauHC-gt0-500.0e-6);
zgradpulse(gzlvl0,gt0);
txphase(one);
delay(500.0e-6);
rgpulse(pw,one,0.0,0.0);
delay(2.0e-6);
zgradpulse(gzlvl3,gt3);
delay(200.0e-6);
decrpulse(pwC,zero,0.0,0.0);

```

```

delay(tauCC-0.5*10.9333*pwC);
shaka6(pwC); /* Shaka composite 180 refocusing and inversion pulse */
decpwrf(rf2);
delay(tauCC-0.5*10.9333*pwC);
decrpulse(pwC90,one,0.0,0.0);
delay(2.0e-6);
decoffset(dof+(180-35)*dfrq);
delay(10.0e-6);
decrpulse(pwC90, t11, 0.0, 0.0); /* CO SQC */

if (d2 > pwC180)
{
    delay(d2/2);
    simshaped_pulse("", "cocb180", 2*pw, pwC180, t3, t3, 0.0, 0.0);
    delay(d2/2);
}
else
    delay(d2);

decrpulse(pwC90, t1, 0.0, 0.0);
delay(2.0e-6);
decoffset(dof);
delay(10.0e-6);
decrpulse(pwC90, one, 0.0, 0.0);
decpwrf(rf0);
decphase(zero);
delay(tauCC-0.5*10.9333*pwC);
shaka6(pwC); /* Shaka composite 180 refocusing and inversion pulse */
delay(tauCC-0.5*10.9333*pwC);

simpulse(pw, pwC, one, t2, 0.0, 0.0);
delay(2.0e-6);
txphase(zero);
decphase(zero);
zgradpulse(gzlv15, gt5);
delay(tauHC-pwC -gt5 - 2.0e-6);
simpulse(2*pw, 2*pwC, zero, zero, 0.0, 0.0);
delay(tauHC-pwC-gt5-500.0e-6-2.0*POWER_DELAY);
zgradpulse(gzlv15, gt5);
decpower(dpwr);
dec2power(dpwr2);
delay(500.0e-6);

status(C); /* Acquire the fid in t3 */
rcvtron();
delay(rof2);
setreceiver(t60);
}

```

## APPENIDIX B

### Buffer components as internal pH indicators for biomolecular NMR\*

#### Outline

During NMR experiments, pH of the sample can be determined using a non-invasive convenient technique where the components of the buffer serve as internal pH indicators. Formate, tris, piperazine, and imidazole, whose  $^1\text{H}$  NMR chemical shifts are sensitive to pH in a range from 2.5 to 9.8, were used to report pH of the sample in this work. This method is suitable for a wide range of applications where samples are handled intensively during NMR titrations or in high throughput analysis in structural genomics or metabolomics.

#### Introduction

The vast majority of biological processes occur under a stringent control of pH, which should be reproduced during *in vitro* experiments. The most common method to report pH *in vitro* employs a pH sensitive glass electrode. This method, although well established and widely used, has significant disadvantages in solution state biological NMR experiments. First, positively charged proteins and peptides are frequently nonspecifically absorbed by a negatively charged glass surface of an electrode, which leads to the loss of a protein, contamination of the glass surface, and as a result to the distortion of pH calibration. Second, due to the moistening of the electrode surface, a fraction of protein solution is retained on the walls of an electrode, leading to an error in the determination of sample volume, which is especially significant when amplified during repeated pH measurements. Third, the measurement of pH with an electrode is an invasive technique, during which the impurities affecting equilibria and promoting sample degradation might be introduced. This is especially important for rare samples or for the samples particularly sensitive to denaturants and degradation. Fourth, in the case of high throughput analysis, pH measurements with an electrode amount to a substantial time cost.

---

\* The version of this appendix was published. Baryshnikova O.K., Williams T.C., and Sykes B.D. (2008) *J Biomol NMR*. 41(1): 5-7.

The idea of monitoring pH using NMR chemical shifts of  $^{19}\text{F}$ ,  $^{13}\text{C}$ , and  $^{31}\text{P}$  has found numerous biological (1-4) and chemical applications (5). Several compounds were proposed as pH indicators with their  $^1\text{H}$  NMR chemical shifts sensitive to pH: hydroxylammonium chloride (6), which is a reducing agent that might react with protein solution; phosphoric acid (7), which has limited solubility in the presence of biologically relevant metals, precipitating as a salt; mixture of methylated imidazoles (8), that covers mostly basic pH range; guanidinium chloride (6), which might promote highly undesirable protein unfolding; cacodylate (9), which is an arsenic compound and potentially toxic. While many compounds can be employed, common buffers are the most convenient to use as pH indicators. Tris, formate, piperazine, and imidazole were employed as internal reporters of pH, since they demonstrate a dependence of  $^1\text{H}$  NMR chemical shifts on pH and their 1D  $^1\text{H}$  NMR spectra are relatively simple, showing one or two peaks, so that the resonances of a protein or a peptide of question would be minimally obstructed. These compounds are also biologically inert including the lack of interactions with metal ions such as  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ , which is important in case of  $\text{Ca}^{2+}$  binding proteins for example troponin C. In Dr. Sykes' laboratory, an imidazole based buffer sensitive to pH in a region of 5.6-8.6 with a pKa of imidazole equal to 7.11 (100 mM KCl, 10 mM imidazole, 30°C) has been previously developed (T.C. Williams, unpublished data). This buffer has proved to be useful in a large number of NMR studies of troponin C performed in Dr. Sykes' laboratory over a period of more than a decade (10-12). In order to control pH over a wider range, a combination of four pH reporters (tris, formate, piperazine, and imidazole) has been developed.

## Results and Discussion

The buffer containing 100 mM KCl, 2 mM tris, 2 mM formate, 2 mM piperazine, 2 mM imidazole, 0.2 mM DSS, and 5%  $\text{D}_2\text{O}$  was calibrated using a standard pH-sensitive glass electrode (purchased from Radiometer America Inc.) in a 30°C thermostat. Solutions of 1 M HCl or NaOH were added to adjust the pH of a buffer. 1D  $^1\text{H}$  NMR spectra were recorded on a Varian Inova 500 MHz spectrometer at 30 °C using a standard 1D pulse sequence provided by BioPack. Calibrations were repeated three times with an example calibration shown in Figure B-1.  $^1\text{H}$  chemical shifts for formate were found to range from 8.236 to 8.442 for protonated and deprotonated forms correspondingly (Fig. B-1a, Table B-1), and the value of pKa was found to be  $3.56 \pm 0.05$ , as determined using xcvfit program (<http://www.bionmr.ualberta.ca/bds/software/index.html>).  $^1\text{H}$  chemical

shift for the H2 proton of imidazole ranged from 8.672 to 7.765 for protonated and deprotonated forms (Fig. B-1a, Table B-1) with the pKa of  $7.08 \pm 0.08$ .  $^1\text{H}$  chemical shifts for tris ranged from 3.732 to 3.509 for protonated and deprotonated forms (Fig. B-1b, Table B-1) with the pKa value of  $8.23 \pm 0.08$ .  $^1\text{H}$  chemical shifts for piperazine ranged from 3.585 to 3.113 for protonated and deprotonated forms (Fig. B-1b, Table B-1) with the pKa<sub>1</sub> value of  $5.61 \pm 0.07$ .

The values of pKa's and  $^1\text{H}$  chemical shifts for protonated and deprotonated forms determined during calibration were used in a macro written for VnmrJ and available on-line (<http://www.bionmr.ualberta.ca/mc/>). The macro utilizes the following relationship:  $\text{pH} = \text{pKa} - \lg \left( \frac{\delta_{\text{obs}} - \delta_{\text{HA}}}{\delta_{\text{A}} - \delta_{\text{obs}}} \right)$ , where  $\delta_{\text{HA}}$  and  $\delta_{\text{A}}$ , are the chemical shifts for protonated and deprotonated forms correspondingly, and  $\delta_{\text{obs}}$  is the  $^1\text{H}$  chemical shift observed for the solution with unknown pH. The most precise determination of pH is achieved when  $\delta_{\text{obs}}$  lies within the linear portion of chemical shift dependence on pH (Fig. B-1). The limits within which pH indicators provide the most reliable information were truncated for formate to pH of 2.5 ( $\delta$  of 8.246) and 4.7 ( $\delta$  of 8.427), for piperazine to pH 4.4 ( $\delta$  of 3.542) and 6.4 ( $\delta$  of 3.175), for imidazole to pH 5.8 ( $\delta$  of 8.635) and 8.6 ( $\delta$  of 7.803), and for tris to pH 6.7 ( $\delta$  of 3.725) and 9.8 ( $\delta$  of 3.516) (Table B-1). Thus, the overall pH range covered by these four compounds spans from 2.5 to 9.8, which can be easily extended if needed by adding another buffer component with an appropriate pKa value. Within these limits, the pH of solution can be determined with the error of  $\pm 0.05$  provided that the spectrum was correctly referenced to DSS and the error of a standard glass electrode used during calibration was of the same order or less.

**Table B-1.** pKa values,  $^1\text{H}$  chemical shifts for protonated and deprotonated forms of internal pH reporters and the limits of their applicability.

	pKa	$\sigma_{\text{HA}}$	$\sigma_{\text{A-}}$	applicability regions	
formate	$3.56 \pm 0.05$	8.236	8.442	2.5 ( $\delta$ of 8.246)	4.7 ( $\delta$ of 8.427)
piperazine	$5.61 \pm 0.07$	3.585	3.113	4.4 ( $\delta$ of 3.542)	6.4 ( $\delta$ of 3.175)
imidazole	$7.08 \pm 0.08$	8.672	7.765	5.8 ( $\delta$ of 8.635)	8.6 ( $\delta$ of 7.803)
tris	$8.23 \pm 0.08$	3.732	3.509	6.7 ( $\delta$ of 3.725)	9.8 ( $\delta$ of 3.516)



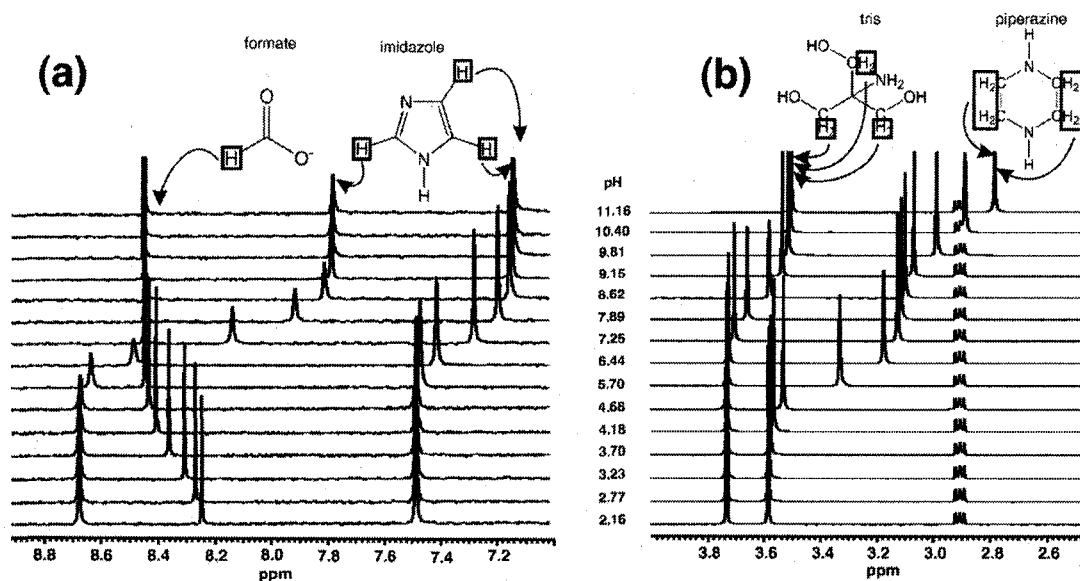
This method requires a calibration that needs to be performed only once for a particular set of conditions, such as temperature, salt concentration, etc. Parameters  $pK_a$ ,  $\delta_{HA}$ , and  $\delta_A$  for the buffers chosen do not change greatly over the typical range of conditions for biological NMR samples. For example, over the temperature of 25 to 40 °C, the  $pK_a$  of piperazine changes by 0.3, the  $pK_a$  of tris changes by 0.2, the  $pK_a$  of formate changes by 0.2, and the  $pK_a$  of imidazole, the most stable among studied compounds, changes by 0.1. The changes in  $pK_a$ 's with ionic strength ranging from ~ 20 mM to 100 mM were negligible within the errors of experiment; however, it is desirable to calibrate against a glass electrode for a particular set of conditions to achieve the best detection accuracy. Once the calibration has been performed, pH can be determined from the  $^1H$  chemical shifts of buffer compounds.

It might seem that the measuring of pH using NMR chemical shifts vs glass electrode is a more time consuming during the initial preparation of the sample due to the necessity for shimming and re-equilibrating each time after adjusting pH. However, the use of a glass electrode also requires that the sample and buffer are equilibrated at all temperatures, and that pH meter calibration curves are done at all temperatures, which is a tedious procedure and often not done properly. Once the sample is prepared, the NMR method is far easier to quickly measure pH within the sample due to any changes that might occur with time, reaction, or degradation, for example.

## Conclusions

The method described here utilized the pH dependence of  $^1H$  chemical shifts of buffer components such as formate, tris, imidazole, and piperazine to determine pH of the NMR sample. This method is devoid of many disadvantages of a glass electrode. It affords a better reproducibility of results, consistency among users, and availability for after the fact validation. It can be conveniently used for measuring pH in kinetic experiments or when monitoring long-term stability of a sample. It is indispensable when proteins are in shigemi style tubes or when pH is measured against temperature. Most importantly this method is highly beneficial during NMR titrations when samples are intensively handled or during high throughput analysis in proteomics, metabolomics, and structural genomics.

**Figure B-1.** 1D 500 MHz  $^1\text{H}$  NMR spectra of the buffer containing four internal pH indicators (formate (a), imidazole (a), tris (b), and piperazine (b)) as a function of pH. Protons that give rise to peaks in  $^1\text{H}$  1D NMR spectra are shown in boxes and connected to the corresponding resonances with arrows. The triplet on panel b corresponds to a  $\text{CH}_2$  resonance arising from DSS that served as a reference compound.



## References

- (1) Taylor, J., and Deutsch, C. (1988)  $^{19}\text{F}$ -nuclear magnetic resonance: measurements of  $[\text{O}_2]$  and pH in biological systems. *Biophys J* 53, 227-33.
- (2) DeFronzo, M., and Gillies, R. J. (1987) Characterization of methylphosphonate as a  $^{31}\text{P}$  NMR pH indicator. *J Biol Chem* 262, 11032-7.
- (3) Pietri, S., Miollan, M., Martel, S., Le Moigne, F., Blaive, B., and Culcasi, M. (2000)  $\alpha$ - and  $\beta$ -phosphorylated amines and pyrrolidines, a new class of low toxic highly sensitive  $^{31}\text{P}$  NMR pH indicators. Modeling of pKa and chemical shift values as a function of substituents. *J Biol Chem* 275, 19505-12.
- (4) Pietri, S., Martel, S., Culcasi, M., Delmas-Beauvieux, M. C., Canioni, P., and Gallis, J. L. (2001) Use of diethyl(2-methylpyrrolidin-2-yl)phosphonate as a highly sensitive extra- and intracellular  $^{31}\text{P}$  NMR pH indicator in isolated organs. Direct NMR evidence of acidic compartments in the ischemic and reperfused rat liver. *J Biol Chem* 276, 1750-8.
- (5) Vistad, O. B., Akporiaye, D. E., Taulelle, F., and Lillerud, K. P. (2003) Morpholine, an in situ  $\text{C-}^{13}$  NMR pH meter for hydrothermal crystallogenesis of SAPO-34. *Chemistry of Materials* 15, 1650-1654.
- (6) Rabenstein, D. L., and Fan, S. Y. (1986) Proton nuclear magnetic resonance spectroscopy of aqueous solutions: complete elimination of the water resonance by spin-spin relaxation. *Anal Chem* 58, 3178-84.
- (7) Szakacs, Z., Hagele, G., and Tyka, R. (2004)  $^1\text{H}/^{31}\text{P}$  NMR pH indicator series to eliminate the glass electrode in NMR spectroscopic pKa determinations. *Analytica Chimica Acta* 522, 247-258.
- (8) Sorkin, D. L., and Miller, A. F. (1997) Spectroscopic measurement of a long-predicted active site pK in iron-superoxide dismutase from *Escherichia coli*. *Biochemistry* 36, 4916-24.
- (9) Valcour, a. a., and Woodworth, R. C. (1986) Internal Proton Magnetic-Resonance Probes for Ph Titration of Proteins. *Journal of Magnetic Resonance* 66, 536-541.
- (10) Sia, S. K., Li, M. X., Spyropoulos, L., Gagne, 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, 18216-21.
- (11) Li, M. X., Spyropoulos, L., and Sykes, B. D. (1999) Binding of cardiac troponin-I147-163 induces a structural opening in human cardiac troponin-C. *Biochemistry* 38, 8289-98.
- (12) 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, 27024-34.

## APPENDIX C

### Structures of peptide agonists derived on the basis of SDF-1 $\alpha$ reveal a common motif possibly implicated in the binding to the CXCR4 receptor\*

#### Outline

CXCR4 is an important pharmacological target due to its involvement in HIV-1 pathogenesis and cancer metastasis. Two recently discovered allosteric agonists (17-mers) that bind and activate CXCR4, the ASLW and RSVM peptides, were analyzed using solution NMR spectroscopy. Both peptides assumed an extended backbone conformation with several well defined local motifs in the regions from residues 5 to 8 and 9 to 12. Similar examples in literature suggest the relevance of nascent structures in peptides to their biologically relevant conformations. The local structure in the region 9 to 12 was adopted by the entire ensemble of the ASLW peptide structures and by the subset of conformations for the RSVM peptide. The same turn was found in full length SDF-1 $\alpha$  and in the small family of the SDF-1 $\alpha$  N-terminal 17-mer. The significance of found local structures and implications for further drug design are discussed.

#### Introduction

Chemokine receptors belong to a family of G-protein coupled receptors (GPCRs) that are especially attractive as therapeutic targets (1-5). The chemokine receptor CXCR4, through interactions with the stromal derived factor, SDF-1 $\alpha$ , mediates trafficking of lymphocytes, memory immune responses (6, 7), vascular and germ cell development (8, 9), haematopoieses and lymphopoiesis (10), and neuronal patterning (11). CXCR4 has a crucial role in metastasis of various cancers (2, 12-17) and in HIV-1 pathogenesis (18), which provides a strong impetus for studying this system on a molecular level.

A number of approaches have been used to address the structural determinants of interactions between GPCRs and their ligands (19). Membrane bound proteins, such as GPCRs, represent a substantial challenge for traditional structural approaches using NMR and X-ray techniques. One approach employs structural studies of isolated biologically active peptides in order to infer information about the interactions with their cognate

---

\* The version of this chapter was published. Baryshnikova O.K., Rainey J.K., and Sykes B.D. (2005) *J Pep Res* 66 (S1), 12-21.

receptors and possible binding modes (19). In the majority of cases, isolated peptides are considered to be disordered in solution. However, within apparently unstructured conformations it is often possible, especially at lower temperatures, to find regions with local structures that have been demonstrated to be biologically relevant (20, 21). With the decrease of temperature, according to Gibb's equation for free energy, the favorable intramolecular enthalpic contacts that determine the conformation of a peptide will play an increasingly more important role since the entropic component is lowered. In the bound form, the conformation of a peptide is also determined to a considerable degree by enthalpic contacts, including intermolecular contacts. Thus, the conformations obtained for both conditions should demonstrate some similarity (21). We applied this methodology to study at lower temperature the conformations of unbound peptides, which might provide information regarding possible binding modes and a valuable starting point for rational drug design especially for a system with very low structural information available.

A number of compounds have been shown to bind CXCR4 and inhibit the binding of its natural ligand SDF-1 $\alpha$ . Some compounds act as agonists of CXCR4 activity and activate the signalling cascade upon binding to the receptor (22, 23), while others act as antagonists and bind to CXCR4 without triggering the downstream biological response (23). The best known inhibitor, the bicyclam AMD3100, is remarkably specific as a CXCR4 inhibitor, and the potential for its use in clinical applications including AIDS, cancer, rheumatoid arthritis, and stem cell transplantation is very promising (24). In spite of the success of AMD3100, due to its inhibition of CXCR4, concerns have often arisen about possible detrimental effects, such as defects in bone-marrow myelopoiesis, haematopoiesis, and cardiogenesis (24). It is often desirable that an inhibitor could not only bind to a receptor but also induce its biological activity so that the detrimental effects of receptor silencing were minimal during the inhibition. Two peptides were recently reported that act in this manner (22). These peptides, designated as RSVM and ASLW, are 17-mers derived from the SDF-1 $\alpha$  N-terminus and differ only in the four N-terminal residues. It was shown that the RSVM and ASLW peptides bind to CXCR4 and trigger the signalling of a receptor. Noteworthy, they bind to CXCR4 via a different binding site from that of AMD3100 and SDF-1 $\alpha$  thus demonstrating allostereism (22).

We used solution NMR spectroscopy to study the conformations of the ASLW and RSVM peptides at 5°C. The studied peptides adopted several local structures or

motifs that were compared with each other and with motifs described in literature. In the region from residues 5 to 8 the structures of the ASLW and RSVM peptides adopted local structures, which were unique and not comparable to other described motifs or each other. In the region from 9 to 12, the ASLW peptide formed a local structure that was also present in full length SDF-1 $\alpha$ , and in a subset of the RSVM and the SDF-1 $\alpha$  N-terminal 17-mer structures. The structures of two allosteric peptides, their relevance to bound conformations and implications for future drug design are discussed further.

## Results

### *Structural characterization of the ASLW and RSVM peptides: chemical shifts, J-coupling constants, and NOE connectivities*

The sequential assignment of proton resonances for two allosteric peptides agonists was accomplished using standard methods described by Wüthrich (25). NH and H $\alpha$  chemical shifts obtained for both peptides were shown to be similar to each other and to the chemical shifts for the SDF-1 $\alpha$  N-terminal 17-mer (26) (Table C-1).

The signs and values of NMR chemical shift deviations (CSD) from random coil chemical shifts are often characteristic of the secondary structure adopted by a polypeptide (27). CSD values for H $\alpha$  and NH NMR chemical shifts are mostly negative in helices and positive in  $\beta$ -sheets. The alternating signs and small values of CSD for the ASLW and RSVM peptides and the SDF-1 $\alpha$  17-mer (Table C-2) indicate that under experimental conditions all three peptides do not adopt either extensive helical nor  $\beta$ -strand conformations on average.

J-coupling constants are correlated with torsion angles  $\phi$  and therefore provide the information on secondary structure of polypeptides. For the studied peptides, J-coupling measurements from 1D spectra ranged from 5.6 to 8.6 Hz, which suggests that torsion angles can adopt a variety of values. This renders J-coupling constants less useful for structure calculations.

NOE patterns are also useful indicators of the elements of secondary structure adopted by polypeptides. For example the presence of  $\beta$ -turns is often accompanied by connectivities between  $d_{NN}(2,3)$ ,  $d_{NN}(3,4)$ ,  $d_{\alpha N}(2,3)$ ,  $d_{\alpha N}(3,4)$  and  $d_{\alpha N}(2,4)$ , where the numbers refer to the position of a residue in the turn. The summary of NOE connectivities for the ASLW peptide (Fig. C-1a) suggests the possible formation of a local turn-like structure in the region from Tyr<sup>7</sup> to Phe<sup>13</sup>. The  $d_{\alpha N}(i,i+2)$  connectivity between Ser<sup>2</sup> and Trp<sup>4</sup> may also suggest a  $\beta$ -turn in the region from Ala<sup>1</sup> to Trp<sup>4</sup>. The

RSVM peptide is less structured with one medium range connectivity between Ser<sup>6</sup> and Arg<sup>8</sup> (Fig. C-1b). Missing  $d_{\alpha N}(i,i+1)$  connectivities (Figs. C-1a and C-1b) next to Pro<sup>10</sup> in the RSVM peptide and Ser<sup>6</sup> in the ASLW peptide, were not included due to ambiguity in their assignments. In general, both peptides adopt flexible conformations with possible local structures as indicated by NOE pattern.

**Table C-1.** NMR chemical shifts for the ASLW and RSVM peptides and the NH and H $\alpha$  chemical shifts for the SDF-1 $\alpha$  17-mer. NMR data was obtained at 5°C for the ASLW and RSVM peptides in this study and at 8°C for the SDF-1 $\alpha$  N terminal 17-mer as described earlier (26).

NH chemical shifts			H $\alpha$ chemical shifts			Chemical shifts for proton nuclei other than NH and H $\alpha$		Primary sequence for SDF-1 $\alpha$ 17-mer
ASLW peptide	RSVM peptide	SDF-1 $\alpha$ 17-mer	ASLW peptide	RSVM peptide	SDF-1 $\alpha$ 17-mer	ASLW peptide	RSVM peptide	
			4.10	4.08	4.35		H $\beta$ 1.94, 1.93	Lys <sup>1</sup>
8.79	8.95		4.43	4.57	4.53		$\gamma$ CH <sub>2</sub> 1.65, 1.65	Pro <sup>2</sup>
8.58	8.55	8.52	4.26	4.13	4.07		$\delta$ CH <sub>2</sub> 3.22, 3.22	Val <sup>3</sup>
8.19	8.58	8.51	4.63	4.48	4.46		$\epsilon$ NH <sub>3</sub> <sup>+</sup> 7.28	Ser <sup>4</sup>
7.96	8.47	8.56	4.26	4.38	4.37	H $\beta$ 3.72, 3.64	H $\beta$ 3.85, 3.85	Leu <sup>5</sup>
8.08	8.33	8.31	4.28	4.37	4.39	H $\beta$ 1.55, 1.51	H $\beta$ 2.07	Ser <sup>6</sup>
8.08	8.29	8.17	4.49	4.52	4.51	$\gamma$ CH <sub>2</sub> 1.45	$\gamma$ CH <sub>2</sub> 0.95, 0.95	Tyr <sup>7</sup>
8.08	8.16	8.11	4.23	4.25	4.24	$\delta$ CH <sub>2</sub> 0.87, 0.82		Arg <sup>8</sup>
8.37	8.45	8.41	4.61	4.64	4.63	H $\beta$ 3.29, 3.19	H $\beta$ 2.07, 1.97	Cys <sup>9</sup>
			4.41	4.42	4.40	2H 7.20	$\gamma$ CH <sub>2</sub> 2.57, 2.51	Pro <sup>10</sup>
8.48	8.52	8.48	4.38	4.40	4.39	4H 7.60		Cys <sup>11</sup>
8.50	8.53	8.48	4.25	4.24	4.24	5H 7.13		Arg <sup>12</sup>
8.23	8.27	8.23	4.55	4.56	4.55	6H 7.23		Phe <sup>13</sup>
8.22	8.24	8.20	4.53	4.53	4.53	7H 7.48		Phe <sup>14</sup>
8.30	8.31	8.28	4.21	4.21	4.21	indole NH 10.22		Glu <sup>15</sup>
8.42	8.43	8.39	4.33	4.33	4.33	H $\beta$ 1.53, 1.50	H $\beta$ 1.63, 1.63	Ser <sup>16</sup>
8.71	8.72	8.68	4.69	4.69	4.70	$\gamma$ CH <sub>2</sub> 1.41	H $\beta$ 1.49	His <sup>17</sup>

One might expect the formation of a hydrophobic core around residues Tyr<sup>7</sup>, Phe<sup>13</sup>, and Phe<sup>14</sup> similarly to the structural studies reported previously (28). The experiment spectral resolution was sufficient to resolve frequencies for the majority of

aromatic residues for both peptides. Surprisingly, the ring protons for both peptides failed to give rise to medium range inter-residue NOEs (data not shown). Trp<sup>4</sup> in the ASLW peptide also demonstrated no contacts to the aromatic residues Tyr<sup>7</sup>, Phe<sup>13</sup>, and Phe<sup>14</sup> through either its backbone atoms or ring protons.

**Table C-2.** NH and H $\alpha$  NMR chemical shifts deviations. CSD values were calculated as  $\sigma_{\text{obs}} - \sigma_{\text{ref}}$ , where NH  $\sigma_{\text{ref}}$  obtained for 25°C (29) were corrected using NH temperature gradient of -7.6 ppb/°C reported for random coil (30) and  $\sigma_{\text{obs}}$  is the observed chemical shift. H $\alpha$   $\sigma_{\text{ref}}$  were assumed to stay unaltered for temperatures of interest (29). CSD for the SDF-1 $\alpha$  17-mer were taken from previous work (26).

H $\alpha$ CSD			NH CSD		
ASLW peptide	RSVM peptide	SDF-1 $\alpha$ 17-mer	ASLW peptide	RSVM peptide	SDF-1 $\alpha$ 17-mer
-0.21	-0.26	-0.24			
-0.04	0.10	0.12	0.32	0.48	
-0.08	0.01	-0.05	0.27	0.37	0.19
-0.03	0.00	0.00	-0.22	0.15	0.1
-0.08	0.03	0.03	-0.35	0.16	0.19
-0.19	-0.10	-0.08	-0.38	-0.13	-0.11
-0.06	-0.02	-0.04	-0.19	0.01	-0.27
-0.11	-0.09	-0.10	-0.30	-0.22	-0.29
0.06	0.09	-0.18	-0.10	-0.02	0.08
-0.01	0.00	-0.05			
-0.17	-0.15	-0.16	0.01	0.05	0.13
-0.09	-0.10	-0.11	0.12	0.15	0.09
-0.07	-0.06	-0.07	-0.22	-0.18	-0.20
-0.09	-0.09	-0.09	-0.23	-0.21	-0.23
-0.14	-0.14	-0.14	-0.27	-0.26	-0.19
-0.14	-0.14	-0.13	-0.04	-0.03	-0.03
-0.03	-0.04		0.13	0.14	0.28

Due to the presence of *cis*- and *trans*- isomerism of X-Pro<sup>10</sup> peptide bond, the resonances in the proton NMR spectra for peptides containing Pro is often doubled, so that minor and major conformations can be identified. In case of the ASLW and RSVM peptides, the doubling of resonances was negligible, which together with the characteristic NOEs from  $\beta$ -hydrogens of Cys<sup>9</sup> to  $\delta$ -hydrogens of Pro<sup>10</sup> present for both peptides indicates that these were the *trans* isomers of Cys<sup>9</sup>-Pro<sup>10</sup> peptide bond.

#### Structure Calculation and Clustering Analysis

Two regions on the ASLW peptide, residues 5 to 8 and residues 9 to 12, can be superimposed over their backbone atoms with the rmsd less than 0.7 Å for an entire



ensemble (Table C-3 and Figs. C-2a and C-2b). Regions other than 5-8 and 9-12 gave poor superposition statistics, with rmsd values of no less than 1.2 Å. The observed local structures were in agreement with NOE connectivity data except for the region around Ser<sup>2</sup> where, even in the presence of connectivity (Fig. C-2a), local structure was not observed. The RSVM peptide had only one pronounced local feature for an entire ensemble, which is a turn in the region 5-8 with the superposition of 0.64 Å (Table C-3 and Figs. C-2c and C-2d). For both peptides, the regions with low rmsd were further analyzed with NMRCLUST (31) to extract individual subsets of conformations with even lower rmsd values (Table C-3).

**Table C-3.** The results of cluster analysis for the ensembles of structures of the ASLW and RSVM peptides, the SDF-1 $\alpha$  17-mer, and full length SDF-1 $\alpha$ . The data for major families (more than 10% of an entire ensemble) is shown.

	rmsd for backbone atoms for ensemble, Å	Cluster 1		Cluster 2		Cluster 3	
		rmsd for backbone atoms, Å	Relative population of a cluster, %	rmsd for backbone atoms, Å	Relative population of a cluster, %	rmsd for backbone atoms, Å	Relative population of a cluster, %
ASLW peptide 5-8 region	0.71±0.24	0.51±0.15	43	0.51±0.13	25	—	—
ASLW peptide 9-12 region Motif II	0.67±0.21	0.21±0.07	25	0.29±0.12	25	0.24±0.09	18
RSVM peptide 5-8 region	0.64±0.31	0.41±0.14	68	0.99±0.24	23	—	—
RSVM peptide 9-12 region	1.17±0.52	0.70±0.22	33	0.55±0.18	30	0.74±0.20	20
SDF-1 $\alpha$ 17-mer 5-8 region Motif I	0.56±0.27	0.22±0.09	34	0.12±0.09	28	0.13±0.04	15
SDF-1 $\alpha$ 17-mer 9-12 region	1.22±0.40	0.53±0.16	28	0.59±0.23	28	0.51±0.18	26
full length SDF-1 $\alpha$ 5-8 region	1.11±0.30	1.00±0.22	63	1.06±0.16	12	—	—
full length SDF-1 $\alpha$ 9-12 region	0.29±0.08	—	—	—	—	—	—

Clustering analysis for the SDF-1 $\alpha$  N-terminal 17-mer was repeated and the results were essentially the same as reported earlier (26). The ensemble of SDF-1 $\alpha$  17-mer structures at 8°C adopted a local turn in a region 5-8 with rmsd equal to 0.56 Å for an ensemble, which can be clustered into the families with rmsd values smaller than 0.22 Å.

The local structure for Cluster 1 shown to be shared with several other CXCR4 ligands such as the ITAC 1-8 peptide and the vMipII 1-10 peptide (21) is referred here as Motif I. The region 9-12 of the SDF-1 $\alpha$  17-mer was weakly convergent for an entire family; however clustering analysis revealed individual subsets of structures with rmsd lower than 0.59 Å (Table C-3).

The ensemble of full length SDF-1 $\alpha$  structures determined at 30°C (32) was subjected to similar analysis. The region 9-12 was very well defined with the rmsd of 0.27 Å for an ensemble of 30 structures. The region 5-8 was convergent with the rmsd of 1.03 Å for an ensemble and individual clusters obtained with NMRCLUST showed rmsd's no less than 1.00 Å (Table C-3).

#### *Comparison of motifs*

The local structures found for the two allosteric peptides were compared to each other and to the corresponding regions of SDF-1 $\alpha$  and the SDF-1 $\alpha$  17-mer (Tables C-4 and C-5, Figs. C-3a and C-3b). The conformation in the region of residues 5-8 was distinct for every peptide under investigation. The RSVM peptide formed a unique structure, which is quite different from the ASLW peptide and the SDF-1 $\alpha$  17-mer. The same is true about the ASLW peptide, which is well converged in the region of interest, but unique in its conformation (Fig. C-3a). Motif I found for the SDF-1 $\alpha$  N-terminal 17-mer (21) as it is shown in Figure C-3a is not comparable to the conformations of allosteric peptides.

The local structures in the region 9-12 for the ASLW peptide and full length SDF-1 $\alpha$  were identical and will be referred here as Motif II (Fig. C-3b). In full length SDF-1 $\alpha$  this region is constrained not only by nonbonding interactions with the core of the protein but also by two disulphide bonds at Cys<sup>9</sup>-Cys<sup>34</sup> and Cys<sup>11</sup>-Cys<sup>50</sup>. Interestingly, the ASLW peptide adopts the same local structure without long range constraints. The conformations for 20% of the RSVM peptide structures (Table C-5, Cluster 3) and 28% of the SDF-1 $\alpha$  17-mer (Table C-5, Cluster 2) were comparable to Motif II in this region (Fig. C-3b), although the rmsd for these superpositions were higher (0.9 – 1.14 Å).

#### **Discussion**

The structural information regarding CXCR4 inhibitors is of certain value for the development of novel agents capable of blocking and modulating CXCR4 activity. In spite of the great importance of CXCR4 as a pharmacological target, the number of

available agents remains limited. We obtained the NMR structures of two potent allosteric agonists, the ASLW and RSVM peptides, that might be considered for further biochemical characterization and subsequent clinical applications or might serve as a useful platform for designing new inhibitors on their base. The studied peptides adopted an extended conformation suggested by the values of J-couplings, low numbers of medium and long range NOEs, and chemical shifts deviations. However, it was shown theoretically and experimentally that the conformational space sampled by a peptide is significantly restricted (33). It is thought that a polypeptide is never a completely random coil but instead exists as an ensemble of preferential conformations (33, 34). Indeed, the analysis of obtained structures for the ASLW and RSVM peptides revealed the presence of local structuring in several areas of a backbone. In the region of residues 5-8, well-defined local structures were present for each peptide under investigation. Also, in the region of residues 9-12, the ASLW peptide adopted a well-defined motif, which was also present in a subset of structures for the RSVM peptide. The same local structure was found in the region of residues 9-12 for full length SDF-1 $\alpha$  and for a subset of structures of the N-terminal SDF-1 $\alpha$  17-mer.

It was proposed that the local structures found for peptides in solution might indicate the similar structural features present in its bound conformation (21). The thermodynamical rationale behind this assumption is based on the consideration of the enthalpy and entropy interplay during a binding event. The entropic cost to fold a polypeptide is thought to be paid by enthalpic contributions generated upon binding (35). Thus, ligands, which are folded prior to binding, or possess a correct nascent structure, should have an advantage in terms of their binding energy. However, the exact contributions of free energy are not known and numerous examples in literature were reviewed where solution conformations are similar to their bound states (36) and where conformations are globally different (37). The insights from protein folding studies stresses the importance of the correct minimal structure in the denatured state that propagates the correct folding (38-40) and in our own experience local motifs in the Hox peptide, in solution, were demonstrated to be analogous to the bound conformations (20). If the structures of the RSVM and ASLW peptides in solution are indeed related to the conformations in their bound states one might propose that the dissimilarity in local structures in the regions of residues 5-8 indicates the dissimilarity in their binding modes. The formation of the same structure in the region 9-12 for allosteric peptides and natural ligands (Motif II) argues for at least partial overlapping of the binding interfaces.

In the context of the full length SDF-1 $\alpha$  structure, Motif II formed between two cysteine bonds is concave, which renders the side chains of the residues composing the motif (Cys<sup>9</sup>, Pro<sup>10</sup>, Cys<sup>11</sup>) unavailable for binding (Fig. C-4). It is also known that those residues are rarely involved in the formation of protein - protein interfaces (41). However, arginine side chains often form electrostatic interactions on interfaces and the formation of a rigid motif similar to Motif II can position the side chains of Arg<sup>8</sup> and Arg<sup>12</sup>, flanking the region 9-12, more rigidly by constraining the backbone between them. A number of compounds have been shown to bind CXCR4 through electrostatic interactions. AMD3100, a well known inhibitor, binds to CXCR4 through the interaction of positively charged cyclam rings with carboxylic moieties on Asp<sup>171</sup>, Asp<sup>262</sup>, and Glu<sup>288</sup> (42, 43). ALX40-4C is another highly charged and efficient inhibitor of CXCR4 that consists of nine arginine residues (44). Since arginine residues have a propensity to appear frequently on protein - protein interfaces (41) and since electrostatic interactions play a significant role in binding of CXCR4 inhibitors we believe that Motif II might serve to position the two arginine side chains in close proximity to each other against their electrostatic repulsion and to orient them favorably for the consequent binding to CXCR4 (Fig. C-4). The fact that the RSVM peptide is eight times less potent than the ASLW peptide as an inhibitor of CXCR4 (22) might be also relevant to the formation of Motif II prior to binding since the initial structuring of the RSVM peptide is less pronounced.

## Materials and Methods

### *Peptide Synthesis and NMR Sample Preparation*

The peptides H-ASLWLSYRCPFRFFESH-NH<sub>2</sub> and H-RSVMLS YRCPFRFFESH-NH<sub>2</sub> were synthesized by stepwise solid phase peptide synthesis using t-butoxycarbonyl protection chemistry. NMR samples contained ~3 mM of a peptide, 20 mM CD<sub>3</sub>COONa, 15 mM deuterated DTT, 1 mM NaN<sub>3</sub>, 1 mM DSS in 500  $\mu$ l of 90% H<sub>2</sub>O/10% D<sub>2</sub>O. The pH was adjusted to 5.0 $\pm$ 0.05.

### *NMR Spectroscopy*

<sup>1</sup>H 1D and TOCSY 2D NMR spectra for both peptides were acquired at Varian Unity 600 spectrometer. 2D NOESY NMR spectra were acquired at 5°C with the mixing time 500 ms at Varian Unity 600 spectrometer for the RSVM peptide and at Varian Inova 800 spectrometer for the ASLW peptide. NOESY spectra were used for integration of NOE cross peaks, which were consequently converted into the distance restraints using Bin module in NMRVIEW (45). NOE connectivities were grouped as weak, medium, and strong corresponding to upper distance restraints of 5.0, 3.4, and 2.8 Å respectively. In total, to calculate structures we used 79 inter-residue and 59 intra-residue connectivities for the ASLW peptide and 37 inter-residue and 63 intra-residue connectivities for the RSVM peptide.

### *Structure Calculations and Clustering Analysis*

In case of partially unfolded proteins, the procedure commonly used for calculating structures of folded proteins is not applicable since long distance NOE restraints observed in different conformations cannot be satisfied simultaneously in one structure (38, 46). This generally results in a high total energy of structures and significant number of NOE violations. There are also examples in literature when the structures of smaller unfolded proteins or peptides were successfully calculated using standard NOE based approach where the entire set of NOE restraints is used to generate the family of conformers (21, 26, 28, 47, 48). In our case, the structures obtained using standard simulated annealing protocol implemented in CNS version 1.1 (49) demonstrated neither dihedral violations more than 5° nor NOE violations higher than 0.2 Å and the energies for the calculated structures did not exceed 35 kcal/mol and 15 kcal/mol for the ASLW and RSVM peptides respectively. The calculations were performed in torsion angle space using square well potential and the force constant for the NOE term equal to 1.0. The scaling factors for high temperature annealing stage (T=50000K) were 150 for NOE energy term and 100 for dihedral energy term. The scaling factors for the slow-cool annealing stage (T=50000K) were 150 for NOE energy term and 200 for dihedral energy term. For the final minimization stage, the scaling factor for NOE energy term was 75 and for dihedral energy term the scale factor was 400. PROCHECK (50) test showed 0.4% and 4.3% of residues in disallowed regions for the ASLW and RSVM peptides correspondingly.

For further analysis, the 40 structures with the lowest energies were chosen out of 80 calculated structures. In order to locate regions that adopt similar conformations the families of 40 structures were superimposed in regions of 3-6 residues in length scanning the entire peptide and examined. Clustering was performed using NMRCLUST (31) software. Both superposition and clustering included backbone atoms for specific regions.

**Table C-4.** The convergence of individual clusters between each other in the region of residues from 5 to 8. Representative structures for individual clusters of peptides were obtained with NMRCCLUS program and superimposed with their backbone atoms in the region of interest. The relative population of a cluster is indicated as a percent of the total number of structures in ensemble.

	RSVM peptide Cluster 1 67.5%	RSVM peptide Cluster 2 22.5%	ASLW peptide Cluster 1 42.5%	ASLW peptide Cluster 2 25%	SDF-1 17- mer Cluster 1 34.0% <b>Motif I</b>	SDF-1 17- mer Cluster 2 27.7%	SDF-1 17-mer Cluster 3 14.9%
RSVM peptide Cluster 1 67.5%	—	—	—	—	—	—	—
RSVM peptide Cluster 2 22.5%	1.12 Å	—	—	—	—	—	—
ASLW peptide Cluster 1 42.5%	1.89 Å	1.63 Å	—	—	—	—	—
ASLW Cluster 2 25%	1.73 Å	1.35 Å	1.18 Å	—	—	—	—
SDF-1 17- mer Cluster 1 <b>Motif I</b> 34.0%	1.19 Å	1.37 Å	1.51 Å	1.73 Å	—	—	—
SDF-1 17- mer Cluster 2 27.7%	1.58 Å	1.32 Å	1.60 Å	1.35 Å	1.05 Å	—	—
SDF-1 17- mer Cluster 3 14.9%	1.46 Å	1.71 Å	1.33 Å	1.79 Å	0.69 Å	1.47 Å	—

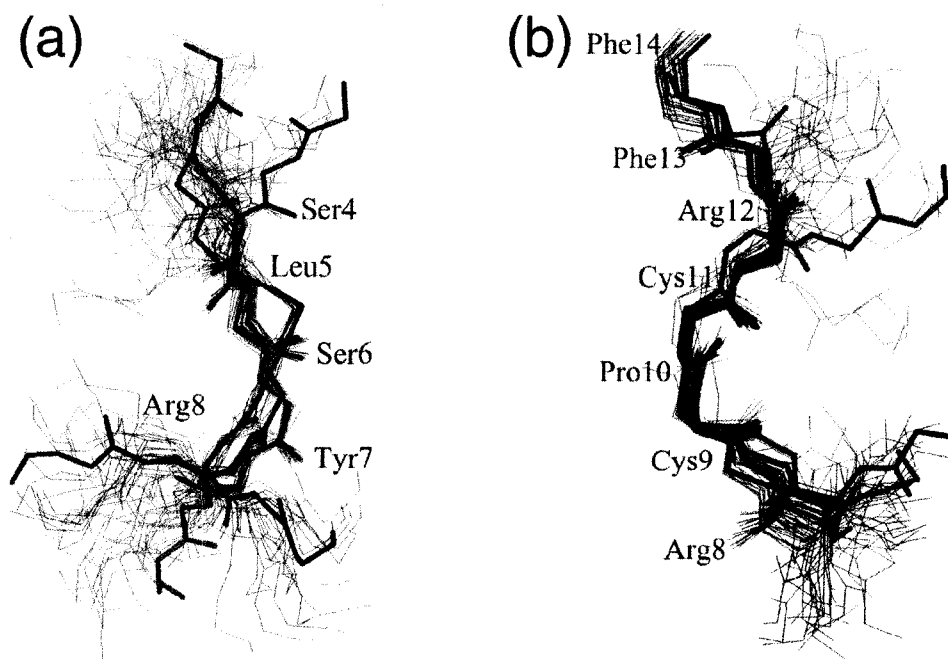
**Table C-5.** Convergence of individual clusters between each other in the region of residues from 9 to 12. Representative structures for every individual cluster of peptides were obtained with NMRCLUST. Rmsd values were calculated for backbone atoms of representative structures superimposed in the region of interest. The relative population of a cluster is indicated as a percentage of the total number of structures. The values in bold were considered as good or moderately good convergence.

	RSVM peptide Cluster 1 32.5%	RSVM peptide Cluster 2 30.0%	RSVM peptide Cluster 3 20.0%	ASLW peptide Cluster 1 25% <b>Motif II</b>	ASLW peptide Cluster 2 25% <b>Motif II</b>	ASLW peptide Cluster 3 17.5%	SDF-1 $\alpha$ 17-mer Cluster 1 27.7%	SDF-1 $\alpha$ 17-mer Cluster 2 27.7%	SDF-1 $\alpha$ 17-mer Cluster 3 25.5%
RSVM peptide Cluster 1, 32.5%	—	—	—	—	—	—	—	—	—
RSVM peptide Cluster 2, 30.0%	1.36 Å	—	—	—	—	—	—	—	—
RSVM peptide Cluster 3, 20.0%	2.12 Å	2.19 Å	—	—	—	—	—	—	—
ASLW peptide Cluster 1 25% <b>Motif II</b>	1.53 Å	2.00 Å	1.42 Å	—	—	—	—	—	—
ASLW peptide Cluster 2 25% <b>Motif II</b>	1.66 Å	1.89 Å	<b>1.12 Å</b>	<b>0.52 Å</b>	—	—	—	—	—
ASLW peptide Cluster 3, 17.5%	1.85 Å	1.98 Å	1.45 Å	0.96 Å	1.09 Å	—	—	—	—
SDF-1 $\alpha$ 17- mer Cluster 1, 27.7%	1.39 Å	1.37 Å	1.91 Å	<b>1.05 Å</b>	1.36 Å	1.39 Å	—	—	—
SDF-1 $\alpha$ 17- mer Cluster 2 27.7%	1.43 Å	2.06 Å	1.36 Å	1.18 Å	1.14 Å	1.35 Å	1.47 Å	—	—
SDF-1 $\alpha$ 17- mer Cluster 3, 25.5%	1.20 Å	1.06 Å	2.07 Å	1.89 Å	1.89 Å	2.05 Å	1.43 Å	1.69 Å	—
SDF-1 $\alpha$ full length ensemble	1.80 Å	2.06 Å	<b>0.90 Å</b>	<b>0.78 Å</b>	<b>0.49 Å</b>	1.14 Å	1.48 Å	<b>1.04 Å</b>	1.93 Å

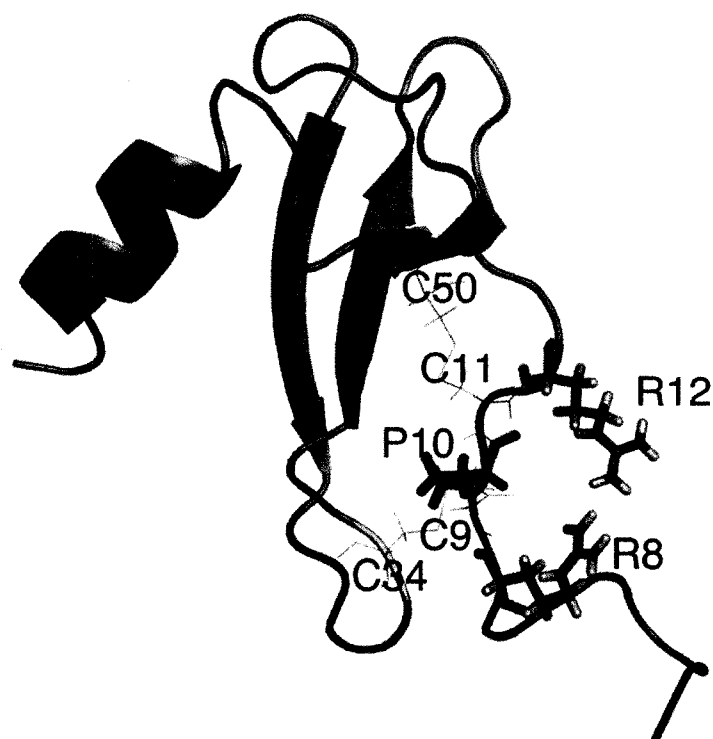
[illegible]



**Figure C-3.** The superposition of the ASLW and RSVM peptides with the SDF-1 $\alpha$  17-mer in the region 5-8 (a) and with the full length SDF-1 $\alpha$  in the region 9-12 (b). (a), The structures shown are Cluster 1 for the SDF-1 $\alpha$  17-mer (Motif I) and the most convergent families of the ASLW peptide (Cluster 1) and the RSVM peptide (Cluster 1). (b), The entire ensemble of full length SDF-1 $\alpha$  structures is shown together with the most convergent families of the ASLW peptide (Cluster 1 and 2) and the RSVM peptide (Cluster 3). Backbone atoms for residues 4-9 (a) and 8-14 (b) are shown in black for SDF-1 $\alpha$  17-mer (a) or SDF-1 $\alpha$  (b), in red for the RSVM peptide, and in blue for the ASLW peptide.



**Figure C-4.** Motif II in the context of SDF-1 $\alpha$  structure. Motif II, formed by the backbone atoms of residues from 9 to 12 is shown in red. Pro<sup>10</sup> side chain in the middle of Motif II is also shown in red in sticks representation. Residues Cys<sup>9</sup>, Cys<sup>11</sup>, Cys<sup>34</sup>, and Cys<sup>50</sup> that form disulphide bonds constraining Motif II are shown in yellow. Side chain arginines, Arg<sup>8</sup> and Arg<sup>12</sup>, are shown in stick representation and colored according to their atoms.



## Reference

- (1) Onuffer, J. J., and Horuk, R. (2002) Chemokines, chemokine receptors and small-molecule antagonists: recent developments. *Trends Pharmacol Sci.* 23, 459-67.
- (2) Murphy, P. M. (2001) Chemokines and the molecular basis of cancer metastasis. *N Engl J Med.* 345(11), 833-5.
- (3) Fernandez, E. J., and Lolis, E. (2002) Structure, function, and inhibition of chemokines. *Annu Rev Pharmacol Toxicol.* 42, 469-99.
- (4) Baggiolini, M. (2001) Chemokines in pathology and medicine. *J Intern Med.* 250(2), 91-104.
- (5) Horuk, R. (1998) Chemokines beyond inflammation. *Nature* 393, 524-5.
- (6) Baggiolini, M. (1998) Chemokines and leukocyte traffic. *Nature* 392(6676), 565-8.
- (7) Sallusto, F., Mackay, C. R., and Lanzavecchia, A. (2000) The role of chemokine receptors in primary, effector, and memory immune responses. *Annu Rev Immunol* 18, 593-620.
- (8) Tachibana, K., Hirota, S., Iizasa, H., Yoshida, H., Kawabata, K., Kataoka, Y., Kitamura, Y., Matsushima, K., Yoshida, N., Nishikawa, S., Kishimoto, T., and Nagasawa, T. (1998) The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract. *Nature* 393, 591-4.
- (9) Knaut, H., Werz, C., Geisler, R., Nusslein-Volhard, C., and Consortium., T. S. (2003) A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor. *Nature* 421, 279-82.
- (10) Ma, Q., Jones, D., Borghesani, P. R., Segal, R. A., Nagasawa, T., Kishimoto, T., Bronson, R. T., and Springer, T. A. (1998) Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proc Natl Acad Sci U S A* 95, 9448-53.
- (11) Zou, Y. R., Kottmann, A. H., Kuroda, M., Taniuchi, I., and Littman, D. R. (1998) Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* 393, 595-9.
- (12) Muller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M. E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S. N., Barrera, J. L., Mohar, A., Verastegui, E., and Zlotnik, A. (2001) Involvement of chemokine receptors in breast cancer metastasis. *Nature* 410, 50-6.
- (13) Burger, J. A., Burger, M., and Kipps, T. J. (1999) Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells. *Blood* 94, 3658-67.
- (14) Juarez, J., Bradstock, K. F., Gottlieb, D. J., and Bendall, L. J. (2003) Effects of inhibitors of the chemokine receptor CXCR4 on acute lymphoblastic leukemia cells in vitro. *Leukemia* 17, 1294-300.
- (15) Rubin, J. B., Kung, A. L., Klein, R. S., Chan, J. A., Sun, Y., Schmidt, K., Kieran, M. W., Luster, A. D., and Segal, R. A. (2003) A small-molecule antagonist of CXCR4 inhibits intracranial growth of primary brain tumors. *Proc Natl Acad Sci U S A* 100, 13513-8.
- (16) Phillips, R. J., Burdick, M. D., Lutz, M., Belperio, J. A., Keane, M. P., and Strieter, R. M. (2003) The stromal derived factor-1/CXCL12-CXC chemokine receptor 4 biological axis in non-small cell lung cancer metastases. *Am J Respir Crit Care Med.* 167(12), 1676-86.

- (17) Cooper, C. R., Chay, C. H., Gendernalik, J. D., Lee, H. L., Bhatia, J., Taichman, R. S., McCauley, L. K., Keller, E. T., and Pienta, K. J. (2003) Stromal factors involved in prostate carcinoma metastasis to bone. *Cancer*. 97(3 Suppl), 739-47.
- (18) Berson, J. F., Long, D., Doranz, B. J., Rucker, J., Jirik, F. R., and Doms, R. W. (1996) A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *J Virol*. 70(9), 6288-95.
- (19) Marshall, G. R. (2001) Peptide interactions with G-protein coupled receptors. *Biopolymers* 60(3), 246-77.
- (20) Slupsky, C. M., Sykes, D. B., Gay, G. L., and Sykes, B. D. (2001) The HoxB1 hexapeptide is a prefolded domain: implications for the Pbx1/Hox interaction. *Protein Sci*. 10(6), 1244-53.
- (21) Booth, V., Slupsky, C. M., Clark-Lewis, I., and Sykes, B. D. (2003) Unmasking ligand binding motifs: identification of a chemokine receptor motif by NMR studies of antagonist peptides. *J Mol Biol*. 327, 329-34.
- (22) Sachpatzidis, A., Benton, B. K., Manfredi, J. P., Wang, H., Hamilton, A., Dohlman, H. G., and Lolis, E. (2003) Identification of allosteric peptide agonists of CXCR4. *J Biol Chem*. 278, 896-907.
- (23) Heveker, N., Montes, M., Germeroth, L., Amara, A., Trautmann, A., Alizon, M., and Schneider-Mergener, J. (1998) Dissociation of the signalling and antiviral properties of SDF-1-derived small peptides. *Curr Biol*. 8, 369-76.
- (24) De Clercq, E. (2003) The bicyclam AMD3100 story. *Nat Rev Drug Discov.*, 581-7.
- (25) Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
- (26) Elisseeva, E. L., Slupsky, C. M., Crump, M. P., Clark-Lewis, I., and Sykes, B. D. (2000) NMR studies of active N-terminal peptides of stromal cell-derived factor-1. Structural basis for receptor binding. *J Biol Chem*. 275, 26799-805.
- (27) Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J Mol Biol*. 222, 311-33.
- (28) Fairbrother, W. J., McDowell, R. S., and Cunningham, B. C. (1994) Solution conformation of an atrial natriuretic peptide variant selective for the type A receptor. *Biochemistry* 33, 8897-904.
- (29) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J Biomol NMR*. 5, 67-81.
- (30) Merutka, G., Dyson, H. J., and Wright, P. E. (1995) 'Random coil' <sup>1</sup>H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J Biomol NMR*. 5, 14-24.
- (31) Kelley, L. A., Gardner, S. P., and Sutcliffe, M. J. (1996) An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Eng*. 9, 1063-5.
- (32) Crump, M. P., Gong, J. H., Loetscher, P., Rajarathnam, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J. L., Baggiolini, M., Sykes, B. D., and Clark-Lewis, I. (1997) Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J*. 16, 6996-7007.
- (33) Pappu, R. V., Srinivasan, R., and Rose, G. D. (2000) The Flory isolated-pair hypothesis is not valid for polypeptide chains: implications for protein folding. *Proc Natl Acad Sci U S A* 97, 12565-70.

- (34) Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchic? I. Local structure and peptide folding. *Trends Biochem Sci* 24, 26-33.
- (35) Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6, 197-208.
- (36) Baldwin, R. L., and Rose, G. D. (1999) Is protein folding hierarchic? II. Folding intermediates and transition states. *Trends Biochem Sci* 24, 77-83.
- (37) Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12, 54-60.
- (38) Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) Conformation of peptide fragments of proteins in aqueous solution: implications for initiation of protein folding. *Biochemistry* 27, 7167-75.
- (39) Daggett, V., and Fersht, A. (2003) The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* 4, 497-502.
- (40) Dobson, C. M., and Karplus, M. (1999) The fundamentals of protein folding: bringing together theory and experiment. *Curr Opin Struct Biol* 9, 92-101.
- (41) Jones, S., and Thornton, J. M. (1996) Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* 93, 13-20.
- (42) Rosenkilde, M. M., Gerlach, L. O., Jakobsen, J. S., Skerlj, R. T., Bridger, G. J., and Schwartz, T. W. (2004) Molecular mechanism of AMD3100 antagonism in the CXCR4 receptor: transfer of binding site to the CXCR3 receptor. *J Biol Chem* 279, 3033-41.
- (43) Trent, J. O., Wang, Z. X., Murray, J. L., Shao, W., Tamamura, H., Fujii, N., and Peiper, S. C. (2003) Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists. *J Biol Chem* 278, 47136-44.
- (44) Doranz, B. J., Grovit-Ferbas, K., Sharron, M. P., Mao, S. H., Goetz, M. B., Daar, E. S., Doms, R. W., and O'Brien, W. A. (1997) A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 coreceptor. *J Exp Med* 186, 1395-400.
- (45) Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* 4, 603-614.
- (46) Choy, W. Y., and Forman-Kay, J. D. (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain. *J Mol Biol* 308, 1011-32.
- (47) Tamamura, H., Hiramatsu, K., Miyamoto, K., Omagari, A., Oishi, S., Nakashima, H., Yamamoto, N., Kuroda, Y., Nakagawa, T., Otaka, A., and Fujii, N. (2002) Synthesis and evaluation of pseudopeptide analogues of a specific CXCR4 inhibitor, T140: the insertion of an (E)-alkene dipeptide isostere into the betaII'-turn moiety. *Bioorg Med Chem Lett* 12, 923-8.
- (48) Crump, M. P., Elisseeva, E., Gong, J., Clark-Lewis, I., and Sykes, B. D. (2001) Structure/function of human herpesvirus-8 MIP-II (1-71) and the antagonist N-terminal segment (1-10). *FEBS Lett* 489, 171-5.
- (49) 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., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54, 905-21.
- (50) Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26P, 283-291.

## APPENDIX D

### Dimerization of SDF-1 $\alpha$ studied by NMR relaxation methods\*

#### Outline

The structures of SDF-1 $\alpha$  determined by NMR spectroscopy have been shown to be monomeric while X-ray structures are dimeric. Biochemical data and *in vivo* studies suggest that dimerization is likely to be important for the function of chemokines. We determined the dynamics of SDF-1 $\alpha$  using the backbone  $^{15}\text{N}$  NMR relaxation methods. Data were obtained at several concentrations of SDF-1 $\alpha$  and used to determine  $K_d$  of  $\sim 5$  mM for a monomer-dimer equilibrium.  $K_d$  obtained was used to extrapolate relaxation data to monomeric SDF-1 $\alpha$ . Experimental and extrapolated data were analyzed using Model Free approach. Effects of protein dimerization on relaxation parameters are discussed. Our results help rationalize the apparent discrepancies between solution NMR structures obtained under different conditions, clearly indicating the importance of the knowledge of oligomerization degree in the calculation of solution structures.

#### Introduction

SDF-1 $\alpha$  is a member of the chemokine family, which is responsible for the trafficking of lymphocytes and numerous reactions in the immune system (1, 2). SDF-1 $\alpha$  is particularly interesting due to its involvement in HIV pathogenesis (3) and in a variety of cancers (4-8). It is generally expected that the inhibition of the interaction of SDF-1 $\alpha$  with its receptor CXCR4 will be an important therapeutic breakthrough that will reduce the propagation rate of a variety of cancers and HIV-1 (9). SDF-1 $\alpha$  exerts its biological function through interaction with the G-protein coupled receptor, CXCR4, in two steps. The first step includes the recruitment of a chemokine from solution and the second step involves the insertion of the SDF-1 $\alpha$  N-terminus into the seven-helix bundle of the receptor and the initiation of a signaling cascade (10).

The interaction of a chemokine with its receptor can be modulated in several ways. One of them is through the interaction of a chemokine with the components of extracellular matrix such as glycosaminoglycans. These have a noticeable affinity for chemokines and help to bring the chemokines from bulk solution and concentrate them in

---

\* The version of this chapter was published. Baryshnikova O.K., and Sykes B.D. (2006) *Prot Sci* 15, 2568-2578.

the proximity to membrane surface (11). Another possible way to modulate chemokine signaling is through multivalency, which is a well-known strategy employed by nature (12). Multivalency leads to the increase in the local concentration of a ligand on a surface and hence increases the observed affinity for its receptor. Several chemokines are known to exist as oligomers in concentrated solutions (13). If chemokines are also oligomeric *in vivo* at physiological concentrations, this might enhance the efficiency of recruitment and bring an additional level of regulation of signaling. Some *in vivo* studies suggested that dimerization is not required for the activation of a receptor (for example, IL-8) (14). Other studies demonstrated that at least in certain cases the ability of chemokines to dimerize influences the magnitude of signaling efficiency (for example, CCL2, CCL4 and CCL5) (15).

The oligomerization behavior of SDF-1 $\alpha$  *in vitro* in terms of monomer-dimer equilibrium is controversial. The first NMR structure reported by Crump and coworkers characterized SDF-1 $\alpha$  as a monomer under experimental conditions based on the following observations: no NOE contacts corresponding to a dimer interface were found; no slowly exchanging protons at the dimer interface were observed; and the molecular weights obtained with sedimentation equilibrium centrifugation were consistent with that expected for the monomeric species (10). In the two reported crystal structures SDF-1 $\alpha$  was dimeric, with the dimer interface similar to the interface formed by IL-8 (16, 17). The original NMR structure of SDF-1 $\alpha$  determined by Crump and coworkers (10) was slightly different from the structures of determined by crystallography, especially in the position of the C-helix. The most recent NMR structure of SDF-1 $\alpha$  (18) determined under different experimental conditions was instead very similar to the individual subunits of SDF-1 $\alpha$  in the crystal structure. The dimerization of SDF-1 $\alpha$  can account for differences between these structures.

Monomer-dimer equilibrium for SDF-1 $\alpha$  has been characterized by several biophysical methods, including solution state NMR, with  $K_d$  values in the range of ~150  $\mu$ M to ~3 mM (19, 20). In the study by Veldkamp et al., special emphasis was drawn to the role of experimental conditions, e.g. pH and buffer content, on the monomer-dimer equilibrium (20). Given that oligomerization of chemokines may be crucial for enhancement of signaling by increasing the affinity of the chemokine-receptor interaction, a thorough investigation of the nature of SDF-1 $\alpha$  oligomerization is important.

## Results

### *<sup>15</sup>N- $T_1$ , $T_2$ , and NOE data*

Backbone amide <sup>15</sup>N relaxation parameters were obtained using 2D {<sup>1</sup>H, <sup>15</sup>N} correlation NMR spectroscopy with a typical spectrum of 1.5 mM SDF-1 $\alpha$  shown in Figure D-1. Backbone <sup>15</sup>N and <sup>1</sup>H chemical shifts for SDF-1 $\alpha$  were assigned using 3D NOESY-HSQC and 3D TOCSY-HSQC spectra (21), and were in good agreement with <sup>1</sup>H chemical shifts obtained previously using homonuclear 2D NMR spectroscopy (10). <sup>1</sup>H chemical shifts were found to be dependent on protein concentration. Relaxation properties were determined for 56 of the 67 backbone amide <sup>1</sup>H-<sup>15</sup>N pairs. Residues that could not be analyzed include: Lys<sup>1</sup>, due to exchange with solvent; Phe<sup>13</sup> and Cys<sup>34</sup>, whose chemical shifts overlap; Pro<sup>2</sup>, Pro<sup>10</sup>, Pro<sup>32</sup> and Pro<sup>53</sup> lacking backbone amide protons; Leu<sup>26</sup>, due to line broadening; and Gln<sup>37</sup>, whose chemical shifts were not assigned. In several cases, Tyr<sup>61</sup>, Leu<sup>62</sup>, Phe<sup>14</sup>, and Val<sup>4</sup> were excluded due to peak overlap. In most diluted samples, resonances for Phe<sup>13</sup> and Cys<sup>34</sup> could be discriminated and were therefore included in the analysis.

<sup>15</sup>N- $T_1$ ,  $T_2$ , and NOE data for a sample of 1.5 mM recombinant <sup>15</sup>N labeled SDF-1 $\alpha$  were obtained at 500 and 600 MHz, and 30 °C (Table D-1, Fig. D-2). The average NOE<sup>600</sup>/NOE<sup>500</sup> ratio was  $1.04 \pm 0.20$  (Fig. D-2d).  $T_1^{500}$  and  $T_1^{600}$  followed similar per residue patterns with average  $T_1^{600}/T_1^{500} = 1.29 \pm 0.07$  (Fig. D-2e).  $T_2^{500}$  and  $T_2^{600}$  also showed similar per residue patterns with average  $T_2^{600}/T_2^{500} = 0.87 \pm 0.06$  (Fig. D-2f).

The N- and C-terminal regions (residues 3-8 and 62-67) of SDF-1 $\alpha$  were more flexible than the protein core (Fig. D-2). This was in agreement with the unstructured N-terminus in NMR structure and the lack of electron density in X-ray structures (10, 17). Residues 23-27 and 30-31 had NOE's and  $T_2$ 's lower than the average and may be more flexible than the protein core. However, the per residue profile for  $T_1/T_2$  ratios (Fig. D-2g) shows an increase in this region, suggesting the presence of chemical exchange. Regions of residue 36-42, 49-51, and 65-67 had also high  $T_1/T_2$ 's and low  $T_2$ 's indicative of chemical exchange.

### *Concentration dependence of <sup>15</sup>N- $T_2$ values*

Theoretical values for monomeric SDF-1 $\alpha$  are  $T_2^{500}=176$  ms and  $T_2^{600}=181$  ms ( $S^2 = 0.85$  (22), and  $\tau_m = 3.9$  ns) (23) (in-house script written by Stéphane Gagné). These values are significantly larger than experimental  $T_2^{500}$  of 109 ms and  $T_2^{600}$  of 130 ms. For



dimeric SDF-1 $\alpha$ , theoretical values were  $T_2^{500}$ =102 ms and  $T_2^{600}$ =107 ms. Thus in fast exchange limit, SDF-1 $\alpha$  in solution is likely to be partially dimeric, which resulted in a decrease of experimental  $T_2$ 's. To explore this further, we measured  $T_2^{500}$  for a range of concentrations (2.6, 1.5, 0.9, 0.5, and 0.3 mM). The data shows a clear dependence of  $T_2^{500}$  on concentration, with the lowest  $T_2^{500}$  value observed for the most concentrated sample (Fig. D-3). This dependence is also highlighted in the plot of average  $R_2^{500}$  ( $1/T_2^{500}$ ) for the core residues of SDF-1 $\alpha$  as a function of concentration (Fig. D-4). Core residues included 15, 16, 18-22, 33, 35, 39, 41-43, 46, 50, 52, and 54-60.

**Table D-1.** NMR relaxation measurements obtained for 1.5 mM SDF-1 $\alpha$ .

Frequency (MHz)	$T_1$ (all residues)	$T_2$ (all residues)	$T_1$ (protein core) <sup>a</sup>	$T_2$ (protein core)	NOE (all residues)	NOE (protein core)
500	446 $\pm$ 71 <sup>b</sup>	130 $\pm$ 91 <sup>d</sup>	432 $\pm$ 41 <sup>f</sup>	107 $\pm$ 22 <sup>h</sup>	0.50 $\pm$ 0.24	0.68 $\pm$ 0.05
600	576 $\pm$ 97 <sup>c</sup>	109 $\pm$ 56 <sup>e</sup>	550 $\pm$ 50 <sup>g</sup>	90 $\pm$ 22 <sup>i</sup>	0.57 $\pm$ 0.24	0.74 $\pm$ 0.04

<sup>a</sup> The excluded residues were the following: 3-14, 17, 23-25, 27, 29, 31, 38, 45, 48, 66-67 for the data set acquired at 500 MHz and 3-9, 12, 18, 29-31, 35-36, 64-67 for the data set acquired at 600 MHz. Residues were excluded if NOE<sup>500</sup><0.60 and NOE<sup>600</sup><0.65. <sup>b</sup> An average error of fitting was 15 ms. <sup>c</sup> An average error of fitting was of 32 ms. <sup>d</sup> An average error fitting was of 3 ms. <sup>e</sup> An average error of fitting was 6 ms. <sup>f</sup> An average error of fitting was 12 ms. <sup>g</sup> An average error of fitting was 25 ms. <sup>h</sup> An average error of fitting was 2 ms. <sup>i</sup> An average error of fitting was 4 ms.

Observed transverse relaxation rate  $R_2^{obs}$  can be written in terms of dimerization as following:

$$R_2^{obs} = f_m R_2^m + 2f_d R_2^d \quad (1),$$

where  $R_2^m$  and  $R_2^d$  are related to pure monomer and dimer, respectively, and the fractions of monomers and dimers are:

$$f_m = \frac{[mon]}{P_{total}} \quad \text{and} \quad f_d = \frac{[dim]}{P_{total}} \quad (2).$$

Total concentration of the protein is:

$$P_{total} = [mon] + 2[dim] \quad (3).$$

Substitution of eq.2 and eq.3 into eq.1 yields:

$$R_2^{obs} = \frac{[mon]}{P_{total}} R_2^m + 2 \frac{[dim]}{P_{total}} R_2^d \quad (4).$$

Concentrations of monomers and dimers can be derived by solving eq. 3 and the equation for  $K_d$ :

$$[mon] = \frac{1}{4}(-K_d + \sqrt{K_d} \sqrt{K_d + 8P_{total}}) \quad (5),$$

$$[dim] = \frac{1}{8}(K_d + 4P_{total} - \sqrt{K_d} \sqrt{K_d + 8P_{total}}) \quad (6).$$

Substitution of eq.6 and eq.7 into eq.4 and simplification yields

$$R_2^{obs} = \frac{(K_d(R_2^d - R_2^m) + 4R_2^d P_{total} + \sqrt{K_d}(R_2^m - R_2^d)\sqrt{K_d + 8P_{total}})}{4P_{total}} \quad (7).$$

Experimental  $R_2$  data were fit to eq.7 ([www.pence.ca/~software](http://www.pence.ca/~software)) yielding the following best fit parameters:  $K_d = 9$  mM,  $R_2^m = 7$  s<sup>-1</sup> and  $R_2^d = 18$  s<sup>-1</sup> (Fig. D-4). The quality of the fit at lower concentrations was significantly better. We also evaluated  $K_d$  using other methods as described below to obtain a more reliable value.

#### *Dimerization constant calculated using 1D <sup>1</sup>H NMR spectroscopy*

Chemical shifts are sensitive to changes in molecular structure and local environment. Chemical shifts obtained from 1D <sup>1</sup>H spectra offer advantages of being very accurate, since the resolution in 1D <sup>1</sup>H experiment at 500 MHz is normally below 1 Hz. However, due to peak overlap, only a few resonances can be easily traced. Changes in chemical shifts as a function of SDF-1 $\alpha$  concentration for clearly distinguishable peaks (indole NH of Trp<sup>57</sup>, 2,6H and 3,5H of Tyr<sup>7</sup>, and 2,6H of Phe<sup>13</sup> and Phe<sup>14</sup>) were fit to eq.8 derived similarly to eq.7:

$$\sigma = \sigma_m f_m + 2\sigma_d f_d = \sigma_m \frac{[mon]}{P_{total}} + 2\sigma_d \frac{[dim]}{P_{total}} \quad (8),$$

where  $\sigma_m$  and  $\sigma_d$  are chemical shifts for monomer and dimer. The average  $K_d$  was  $5 \pm 1$  mM (Table D-2, Fig D-5a)

#### *Dimerization constant calculated using 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectroscopy*

Chemical shift changes as a function of SDF-1 $\alpha$  concentration were also followed with 2D {<sup>1</sup>H, <sup>15</sup>N} HSQC NMR spectra. The resolution for the <sup>1</sup>H dimension in these spectra was ~20 Hz, which is significantly less than in 1D <sup>1</sup>H spectra. However, chemical shifts for a larger number of peaks could be traced. Data was fit to eq.8 (Table D-2, Fig. D-5b) with the average  $K_d$  of  $7 \pm 4$  mM.

**Table D-2.** The values of dissociation constant as determined by chemical shift changes in 1D and 2D HSQC spectroscopy for indicated residues.

Residue	K <sub>d</sub> values as estimated by chemical shifts changes in 1D spectra	Residue	K <sub>d</sub> values as estimated by chemical shifts changes in 2D HSQC spectra
2,6 H Tyr <sup>7</sup>	4.7 ± 3.9	Ala <sup>19</sup>	10.9 ± 3.7
3,5 H Tyr <sup>7</sup>	4.7 ± 3.9	Val <sup>23</sup>	9.8 ± 12.8
2,6 Phe <sup>13</sup>	7.3 ± 5.8	Ile <sup>38</sup>	3.1 ± 3.6
2,6 Phe <sup>14</sup>	4.7 ± 3.9	Val <sup>39</sup>	9.7 ± 7.4
indole H of Trp <sup>57</sup>	5.2 ± 3.1	Leu <sup>42</sup>	2.1 ± 2.3
		Cys <sup>50</sup>	13.4 ± 8.5
		Ile <sup>51</sup>	3.0 ± 1.2
		Asp <sup>52</sup>	6.1 ± 5.1
		Leu <sup>55</sup>	5.5 ± 7.6
		Trp <sup>57</sup>	6.0 ± 1.7
		Glu <sup>60</sup>	6.1 ± 6.9
	average 5 ± 4		average 7 ± 5

*Extrapolation of per residue T<sub>1</sub><sup>500</sup> and T<sub>2</sub><sup>500</sup> values to infinite dilution*

Per residue T<sub>2</sub><sup>500</sup> and T<sub>1</sub><sup>500</sup> at infinite dilution were obtained using eq.7 and K<sub>d</sub> of 5 mM. Datasets used for extrapolation included T<sub>2</sub><sup>500</sup> data for 2.6, 1.5, 0.9, 0.5, and 0.3 mM SDF-1α, and T<sub>1</sub><sup>500</sup> data for 2.6, 1.5, 0.9 and 0.3 mM SDF-1α. The average for extrapolated data for protein core was: T<sub>2</sub><sup>500</sup> = 158 ± 24 ms (Fig. D-6a) and T<sub>1</sub><sup>500</sup> = 364 ± 40 ms (Fig. D-6b), which is close within determined errors to the theoretical values for monomeric SDF-1α, which are 388 ms for T<sub>1</sub><sup>500</sup> and 181 ms for T<sub>2</sub><sup>500</sup>. Extrapolated T<sub>2</sub> and T<sub>1</sub> followed the same per residue trend as the experimentally obtained T<sub>2</sub><sup>500</sup> and T<sub>1</sub><sup>500</sup> (Fig. D-2).

*Model free analysis of backbone dynamics*

Backbone amide <sup>15</sup>N-T<sub>1</sub>, T<sub>2</sub>, and NOE datasets acquired at 500 and 600 MHz for 1.5 mM and 0.3 mM SDF-1α were examined by model free analysis. This approach affords the interpretation of relaxation data in terms of the overall correlation time, generalized order parameters, and local motions (23, 24). The overall tumbling of SDF-1α was assumed to be isotropic, given that rotational diffusion anisotropy was found to be negligible in previous backbone <sup>15</sup>N relaxation studies of monomeric chemokines (25, 26) and the dimeric chemokine, IL-8 (27). Overall correlation times determined using T<sub>1</sub>, T<sub>2</sub>, and NOE data (28) were the following: τ<sub>m</sub><sup>500</sup> = 6.79 ns and τ<sub>m</sub><sup>600</sup> = 7.24 ns for 1.5 mM

SDF-1 $\alpha$ , and  $\tau_m^{500} = 5.04$  ns and  $\tau_m^{600} = 7.17$  ns for 0.3 mM SDF-1 $\alpha$ . These values appeared to be overestimated in comparison to the expected values for monomeric SDF-1 $\alpha$  (7.8 kDa), even for the most diluted sample where the fraction of dimer is ~5%. This observation is consistent with previous studies, where even the slight dimerization/aggregation can influence the interpretation of relaxation data (29).  $\tau_m^{500}$  was also determined using extrapolated  $T_1^{500}$  and  $T_2^{500}$ , and  $\text{NOE}^{500}$  data for 1.5 mM SDF-1 $\alpha$  (used due to unchanged  $\text{NOE}^{500}$  spectra and minimal error).  $\tau_m^{500}$  obtained in this manner was 4.42 ns.

Generalized order parameters ( $S^2$ ), chemical exchange terms ( $R_{\text{ex}}$ ), and correlation times ( $\tau_m$  and  $\tau_s$ ) are obtained from five different dynamic models in Lipari-Zhabo formalism (23, 30). Average  $S^2$  was equal to  $0.77 \pm 0.04$  (500 MHz) and  $0.81 \pm 0.05$  (600 MHz) for 1.5 mM SDF-1 $\alpha$ ;  $0.79 \pm 0.09$  (500 MHz) and  $0.86 \pm 0.07$  (600 MHz) for 0.3 mM SDF-1 $\alpha$ ; and  $0.77 \pm 0.09$  for the extrapolated dataset (Fig. D-7).  $S^2$  for the core residues were  $0.81 \pm 0.03$  (500 MHz) and  $0.87 \pm 0.03$  (600 MHz) for 1.5 mM SDF-1 $\alpha$ ;  $0.87 \pm 0.07$  (500 MHz) and  $0.91 \pm 0.06$  (600 MHz) for 0.3 mM SDF-1 $\alpha$ ; and  $0.83 \pm 0.07$  for the extrapolated dataset. Residues in the unstructured regions, the N- and C-termini and loops, displayed smaller values of  $S^2$ , indicating larger amplitudes of motions. Experimental error in relaxation measurements increased with SDF-1 $\alpha$  dilution as a result of decreasing signal to noise ratio. Average  $S^2$  was slightly smaller for the extrapolated dataset, but similar to average experimental  $S^2$  for 1.5 and 0.3 mM SDF-1 $\alpha$  (Fig. D-7).

Values for  $R_{\text{ex}}$  typically indicate the presence of chemical exchange. The trend for  $R_{\text{ex}}$  shows the highest values for regions of residues 23-31, 36-40, 50-51, and at the C-terminus (Fig. D-8). Residues with the largest  $R_{\text{ex}}$  are directly involved in the formation of dimer interface (16, 17).  $R_{\text{ex}}$  terms decreased upon dilution from  $\sim 20$  s $^{-1}$  for 1.5 mM SDF-1 $\alpha$  to  $\sim 10$  s $^{-1}$  for 0.3 mM SDF-1 $\alpha$ .  $R_{\text{ex}}$  terms were also present in extrapolated case for residues 17, 20 (the N-terminus), 24, 27, 28, 30 (the first  $\beta$  strand), 36, 38, 40 (the second  $\beta$  strand), 50-51 (the third  $\beta$  strand) and 64 (the C-terminal helix). If the rate of interconversion of exchanging species is fast relative to the chemical shift difference, the value of  $R_{\text{ex}}^{600}/R_{\text{ex}}^{500}$  should be equal to  $600^2/500^2$ , or 1.44. Residues with  $R_{\text{ex}}^{600}/R_{\text{ex}}^{500}$  values within the range  $1.44 \pm 0.28$  (20% of error) included 23, 25, 27, 28, 50, and 65 for 1.5 mM SDF-1 $\alpha$ , and 25, 27, 28, 29, 38, 41, and 65 for 0.3 mM SDF-1 $\alpha$ .

## Discussion

### *SDF-1 $\alpha$ exists in monomer-dimer equilibrium*

Experimental relaxation data demonstrates the existence of monomer-dimer equilibrium for SDF-1 $\alpha$ . Proton chemical shifts observed in both 1D and 2D NMR experiments, and per residue transverse relaxation rates,  $R_2$ , demonstrated a dependence on the concentration of SDF-1 $\alpha$ . Determined  $K_d$  of  $5 \pm 1$  mM indicates a weaker self-association than previously reported (19), and agrees with the  $K_d$  dependence on buffer composition (20). In our study, the buffer was of low ionic strength, containing only 20 mM acetate and negligibly small amounts of azide and chloride. This buffer was optimized previously to minimize dimerization (10). High ionic strength buffers are expected to neutralize the positively charged residues on the first  $\beta$ -sheet of SDF-1 $\alpha$  (Lys<sup>24</sup>, His<sup>25</sup>, Lys<sup>27</sup>), reducing the repulsion between monomers, and favoring dimer formation (20). In our experiments, the percentage of dimer was 27.4% in the most concentrated sample (2.6 mM SDF-1 $\alpha$ ), 19.2% and 5.2% in the sample containing 1.5 mM and 0.3 mM SDF-1 $\alpha$ , respectively.

### *Model free analysis*

Protein dimerization often complicates the interpretation of relaxation measurements (31, 32), causing significant overestimations of the generalized order parameter and internal motions, as shown by simulations (29). The model free analysis, appropriate for the interpretation of relaxation data of purely monomeric or dimeric species, is not strictly suitable if both species are present in substantial quantities. There are a number of approaches that account for monomer-dimer equilibrium in model free analysis (33-36). One of them presents relaxation data only qualitatively (36). Another requires approximations such as equivalent local motions for both dimeric and monomeric species (33). We have extrapolated relaxation parameters to infinite dilution to obtain  $T_1$  and  $T_2$  for monomeric species, similarly to the approach of Pfuhl and coworkers, who extrapolated their data to infinite dilution where only a pure dimer existed (34). Model free analysis was performed on datasets obtained for two different concentrations of SDF-1 $\alpha$  and on a dataset extrapolated to infinite dilution. The value of the overall correlation time for SDF-1 $\alpha$  decreased with the decreasing fraction of dimer, as previously observed (33, 34, 37). The value of 4.42 ns for  $\tau_m^{500}$  obtained using extrapolated dataset was reasonable for a 7.8 kDa protein containing flexible termini. This is consistent with hydrodynamic calculations performed on the dynamin pleckstrin

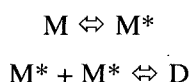
homology (PH) domain, where the presence of flexible tail slightly decreased the rate of overall rotational tumbling (33).

Generalized order parameters obtained for 1.5 and 0.3 mM SDF-1 $\alpha$ , and for the extrapolated dataset showed similar trends with the N-terminus (residues 1-9) and the C-terminus (residues 63-67) being flexible.  $S^2$  can be substantially overestimated for proteins that self-associate (29). However, the results presented in our work were similar to the results of previous experiments (34) and showed that the increases in  $S^2$  in the presence of dimerization was not drastic.

Interestingly, residues 8-12 have been suggested to interact with the cognate SDF-1 $\alpha$  receptor (16, 38). The N-terminus of SDF-1 $\alpha$  (residues 1-17) adopts an extended conformation with two cysteines, Cys<sup>9</sup> and Cys<sup>11</sup>, constraining the region 8-12 by the two disulphide bonds Cys<sup>9</sup>-Cys<sup>34</sup> and Cys<sup>11</sup>-Cys<sup>50</sup>.  $S^2$  values of  $0.91 \pm 0.09$  and  $0.87 \pm 0.1$  for Cys<sup>11</sup> and Arg<sup>12</sup> indicate that the flexibility of this region is restricted on the ps timescale. For comparison, the average  $S^2$  value for Arg<sup>8</sup> was found to be  $0.43 \pm 0.05$ . Restriction of flexibility within this region may contribute favorably to the energy of the interaction with a receptor.

The number of residues requiring  $R_{ex}$  terms to properly account for relaxation data and the magnitude of  $R_{ex}$  terms showed a clear dependence on protein concentration similarly to reported data (34). This dependence was not linear, as expected from theoretical (39) and experimental (40) studies. Given that relaxation data for residues 25, 26, and 29 were missing due to the line broadening or due to the poor quality of extrapolations, it is possible that the entire first  $\beta$ -strand is involved in the conformational exchange. The second set of residues with  $R_{ex}$  terms was mapped to the second  $\beta$ -strand, the third  $\beta$ -strand, and the C-terminus. Residues belonging to the first  $\beta$ -strand and the C-helix directly participate in the formation of dimer interface. Residues with  $R_{ex}$  terms in the second and third  $\beta$ -strands interact directly with the residues comprising interface.

In the absence of monomer-dimer exchange, e.g. in the purely monomeric case, some residues still required  $R_{ex}$  terms in the model free analysis. These are generally the same residues involved in dimerization (Fig. D-9), and must reflect conformational flexibility within the monomer. The mechanism consistent with this data is:



where M represents a conformation of the monomer less prone to dimerization, M\* a conformation more prone to dimerization, and D is the dimer. Such a mechanism would help explain the apparent discrepancies in the solution structures of SDF-1 $\alpha$  determined by NMR. The NMR structure determined by Gozansky et al. (18) is similar to dimeric structures determined using X-ray crystallography (16, 17), and would represent the species M\*; whereas the Crump et al. structure (10) would represent the species M. The phosphate buffer used by Gozansky et al. is expected to favor dimerization as compared to the acetate buffer used by Crump et al. (20).

#### *Comparison with other chemokines and implications for receptor interaction*

Chemokines were found to undergo dimerization using two major subclasses of binding modes. The first group dimerizes in a fashion similar to IL-8, utilizing residues on the first  $\beta$ -strand to establish the dimer interface, and the second group dimerizes similar to vMIP-I, utilizing residues from N-terminal loops (41). It would be interesting to compare the parameters obtained from relaxation measurements, especially  $R_{ex}$  terms, with respect to the different dimerization modes utilized by chemokines. In case of IL-8, which exists predominantly in dimeric form, the existence of monomer-dimer equilibrium resulted in line broadening in the regions of residues that make intersubunit contacts, even if the concentration of a monomer was negligible (27). In the case of eotaxin, eotaxin-2, and eotaxin-3 (25, 42, 43), the information about dimer interface is lacking and therefore,  $R_{ex}$  terms cannot be ascribed to self-association. In case of fractalkine, residues involved in crystal packing and those with  $R_{ex}$  terms are partially overlapped (26). These studies suggest a correlation between oligomerization behavior of proteins and the requirement for  $R_{ex}$  terms to account for the relaxation data. The detailed studies of relaxation and its concentration dependence for larger protein groups might be interesting.

The biological relevance of chemokine oligomerization may differ among the various members of the chemokine family that includes more than 60 members with a broad range of oligomerization behavior. The dependence of SDF-1 $\alpha$  dimerization on pH and buffer composition was studied recently (20). SDF-1 $\alpha$ , with a physiological concentration of 25-50 ng/ml (44) and a dimerization constant from  $\sim 150 \mu\text{M}$  in PBS buffer (19) to  $\sim 5 \text{ mM}$  in acetate buffer (in our work), is unlikely to dimerize under physiological conditions. However, high ionic strength and interactions between SDF-1 $\alpha$  and glycosaminoglycans at the cell surface can increase the fraction of dimeric SDF-1 $\alpha$ .

Therefore it is reasonable to expect that SDF-1 $\alpha$  *in vivo* can function both as a monomer and as a dimer under certain conditions. The intrinsic ability of SDF-1 $\alpha$  to dimerize, and the interplay with the cell surface environment may play role in modulating the interaction of SDF-1 $\alpha$  with its cognate receptor.

## Materials and Methods

### *Protein expression and purification*

The DNA encoding human SDF-1 $\alpha$  (residue 1-67) was purchased from InvivoGen and subcloned into the pET-3a expression vector from Novagen. *E. coli* strain BL21(DE3)pLysS was transformed with the expression vector and incubated at 37 °C to an OD<sub>600</sub> of 0.6-0.8. Cell cultures were induced with IPTG and harvested after 3 hours. Uniformly <sup>15</sup>N-labelled SDF-1 $\alpha$  was expressed similarly in a media enriched with <sup>15</sup>N ammonium sulfate (45). Cell pellet was lysed using French press, centrifuged (10000 rpm, 10 min), homogenized, and solubilized in 8M urea buffer (50 mM Tris, pH 8.0). Homogenate was centrifuged again to remove cell walls and membrane debris (18000 rpm, 1 hour), and applied to a cation exchange CM-Sepharose column (50 mM Tris, pH 8.0). Protein was refolded by overnight dialysis against 10 mM NH<sub>4</sub>HCO<sub>3</sub> and the first methionine residue was cleaved with CNBr as described earlier (46). The protein was purified on a size exclusion column (Superdex 75), dialyzed against 10 mM NH<sub>4</sub>HCO<sub>3</sub>, and lyophilized. The molecular weight determined by MALDI mass-spectroscopy was equal to the expected (7831  $\pm$  1 for reduced species). The amino acid composition was confirmed by the amino acid analysis. Biological activity of recombinant SDF-1 $\alpha$  was tested in the chemotaxis assay using mice splenocytes. Recombinant SDF-1 $\alpha$  demonstrated chemotactic response similar to the commercially available control SDF-1 $\alpha$  sample (data not shown).

### *Sample preparation*

NMR samples were prepared in 500  $\mu$ l of buffer containing 20 mM CD<sub>3</sub>COONa, 1 mM NaN<sub>3</sub>, and 1 mM DSS in 90% H<sub>2</sub>O/10% D<sub>2</sub>O. The pH was adjusted to 4.9. The two most concentrated samples were prepared by dissolving the lyophilized powder of recombinant SDF-1 $\alpha$ . The rest of the samples were serially diluted from the sample containing 1.5 mM of SDF-1 $\alpha$ , where the concentration was verified by amino acid analysis. The concentrations of recombinant SDF-1 $\alpha$  analyzed in this work were equal to 2.6, 1.5, 0.9, 0.5, 0.3, and 0.1 mM.

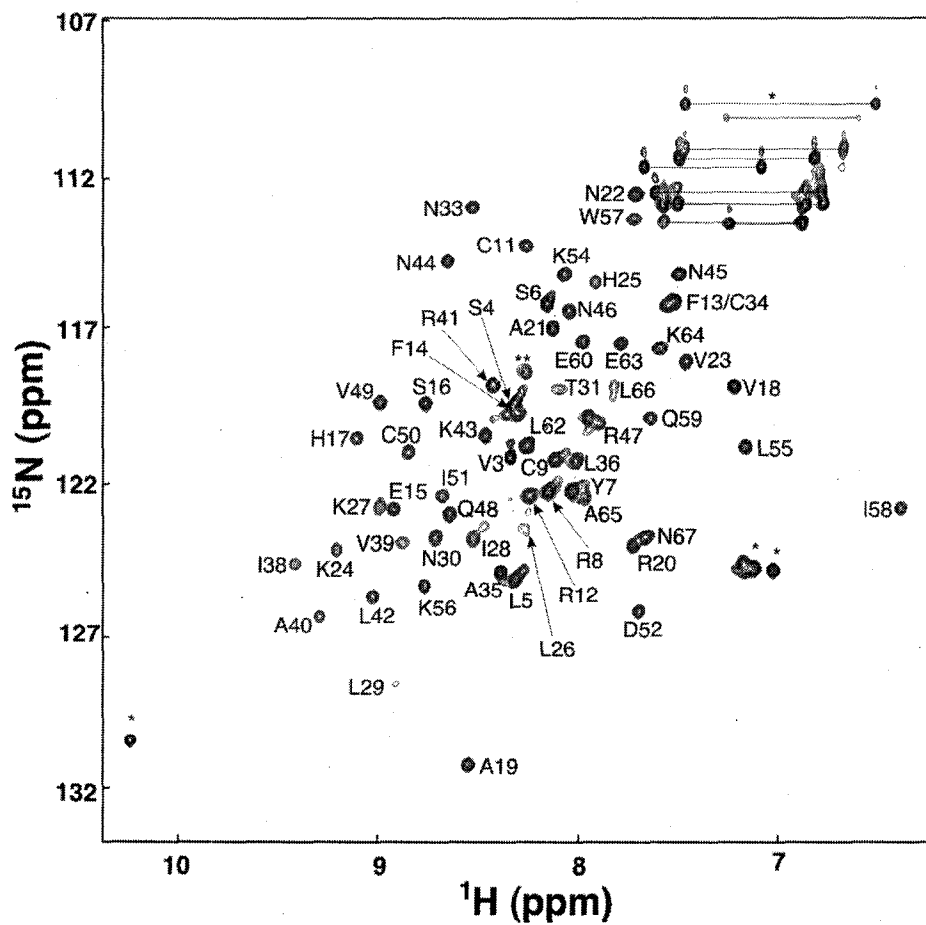
### *NMR spectroscopy (data acquisition and processing)*

All NMR data were acquired at 30 °C on a Varian Inova 500 MHz spectrometer and a Varian Unity 600 MHz spectrometer equipped with z-axis pulsed field gradient triple resonance probes. In order to assign the backbone and side chain <sup>1</sup>H and <sup>15</sup>N chemical shifts, 3D NOESY-HSQC, 3D TOCSY-HSQC, and 2D {<sup>1</sup>H,<sup>15</sup>N} HSQC were acquired at 500 MHz for 1.5 mM SDF-1 $\alpha$ . Relaxation data were obtained from <sup>15</sup>N-T<sub>1</sub>, <sup>15</sup>N-T<sub>2</sub>, and <sup>1</sup>H-<sup>15</sup>N NOE experiments conducted at 500 and 600 MHz for samples containing 2.6, 1.5, 0.8, 0.5, and 0.3 mM SDF-1 $\alpha$ . At 500 and 600 MHz, T<sub>1</sub> data were acquired using relaxation delays of 10, 50, 100, 200, 300, and 400 ms, and T<sub>2</sub> data were acquired using relaxation delays equal to 10, 30, 50, 70, 90, and 110 ms. The delays between transients in T<sub>1</sub> and T<sub>2</sub> experiments were set to 3.5 s for 600 MHz and 2.5 s and 3.5 s in T<sub>1</sub> and T<sub>2</sub> experiments respectively for 500 MHz. A sufficiently long delay

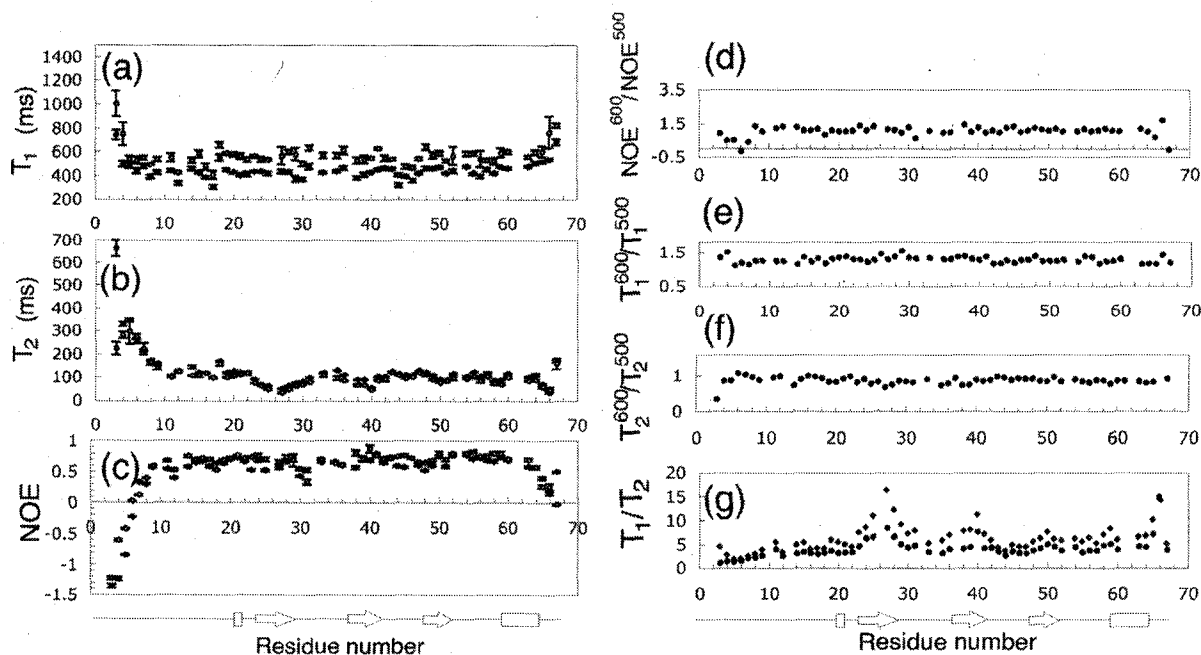


between transients was necessary to prevent sample from heating as previously shown (47).  $\{^1\text{H}, ^{15}\text{N}\}$  NOE experiments were conducted using delays of 5 s for spectra recorded without proton saturation at both fields and delays of 2 s for spectra recorded with proton saturation. The proton saturation time was equal to 3 s and the total time between acquisitions was equal to 5 s. In addition, to follow the chemical shift changes as a function of SDF-1 $\alpha$  concentration, 1D  $^1\text{H}$  spectra and 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC spectra were acquired for every sample. Processing of NMR data was accomplished with NMRPipe software. Relaxation parameters  $T_1$  and  $T_2$  were obtained using NMRview and were in agreement with  $T_1$  and  $T_2$  values calculated using Mathematica scripts provided by Leo Spyropoulos.

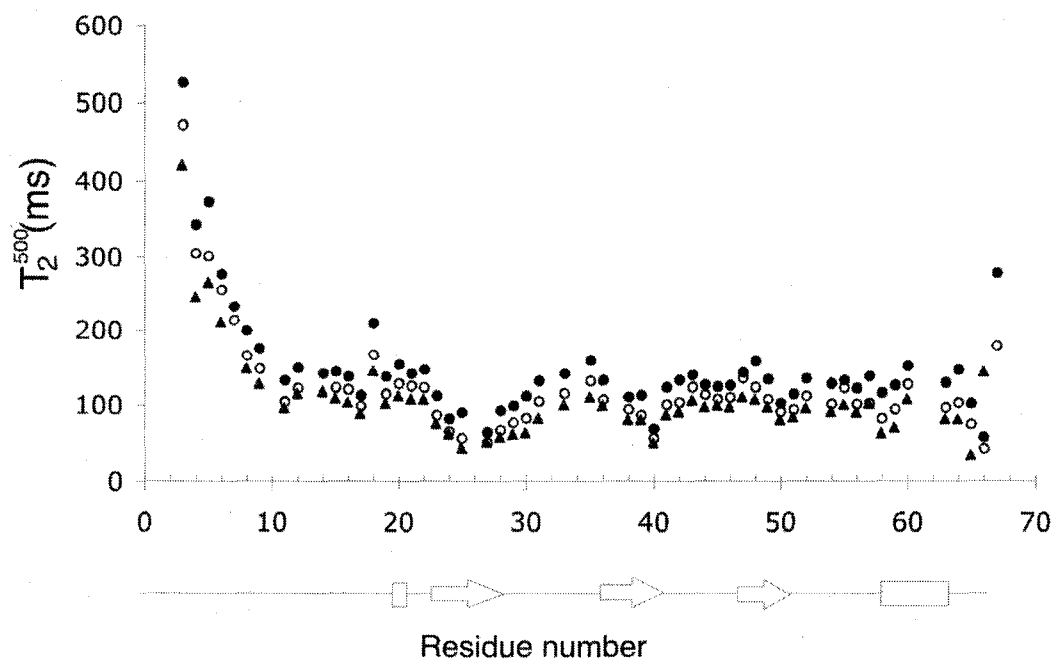
**Figure D-1.** 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC correlation spectrum of 1.5 mM SDF-1 $\alpha$ . Assigned residues are indicated on the spectrum with labels. Arginine and glutamine side chain amide protons as well as the indole NH of Trp<sup>57</sup> are indicated with an asterisk. The resonance marked with a double asterisk was not assigned.



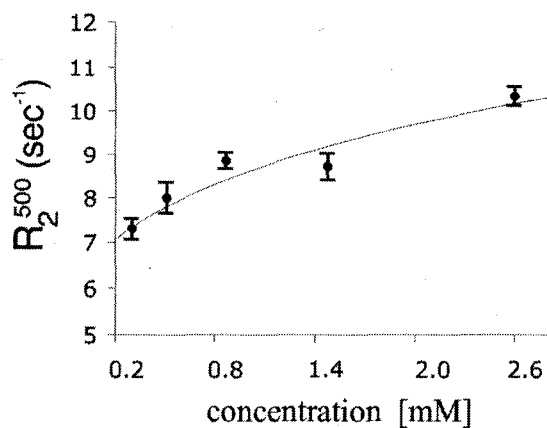
**Figure D-2.** Relaxation data obtained at two magnetic fields for 1.5 mM SDF-1 $\alpha$ . (a)  $T_1^{600}$  (○) and  $T_1^{500}$  (●), (b)  $T_2^{600}$  (○) and  $T_2^{500}$  (●), (c)  $\text{NOE}^{600}$  (○) and  $\text{NOE}^{500}$  (●), (d)  $\text{NOE}^{600}/\text{NOE}^{500}$  ratio, (e)  $T_1^{600}/T_1^{500}$  ratio, (f)  $T_2^{600}/T_2^{500}$  ratio, and (g)  $T_1^{600}/T_2^{600}$  ratio (◇) and  $T_1^{500}/T_2^{500}$  ratio (●). Elements of secondary structure were determined from the coordinates of the crystal structure (17).



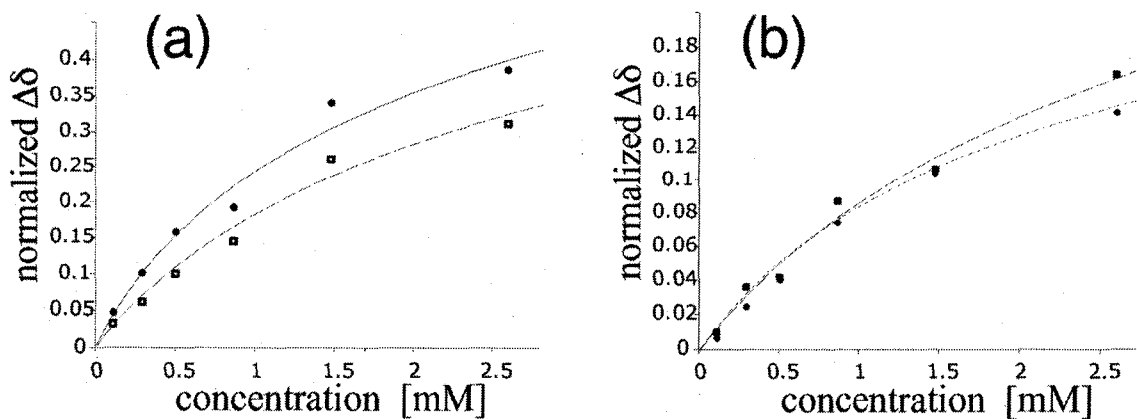
**Figure D-3.**  $^{15}\text{N}$ - $T_2^{500}$  relaxation data for SDF-1 $\alpha$  samples containing different concentrations.  $^{15}\text{N}$ - $T_2^{500}$  relaxation data is shown for 2.6 mM (●), 0.9 mM (○), and 0.3 mM (▲) SDF-1 $\alpha$ . Only three  $T_2^{500}$  datasets from a total of five are shown for clarity.



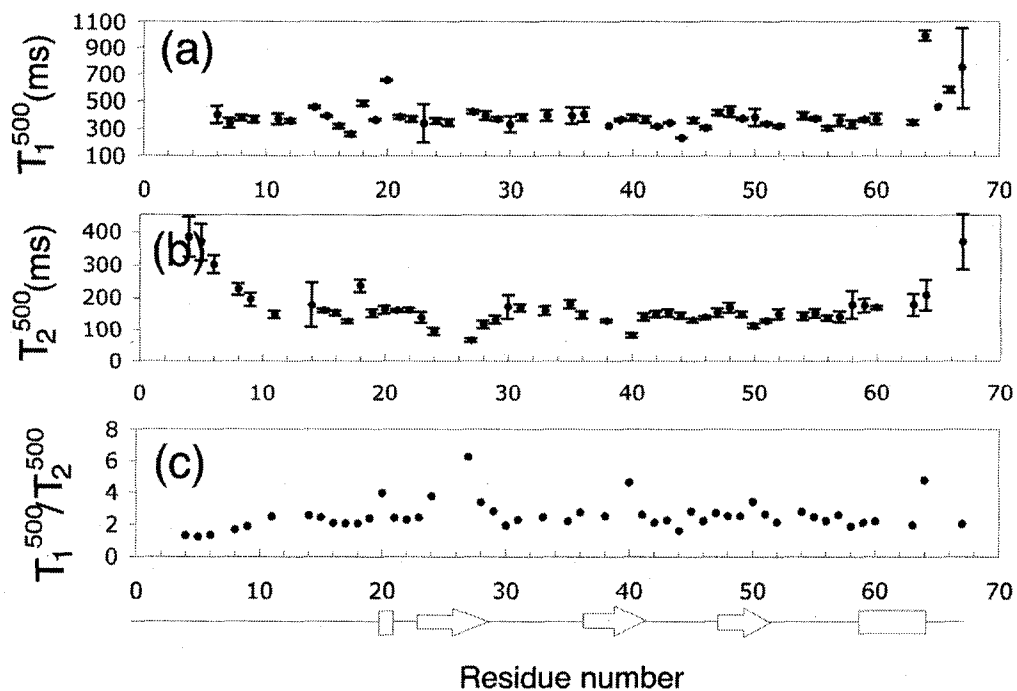
**Figure D-4.**  $^{15}\text{N}$ - $R_2^{500}$  relaxation data obtained for residues in the structured core of SDF-1 $\alpha$ .  $R_2$  values were averaged and plotted as a function of the concentration of SDF-1 $\alpha$ . The best fit to eq.7 yielded  $K_d = 9 \pm 24$  mM,  $R_2^m = 7 \pm 1$  s $^{-1}$  and  $R_2^d = 18 \pm 15$  s $^{-1}$ .



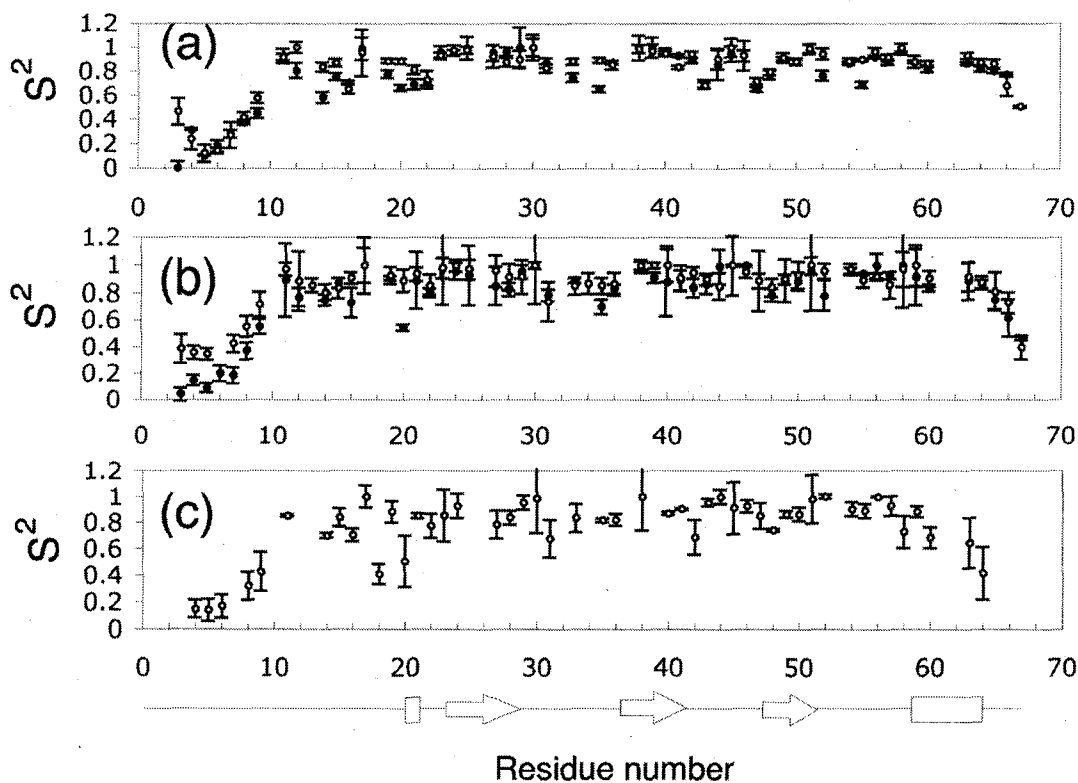
**Figure D-5.** Chemical shift changes obtained using 1D  $^1\text{H}$  (a) and 2D  $\{^1\text{H}, ^{15}\text{N}\}$  HSQC NMR spectroscopy (b) plotted as a function of SDF-1 $\alpha$  concentration for 3,5H of Tyr $^7$  (●) and 2,6H of Phe $^{14}$  (□) (a) and for Val $^{39}$  (■) and Trp $^{57}$  (●) (b).



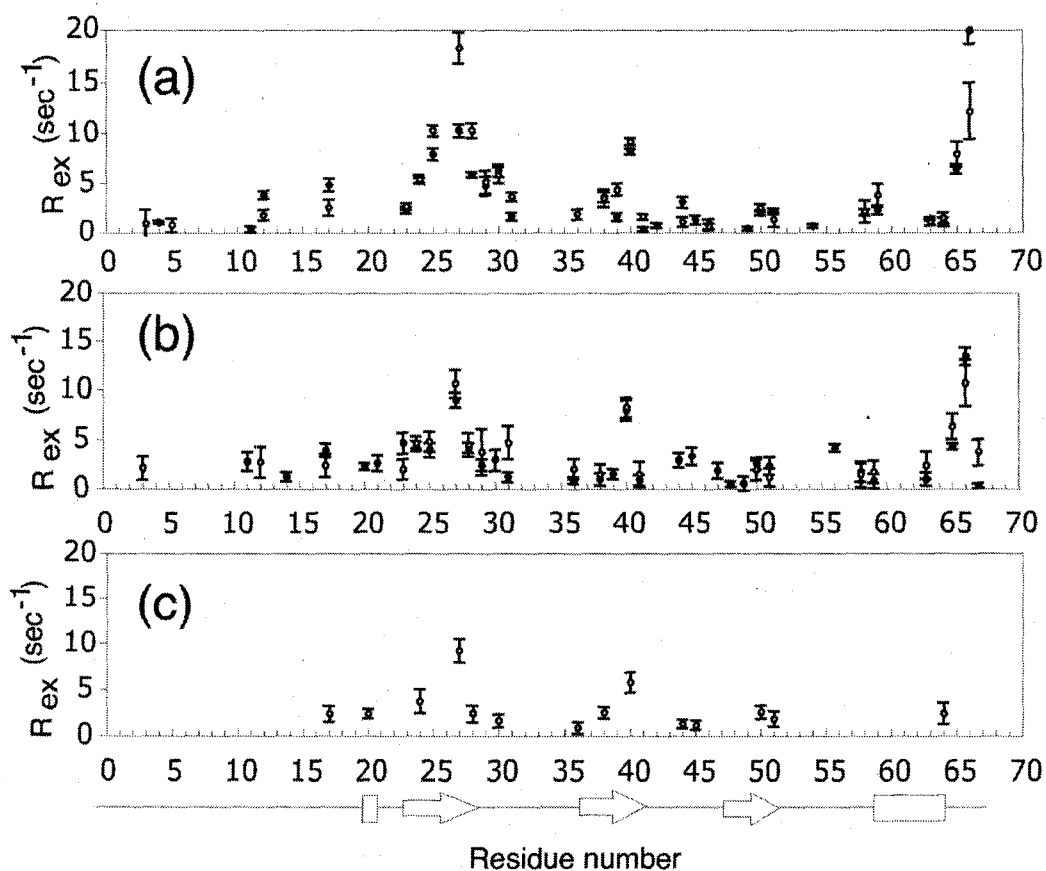
**Figure D-6.** Relaxation parameters,  $T_1^{500}$  (a),  $T_2^{500}$  (b), and  $T_1^{500}/T_2^{500}$  (c), extrapolated to infinite dilution. Elements of secondary structure were determined from the coordinates of the crystal structure (17).



**Figure D-7.** Generalized order parameters,  $S^2$ , for samples containing different concentrations of SDF-1 $\alpha$ , and relaxation data extrapolated to infinite dilution. Generalized order parameters for 1.5 mM SDF-1 $\alpha$  at 500 (●) and 600 (○) MHz (a). Generalized order parameters for 0.3 mM SDF-1 $\alpha$  at 500 (●) and 600 (○) MHz (b). Generalized order parameters obtained using  $T_1^{500}$  and  $T_2^{500}$  values extrapolated to infinite dilution (c).

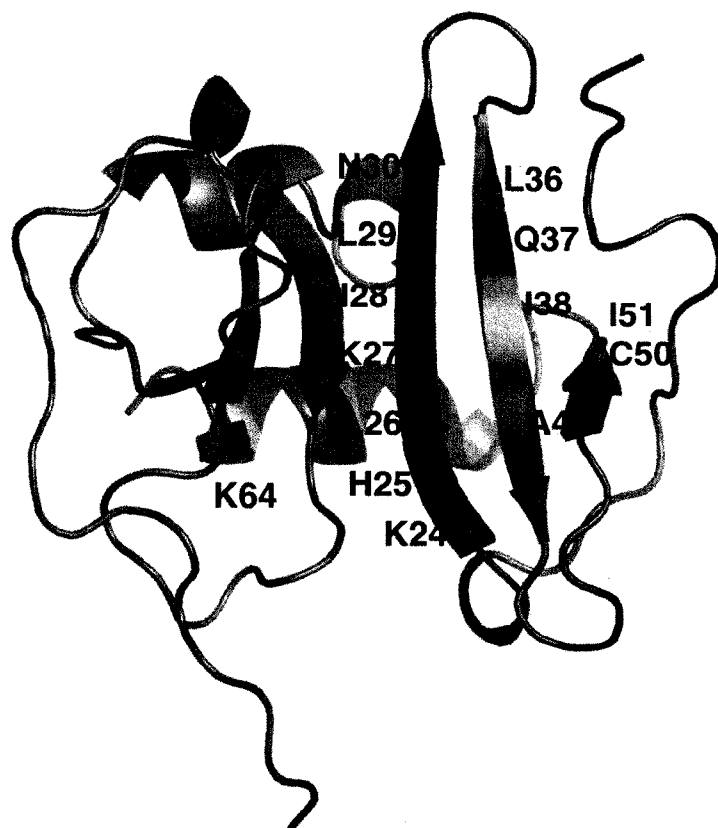


**Figure D-8.**  $R_{ex}$  parameters obtained for samples containing different concentrations of SDF-1 $\alpha$  and for the case where relaxation parameters extrapolated to infinite dilution.  $R_{ex}$  parameters for 1.5 mM SDF-1 $\alpha$  at 500 (●) and 600 (○) MHz (a).  $R_{ex}$  parameters for 0.3 mM SDF-1 $\alpha$  at 500 (●) and 600 (○) MHz (b).  $R_{ex}$  parameters obtained using  $T_1^{500}$  and  $T_2^{500}$  values extrapolated to infinite dilution (c).





**Figure D-9.** Mapping of the residues requiring  $R_{ex}$  terms on the structure of SDF-1 $\alpha$  dimer. The structure of dimeric SDF-1 $\alpha$  is taken from Ohnishi *et al.* (16), (PDB code 1QG7). The two monomers are shown in grey and blue, with the residues suspected to undergo chemical exchange indicated with labels on the blue monomer. Regions colored in red are directly involved in the formation of dimer. Regions colored in pink may be indirectly affected by the formation of dimer. Residues in black are missing from the model free analysis due to line broadening or inadequate quality of the extrapolations.



## References

- (1) Baggiolini, M. (1998) Chemokines and leukocyte traffic. *Nature* 392(6676), 565.
- (2) Kunkel, S. L., and Godessart, N. (2002) Chemokines in autoimmunity: from pathology to therapeutics. *Autoimmun Rev* 1, 313-20.
- (3) Berson, J. F., Long, D., Doranz, B. J., Rucker, J., Jirik, F. R., and Doms, R. W. (1996) A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *J Virol.* 70(9), 6288.
- (4) Murphy, P. M. (2001) Chemokines and the molecular basis of cancer metastasis. *N Engl J Med.* 345(11), 833-5.
- (5) Rubin, J. B., Kung, A. L., Klein, R. S., Chan, J. A., Sun, Y., Schmidt, K., Kieran, M. W., Luster, A. D., and Segal, R. A. (2003) A small-molecule antagonist of CXCR4 inhibits intracranial growth of primary brain tumors. *Proc Natl Acad Sci U S A* 100, 13513-8.
- (6) Cooper, C. R., Chay, C. H., Gendernalik, J. D., Lee, H. L., Bhatia, J., Taichman, R. S., McCauley, L. K., Keller, E. T., and Pienta, K. J. (2003) Stromal factors involved in prostate carcinoma metastasis to bone. *Cancer.* 97(3 Suppl), 739-47.
- (7) Phillips, R. J., Burdick, M. D., Lutz, M., Belperio, J. A., Keane, M. P., and Strieter, R. M. (2003) The stromal derived factor-1/CXCL12-CXC chemokine receptor 4 biological axis in non-small cell lung cancer metastases. *Am J Respir Crit Care Med.* 167(12), 1676-86.
- (8) Muller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M. E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S. N., Barrera, J. L., Mohar, A., Verastegui, E., and Zlotnik, A. (2001) Involvement of chemokine receptors in breast cancer metastasis. *Nature* 410, 50-6.
- (9) Epstein, R. J. (2004) The CXCL12-CXCR4 chemotactic pathway as a target of adjuvant breast cancer therapies. *Nat Rev Cancer* 4, 901-9.
- (10) Crump, M. P., Gong, J. H., Loetscher, P., Rajarathnam, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J. L., Baggiolini, M., Sykes, B. D., and Clark-Lewis, I. (1997) Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J.* 16, 6996-7007.
- (11) Kuschert, G. S., Coulin, F., Power, C. A., Proudfoot, A. E., Hubbard, R. E., Hoogewerf, A. J., and Wells, T. N. (1999) Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses. *Biochemistry* 38, 12959-68.
- (12) Mammen, M., Choi, S. K., and Whitesides, G. M. (1998) Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angewandte Chemie International Edition* 37, 2754-2794.
- (13) Fernandez, E. J., and Lolis, E. (2002) Structure, function, and inhibition of chemokines. *Annu Rev Pharmacol Toxicol.* 42, 469-99.
- (14) Rajarathnam, K., Sykes, B. D., Kay, C. M., Dewald, B., Geiser, T., Baggiolini, M., and Clark-Lewis, I. (1994) Neutrophil activation by monomeric interleukin-8. *Science* 264, 90-2.
- (15) Proudfoot, A. E., Handel, T. M., Johnson, Z., Lau, E. K., LiWang, P., Clark-Lewis, I., Borlat, F., Wells, T. N., and Kosco-Vilbois, M. H. (2003) Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines. *Proc Natl Acad Sci U S A* 100, 1885-90.
- (16) Ohnishi, Y., Senda, T., Nandhagopal, N., Sugimoto, K., Shioda, T., Nagai, Y., and Mitsui, Y. (2000) Crystal structure of recombinant native SDF-1alpha with additional mutagenesis studies. *J Interferon Cytokine Res* 20, 691-700.

- (17) Dealwis, C., Fernandez, E. J., Thompson, D. A., Simon, R. J., Siani, M. A., and Lolis, E. (1998) Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1 $\alpha$ , a potent ligand for the HIV-1 "fusin" coreceptor. *Proc Natl Acad Sci U S A* 95, 6941-6.
- (18) Gozansky, E. K., Louis, J. M., Caffrey, M., and Clore, G. M. (2005) Mapping the binding of the N-terminal extracellular tail of the CXCR4 receptor to stromal cell-derived factor-1 $\alpha$ . *J Mol Biol* 345, 651-8.
- (19) Holmes, W. D., Consler, T. G., Dallas, W. S., Rocque, W. J., and Willard, D. H. (2001) Solution studies of recombinant human stromal-cell-derived factor-1. *Protein Expr Purif* 21, 367-77.
- (20) Veldkamp, C. T., Peterson, F. C., Pelzek, A. J., and Volkman, B. F. (2005) The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci* 14, 1071-81.
- (21) Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone <sup>1</sup>H and <sup>15</sup>N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J Biomol NMR* 4, 845-58.
- (22) Goodman, J. L., Pagel, M. D., and Stone, M. J. (2000) Relationships between protein structure and dynamics from a database of NMR-derived backbone order parameters. *J Mol Biol* 295, 963-78.
- (23) 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.
- (24) Lipari, G., and Szabo, a. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules.2. Analysis of Experimental Results. *Journal of the American Chemical Society* 104, 4559.
- (25) Crump, M. P., Spyrapoulos, L., Lavigne, P., Kim, K. S., Clark-lewis, I., and Sykes, B. D. (1999) Backbone dynamics of the human CC chemokine eotaxin: fast motions, slow motions, and implications for receptor binding. *Protein Sci* 8, 2041-54.
- (26) Mizoue, L. S., Bazan, J. F., Johnson, E. C., and Handel, T. M. (1999) Solution structure and dynamics of the CX3C chemokine domain of fractalkine and its interaction with an N-terminal fragment of CX3CR1. *Biochemistry* 38, 1402-14.
- (27) Grasberger, B. L., Gronenborn, A. M., and Clore, G. M. (1993) Analysis of the backbone dynamics of interleukin-8 by <sup>15</sup>N relaxation measurements. *J Mol Biol* 230, 364-72.
- (28) Spyrapoulos, L., Lavigne, P., Crump, M. P., Gagne, S. M., Kay, C. M., and Sykes, B. D. (2001) Temperature dependence of dynamics and thermodynamics of the regulatory domain of human cTn C. *Biochemistry* 40, 12541-51.
- (29) Schurr, J. M., Babcock, H. P., and Fujimoto, B. S. (1994) A test of the model-free formulas. Effects of anisotropic rotational diffusion and dimerization. *J Magn Reson B* 105, 211-24.
- (30) Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the Simple 2-Parameter Model-Free Approach to the Interpretation of N-15 Nuclear Magnetic-Relaxation of Proteins. *J Am Chem Soc* 112, 4989-4991.
- (31) Spyrapoulos, L., and Sykes, B. D. (2001) Thermodynamic insights into proteins from NMR spin relaxation studies. *Curr Opin Struct Biol* 11, 555-9.
- (32) Korchuganov, D. S., Nolde, S. B., Reibarkh, M. Y., Orekhov, V. Y., Schulga, A. A., Ermolyuk, Y. S., Kirpichnikov, M. P., and Arseniev, A. S. (2001) NMR study of monomer-dimer equilibrium of barstar in solution. *J Am Chem Soc* 123, 2068.

- (33) Fushman, D., Cahill, S., and Cowburn, D. (1997) The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: analysis of <sup>15</sup>N relaxation with monomer/dimer equilibration. *J Mol Biol* 266, 173-94.
- (34) Pfuhl, M., Chen, H. A., Kristensen, S. M., and Driscoll, P. C. (1999) NMR exchange broadening arising from specific low affinity protein self-association: analysis of <sup>15</sup>N nuclear relaxation for rat CD2. *J Biomol NMR* 14, 307-20.
- (35) Gryk, M. R., Abseher, R., Simon, B., Nilges, M., and Oschkinat, H. (1998) Heteronuclear relaxation study of the PH domain of beta-spectrin: restriction of loop motions upon binding inositol trisphosphate. *J Mol Biol* 280, 879-96.
- (36) Grzesiek, S., Bax, A., Hu, J. S., Kaufman, J., Palmer, I., Stahl, S. J., Tjandra, N., and Wingfield, P. T. (1997) Refined solution structure and backbone dynamics of HIV-1 Nef. *Protein Sci* 6, 1248-63.
- (37) Mercier, P., Spyrapoulos, L., and Sykes, B. D. (2001) Structure, dynamics, and thermodynamics of the structural domain of troponin C in complex with the regulatory peptide 1-40 of troponin I. *Biochemistry* 40, 10063-77.
- (38) Baryshnikova, O. K., Rainey, J. K., and Sykes, B. D. (2005) Nuclear magnetic resonance studies of CXC chemokine receptor 4 allosteric peptide agonists in solution. *J Pept Res* 66, 12-21.
- (39) Luz, Z., and Meiboom, S. (1963) NMR Study of Protolysis of Trimethylammonium Ion in Aqueous Solution - Order of Reaction with Respect to Solvent. *J Chem Phys* 39, 366-&.
- (40) Baldo, J. H., Halford, S. E., Patt, S. L., and Sykes, B. D. (1975) The stepwise binding of small molecules to proteins. Nuclear magnetic resonance and temperature jump studies of the binding of 4-(N-acetylaminoglucosyl)-N-acetylglucosamine to lysozyme. *Biochemistry* 14, 1893-9.
- (41) Luz, J. G., Yu, M., Su, Y., Wu, Z., Zhou, Z., Sun, R., and Wilson, I. A. (2005) Crystal Structure of Viral Macrophage Inflammatory Protein 1 Encoded by Kaposi's Sarcoma-associated Herpesvirus at 1.7Å. *J Mol Biol* 352, 1019-28.
- (42) Ye, J., Mayer, K. L., Mayer, M. R., and Stone, M. J. (2001) NMR solution structure and backbone dynamics of the CC chemokine eotaxin-3. *Biochemistry* 40, 7820-31.
- (43) Mayer, K. L., and Stone, M. J. (2003) Backbone dynamics of the CC-chemokine eotaxin-2 and comparison among the eotaxin group chemokines. *Proteins* 50, 184-91.
- (44) Liao, T. S., Yurgelun, M. B., Chang, S. S., Zhang, H. Z., Murakami, K., Blaine, T. A., Parisien, M. V., Kim, W., Winchester, R. J., and Lee, F. Y. (2005) Recruitment of osteoclast precursors by SDF-1 in giant cell tumor of bone. *J Orthop Res* 23, 203-9.
- (45) Li, M. X., Corson, D. C., and Sykes, B. D. (2002) Structure determination by NMR. Isotope labeling. *Methods Mol Biol* 173, 255-65.
- (46) Smith, B. J. (2003) Chemical cleavage of polypeptides. *Methods Mol Biol* 211, 63-82.
- (47) Gagne, S. M., Tsuda, S., Spyrapoulos, L., Kay, L. E., and Sykes, B. D. (1998) Backbone and methyl dynamics of the regulatory domain of troponin C: anisotropic rotational diffusion and contribution of conformational entropy to calcium affinity. *J Mol Biol* 278, 667-86.