

University of Alberta

*Structure and dynamics of biomolecules: probing muscle regulation,
prion protein unfolding and drug insertion into DNA
by nuclear magnetic resonance spectroscopy*

by

Olivier Julien

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

© Olivier Julien
Spring 2011
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Examining Committee

Brian D. Sykes, Department of Biochemistry

Leo Spyropoulos, Department of Biochemistry

Howard S. Young, Department of Biochemistry

David Westaway, Department of Medicine

Frank Sönnichsen, University of Kiel

Education is what remains after one has forgotten everything he learned in school.

- Albert Einstein

ABSTRACT

Nuclear magnetic resonance (NMR) spectroscopy is a powerful approach to study the structure and dynamics of macromolecules in a *close-to-native* solution environment. In the present thesis I present my investigation of protein and nucleic acid structure and dynamics in a wide variety of biological systems using NMR spectroscopy. The center of attention of the Sykes laboratory for the last 35 years has been the role of the Troponin complex in the regulation of muscle contraction. Accordingly, the main focus of this thesis is the study of this important nano-machine, and how its structure and dynamics regulate its biological function. In Chapter II, the perturbation of Troponin C's structure and dynamics by the attachment of two different bifunctional rhodamine probes is investigated. In Chapter III, the dynamics and position of the bifunctional rhodamine probe when attached on the C helix of Troponin C is studied. In Chapter IV, the structure and dynamics of tryptophan mutants of Troponin C is reported. In Chapter V, the effect of the co-solvent trifluoroethanol on the tryptophan side chain position of mutant F77W of the N-domain of Troponin C is examined. In the following chapter, Chapter VI, the structure and dynamics of a Troponin C – Troponin I chimera is studied using NMR spectroscopy and molecular dynamics simulations to assess the presence or absence of an intrinsically disordered region in Troponin I, and to assess the validity of the fly-casting mechanism proposed to regulate muscle contraction. In Chapter VII and VIII, a different topic is introduced. The structural changes occurring during the denaturation process of the bovine prion protein are monitored using NMR

spectroscopy to gain insights into the protein misfolding process that causes diseases. In Chapter IX, the structural impact of inserting nucleoside phosphonates into DNA are examined by reporting the NMR structure of a DNA dodecamer duplex containing the modified nucleoside Cidofovir at position 7.

ACKNOWLEDGMENTS

I want to start this long list of acknowledgments by thanking my supervisor Brian Sykes. Brian, you are a great mentor and a source of inspiration for the rest of my life. Everyone who knows me knows how difficult I can be at times. Thank you for your legendary patience, and all the great memories that include many golf games, curling seasons, several conferences all around the world, for being present at my wedding, and so many memorable lab events. I can't thank you enough for the last six years of my life spent in your lab. I will miss you. I thank both you and Nancy for being a part of these great memories I have of Edmonton. You won't be forgotten.

Thank you to the current and past members of the Sykes lab: Olga Baryshnikova, Ryan Hoffman (thank you for your friendship and for teaching me English!), Ian Robertson (thanks for supporting/enduring me all these years), Stacey Reinke (I am not sure why... haha you'd think I'd give you a break here ☺), Brian Lee, Gump Zhao, Marta Oleszczuk, Tharin Blumenschein, Steffen Graether, Jan Rainey, Elia Fong, Robert Boyko (thanks for everything: help with computers, friendship, curling, golf, stock discussions, parties, darts at Ave, etc.), Nick Shaw, Monica Li, Xu Wang, Darrin Lindhout, Erik Saude, and Sandra Pineda. An enormous thank you to Dave Corson, Melissa Rakovszky (go Habs go!) and Angela Thiessen, whom expressed and purified such great proteins during my PhD, that were the foundation of my thesis. Merci à Pascal Mercier pour les six

années de travail collaboratif qui m'ont été vraiment bénéfiques. Je n'oublierai pas toutes ces parties de football non plus. Go Alouettes go! Bonne chance avec les Albertains! Thanks to my precious friends outside the lab: Jonny, JP, Dustin, Q, Rory, Mim (and Maksim), and the entire *muffins* ultimate Frisbee team. I hope we will keep in touch. Love you guys.

Thank you to some of the past and current employees of the Department of Biochemistry for their help; Colleen, Christina and company, and a special thanks to Sue Smith for helping with scholarships and many other things. Thanks to the people at Nanuc: Ryan McKay, Bruce Lix, Chris Skappak and Derryck Webb. Thanks Leo (Spyracopoulos) for all your assistance during my PhD. Your door was always open at any time I needed help, from the first day I met you. I really appreciate it. Thanks to Howard Young for telling me so many times “*Get out of my lab!*” (Go Habs go!), but also for the lunches and scientific discussions. Thanks to David Westaway for being a great committee member, and to Frank Sönnichsen for coming all the way from Germany (in March... sorry about that) for my defense! Merci aussi à Bernard Lemire pour toutes ces inombrables discussions à propos de tout et de rien. Ton support, idées et opinions sont toujours appréciées. Thanks to our collaborators in England: Dr. David Trentham, Dr. John Corrie, Dr. Malcom Irving and Dr. Yin-Bao Sun for the BR labeling project that ended up being such an interesting and successful story. I also thank David Evans and Wendy Magee in the department of MMI for the successful DNA collaboration.

My former laboratory companions back at Université Laval: Pierre-Yves Savard, Prof. Patrick Lagüe, Sébastien Morin, Olivier Fisette, and Leigh Willard. A special thanks to Prof. Stéphane M. Gagné, my undergraduate supervisor who introduced me to linux, NMR and pool, and for sending me to, and letting me stay in Alberta. Hope we will still keep in touch. Thanks to Dr. Pierre Lavigne at the Université de Sherbrooke for his continuous support, it was a pleasure to see you in Australia again.

Merci à mes parents pour leur appuie durant toutes ces années passées si loin de la maison. Je vous aime fort. Finally, thank you Sue-Ann for your love, support and patience. Living with me is not an easy task, I know, but you still love me for who I am. You are my best friend and I love you.

Finally, an incredible thank you to the different funding agencies that directly or indirectly supported my research: le Fonds de la recherche en santé du Québec (FRSQ), the Canadian Institutes of Health Research (CIHR) for a Frederick Banting and Charles Best Canada Graduate Scholarship, the Alberta Heritage Foundation for Medical Research (AHFMR) for a studentship and the Lionel E. McLeod scholarship, and the University of Alberta for many awards. Thanks to the Alberta Prion Research Institute, funded by Alberta Ingenuity, and PrioNet Canada for their financial support as well.

TABLE OF CONTENTS

CHAPTER I - Introduction	1
X-ray Crystallography – The Structure	2
Advantages	2
Disadvantages.....	3
NMR Spectroscopy – An Ensemble of Structures	4
Advantages	4
Disadvantages.....	5
NMR studies of biological molecules in the 21st century	7
FIGURES.....	9
REFERENCES	16
CHAPTER II – Two bifunctional rhodamine probes	19
OVERVIEW.....	19
INTRODUCTION	20
MATERIAL AND METHODS	22
Chemicals and biochemicals	22
Labeling of E56C, E63C-mutant sTnC with BSR-I ₂	23
Labeling of E56C, E63C, C101A-mutant sTnC with BSR-I ₂	24
Competitive C helix labeling with BR-I ₂ and BSR-I ₂	25
NMR sample preparation	25
NMR spectroscopy	25
Calcium Titration	26
Muscle fibers and solutions.....	27
TnC extraction and reconstitution	28
Fluorescence polarization and physiological measurements.....	28
RESULTS	31
Diastereoisomers of labeled proteins	31
Relative reactivity of BR-I ₂ and BSR-I ₂	31
Backbone chemical shift comparison between sTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃ and sTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BR ₅₆₋₆₃	31
Transverse relaxation measurement of sTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃	33
Ca ²⁺ -regulation of isometric force in muscle fibers	34

Orientation of the rhodamine dipoles on TnC in muscle fibers	35
Calcium Titration	38
DISCUSSION	39
FIGURES AND TABLES	45
REFERENCES	55
CHAPTER III – Bifunctional rhodamine dynamics	59
OVERVIEW	59
INTRODUCTION	60
THEORETICAL ASPECTS	62
MATERIALS AND METHODS	65
RESULTS	67
DISCUSSION	70
CONCLUSION	73
FIGURES AND TABLES	74
REFERENCES	81
CHAPTER IV – Trp mutants of TnC	85
OVERVIEW	85
INTRODUCTION	86
MATERIALS AND METHODS	88
Protein and Peptide Nomenclature	88
Protein Expression and Purification of TnC and cNTnC Mutants	89
Calcium Titrations	89
cSp Titrations	89
¹⁵ N-Relaxation Measurements	90
Data acquisition, Chemical Shift Assignments and NMR restraints	90
Structure Calculation	91
Muscle fibers and solutions	91
TnC extraction and reconstitution	92
RESULTS	92
Dimerization of the N-domain of Troponin C	93
Calcium and cSp affinities of F77W mutants	94
Characterization of the dynamics of F77W-V82A-cNTnC•Ca ²⁺	95
Structure of F77W-V82A-cNTnC•Ca ²⁺ in TFE	95

Regulation of active force by Trp mutants in skeletal muscle fibers	96
DISCUSSION	97
Ca ²⁺ and cSp affinity measurements	98
Structure of F77W-V82A-cNTnC•Ca ²⁺ and Trp orientation	98
Structural Consequences	99
Force Recovery of different Phe-to-Trp mutants	100
FIGURES AND TABLES	102
REFERENCES	109
CHAPTER V – TFE	114
OVERVIEW	114
INTRODUCTION	115
MATERIAL AND METHODS	118
Protein Expression and Purification	118
NMR sample preparation	119
NMR data acquisition and processing.....	120
NMR Structure determination of F77W-cNTnC•cTnI ₁₄₄₋₁₆₃	121
RESULTS	121
DISCUSSION	125
FIGURES AND TABLES	128
CHAPTER VI – Chimera	141
OVERVIEW	141
INTRODUCTION	142
MATERIALS AND METHODS	145
Sample preparation.....	145
NMR spectroscopy	145
Backbone amide relaxation data.....	146
Computational methods.....	147
RESULTS	149
Chemical Shift Assignment.....	149
NMR relaxation data	150
Molecular dynamics simulations.....	151
DISCUSSION	152
CONCLUSION	154

FIGURES.....	156
Chapter VII – bPrP^C.....	167
OVERVIEW.....	167
INTRODUCTION.....	168
MATERIALS AND METHODS.....	171
RESULTS.....	173
DISCUSSION.....	175
FIGURES.....	177
REFERENCES.....	180
CHAPTER VIII – bPrP^C.....	183
OVERVIEW.....	183
INTRODUCTION.....	184
MATERIALS AND METHODS.....	186
Protein purification.....	186
NMR samples.....	187
NMR spectroscopy.....	187
Data processing.....	188
Data analysis.....	188
RESULTS.....	189
DISCUSSION.....	195
FIGURES AND TABLES.....	200
REFERENCES.....	209
CHAPTER IX – Cidofovir DNA duplex.....	214
OVERVIEW.....	214
INTRODUCTION.....	215
EXPERIMENTAL SECTION.....	217
General Synthetic Chemistry.....	217
Synthesis of the CDV Monomer.....	218
Oligonucleotide Synthesis.....	219
CD Spectroscopy.....	219
Structure Calculations.....	221
RESULTS.....	222
Synthesis of a CDV-Containing Oligonucleotide.....	222

CD Spectroscopy of the CDV DNA Duplex.....	223
NMR Spectra.....	224
Exchangeable Imino Protons.....	224
Temperature Series of the CDV and Control DNA Duplexes	225
Chemical Shift Assignments of the CDV and Incorporation Into DNA	225
Sequential Assignment of the Non-exchangeable Protons in the Control and CDV DNA Duplexes.....	226
Solution Structures of the Control and CDV DNA Duplexes.....	227
DISCUSSION.....	228
FIGURES.....	233
REFERENCES	243
CHAPTER X - Conclusions	249
GENERAL DISCUSSION AND CONCLUSIONS	249
REFERENCES	255
Appendix A – Supporting Information for Chapter II	257
MATERIAL AND METHODS	257
RESULTS	257
REFERENCES	259
Appendix B – Supporting Information for Chapter III	260
MATERIAL AND METHODS	260
¹⁵ N-Transverse Relaxation vs Concentration of F77W-V82A-cNTnC	260
RESULTS	261
Dimerization of F77W-V82A-cNTnC•Ca ²⁺	261
DISCUSSION.....	268
REFERENCES	269
Appendix C – Supporting Information for Chapter VIII.....	270
Lineshape Analysis of the 1D ¹ H NMR spectra.....	270
NOESY Building Curve Analysis.....	275
REFERENCES	279
Appendix D – Supporting Information for Chapter IX.....	280
Primer Extension Analysis	280
FIGURES.....	281

REFERENCES	285
Appendix E– Troponin Complex.....	286
MATERIAL AND METHODS	286
NMR sample preparation: sample TnI-1	286
NMR sample preparation: sample TnI-2.....	286
NMR sample preparation: sample TnI-3.....	286
NMR sample preparation: sample TnI-7.....	286
RESULTS	287
Sample TnI-7.....	289
CONCLUSION	291

LIST OF TABLES

Table II-1. Fiber force and fluorescence polarization measurements.....	54
Table III-1. NMR Parameters Used to Acquire the Different ^{13}C Relaxation Experiments at 500, 600 and 800 MHz	79
Table III-2. Longitudinal ^{13}C Methyl Relaxation Rates and $\{^1\text{H}\}$ - ^{13}C NOE.....	80
Table IV-1. Structural Statistics of F77W-V82A-cNTnC•Ca $^{2+}$ (30 structures) .	107
Table IV-2. Active force at pCa 4.5 after reconstitution with cTnC (% of the maximal force observed prior to extraction).....	108
Table V-1. Structural statistics for F77W-cNTnC in the calcium saturated F77W-cNTnC•cTnI $_{144-163}$ complex for the 20 best target-function structures.....	133
Table V-2. Comparison of the W77 aromatic protons inter-residual NOEs observed in TFE and H $_2$ O	134
Table VIII-1. Thermodynamic properties of different residues of recombinant bPrP $_{121-230}$ as measured by 1D ^1H NMR.....	208

LIST OF FIGURES

Fig. I-1. X-ray diagram showing the helical properties of sodium deoxyribose nucleate and structural model of the double helix of DNA	9
Fig. I-2. First protein structures of myoglobin and haemoglobin	10
Fig. I-3. Schematic representation of the hanging drop method	11
Fig. I-4. The <i>H. marismortui</i> large ribosomal subunit in the rotated crown view	12
Fig. I-5. First protein structure determined by NMR spectroscopy	13
Fig. I-6. Representation of the steps involved during the process of characterization of a protein structure	14
Fig. I-7. Schematic representation of the work presented in this thesis	15
Fig. II-1. Structures of the bifunctional rhodamines	45
Fig. II-2. Electrospray mass spectrum of the unfractionated E56C, E63C, C101A sTnC mutant after competitive labeling with a mixture of BR-I ₂ and BSR-I ₂	46
Fig. II-3. Selected region of the 2D ¹ H- ¹⁵ N HSQC NMR spectrum of sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃	47
Fig. II-4. Difference in ¹⁵ N, ¹ HN, ¹³ Cα and ¹³ Cβ chemical shifts between sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃ and sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BR ₅₆₋₆₃	48
Fig. II-5. Secondary structure prediction of sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃ ...	49
Fig. II-6. Backbone amide ¹⁵ N-R ₂ NMR relaxation rates for each residue of sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BSR ₅₆₋₆₃	50
Fig. II-7. The Ca ²⁺ dependence of force in control fibers	51
Fig. II-8. Ca ²⁺ dependence of dipole order parameters <P ₂ > and <P ₄ >	52
Fig. II-9. Distributions of angles θ between the bifunctional rhodamine dipole and the fiber axis in relaxed muscle	53
Fig. III-1. Structure of the molecular complex sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BR ₅₆₋₆₃ .	74
Fig. III-2. NMR spectra for sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BR ₅₆₋₆₃	75
Fig. III-3. ¹³ C-R ₂ values calculated from off-resonance R _{1ρ} measurements of the BR probe as a function of the effective field (ω _{eff} ²)	76

Fig. III-4. Contour plot showing calculated R_1 (---), R_2 (—) and NOE (•••) as a function of S_s^2 and τ_f	77
Fig. III-5. Main chain amide hydrogen exchange for sNTnC•2Ca ²⁺ •TnI ₁₁₅₋₁₃₁ •BR ₅₆₋₆₃	78
Fig. IV-1. 2D ¹ H- ¹⁵ N HSQC NMR spectra at 600 MHz of F77W-cNTnC•Ca ²⁺ and F77W-V82A-cNTnC•Ca ²⁺	102
Fig. IV-2. cSp titrations of F77W and F77W-V82A-cNTnC•Ca ²⁺ performed at a low concentration of protein (~0.1 mM).....	103
Fig. IV-3. ¹⁵ N-Relaxation of F77W-V82A-cNTnC•Ca ²⁺ at ~0.2 mM.....	104
Fig. IV-4. Comparison of the Trp orientation of F77W-V82A-cNTnC•Ca ²⁺ with cNTnC, silver hake parvalbumin and F153W-cCTnC	105
Fig. IV-5. A) Superposition of F77W-V82A-cNTnC•Ca ²⁺ and cNTnC•Ca ²⁺ •cSp structures both determined by NMR.....	106
Fig. V-1. Comparison of the side chain orientation of W77 for two Phe-to-Trp mutants of cNTnC.....	128
Fig. V-2. W77 CD1 planes of 3D ¹³ C NOESYHSQCs acquired at 600 MHz ...	129
Fig. V-3. Titration of F77W-cNTnC with the co-solvent TFE monitored by ¹³ C-HSQC on a 500 MHz NMR spectrometer.....	130
Fig. V-4. TFE titration of F77(5fW)-cNTnC from 0 to 25% TFE monitored by 1D ¹⁹ F NMR spectra.	131
Fig. V-5. Tryptophan side chain χ_1 x χ_2 conformations found in various calcium binding proteins	132
Fig. VI-1. Four plausible models for the skeletal TnC-TnI chimera based on different TnI conformations.....	156
Fig. VI-2. 1D NMR spectra of the skeletal TnC-TnI chimera acquired at different temperatures.....	157
Fig. VI-3. ¹ H- ¹⁵ N-HSQC NMR spectrum of the skeletal TnC-TnI chimera	158
Fig. VI-4. Chemical shifts of the skeletal TnC-TnI chimera	159
Fig. VI-5. Backbone amide NMR relaxation parameters ¹⁵ N-T ₁ , ¹⁵ N-T ₂ and { ¹ H}- ¹⁵ N-NOE of the skeletal TnC-TnI chimera.....	160

Fig. VI-6. Correlation functions extracted from molecular dynamic simulations	161
Fig. VI-7. Comparison of the experimental NMR relaxation parameters $^{15}\text{N-T}_1$, $^{15}\text{N-T}_2$, $\{^1\text{H}\}$ - ^{15}N -NOE with the one calculated from the MD trajectories.	162
Fig. VII-1. Superimposition of 28 three-dimensional structures of the C-terminus of PrP ^C	177
Fig. VII-2. Aromatic region of 1D ^1H NMR spectra of bPrP ₍₁₂₁₋₂₃₀₎ in D ₂ O acquired at 600 MHz as a function of [urea]	178
Fig. VII-3. 2D ^1H - ^1H NOESY NMR spectrum of bPrP ₍₁₂₁₋₂₃₁₎ acquired at 800 MHz	179
Fig. VIII-1. Titration from 0 to 10 M urea of recombinant bPrP ₁₂₁₋₂₃₀ monitored by 1D ^1H NMR recorded at 800 MHz	200
Fig. VIII-2. Peak area of the eight resonances identified in Fig. VIII-1, plotted as a function of the urea concentration	201
Fig. VIII-3. Fraction of the native conformation of different residues of bPrP ₁₂₁₋₂₃₀ as a function of the urea concentration and their location in bPrP ^C	202
Fig. VIII-4. Four 2D ^1H - ^1H NOESY spectra of bPrP ^C acquired at 800 MHz	203
Fig. VIII-5. Inter-residue NOEs found in the β 1-strand region of bPrP ^C	204
Fig. VIII-6. Inter-residue NOEs found in the β 2-strand region of bPrP ^C	205
Fig. VIII-7. Inter-residue NOEs found in the α 1-helix region of bPrP ^C	206
Fig. VIII-8. Inter-residue NOEs of bPrP ^C identified at 8 M of urea.....	207
Scheme IX-1. Synthesis of the monomer used to prepare a CDV-containing oligonucleotide.....	233
Fig. IX-1. Structures of deoxycytidine monophosphate and CDV, sequences of the control and CDV DNA duplexes and their respective 1D ^1H NMR spectra of for the imino proton regions	234
Fig. IX-2. Melting profiles for CDV- and dCMP-containing duplexes.....	235
Fig. IX-3. 2D ^1H - ^1H NOESY spectrum of the imino proton regions of the CDV DNA duplex at 10°C.....	236
Fig. IX-4. Temperature series of the control and CDV DNA duplexes monitored by 1D ^1H NMR spectra	237

Fig. IX-5. Natural abundance ^1H - ^{13}C HSQC and DQF-COSY for the CDV drug, CDV DNA duplex and control DNA duplex	238
Fig. IX-6. 2D ^1H - ^1H NOESY NMR spectra showing the sequential assignment of the base H6/H8 protons with the H1'-ribose protons for the control and CDV DNA duplexes.....	239
Fig. IX-7. NMR structures of the control and CDV DNA duplexes	240
Fig. IX-8. Comparison of the helical parameters for the ensemble of 10 structures obtained for the control and CDV DNA duplexes.....	241
Fig. IX-9. Structural comparison of C7 and X7 in the middle of the DNA duplexes	242

ABBREVIATIONS

NMR	nuclear magnetic resonance
MD	Molecular dynamic
CD	Circular Dichroism
TnC	Troponin C
TnI	Toponin I
TnT	Toponin T
cCTnC	C-domain of human cardiac TnC (sCTnC for skeletal)
cNTnC	N-domain of human cardiac TnC (sNTnC for skeletal)
cSp or cTnI ₁₄₇₋₁₆₃	Synthetic switch peptide (residues 147-163) of human cardiac TnI
sTnI ₁₁₅₋₁₃₁	synthetic ‘switch’ peptide (residues 115-131) of skeletal TnI
BR	bifunctional carborhodamine
BSR	bifunctional sulforhodamine
TSE	transmissible spongiform encephalopathies
PrP ^C	cellular, proteinase K-sensitive prion protein isoform
PrP ^{Sc}	disease-associated, proteinase K-resistant prion protein isoform
HMQC	heteronuclear multi-quantum coherence.
HSQC	heteronuclear single-quantum coherence
TOCSY	total correlation spectroscopy
NOE	nuclear Overhauser effect
T_1	longitudinal relaxation time
T_2	transverse relaxation time
R_1	longitudinal relaxation rate (= $1/T_1$)
R_2	transverse relaxation rate (= $1/T_2$)
RMSD	root-mean-square standard deviation
s^2	order parameter (square)
$\langle P_{2d} \rangle$	order parameter from fluorescence measurements (not square)
S_f^2	order parameter for methyl rotation

S_s^2	order parameter for motion of the ^{13}C - ^{14}N bond
τ_f	correlation time for rotation of the methyl group
τ_s	internal correlation time for the ^{13}C - ^{14}N bond
τ_m	global rotational correlation time
CSI	Chemical Shift Index
FID	Free Induction Decay
CSA	Chemical Shift Anisotropy
K_d	dissociation constant
K_{dimer}	dissociation constant for the formation of dimer
DSS	2,2-dimethyl-2-silapentane-5-sulfonic acid.
DTT	threo-1,4-dimercapto-2, 3-butanediol
TFA	Trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
4fW	4-fluorotryptophan
5fW	5-fluorotryptophan

CHAPTER I - Introduction*After Sixty Years of Structural Biology, Where Are We?*

The last sixty years have been incredible; full of mind-blowing discoveries in the field of structural biology. Historically, X-ray crystallography evolved first and became the method of choice to determine the structure of molecules at the atomic level. It all started with the determination of the first protein sequence for chain B of insulin (1), the identification of α -helices and β -sheets as secondary structures in proteins (2,3) and the report of the DNA double helix (4-6). Shortly thereafter, the first three-dimensional protein structures for myoglobin and haemoglobin were determined (7,8), followed by the structure of lysozyme a few years later, the first three-dimensional structure of an enzyme (9,10). One could consider these discoveries as milestones in structural biology research.

Today, a multitude of methods are available to study the structural properties of macromolecules. A variety of spectroscopies exploit the properties of light at different wavelengths – Circular Dichroism (CD), infra-red, UV-visible and Raman spectroscopy – while others take advantage of the nuclear magnetic moments of elements, like Nuclear Magnetic Resonance (NMR) spectroscopy and Magnetic Resonance Imaging (MRI). Some techniques use the diffraction of electromagnetic radiation when applied to crystallised or solution compounds, like X-ray crystallography and small-angle X-ray scattering (SAXS). Other methods use a beam of electron to visualize objects at high-resolution, like electron microscopy (EM). Each of these techniques has advantages and disadvantages, but thus far, only X-ray crystallography and NMR spectroscopy has been routinely used to determine the complete structures of molecules at the atomic level. Although the main focus of this thesis is the study of biological macromolecules using NMR spectroscopy, it is always good to present a yin and yang comparison with its partner technique, X-ray crystallography.

X-ray Crystallography – The Structure

Since the determination of the DNA double helix structure by Franklin, Watson, Wilkins and co-workers (4-6) (Fig. I-1), X-ray crystallography has been the method of choice to study the structure of biological macromolecules. It took only a few more years for the determination of the first protein structures by X-ray diffraction methods to make their appearance, when Perutz and Kendrew solved the structures of hemoglobin and myoglobin, respectively (7,8) (Fig. I-2). These great breakthroughs were made possible, in large part, by the work of William and Lawrence Bragg, father and son, leading to the formulation of the Bragg's law ($n\lambda = 2d\sin\theta$), and the development of isomorphous replacement by John M. Robertson as the method of choice to solve the phase problem in diffraction patterns. To this day, more than 61,000 deposited structures in the PDB have been determined using X-ray crystallography (11).

The principles behind the method consist of first crystallising the compound of interest. In biology, the hanging drop technique is the most common method used to obtain protein crystals (see Fig. I-3). The crystal is then mounted on a goniometer and gradually rotated while positioned in the path of an X-ray beam. The arrangement of the atoms in the crystal scatters the X-ray waves, and the resulting reflections are detected for each of the different crystal orientations. The diffraction patterns are then converted into electronic density using a Fourier transform once the initial phases have been determined (using molecular replacement or multiple isomorphous replacement for example). Positioning each atom into their respective electronic density produces a three-dimensional model of the crystallized compound, often referred to as “The Structure”.

Advantages

The main advantage of using X-ray crystallography is the possibility to determine the structure of chemical compounds at a sub-atomic resolution (i.e. approximately 1 Å). Also, size limitation is not really an issue compared to NMR.

X-ray crystallography permits the study of large multimeric complexes composed of proteins, nucleic acids and other molecules. A good example of this is the recent determination of the structure of the ribosome at high-resolution (12-14) (Fig. I-4). Another advantage of this method is that once the structure of a macromolecule is known, it is relatively easy and simple to study the effect of various mutations and the effect of ligands binding to the structure by using molecular replacement.

Disadvantages

With great power comes great limitations; the first limitation of X-ray crystallography is the difficulty in crystallising a new molecule. This problem often leads to the use of non-biological crystallisation conditions, the consequences of which are too often disregarded. Another difficulty is that obtaining a pure solution at fairly high concentration is required (e.g. 2-50 mg/ml of protein). Once crystals are finally obtained, there is no guarantee that they will diffract at high resolution. And even after a structure is obtained, one has to be careful of the presence of crystal packing artefacts that could affect the structure. Another important limitation of the method is that protons are not observable because they have only one electron.

Since the crystal is a solid form, one cannot study the behaviour of the molecule of interest in a solution environment. Structure determination by X-ray crystallography has brought us to the perception that protein and DNA structures are rigid entities. The term “The Structure” is too often used to describe biological molecules, despite the recognized dynamics properties of macromolecules. Another major problem of this technique is the absence of electronic density for flexible regions, making the study of intrinsically disordered regions (IDR) in proteins practically impossible.

NMR Spectroscopy – An Ensemble of Structures

The first *de novo* protein structure determination solved by NMR spectroscopy was published in 1985, when Kurt Wüthrich's group reported an ensemble of 5 structures (PDB ID 1BUS) for the proteinase inhibitor IIA (15) (Fig. I-5). This achievement was made possible by new technological advances in the NMR field; (i) the use of high-field NMR spectrometers (i.e. 300 MHz at the time), (ii) the development of Fourier transform NMR spectroscopy allowing the acquisition of 2D COSY and 2D TOCSY ^1H NMR experiments for chemical shifts sequential assignment and 2D NOESY ^1H NMR experiments for the measurement of NOE distances and (iii) the use of computational tools for structural interpretation of the NMR data and structure calculations. As of today, more than 9,000 NMR structures have been deposited in the Protein Data Bank (11).

Advantages

NMR spectroscopy has numerous advantages over X-ray crystallography. First, biological molecules can be studied in solution. This allows biological molecules to be studied in a *close to native* environment (i.e. pH, temperature, ionic strength, etc.). Like X-ray crystallography, NMR spectroscopy can be used to determine the three-dimensional structure of macromolecules. However, instead of trying to find a single structure that fits best the electronic densities (X-ray), a NMR ensemble of 5 to 50 structures that fit the experimental data is reported. The differences between the structures in the ensemble reflect the fact that biological molecules like proteins and nucleic acids are dynamic molecules (although structural heterogeneity can sometimes suggest a lack of experimental data).

Other than structure determination, multiple NMR experiments can be performed in order to obtain information on the structure and dynamics of a molecule. The measurement of backbone chemical shifts can provide information about the presence of secondary structures in proteins using the Chemical Shift Index (16). The measurement of coupling constants can be used to measure dihedral angles.

For example, the value of $^3J_{\text{HN-H}\alpha}$ in proteins is directly correlated with the Phi angle as described by the Karplus equation (17,18). In a NOESY experiment (Nuclear Overhauser Effect Spectroscopy), the peak intensities in the spectrum are proportional to ^1H - ^1H distances and can provide important information about protein folding. The measurement of relaxation parameters (T_1 , T_2 , NOE) can provide information about internal and external motions in molecules, and can be used to calculate the order parameters (S^2) in a given protein on a per-residue basis. Domain motions of biological molecules can be investigated using relaxation data. Protein motions and rigidity can also be measured by hydrogen/deuterium exchange experiments. The possibility to study protein dynamics is probably the most important advantage that NMR spectroscopy has over any other method. For example, the work of Lewis Kay and his group on the dynamics of the proteasome and the characterisation of low-populated protein intermediates is at the frontier of proteins dynamics studied by NMR (19,20). Since, the molecules of interest are free to move in solution and interact with each other and with other compounds, this allows quantitative measurement of protein-protein interaction (e.g. Troponin C and Troponin I) (21), measurement of dissociation constants of small molecules (e.g. Troponin C and drugs) (22) and pKa measurement of exchangeable groups (e.g. Imidazole) (23). All together, these possibilities make NMR spectroscopy a powerful approach to study biological molecules and can reveal much more than a single structure can offer.

Disadvantages

Like any other method, NMR spectroscopy is not without flaws. Like X-ray crystallography, NMR is not a sensitive technique and requires a high concentration (typically 1-10 mg/ml) of the compound of interest. Accordingly, solubility is often an issue. The size of the molecule is also a limitation; the bigger the molecule is, the slower it tumbles in solution, the shorter the relaxation times, and the lower the NMR signal becomes. Large molecules also increase the complexity of the NMR spectra and increase the number of peak overlaps. Depending on the complexity of the system under study, one often needs

uniformly ^{13}C , ^{15}N and/or ^2H labeled proteins. Due to the long time required to acquire multidimensional NMR experiments (days to sometimes weeks) and the multiple number of spectra required for sequential assignment and structure determination, NMR samples usually need to be stable for several consecutive days.

In addition to the several days of data acquisition, several days (if not weeks and months) of computational analysis are required once the NMR data is acquired. To make things worse, no single, uniform, agreed upon protocol exists to determine a protein structure by NMR! This makes the structure determination process painfully hard to learn, but most importantly, makes it very difficult to compare the quality between structures. For every step in the process; data acquisition (i.e. Bruker or Varian), data processing (NMRPipe, Biospin, VnmrJ), chemical shift assignment (NMRView, NMRViewJ, Sparky), atom nomenclature used, NOE calibration (cyana, xplor-nih, aria, MARDIGRAS, Olivia), structure calculation (cyana, xplor-nih, CNS, aria, amber, charmm), and structure validation (procheck, PSVS, CORMA, whatif), no one seems to agree on a single protocol and a single set of software to use. An example of this is presented in (Fig. I-6). This is without mentioning the thousands of in-house scripts that are being used by NMR spectroscopists in the field. All of these methods are being used differently from one location to another (i.e. universities), but most impressively, differently between members of a given laboratory!!!

What criteria should be used to validate NMR ensembles? What criteria should be used to evaluate the accuracy and precision of NMR structures? Is a low RMSD good, or bad? On one hand, if a NMR ensemble has a high RMSD, one can argue that the ensemble is not well defined, but if the RMSD is too low, one could argue that the ensemble is over-defined. What is the threshold for a too high, or too low RMSD? How many restraints violations are acceptable? In the end, one can only rely on a trial and error approach and some common sense.

To summarize, X-ray crystallography and NMR spectroscopy are two powerful methods with advantages and disadvantages. These two techniques possess complimentary strength and weakness, and one can only benefit from using both in order to study biological molecules.

NMR studies of biological molecules in the 21st century

In the present thesis, the investigation of protein and nucleic acid structure and dynamics in quite a variety of systems using NMR spectroscopy is the primary focus. Each chapter (II-VIII) contains an article that has been published in a peer-reviewed journal. Because the systems are quite different I have left the papers intact, including the introductions since they each develop an individual line of research. The reference for the journal publication is given at the beginning of each section. Most importantly, the contribution of each author to the work, focusing on my contributions to the often multi-authored papers, is indicated. One could wonder what was the reason for studying so many topics; some may think that I can't stay focused on one task, but I rather think that it shows I have a wide range of interests. That said, the underlying theme between all of these areas of research is the relationship between the structure and dynamics of biomolecules and how it impacts their biological functions.

A schematic summary of the work presented in this thesis is shown in Fig. I-7. In Chapter II, the perturbation of Troponin C's structure and dynamics by the attachment of two different bifunctional rhodamine (BR) probes is investigated, to understand the structural conclusions made when these labels are used for *in situ* biophysical studies. In Chapter III, the dynamics and position of the BR probe when attached on the C helix of Troponin C is studied, to elucidate the role of dynamics when these fluorescent labels are used for *in situ* biophysical studies. In Chapter IV, the structure and dynamics of tryptophan (Trp) mutants of Troponin C is reported, developed as potential tools for *in situ* solid-state NMR studies. In Chapter V, the effect of the co-solvent trifluoroethanol (TFE) on the Trp side

chain position of mutant F77W of the N-domain of Troponin C is examined, since co-solvents are often used in biophysical studies of peptides and proteins. In the following chapter, Chapter VI, the structure and dynamics of a Troponin C – Troponin I chimera is studied using NMR spectroscopy and molecular dynamics (MD) simulations to assess the presence or absence of an intrinsically disordered region in Troponin I, and to assess the validity of the fly-casting mechanism proposed to regulate muscle contraction. In Chapter VII and VIII, the structural changes occurring during the denaturation process of the bovine Prion protein are monitored using NMR spectroscopy to gain insights into the protein misfolding process that causes diseases. In Chapter IX, the structural impact of inserting nucleoside phosphonates into DNA are examined by reporting the NMR structure of a DNA dodecamer duplex containing the modified nucleoside Cidofovir at position 7. In the last chapter, Chapter X, a global conclusion of the thesis is presented to summarize the work presented in this manuscript.

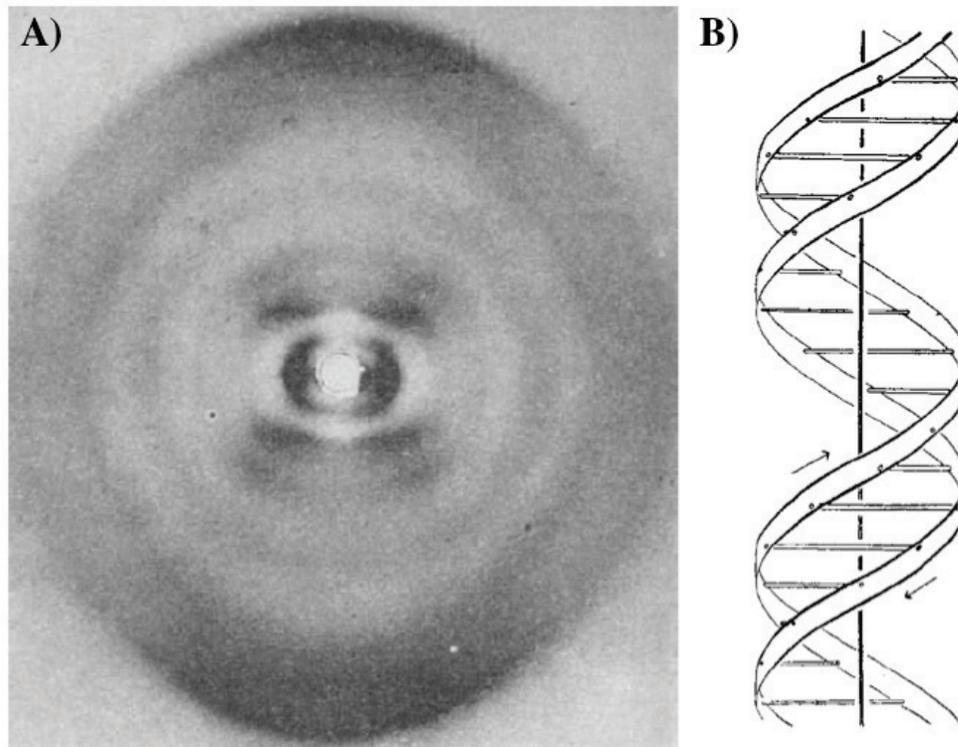
FIGURES

Fig. I-1. A) X-ray diagram showing the helical properties of sodium deoxyribose nucleate and B) the structural model of the double helix of DNA as originally published by Franklin et al. and Watson et al., respectively (4,5).

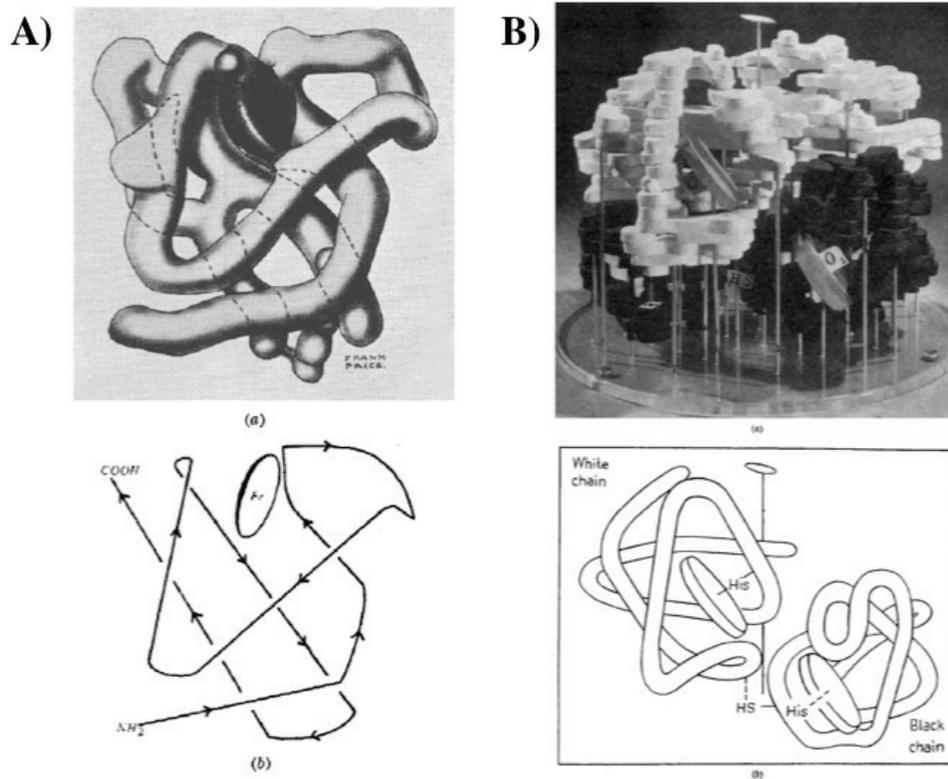


Fig. I-2. First protein structures of A) myoglobin and B) haemoglobin determined by X-ray crystallography as originally published by Kendrew et al. and Perutz et al., respectively (7,8).

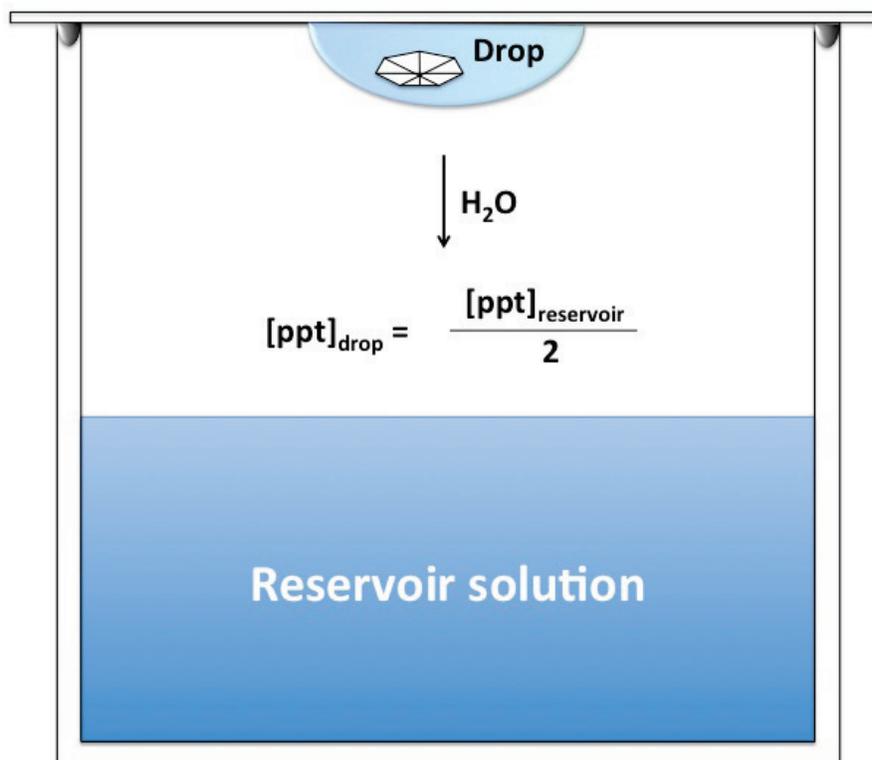


Fig. I-3. Schematic representation of the hanging drop method often used in crystallography, where the precipitant (ppt) concentration in the drop solution is half of the reservoir solution.

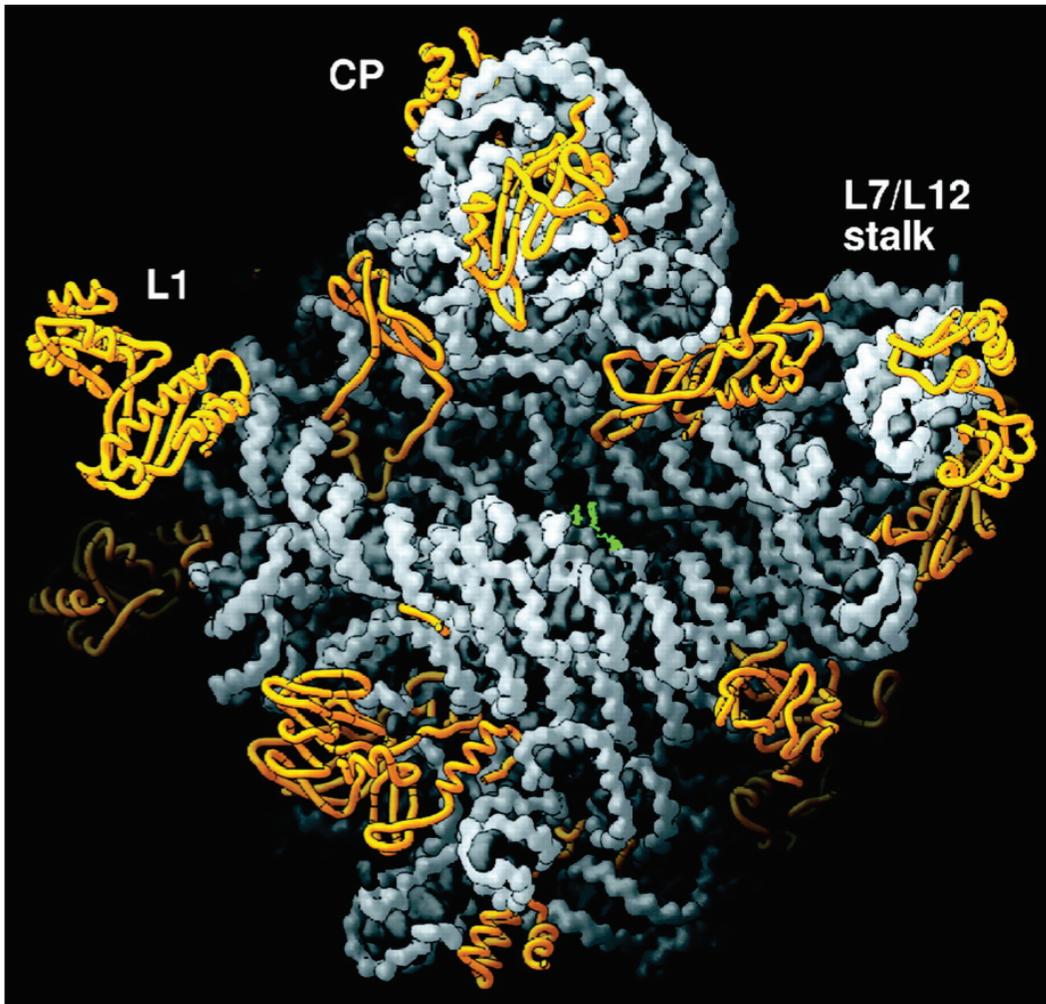


Fig. I-4. The *H. marismortui* large ribosomal subunit in the rotated crown view. RNA is shown in gray in a pseudo-space-filling rendering. The backbones of the proteins visible are rendered in gold. The particle is approximately 250 Å across. As originally published by Ban et al. (12).

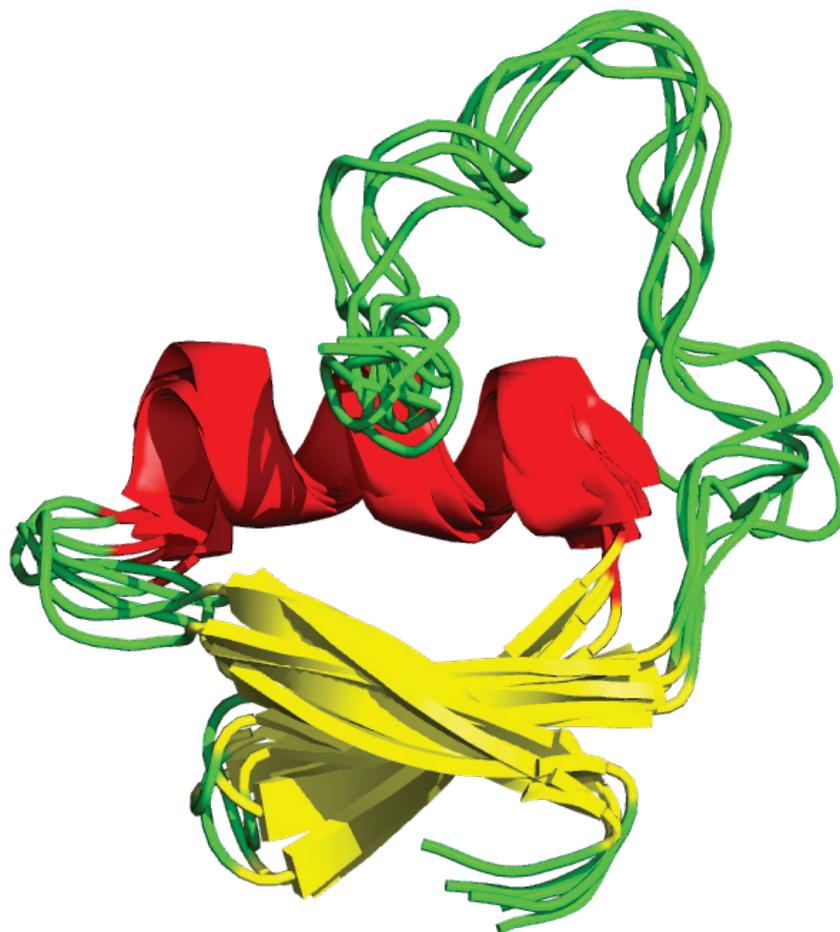


Fig. I-5. First protein structure determined by NMR spectroscopy. The figure shows a cartoon representation of proteinase inhibitor IIA (PDB ID: 1BUS) from bull seminal plasma using ^1H NMR (15). The ensemble of five structures revealed one α -helix colored in red and a small β -sheet colored in yellow.

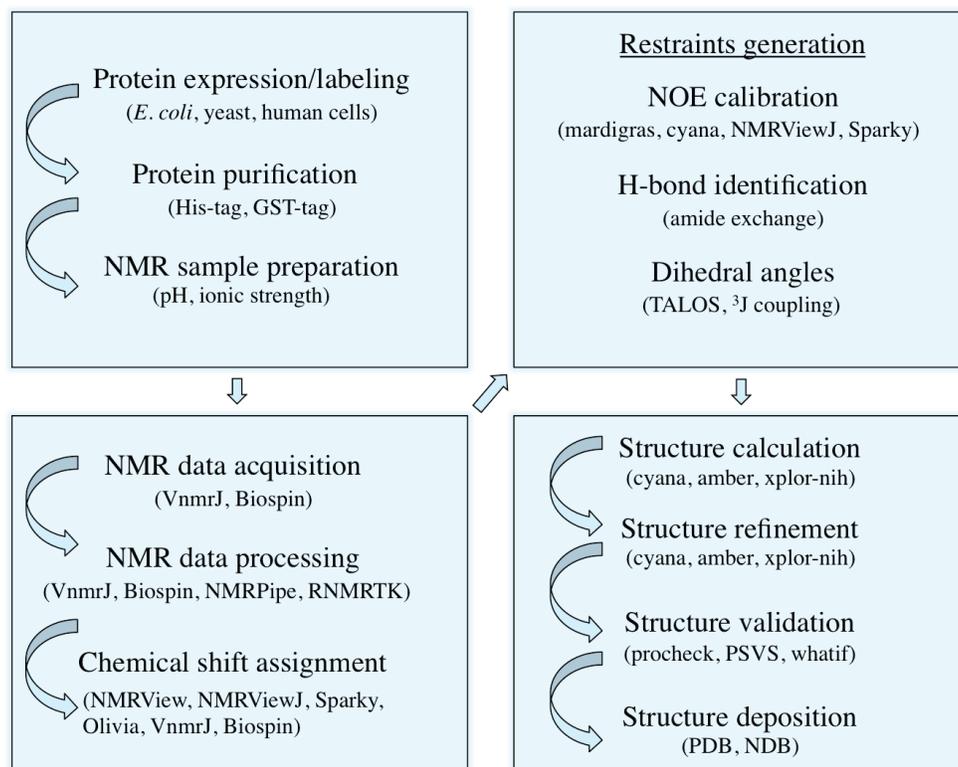


Fig. I-6. Representation of the steps involved in the process of characterizing a protein structure: sample preparation, data acquisition and NMR analysis, restraints generation, and structure calculation and validation.

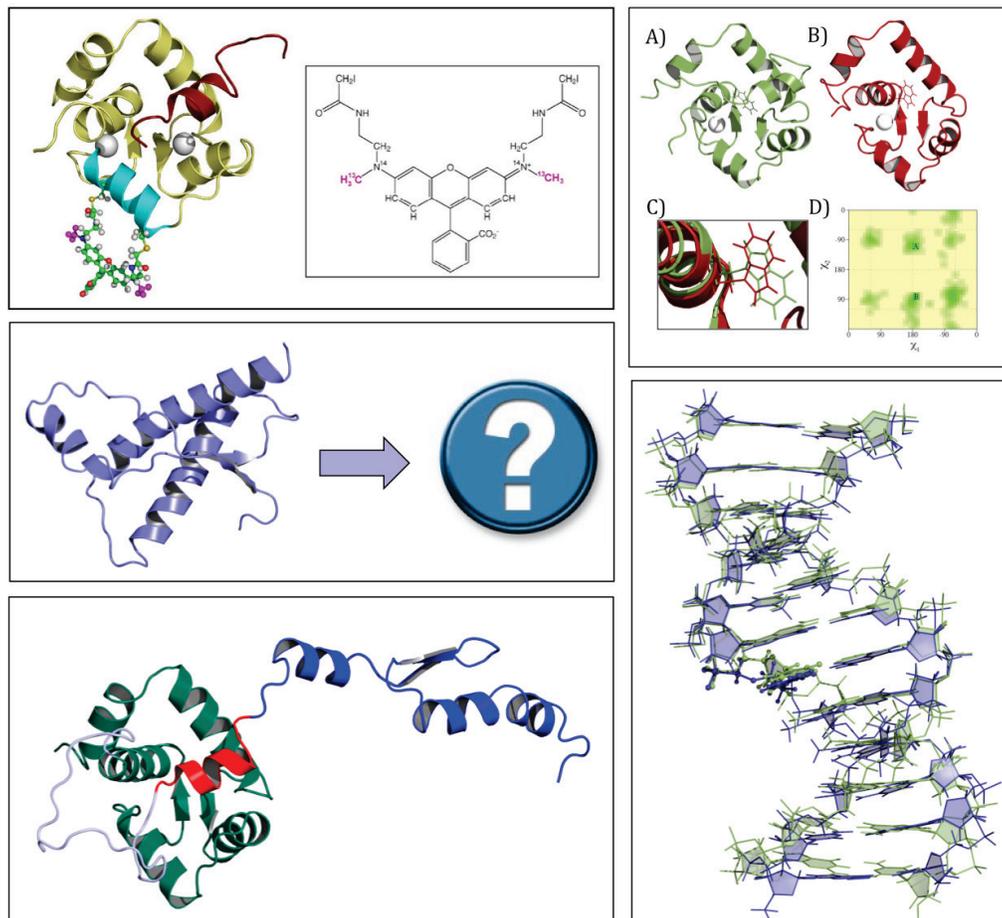


Fig. I-7. Schematic representation of the work presented in this thesis: structure and dynamics of bifunctional rhodamine labeled Troponin-C, Phe-to-Trp mutants of Troponin C, prion protein regional stability and unfolding, structure and dynamics of a Troponin C-Troponin I chimera, and the structure of a dodecamer DNA duplex with the modified base Cidofovir inserted at position 7.

REFERENCES

1. Sanger, F., and Tuppy, H. (1951) The amino-acid sequence in the phenylalanyl chain of insulin. I. The identification of lower peptides from partial hydrolysates, *Biochem. J.* 49, 463-481.
2. Pauling, L., Corey, R., and Branson, H. (1951) The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain, *Proc. Natl. Acad. Sci. U.S.A.* 37, 205-211.
3. Pauling, L., and Corey, R. (1951) The pleated sheet, a new layer configuration of polypeptide chains, *Proc. Natl. Acad. Sci. U.S.A.* 37, 251-256.
4. Watson, J., and Crick, F. (1953) Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid, *Nature* 171, 737-738.
5. Franklin, R., and Gosling, R. (1953) Molecular configuration in sodium thymonucleate, *Nature* 171, 740-741.
6. Wilkins, M., Stokes, A., and Wilson, H. (1953) Molecular structure of deoxypentose nucleic acids, *Nature* 171, 738-740.
7. Perutz, M., Rossmann, M., Cullis, A., Muirhead, H., Will, G., and North, A. (1960) Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis, *Nature* 185, 416-422.
8. Kendrew, J., Dickerson, R., Strandberg, B., Hart, R., Davies, D., Phillips, D., and Shore, V. (1960) Structure of myoglobin: A three-dimensional Fourier synthesis at 2 Å resolution, *Nature* 185, 422-427.
9. Blake, C., Koenig, D., Mair, G., North, A., Phillips, D., and Sarma, V. (1965) Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution, *Nature* 206, 757-761.
10. Johnson, L., and Phillips, D. (1965) Structure of some crystalline lysozyme-inhibitor complexes determined by X-ray analysis at 6 Å resolution, *Nature* 206, 761-763.
11. Protein Data Bank, <http://www.rcsb.org>.

12. Ban, N., Nissen, P., Hansen, J., Moore, P., and Steitz, T. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution, *Science* 289, 905-920.
13. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution, *Cell* 102, 615-623.
14. Wimberly, B., Brodersen, D., Clemons, W. J., Morgan-Warren, R. J., Carter, A., Vornrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature* 407, 327-339.
15. Williamson, M., Havel, T., and Wuthrich, K. (1985) Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ¹H nuclear magnetic resonance and distance geometry, *J. Mol. Biol.* 182, 295-315.
16. Wishart, D., Sykes, B., and Richards, F. (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, *Biochemistry* 31, 1647-1651.
17. Karplus, M. (1959) Contact electron-spin coupling of nuclear magnetic moments, *J. Chem. Phys.* 30, 11-15.
18. Karplus, M. (1963) Vicinal proton coupling in nuclear magnetic resonance, *J. Am. Chem. Soc.* 85, 2870-2871.
19. Ruschak, A., Religa, T., Breuer, S., Witt, S., and Kay, L. (2010) The proteasome antechamber maintains substrates in an unfolded state, *Nature* 467, 868-871.
20. Korzhnev, D., Religa, T., Banachewicz, W., Fersht, A., and Kay, L. (2010) A transient and low-populated protein-folding intermediate at atomic resolution, *Science* 329, 1312-1316.
21. 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, 8289-8298.
22. Wang, X., Li, M., Spyropoulos, L., Beier, N., Chandra, M., Solaro, R., and Sykes, B. (2001) Structure of the C-domain of human cardiac troponin

C in complex with the Ca²⁺ sensitizing drug EMD 57033, *J. Biol. Chem.* 276, 25456-25466.

23. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J. Biomol. NMR* 41, 5-7.

CHAPTER II – Two bifunctional rhodamine probes*Towards Protein Structure in situ: Comparison of Two Bifunctional Rhodamine Adducts of Troponin C*

Olivier Julien,¹ Yin-Biao Sun,² Andrea C. Knowles,³ Birgit D. Brandmeier,³
Robert E. Dale,² David R. Trentham,³ John E. T. Corrie,³ Brian D. Sykes¹ and
Malcolm Irving²

¹ Department of Biochemistry, University of Alberta, Edmonton, Canada.

² Randall Division of Cell and Molecular Biophysics, King's College London, London, United Kingdom.

³ MRC National Institute for Medical Research, London, United Kingdom.

Publication: Julien et al. 2007. *Biophysical Journal*, 93:1008-1020.

Contributions: OJ and YBS contributed equally to this work; OJ and BDS acquired the NMR spectra, OJ assigned the protein chemical shifts, analyzed the relaxation data, performed the calcium titration, and wrote the corresponding portion of the manuscript with BDS. OJ made figures III-3 to III-6, and wrote the supplementary material.

OVERVIEW

As part of a program to develop methods for determining protein structure *in situ*, troponin C from skeletal muscle (sTnC) was labeled with a bifunctional rhodamine (BR or BSR), crosslinking residues 56 and 63 of its C helix. NMR spectroscopy of the N-terminal domain of BSR-labeled sTnC in complex with Ca²⁺ and the troponin I switch peptide (residues 115-131) showed that BSR labeling does not significantly affect the secondary structure of the protein or its dynamics in solution. BR-labeling was previously shown to have no effect on the solution structure of this complex (Mercier et al. 2003. *Biochemistry* 42:4333-

4348). Isometric force generation in isolated demembrated fibers from rabbit psoas muscle into which BR- or BSR-labeled sTnC had been exchanged had reduced Ca^{2+} -sensitivity, and this effect was larger with the BSR label. The orientation of rhodamine dipoles with respect to the fiber axis was determined by polarized fluorescence. The mean orientations of the BR and BSR dipoles were almost identical in relaxed muscle, suggesting that both probes accurately report the orientation of the C helix to which they are attached. The BSR dipole had smaller orientational dispersion, consistent with less flexible linkers between the rhodamine dipole and cysteine-reactive groups.

INTRODUCTION

Fluorescence polarization data from the bifunctional carborhodamine **1** (BR-I₂, Fig. II-1) attached to two suitably-located cysteine residues on muscle proteins (myosin regulatory light chain and troponin C) have been used to study orientations and/or dynamics of the labeled proteins in their native environment in a muscle fiber (1-4). The same rhodamine has been used in single-molecule fluorescence polarization and imaging measurements when attached to one of the calmodulin light chains of myosin V (5,6). Evidence obtained in those studies suggests that the two-site attachment of the BR probe does not alter the native structure or function of the protein. This was supported by an NMR structural study of the N-lobe of the E56C/E63C mutant of chicken skeletal troponin C (sNTnC) labeled with BR-I₂ on cysteine residues at positions 56 and 63 of its C helix, which showed that the labeled protein maintained its native structure (7). In addition, physiological experiments that imply conservation of function are described in several of the above papers. There is thus a body of evidence to suggest that this approach is a reliable and robust way to determine protein and protein domain orientation and motion *in situ*.

BR-I₂ was deliberately synthesized with flexible arms between the reactive iodoacetamido groups and the fluorophore system (8) in the expectation that it

would be able to span cysteines separated over a range of distances. In practice, it has been possible with this reagent to cross-link cysteines located along α -helices, between two helices and between a helix and a random coil, with separations between their paired β -carbon ($C\beta$) atoms in the range 1.0-1.6 nm (4). A second, commercially-available bifunctional sulforhodamine reagent **2** (BSR-I₂, Fig. II-1) has been used in related fluorescence polarization and single-molecule studies on kinesin (9,10). The ring structure of the linkers between the iodoacetamido groups and the fluorophore in BSR-I₂ is more rigid than the open-chain linkers in BR-I₂, and it might be expected that BSR-I₂ would be less able to accommodate different spacings of the two cysteine residues. In practice, the $C\beta$ distances between cysteines linked by the reagent do span a similar range of distances (1.1-1.6 nm) to those spanned by BR-I₂, but in no case yet described have these pairs of linked residues both been located on the same α -helix. Thus, the residues are either between 2 helices (N64C, V71C; 1.1 nm), between the end of a β -sheet and a random coil (169C (a native cysteine), V174C; 1.5 nm), or between two β -sheets connected by a random coil (T330C, V335C; 1.5 nm) (10). Sequence numbers and distances are from the rat kinesin structure (pdb 2kin). The 169-174 $C\beta$ distance in human kinesin (pdb 1bg2) is slightly longer (1.6 nm) (9).

To date there have been no published studies comparing the efficacy of BSR-I₂ and BR-I₂ for determining the *in situ* orientation of the vector joining the same pair of target cysteines. The ideal bifunctional probe reagent for this type of *in situ* structural measurement would have three key properties: 1) efficient cross-linking of a suitably placed pair of cysteine residues in a target protein domain; 2) no alteration of the native structure and function of the target domain as a result of such cross-linking; and 3) accurate reporting of the *in situ* orientation of the vector joining the two cysteines to which the probe is attached. These properties may be to some extent mutually exclusive. For example, flexible linkers between probe and protein may be desirable with respect to the first two properties, but detrimental to the third. Moreover, the third property cannot in general be assessed directly in the absence of alternative techniques that can determine the

orientation of the cysteine-cysteine vector *in situ*. In these circumstances, the most powerful available approach is to compare the orientations of different probes attached to the same pair of cysteines and the structural and functional properties of the different labeled proteins.

In the present work we labeled the C helix of chicken skeletal troponin C (sNTnC) separately with either BR-I₂ and BSR-I₂ via its E56C/E63C mutant, in which the β -carbon atoms of the two cysteines are approximately 1.1 nm apart. We used a kinetic competition experiment to determine the relative chemical reactivity of the two reagents. We studied BSR-labeled sNTnC in complex with Ca²⁺ and the troponin I switch peptide (TnI₁₁₅₋₁₃₁) by NMR for comparison with the published structure of the corresponding BR-labeled complex (7). We labeled the same C helix residues of whole sTnC with either BR-I₂ or BSR-I₂, exchanged the labeled proteins into permeabilized muscle fibers, and determined the effect of each probe on active force generation and its regulation by Ca²⁺. Finally, we measured the *in situ* orientation and mobility of both the BR and BSR probes, using a novel analytic approach to determine these parameters from the polarized fluorescence intensities.

MATERIAL AND METHODS

Chemicals and biochemicals

BR-I₂ was prepared as described previously (8) and BSR-I₂ was purchased from Molecular Probes (Eugene, OR). The double cysteine mutant (E56C, E63C) sNTnC, isotopically labeled with ¹³C and ¹⁵N, was expressed and purified as described previously (7). The isotopic enrichment, estimated from mass spectrometry, was ~95 % for ¹³C and ~93% for ¹⁵N. The full length E56C, E63C, C101A mutant chicken skeletal troponin C, and the same protein labeled with BR-I₂, were prepared as previously described (3). Chemical constituents for the solutions used in the muscle fiber experiments were obtained from Sigma (Poole, Dorset, U.K.) except where noted.

Labeling of E56C, E63C-mutant sNTnC with BSR-I₂

A solution of 8 mg sNTnC at 1 mg/ml in labeling buffer (25 mM Tris/HCl, 100 mM NaCl, 1 mM MgCl₂, pH 7.4) was incubated overnight on ice with 2 mM DTT (final concentration) to reduce disulfides that formed during storage. Reduction was monitored by reverse phase HPLC [(C18 VYDAC column, #218TP54, guard column #218GCC54), eluting at 1 ml/min with a linear gradient of 60% solvent A (H₂O, 0.1% TFA) and 40% solvent B (acetonitrile, 0.082% TFA) to 40% solvent A and 60 % solvent B over 20 min], and was characterized by a change from an initially complex elution profile to a principal species that eluted at 52.3% solvent B. The protein was then gel filtered using PD-10 columns (GE Healthcare, Chalfont St. Giles, U.K.), which were equilibrated with labeling buffer that also contained 52.5 μM tris(carboxyethyl)phosphine. After elution, the protein was diluted to a concentration of 0.55 mg/ml (52.5 μM) using the same buffer. BSR-I₂ was then added (final concentration 105 μM) from a 25 mM stock solution in DMF. The reaction mixture was incubated in the dark at 20 °C for 40 min and the course of the reaction was monitored (after quenching an aliquot with 2-mercaptoethanesulfonate - see below) by analytical reverse-phase HPLC as specified above. Protein elution was monitored by absorbance at 215 nm and by rhodamine fluorescence (λ_{ex} 549 nm, λ_{em} >580 nm). sNTnC in which both cysteines had been cross-linked by BSR (i.e. the desired product) eluted as a double peak at 50.3 and 51.1 % solvent B. Unreacted sNTnC eluted at 52.3 % solvent B, as above. All assignments were made retrospectively, after analysis of fractions by electrospray mass spectrometry. We did not seek to identify material in other fractions, which may have contained species labeled with two BSR moieties. After the 40-min incubation, the labeling reaction was quenched with sodium 2-mercaptoethanesulfonate (final concentration 3.4 mM) and kept in the dark for a further 30 min at 20 °C. Aliquots of the solution were filtered through PD-10 columns (2.5 ml per column) into FPLC buffer (10 mM potassium phosphate, 1 mM MgCl₂, pH 7.5) to remove unconjugated rhodamine. The labeled sNTnC (0.31 mg/ml) was purified at 4 °C on a 16/10 Mono-Q ion

exchange column (GE Healthcare), using a linear gradient of 0.25-0.35 M NaCl in FPLC buffer at a flow rate of 2 ml/min. The protein eluted as two peaks with near-baseline resolution at 0.31 M and 0.32 M NaCl. 1-ml fractions were collected and assayed for purity by analytical HPLC (as above) and electrospray mass spectrometry (measured mass for each peak, 11005.6 ± 1.8 Da). Fractions containing both peaks of the pure sTnC•BSR₅₆₋₆₃ (>90 % homogeneity) were pooled and dialyzed against 10 mM KCl, 0.42 mM CaCl₂ (2 × 5 liter, each for 2 h, then 1 × 3 liter overnight, all at 4°C). The contents of the dialysis bag (2 ml) were then concentrated to 1 ml using a Vivaspin 20 concentrator (Vivascience, Epsom, U.K.) at 1140 × g. The concentrated solution of labeled protein [1.04 mg/ml, based on an extinction coefficient of 52,000 M⁻¹cm⁻¹ at 528 nm (3)] was flash frozen in liquid nitrogen and stored at -20 °C.

Labeling of E56C, E63C, C101A-mutant sTnC with BSR-I₂

Labeling of full-length sTnC with BSR-I₂ was performed by a similar protocol to that previously described for BR-I₂ (3). HPLC analysis was as described above. sTnC•BSR₅₆₋₆₃ eluted as a partly resolved double peak at ~53.0 and 53.3 % acetonitrile. The species in each peak had the same mass (± 1 Da). The labeled protein was purified in 1-mg aliquots on a C4 VYDAC #214TP510 column with a linear gradient from 60% solvent A (H₂O, 0.1% TFA) and 40% solvent B (acetonitrile, 0.082% TFA) to 40% A and 60% B, run at 2 ml/min over 1 h. The two peaks, containing the separate diastereoisomers of the labeled protein and referred to as sTnC•BSR1 and sTnC•BSR2, were collected manually and immediately stored on ice. The combined fractions of each peak were dialyzed as soon as practicable into 10 mM Tris/HCl, 1 mM MgCl₂, 100 mM NaCl, pH 7.5 (each 2 × 5 liter for 2 h, then 1 × 5 liter overnight, all at 4 °C). The dialysis bag was placed on a bed of solid sucrose and concentrated to a final protein concentration of 1-2 mg/ml. The labeled protein was flash frozen in liquid nitrogen and stored at -80 °C.

Competitive C helix labeling with BR-I₂ and BSR-I₂

A 50- μ M solution of reduced mutant sTnC in labeling buffer (as above) was treated with a DMF solution of BR-I₂ and BSR-I₂ (final concentrations each 100 μ M) and incubated for 40 min at 20 °C. Unreacted rhodamines were quenched by addition of sodium 2-mercaptoethanesulfonate (final concentration 6 mM) and, after 40-min incubation at 20°C, unconjugated rhodamine was removed by gel filtration in labeling buffer (PD-10 column). The eluted protein mixture was analyzed by electrospray mass spectrometry.

NMR sample preparation

One ml of the solution prepared above of 1.04 mg/ml [¹³C, ¹⁵N]-sNTnC•BSR₅₆₋₆₃ in 0.42 mM CaCl₂ and 10 mM KCl was concentrated by centrifugation at 5000 \times g using Centricon YM-10 tubes (Millipore Corp., Billerica, MA). 5 μ l of 10 mM DSS and 5 μ l of 100 mM imidazole in D₂O were added to this solution. Afterwards, 50 μ l D₂O was added to obtain a volume of 500 μ l in 90% H₂O/10% D₂O. The pH was adjusted to 6.78 (using 1 M HCl) by following the imidazole peak in the 1D NMR spectrum (11). The final protein concentration was 0.1 mM (by amino acid analysis) in 20 mM KCl. The salt concentration was kept as low as possible to obtain better results using a spectrometer equipped with a cryogenic probe. Later 0.4 mg of the solid switch peptide of TnI (TnI₁₁₅₋₁₃₁; Ac-RMSADAMLKALLGSKHK-NH₂) was added to the sample. The pH dropped to 5.98 and was readjusted to 6.62 using 1 M NaOH. The complex so formed is referred to as sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃.

NMR spectroscopy

All NMR spectra were acquired on Varian INOVA 600-MHz and 800-MHz spectrometers (the latter equipped with a cryogenic probe) using BioPack pulse sequences (Varian Inc., Palo Alto, CA). A two-dimensional ¹H-¹⁵N HSQC spectrum was collected at 800 MHz with 1024 (¹H) \times 512 (¹⁵N) complex points, 32 transients and spectral widths of 11,990 Hz and 3242 Hz for the first and second dimension respectively. The three-dimensional spectra were acquired at

600 MHz. A 3D CBCA(CO)NH spectrum was acquired with 1024 (^1H) \times 128 (^{13}C) \times 64 (^{15}N) complex points and 24 transients. On the same spectrometer, a 3D HNCACB spectrum was collected with 1024 (^1H) \times 128 (^{13}C) \times 64 (^{15}N) complex points and 32 transients. The spectral width for the 3D spectra was 8398 Hz in the ^1H -dimension, 2431 Hz in the ^{15}N -dimension and 12,068 Hz in the ^{13}C -dimension. All spectra were acquired at 30°C.

To calculate a per residue backbone amide ^{15}N transverse relaxation rate, a set of six 2D ^1H - ^{15}N HSQC spectra were acquired at 600 MHz with different relaxation time delays ($\tau = 10, 30, 50, 70, 90$ and 110 ms). Spectra were collected with 1024 (^1H) \times 512 (^{15}N) complex points, 32 transients, an equilibrium delay of 3.5 s, and spectral widths of 11,990 Hz and 3242 Hz. The relaxation rate per residue was calculated with the Rate Analysis function of NMRView 5.0.4 (12) for all assigned peaks of the HSQC spectrum, except for those peaks showing strong overlap (to avoid integration of two resonances into one).

The spectra were processed using NMRPipe (13). In general, a sine-bell function shifted by 60° or 90° was applied to the FIDs, after a linear prediction limited in length to half the number of experimental points collected. The FIDs were zero-filled with a maximum of twice the number of complex points acquired before analysis with NMRView and Smartnotebook (14). The secondary structure prediction based on a $\text{C}\alpha$ chemical shifts homology was performed using the program CSI (15).

Calcium Titration

A calcium titration was performed to measure the calcium binding affinity of sNTnC•BSR₅₆₋₆₃. The NMR sample containing sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃ was recycled into an apo sNTnC•BSR₅₆₋₆₃ sample. 2 mM of EDTA were added to remove the calcium from the two binding sites of the sNTnC. The sample was passed through a G-25 column to separate the sNTnC•BSR₅₆₋₆₃ from EDTA•2Ca²⁺ and TnI₁₁₅₋₁₃₁. The protein solution obtained was then lyophilized overnight. To

minimize loss of sTnC•BSR₅₆₋₆₃, 475 μ l of NMR buffer (100 mM KCl and 10 mM imidazole), 6 μ l of protease inhibitor and 25 μ l of a solution containing 5 mM DSS-D6 and 0.2 % NaN₃ in D₂O were added directly into the lyophilized tube. The solution obtained was then filtered by centrifugation with 0.22 μ m centrifuge tube filter (Corning, NY) and finally transferred into a 5mm NMR tube. During the calcium titration, eight 2D ¹H-¹⁵N HSQC spectra were acquired on a Varian INOVA 600-MHz for different calcium concentrations (0 to 1.88 mM). The spectra were acquired with 1024 (¹H) \times 512 (¹⁵N) complex points, 32 transients and spectral widths of 11,990 Hz and 3242 Hz for the first and second dimension respectively.

Muscle fibers and solutions

All solutions used in muscle fiber experiments contained 25 mM imidazole, 5 mM MgATP, 1 mM free Mg²⁺ (added as magnesium acetate) and 10 mM EGTA, with the exception of pre-activating solution, which had 0.2 mM EGTA. Ionic strength was adjusted to 150 mM by addition of potassium propionate (KPr), and the pH was 7.1 at 10 °C. Ca²⁺ concentration (expressed as pCa, the negative log₁₀ of the molar value) was adjusted by varying the relative amounts of K2EGTA and CaEGTA, keeping total [EGTA] at 10 mM in all solutions except the pre-activating solution. No CaEGTA was added to the relaxing solution (pCa 9). TnC-extraction solution contained 0.5 mM trifluoperazine (TFP, Fluka, Poole, U.K.), 20 mM MOPS, 5 mM EDTA and 130 mM KPr, pH 7.1 at 10 °C.

Adult New Zealand white rabbits were killed by sodium pentobarbitone injection (200 mg kg⁻¹). Small fiber bundles were dissected from the psoas muscle, demembrated, and stored for up to four weeks in relaxing solution containing 50 % (v/v) glycerol at -20 °C (16). Single fiber segments 2.5–3.5 mm long were dissected in the above storage solution on a cooled microscope stage and mounted, via aluminum T-clips, at sarcomere length 2.4 μ m between a force transducer (AE801, Memscap, Bernin, France) and a fixed hook in a 60- μ l glass

trough containing relaxing solution. The experimental temperature was $10.0 \pm 0.5^\circ\text{C}$.

TnC extraction and reconstitution

Native TnC was selectively extracted from single glycerinated rabbit psoas muscle fibers by repeated 30-s incubations in TnC-extraction solution followed by 30 s in relaxing solution. After a total of 20 such cycles, no Ca^{2+} -activated force could be detected at pCa 4.5. Fibers were reconstituted with labeled sTnC by bathing them in relaxing solution containing ~ 1.0 mg/ml labeled sTnC for 60 min at 10°C .

Fluorescence polarization and physiological measurements

Force and polarized fluorescence intensities from BR- or BSR-labeled sTnC were measured as described (17). The central 1.5-mm segment of a fiber mounted horizontally in a temperature-controlled trough was briefly illuminated from below with 532 nm light polarized either parallel or perpendicular to the fiber axis. The fluorescent light emitted from the fiber and propagating in the vertical direction (in line with the illuminating beam) was collected by a $10\times$ 0.3 N.A. objective (Nikon Plan Fluor). A similar objective collected fluorescence in the horizontal direction (at 90° to both illuminating beam and fiber axis). In each channel the fluorescence was selected by a 610 nm filter with 75 nm bandpass, and separated into parallel and perpendicular components by a Wollaston prism. The resultant polarized fluorescence intensities were measured simultaneously by four photomultipliers (Hamamatsu R4632) for each excitation polarization. Relative transmittances of the excitation and emission channels and photomultiplier sensitivities were measured using an isotropic film of rhodamine in polyvinylalcohol sandwiched between two 45° prisms (18), and appropriate correction factors applied to the fiber data. Three independent order parameters, $\langle P_{2d} \rangle$, $\langle P_2 \rangle$ and $\langle P_4 \rangle$, were calculated from the eight corrected intensities (19). The orientation of the BR and BSR dipoles with respect to the fiber axis, averaged over timescales that are long compared with the lifetime of the excited state, was

estimated from $\langle P_2 \rangle$ and $\langle P_4 \rangle$, by fitting model orientation distributions with a Gaussian shape (3), and by maximum entropy analysis (20). Each pair of $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values was fitted by a Gaussian orientation distribution with peak angle θ_g and standard deviation σ according to the formulae

$$\langle P_2 \rangle = \int_{-\pi/2}^{\pi} P_2(\theta) \cdot (\exp [-(\theta - \theta_g)^2 / 2\sigma^2]) d\theta / \int_{-\pi/2}^{\pi} \exp [-(\theta - \theta_g)^2 / 2\sigma^2] d\theta$$

and

$$\langle P_4 \rangle = \int_{-\pi/2}^{\pi} P_4(\theta) \cdot (\exp [-(\theta - \theta_g)^2 / 2\sigma^2]) d\theta / \int_{-\pi/2}^{\pi} \exp [-(\theta - \theta_g)^2 / 2\sigma^2] d\theta$$

where the integration limits are chosen to include the possible contribution of tails of the Gaussian distribution outside the region $0 < \theta < \pi/2$ when θ_g is within that range, $P_2(\theta)$ is the function $0.5(3\cos^2\theta - 1)$, and $P_4(\theta)$ is $0.125(35\cos^4\theta - 30\cos^2\theta + 3)$. Note that no $\sin\theta$ weighting terms are included in the integrals. This model, in contrast with those used previously (1-3), corresponds to a Gaussian distribution of axial angles of the probe molecules at each azimuth without truncation at the quadrant boundaries. The symmetries of the muscle fiber structure, the cylindrical symmetry around the fiber axis combined with the bipolar organization of sarcomeres and the dipole nature of the probes, converts the Gaussian distribution in the molecular coordinate frame into a ‘folded Gaussian’ distribution in the fiber coordinate frame by effectively reflecting the $\theta < 0$ and $\theta > \pi/2$ tails into the region $0 < \theta < \pi/2$. The mean angle (θ_f) of this ‘folded Gaussian’ distribution in the range $0 < \theta < \pi/2$ will in general not be equal to θ_g .

The maximum entropy distribution f_{ME} of the angle θ between the probe dipole and the fiber axis is the broadest distribution that is consistent with the measured $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values, calculated to maximize the informational entropy of the

distribution defined as $-\int_0^{\pi} f_{ME} \ln(f_{ME}) d\theta$. Again there is no $\sin\theta$ weighting in the

integral, in contrast with the previous formulation (20), so here f_{ME} represents the

maximum entropy estimate of the total number of probes at angle θ , independent of space-filling or other assumptions related to the azimuthal distribution of probe dipoles in the cylindrically symmetrical coordinate system of the fiber. f_{ME} is proportional to $\exp[\lambda_2.P_2(\theta) + \lambda_4.P_4(\theta)]$, where λ_2 and λ_4 are Lagrange multipliers, chosen to fit the measured $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values using the equations

$$\langle P_2 \rangle = \int_0^{\pi} P_2(\theta) \cdot (\exp[\lambda_2.P_2(\theta) + \lambda_4.P_4(\theta)]) d\theta / \int_0^{\pi} \exp[\lambda_2.P_2(\theta) + \lambda_4.P_4(\theta)] d\theta$$

and

$$\langle P_4 \rangle = \int_0^{\pi} P_4(\theta) \cdot (\exp[\lambda_2.P_2(\theta) + \lambda_4.P_4(\theta)]) d\theta / \int_0^{\pi} \exp[\lambda_2.P_2(\theta) + \lambda_4.P_4(\theta)] d\theta$$

The mean of the f_{ME} distribution in the range 0 to $\pi/2$, θ_{ME} , was calculated as

$$\int_0^{\pi/2} \theta \cdot f_{ME} d\theta / \int_0^{\pi/2} f_{ME} d\theta$$

Force and polarized fluorescence data were measured during a series of activations at different $[Ca^{2+}]$ in each fiber. Each activation was preceded by a 1-min incubation in pre-activating solution and followed by a 5-min incubation in relaxing solution. Force and polarized fluorescence intensities were measured after a steady force had been established in each activation. Maximum isometric force at pCa 4.5 was recorded before and after each series of activations at sub-maximal $[Ca^{2+}]$; if maximum isometric force had decreased by more than 10%, the fiber was discarded. The dependence of force and of order parameters on $[Ca^{2+}]$ for each fiber were fitted using least-squares regression to the Hill equation:

$$Y = 1/(1 + 10^{n_H(pCa - pCa_{50})})$$

where pCa_{50} is the pCa corresponding to half-maximal change in the parameter Y , and n_H is the Hill coefficient specifying the steepness of the Ca^{2+} -dependence. Force-pCa relationships were described by expressing sub-maximal Ca^{2+} -activated force at each pCa as a fraction of maximum Ca^{2+} -activated force determined at pCa 4.5.

RESULTS

Diastereoisomers of labeled proteins

As previously described for the same proteins labeled with BR-I₂ (3,7), the BSR-labeled N-lobe and full length proteins were produced as a mixture of diastereoisomers. This occurs because there is a barrier to rotation about the bond that joins either the carboxylated or sulfonated phenyl ring to the three coplanar rings that comprise the fluorophore of these rhodamine dyes (Fig. II-1). Quantum mechanics calculations indicate a value of ~27 kcal/mol for this rotational barrier (21). Although the diastereoisomers in this case for the labeled sTnC were resolved both by HPLC and FPLC, they were recombined after purification because of the small amount of protein available.

Relative reactivity of BR-I₂ and BSR-I₂

Mass spectrometric analysis (Fig. II-2) of the unfractionated mixture of products obtained after incubation of the mutant sTnC with a mixture of BR-I₂ and BSR-I₂ (each rhodamine in 2-fold molar concentration with respect to that of the protein) showed the predominant presence of one labeled species, identified as sTnC•BSR. The calculated molecular weights of the various species, from which the structural assignments were made, are given in the Figure legend. sTnC•BR was barely detectable in the unfractionated mixture, and other minor components identified had two rhodamines attached to the protein. The predominance of sTnC•BSR in the product mix shows that BSR-I₂ reacts more rapidly with the thiol groups of the mutant sTnC than BR-I₂.

Backbone chemical shift comparison between sTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃ and sTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BR₅₆₋₆₃

Even with only 1 mg of [¹³C, ¹⁵N]-sTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃, high quality NMR spectra could be obtained using an 800-MHz spectrometer equipped with a cryogenic probe. A portion of the 2D ¹H-¹⁵N HSQC NMR spectrum is shown in

Fig. II-3. $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BR_{56-63}$ was previously found to undergo partial dimerization in solution, with the proportion of dimer decreasing at higher salt concentrations (7). To evaluate the monomer/dimer ratio for $sNTnC \cdot 2Ca^{2+} \cdot BSR_{56-63}$ under the present conditions, the ^{15}N backbone transverse relaxation rate ($^{15}N-R_2$) of $sNTnC \cdot 2Ca^{2+} \cdot BSR_{56-63}$ was measured at 600 MHz before adding $TnI_{115-131}$. A relaxation rate of 9.6 s^{-1} was obtained from the decay of the amide envelope using the 1D trace of the first increment of six 1H - ^{15}N HSQC spectra with different time delays (see Materials and Methods). This relaxation experiment reflects the apparent molecular weight of $sNTnC \cdot 2Ca^{2+} \cdot BSR_{56-63}$, and consequently estimates the monomer/dimer ratio in solution. The $^{15}N-R_2$ value measured at 20 mM KCl was equivalent to that obtained by Mercier et al. (7) for $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BR_{56-63}$ at 320 mM KCl ($R_2 = 9.8 \text{ s}^{-1}$), which indicates that $sNTnC \cdot 2Ca^{2+} \cdot BSR_{56-63}$ has little tendency to dimerize at the concentration used. As a result, no more salt was added as it would have compromised acquisition of NMR spectra on the 800-MHz NMR spectrometer. The subsequent addition of $TnI_{115-131}$ to form the final complex further reduces the tendency to dimerize (22).

The backbone NMR chemical shift assignments were carried out from a 2D 1H - ^{15}N HSQC spectrum for the amides, and 3D HNCACB and 3D CBCA(CO)NH spectra for the $C\alpha$ and $C\beta$ nuclei using Smartnotebook (14). The backbone assignment of $sNTnC$ in the $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BSR_{56-63}$ complex was completed to 97.7% (N, HN, $C\alpha$ and $C\beta$). The only missing atoms are the $C\beta$ of residues 57 and 61. As noted above, two diastereoisomers were generated during the labeling of $sNTnC$ with $BSR-I_2$. As a result, twin peaks were observable for several residues in the HSQC spectrum (Fig. II-3). Twinning was clearly identifiable for residues 54, 56, 58, 59 and 60, situated close to the BSR probe on the C helix. The peak with higher intensity was chosen for the backbone assignment, but relaxation rates were determined for both peaks. These five residues were identified according to the size of the chemical shift difference between the two resonances, the presence of separate $C\alpha$ and $C\beta$ resonances in 3D

spectra, and the relative peak intensity of the corresponding peaks (approximately equal) in 2D and 3D spectra. Other resonances appearing as major and minor peaks (e.g. residues 8, 10, 23, 39, 73) must come from another phenomenon than the presence of the two diastereoisomers. It seems most likely that those peaks are due to the monomer-dimer equilibrium in solution.

The backbone chemical shifts of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$ were compared to $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ for all assigned atoms (Fig. II-4). The overall backbone chemical shifts did not reveal significant changes. On average, the difference of chemical shifts between the two N-domains is 0.26 ppm for the N nuclei, 0.04 ppm for the amide protons, 0.12 for the $\text{C}\alpha$ nuclei and 0.13 ppm for the $\text{C}\beta$ nuclei. The maximum deviations observed for the HN (0.28 ppm), N (0.78 ppm), $\text{C}\alpha$ (0.76 ppm) and $\text{C}\beta$ (0.67 ppm) reflect this high level of similarity. More importantly, the chemical shifts observed for the residues from E56C to E63C do not show significantly higher variation than residues in the rest of the protein, which indicates that the C helix is not perturbed by the presence of this more rigid probe. Moreover, the secondary structure prediction based on $\text{C}\alpha$ chemical shift homology shows the presence of the α -helices and β -sheets in the same positions as for $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ (Fig. II-5). The prediction indicates a secondary structure of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$ essentially identical to that of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ for the well-defined regions (residues 3 to 85), and a perfect match for the C helix.

Transverse relaxation measurement of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$

Having demonstrated that the C helix is intact, we used amide ^{15}N relaxation data to probe the mobility of residues in the protein. The ^{15}N relaxation rates of sNTnC in the $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$ complex were determined on a per residue basis (Fig. II-6). The average ^{15}N - R_2 obtained was $11.6 \pm 2.8 \text{ s}^{-1}$ using every non-overlapping residue (80 peaks). This relaxation rate is higher than the value given here for $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{BSR}_{56-63}$, since $\text{TnI}_{115-131}$ is bound to the N-domain and forms a higher molecular weight complex. The expected ^{15}N - R_2

values for $s\text{TnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$ are $\sim 9\text{ s}^{-1}$ for a monomer and $\sim 18\text{ s}^{-1}$ for a dimer. These were determined from the correlation between the relaxation rates of different TnC complexes and their respective molecular weights (Darrin Lindhout, Ph.D. Thesis, University of Alberta). According to the average $^{15}\text{N-R}_2$, $\sim 75\%$ of the complex is present in the monomeric form. However, this is an underestimation of the amount of monomer in solution, because faster relaxation rates were observed for some residues. It is likely that these faster rates reflect a contribution from exchange broadening due to the monomer:dimer equilibrium. Without these residues ($^{15}\text{N-R}_2 > 13\text{ s}^{-1}$), the average $^{15}\text{N-R}_2$ decreased significantly to $10.4 \pm 1.7\text{ s}^{-1}$ (59 peaks). This indicates that more than 85% of the complex is in the monomeric form, which is consistent with the ratio of the peak intensities for the twin peaks identified above that result from the monomer:dimer equilibrium in solution. Overall, the $^{15}\text{N-R}_2$ relaxation data per residue show the typical characteristics of a folded protein, which indicates that the structure of the N-domain is not perturbed by the presence of the probe. The relaxation rates for residues at the extremities of the sequence were smaller, as is typical for the flexible terminal regions. The relaxation rates observed for residues 56 to 63 of the C helix were similar, but slightly higher, compared to those measured for the other secondary structures with an average $^{15}\text{N-R}_2$ of 12.5 s^{-1} . Therefore the C helix is not perturbed, and is certainly not more flexible with the more rigid BSR probe attached.

Ca^{2+} -regulation of isometric force in muscle fibers

The isometric force produced by single demembrated fibers from rabbit psoas muscles at saturating $[\text{Ca}^{2+}]$ was reduced by replacement of native sTnC by BR- or BSR-labeled recombinant sTnC. After introduction of sTnC•BR, sTnC•BSR1 and sTnC•BSR2, isometric force was $82 \pm 6\%$ (mean \pm SE, $n = 5$), $77 \pm 2\%$ ($n = 4$) and $64 \pm 6\%$ ($n = 4$), respectively, of the pre-exchange value in each fiber (Table II-1). The mean value for sTnC•BR is similar to that reported previously for unlabeled recombinant sTnC (3,23). Isometric force also developed more slowly in fibers containing sTnC•BR, sTnC•BSR1 or sTnC•BSR2 than in control

fibers (Table II-1). As with maximum isometric force, modification of fiber function appears to be greater for BSR than BR, and greater for BSR2 than BSR1.

The Ca^{2+} -sensitivity of isometric force was also reduced by the introduction of the labeled sTnCs (Fig. II-7). The isometric force at each $[\text{Ca}^{2+}]$ was expressed as a fraction of maximum Ca^{2+} -activated force at pCa 4.5, plotted against pCa, and fitted with the Hill equation (see Methods). The pCa value that produced half-maximum force (pCa_{50}) was reduced after introduction of sTnC•BR (squares), sTnC•BSR1 (triangles) or sTnC•BSR2 (diamonds), by 0.40, 0.70 and 1.02 units respectively (Table II-1). A preliminary Ca^{2+} titration of sTnC•BSR₅₆₋₆₃ in solution supports the conclusion that the BSR probe reduces the Ca^{2+} affinity of TnC (see Fig. II-10). The steepness of the force-pCa relationship in muscle fibers (indicated by the Hill coefficient, n_H) in fibers containing sTnC•BR, sTnC•BSR1 or sTnC•BSR2 was similar to that measured in control fibers before sTnC exchange.

Orientation of the rhodamine dipoles on TnC in muscle fibers

The orientations of the BR and BSR probes on sTnC incorporated into muscle fibers were determined by polarized fluorescence measurements and expressed in terms of the order parameters $\langle P_{2d} \rangle$, $\langle P_2 \rangle$ and $\langle P_4 \rangle$ (3,19). $\langle P_{2d} \rangle$ describes the amplitude of the local rotation of the rhodamine dipole with respect to the protein backbone. $\langle P_{2d} \rangle$ was independent of $[\text{Ca}^{2+}]$ and close to 0.91 for each of the three rhodamine-labeled sTnCs (Table II-1). This value corresponds to that expected for uniform wobble in a cone of semi-angle 20° on a timescale that is short compared with the fluorescence lifetime. There were no significant differences between $\langle P_{2d} \rangle$ values for BR- and BSR-labeled sTnC, or between those for the two diastereoisomers of sTnC•BSR. These results suggest that, once attached to the C helix of TnC in a muscle fiber, there is no difference in the fast local motion of the BR and BSR probes relative to the protein backbone. Preliminary measurements of the excited state lifetime of sTnC•BR and sTnC•BSR in isolated

muscle fibers revealed no substantial difference between the lifetimes of the two probes *in situ*; both had values of ~ 4 ns.

The order parameters $\langle P_2 \rangle$ and $\langle P_4 \rangle$ describe the orientation distribution of the BR and BSR dipoles with respect to the muscle fiber axis, averaged over a timescale that is long compared with the fluorescence lifetime, and with the rapid local motion of the probes factored out. Formally, they correspond to the second- and fourth-rank coefficients of the Legendre polynomial expansion of this orientation distribution. $\langle P_2 \rangle$ would be +1 if all the dipoles were parallel to the fiber axis, and -0.5 if they were all perpendicular to that axis. $\langle P_4 \rangle$ gives orientation information at higher resolution. There were reproducible differences between the $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values for the different labeled sTnCs (Table II-1; Fig. II-8), indicating differences between the orientation distributions of the BR and BSR dipoles with respect to the fiber axis.

For all three rhodamine labeled sTnCs, $\langle P_2 \rangle$ decreased substantially when the muscle fibers were activated (Fig. II-8A), signaling a large increase in the angle between the rhodamine dipole and the fiber axis, as reported previously for sTnC•BR (7). $\langle P_4 \rangle$, in contrast, showed different behavior for the three probes (Fig. II-8B). The Ca^{2+} -dependence of the changes in $\langle P_2 \rangle$ and $\langle P_4 \rangle$ were analyzed using the Hill equation (fitted curves in Fig. II-8). The $\langle P_2 \rangle$ data are more reliable as the signal:noise ratio is greater. The pCa_{50} for $\langle P_2 \rangle$ was similar to that for force for each labeled sTnC. The $\langle P_2 \rangle$ - Ca^{2+} relationship was steeper for sTnC•BSR2 than sTnC•BSR1 (Table II-1). These results confirm the conclusion from the force- Ca^{2+} relationship (Fig. II-7), that the introduction of the rhodamine labels on the C helix of sTnC reduces the Ca^{2+} affinity of the regulatory sites, and that this effect is larger for BSR than BR.

The angle (θ) between the probe dipole and the muscle fiber axis was estimated from the measured $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values for each probe by two independent methods: Gaussian model fitting (3) and maximum entropy analysis (20). In

contrast with previous analyses (3) the distribution was not truncated at either $\theta = 0^\circ$ or $\theta = 90^\circ$ (see Materials and Methods, and Fig. II-9A), thus providing a more realistic representation of the orientation distribution in the molecular coordinate frame. The fitted Gaussian distribution for sTnC•BR (Fig. II-9A, continuous line) in relaxed muscle fibers at pCa 9 had $\theta_g = 34.6 \pm 1.9^\circ$ (mean \pm SE, n=5) and standard deviation (σ) $26.0 \pm 1.3^\circ$ (Table II-1). The value of θ_g is larger than that (28°) reported previously using a Gaussian model distribution that was truncated at 0° (3). The fitted full Gaussian distributions for sTnC•BSR1 (Fig. II-9A, dashed line) and sTnC•BSR2 (dotted line) at pCa 9 had θ_g values of $34.9 \pm 0.3^\circ$ and $38.2 \pm 0.2^\circ$, which are not significantly different from that for sTnC•BR. However the distributions for the BSR probes are narrower than that for the BR probe; σ for sTnC•BSR1 and sTnC•BSR2 were $16.4 \pm 0.4^\circ$ and $17.4 \pm 0.2^\circ$ respectively (Table II-1).

The maximum entropy distribution (f_{ME}) of axial angles (θ) (Fig. II-9B) is the broadest orientation distribution consistent with the observed $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values, calculated by maximizing the informational entropy defined as $-f_{ME} \cdot \ln f_{ME}$ (20). The maximum entropy distribution is necessarily symmetrical around $\theta = 90^\circ$, and is plotted in Fig. II-9B only for the region $0^\circ < \theta < 90^\circ$. The f_{ME} distributions for relaxed muscle fibers at pCa 9 for sTnC•BR (continuous line), sTnC•BSR1 (dashed line) and sTnC•BSR2 (dotted line) are centered on about the same axial angle, but the f_{ME} distribution for BR is broader than that for either of the BSR isomers. The mean angles of these distributions (θ_{ME}) were $36.8 \pm 1.1^\circ$, $35.0 \pm 0.2^\circ$ and $38.3 \pm 0.2^\circ$ for sTnC•BR, sTnC•BSR1 and sTnC•BSR2 respectively (Table II-1), and each of the θ_{ME} values is close to the respective peak (θ_g) value for the Gaussian model distributions (Fig. II-9A).

The maximum entropy (f_{ME}) and Gaussian (f_G) distributions can be compared more directly by noting that the cylindrical and bipolar sarcomere symmetry of the muscle fibers, combined with the dipolar nature of the probes, effectively

folds the $\theta < 0^\circ$ and $\theta > 90^\circ$ tails of the molecular Gaussian distribution (Fig. II-9A) into the region $0^\circ < \theta < 90^\circ$, yielding the fiber-level orientation distribution shown in Fig. II-9C. This ‘folded Gaussian distribution’ no longer has a Gaussian shape, rather it is similar in shape to the maximum entropy distribution f_{ME} (Fig. II-9B). The means (θ_f) of the folded Gaussian distributions for sTnC•BR, sTnC•BSR1 and sTnC•BSR2 were $36.2 \pm 0.9^\circ$, $35.0 \pm 0.3^\circ$ and $38.3 \pm 0.2^\circ$ respectively (Table II-1), very close to the corresponding θ_{ME} values. The greater width of the distribution for the BR probe is again apparent in the folded Gaussian distributions (Fig. II-9C).

The Gaussian and maximum entropy analyses were also applied to the $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values obtained for each probe during active isometric contraction (pCa 4.5; Table II-1). θ_g was slightly larger than θ_f and θ_{ME} for each probe at pCa 4.5, but in each case θ_f was very close to θ_{ME} , as observed at pCa 9. All three angles were smaller for sTnC•BSR2 than for sTnC•BSR1, and smaller for sTnC•BSR1 than for sTnC•BR. These differences support the evidence from isometric force measurements (Fig. II-7, Table II-1) that fibers containing sTnC•BSR2, and to a lesser extent those containing sTnC•BSR1, may not be fully activated at pCa 4.5. The same phenomenon is probably responsible for the slightly larger value of σ observed for the BSR probes during active contraction (Table II-1); incomplete activation at pCa 4.5 of fibers containing the BSR probes would lead to a population of TnC molecules in the relaxed orientation, thereby increasing the measured orientational dispersion.

Calcium Titration

The HSQC spectra acquired during the calcium titration showed broad line widths, low signal-to-noise ratio and did not show good peak dispersion. We observed the same phenomena during the preparation of our first NMR sample, which is caused by both the low protein concentration used and the absence of the TnI₁₁₅₋₁₃₁ peptide. For those reasons, only the peaks of Gly35 and Gly43 were used to calculate the calcium binding affinity of sTnC•BSR₅₆₋₆₃. They were

almost the only peaks having unambiguous chemical shifts for both ^1H and ^{15}N dimensions during the calcium titration. The calcium binding constants for the two binding sites were calculated by plotting the chemical shift changes as a function of the ratio of calcium and protein concentration for each step of the titration (See Annexe A, Fig. A-1). The calcium concentrations were adjusted to compensate for residual EDTA in the sample. The curves in red represent the best fit for Gly35 and Gly43. Both curves used the same Ca^{2+} dissociation constants, which were determined to be $K_{d1} = 13 \mu\text{M}$ and $K_{d2} = 145 \mu\text{M}$ for the first and second calcium binding respectively (site II and site I respectively). For comparison, the binding affinity of sTnC has been previously determined to be $0.8 \mu\text{M} \leq K_{d1} \leq 3 \mu\text{M}$ and $5 \mu\text{M} \leq K_{d2} \leq 23 \mu\text{M}$ (24). The calcium binding affinity of sTnC•BSR₅₆₋₆₃ is evidently diminished. Because of the reduced quality of the spectra used to calculate the binding constants, the conclusions should be regarded as qualitative only. However, the results show clearly that the calcium affinity of sTnC is decreased by the presence of the bifunctional sulforhodamine on the C helix, and are qualitatively consistent with the fluorescence data obtained in the muscle fiber studies (Fig. II-8).

DISCUSSION

The experiments described above compared a wide range of properties of two cysteine-directed bifunctional fluorescent probes, BR-I₂ and BSR-I₂, that were used to label the same protein, namely the C helix of chicken skeletal troponin C (sTnC). The results lead to specific conclusions about the efficacy of these two particular probes for determining the orientation of sTnC in muscle fibers. The comparison of the results from the two probes also allows some more general conclusions to be drawn about the power and limitations of the bifunctional fluorophore approach for determining the structure and orientation of protein domains in the cellular environment.

BSR-I₂ was found to react much more rapidly than BR-I₂ with the pair of cysteine residues introduced into sTnC at positions 56 and 63. The greater reactivity of BSR-I₂ is a significant advantage for the convenient production of the desired conjugate, in which the probe cross-links the target cysteines, at the required purity. A similar difference in reactivity between BSR-I₂ and BR-I₂ has been observed for the E8C/L15C, K11C/Q73C, and P70C/N77C mutants of the A2 isoform of the essential light chain of chicken skeletal myosin (25). The difference in cysteine-directed reactivity between BSR-I₂ and BR-I₂ is therefore likely to be a general property of these reagents, which may be due to stereoelectronic influences within the piperazine rings of BSR-I₂ that enhance the reactivity of its iodoacetamido groups.

The native structure of the N-domain of sTnC is not significantly affected by BSR attached to residues E56C and E65C of its C helix. The backbone chemical shifts of sNTnC•BSR₅₆₋₆₃ are similar to those of sNTnC•BR₅₆₋₆₃, even for the C helix region. The identification of protein secondary structure using C α chemical shifts indicates the formation of five α -helices and a small β -sheet. The relaxation data obtained for sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃ are equivalent to those for sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BR₅₆₋₆₃, demonstrating that, like the BR probe (3), the BSR probe does not significantly affect the conformation of the C helix or the overall dynamics of the protein in solution.

Both sTnC•BR and sTnC•BSR were able to support Ca²⁺ regulation of contraction when exchanged into demembranated muscle fibers, but the mean isometric force at pCa 4.5 was smaller for sTnC•BSR than for sTnC•BR, suggesting that fibers containing sTnC•BSR were not fully activated at pCa 4.5. The rate of force development upon transferring fibers from pCa 9 to pCa 4.5 was also lower for sTnC•BSR than for sTnC•BR, and lower for sTnC•BSR2 than for sTnC•BSR1. In each case the rate was lower than in control fibers. These changes in activation kinetics were correlated with the observed changes in Ca²⁺-sensitivity, as indicated by both isometric force and the orientation changes reported by

polarized fluorescence. In each case the change was greater for BSR than for BR, and larger for BSR2 than for BSR1. Thus, although the secondary structure of TnC is not affected by BR- or BSR-labeling, the Ca^{2+} affinity and binding kinetics of the regulatory sites in the N lobe of sTnC are affected, presumably as a direct result of the proximity of the probes to the Ca^{2+} -binding sites. Different Ca^{2+} affinities for the two diastereoisomers of sTnC•BSR might be explained by different orientations of the electric dipole of BSR with respect to the protein. The calculated electric dipole for a carborhodamine such as in BR is 19 Debye, oriented at $\sim 40^\circ$ to the plane of the xanthene ring (21). It would have a similar orientation in a sulforhodamine such as BSR, although the more diffuse charge distribution on the sulfonate may somewhat modify its magnitude. Because the orientation of the dipole depends on the position of the sulfonate group (above or below the plane of the xanthene ring), the dipole's orientation relative to the protein will differ between the two diastereoisomers by $\sim 80^\circ$, assuming the xanthene ring of the BSR has the same orientation in the two diastereoisomers. Much smaller modulations of electric dipoles (changes ≤ 3 Debye) have been shown to have significant effects on the ion channel gating of modified gramicidin (26,27). Since the more rigid BSR probe is likely to hold the dipole of the two diastereoisomers in more defined orientations relative to the protein than is the case for the diastereoisomers of the more flexible BR probe, it is at least feasible that the observed differences in Ca^{2+} affinity of the sTnC•BSR diastereoisomers could be attributed to this differential electric dipolar effect.

The similarity of the observed values of the order parameter $\langle P_{2d} \rangle$ for sTnC•BSR and sTnC•BR in muscle fibers shows that the amplitude of the local orientational motion of the rhodamine fluorescence dipole with respect to the protein backbone was similar for the two probes. $\langle P_{2d} \rangle$ for both probes corresponds to that expected for uniform wobble in a cone of semi-angle $\sim 20^\circ$ on a timescale that is fast compared with the fluorescence lifetime, which is about 4 ns. The fact that the greater flexibility of the linkers between the probe dipole and cysteine attachments in BR does not result in a greater degree of independent rapid motion

of the probe suggests that the BR probe is immobilized to some extent by an interaction with the protein surface.

The orientation of the BR and BSR dipoles with respect to the fiber axis was estimated from the order parameters $\langle P_2 \rangle$ and $\langle P_4 \rangle$ by two independent analytical approaches, based on fitting Gaussian orientation distributions and on maximum entropy analysis. The Gaussian fitting approach was modified from that used previously (1-3), in which the Gaussian fitting was applied in the *fiber* coordinate frame, and truncated at $\theta = 0^\circ$ because negative values of θ are not defined in that frame. Here we applied the Gaussian model in the *molecular* coordinate frame, in which there is no physical reason to exclude negative axial angles, so the distributions were not truncated at $\theta = 0^\circ$. We then transformed this distribution into the fiber frame by folding the $\theta < 0^\circ$ tail of the molecular-frame Gaussian into the $\theta > 0^\circ$ region in the fiber frame. The bipolar symmetry of the muscle sarcomeres and the dipole nature of the probes impose additional mirror symmetry around $\theta = 90^\circ$ in the fiber orientation distributions, and this symmetry was taken into account in previous Gaussian fitting analyses.

The maximum entropy distribution of axial probe angles is the broadest orientation distribution in the fiber co-ordinate frame that is consistent with the observed $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values. It is defined only for positive θ , but has mirror symmetry around $\theta = 90^\circ$. The maximum entropy distributions were similar to the folded Gaussian distributions for each probe in each condition studied here. The means of the two distributions in the region $0^\circ < \theta < 90^\circ$, θ_{ME} and θ_f respectively, were equal to within 1° in each case. The agreement of the mean angles derived from the model-independent maximum entropy analysis and the folded Gaussian analysis suggest that both methods give a reliable estimate of the orientation distribution of the probes in the fiber frame, at least at the relatively low angular resolution implied by knowledge of only the first two second-rank order parameters, $\langle P_2 \rangle$ and $\langle P_4 \rangle$. The folded Gaussian analysis also suggests that these fiber-frame distributions result from Gaussian distributions in the molecular

frame, at the same angular resolution. The isolation of these molecular orientation distributions, characterized by peak angle θ_g and standard deviation σ , is explicitly distinguished from orientation distributions in the fiber frame for the first time by the present analysis.

Applying these general conclusions to the results of the present experiments, the molecular-frame orientations of the sTnC•BR, sTnC•BSR1 and TnC•BSR2 probes are defined by the Gaussian parameters θ_g and σ . The comparison of the parameters for the three probes is most informative for relaxed muscle fibers (pCa 9), in which all the TnC molecules in the fiber are in the ‘off’ conformation. In these conditions, θ_g values for the sTnC•BR, sTnC•BSR1 and TnC•BSR2 probes were almost equal, suggesting that the peak orientation of each of the probes accurately reports that of the C helix residues to which they are attached. However, σ was significantly larger for sTnC•BR than for sTnC•BSR1 and TnC•BSR2; as might be expected on the basis of the greater flexibility of the probe-protein linkers in the BR probe. In this context it is important to note that σ is not related to the independent motion of the probes on timescales shorter than the ~4 ns fluorescence lifetime, which we have shown on the basis of the $\langle P_{2d} \rangle$ measurements is similar for the BR and BSR probes. The difference in σ rather suggests that the greater flexibility of BR allows the probe to adopt a wider range of conformations with respect to the protein backbone, but that motion between these conformations takes place on timescales longer than 4 ns. The value of σ observed for the less flexible BSR probe, about 17° , thus represents a probable upper limit for the standard deviation of the axial orientations of the C helix of TnC in a relaxed muscle fiber.

The present results demonstrate the potential of polarized fluorescence measurements with bifunctional probes to measure the orientation of protein domains in their native environment with minimal disruption of native structure. They also illustrate the sensitivity of the measured orientational and functional parameters to probe chemistry. In the case of cysteines at positions 56 and 63 on

the C helix of sTnC, the bifunctional probes BR and BSR have complementary advantages and disadvantages. BSR has much greater reactivity for this pair of sites, and probably for all cysteine residues, and this facilitates the preparation of probe-protein conjugates at the required purity. Neither probe alters the structure of sTnC in solution, but both reduce the Ca^{2+} affinity of its regulatory sites and this effect is more pronounced for BSR than for BR. The extent of the reduction is different for the two diastereoisomers of sTnC•BSR. Both probes are likely to accurately report the mean *in situ* orientation of the vector joining the two cysteines to which they are attached.

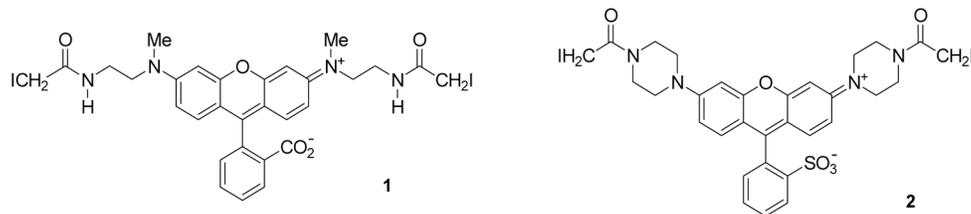
FIGURES AND TABLES

Fig. II-1. Structures of the bifunctional rhodamines 1 (BR-I₂) and 2 (BSR-I₂) used for labeling.

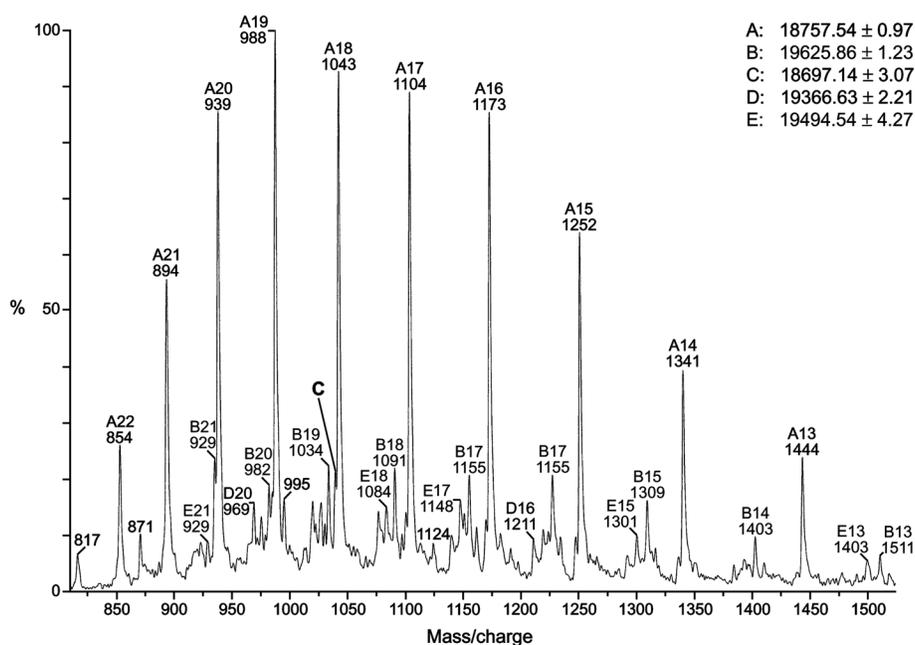


Fig. II-2. Electrospray mass spectrum of the unfractionated E56C, E63C, C101A sTnC mutant after competitive labeling with a mixture of BR-I₂ and BSR-I₂. The major series (A-peaks) corresponds to the expected mass of sTnC•BSR (calculated mass 18757.4 Da). The single peak marked C is the only one resolved that corresponds to sTnC•BR (calculated mass 18697.3 Da). Of the remaining species (B, D and E) two correspond to TnC labeled with 2 rhodamine moieties. B is from TnC carrying 2 BSR moieties, with the iodoacetamide functions on their distal ends having both been displaced by 2-mercaptoethanesulfonate (calculated mass 19624.5 Da). E represents double labeling by BR, but with only one of the iodoacetamides on the side chains not attached to cysteine residues having been displaced by 2-mercaptoethanesulfonate (calculated mass 19491.0 Da). The exact nature of species D could not be assigned.

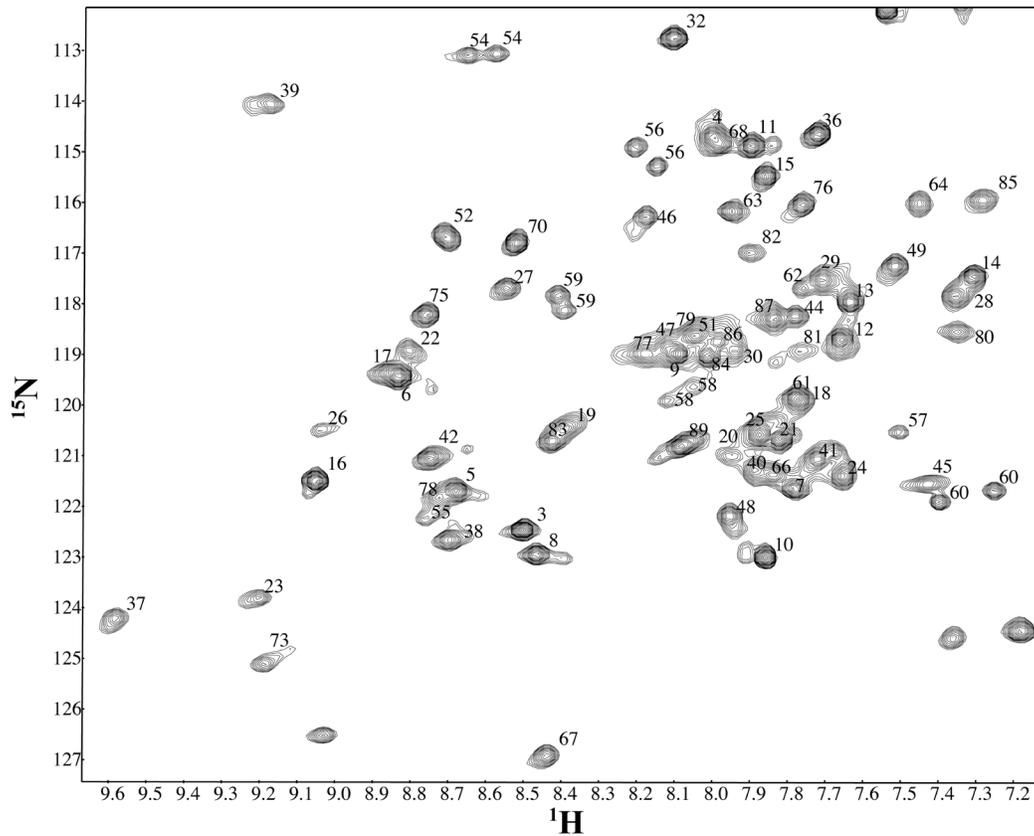


Fig. II-3. Selected region of the 2D ^1H - ^{15}N HSQC NMR spectrum of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BSR}_{56-63}$ acquired at 800 MHz using a cryogenic probe. Twin peaks are clearly observed for residues 54, 56, 58, 59 and 60, reflecting the presence of two diastereoisomers in solution.

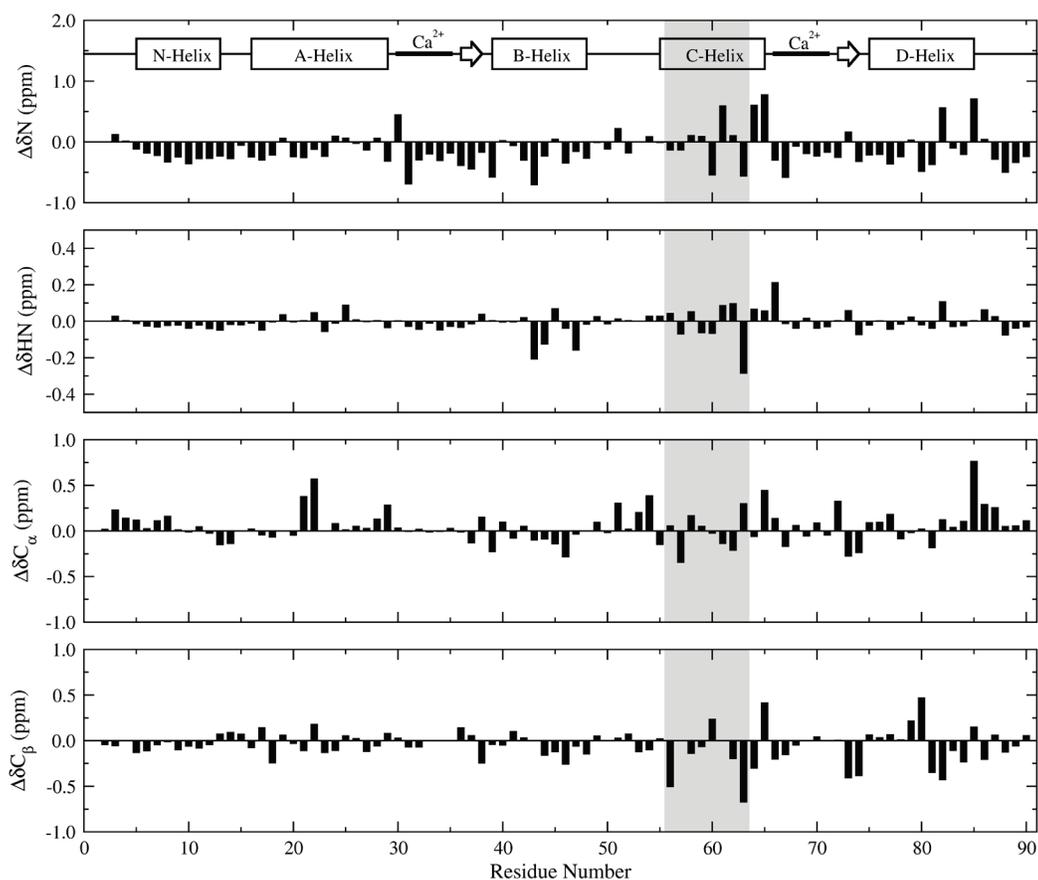


Fig. II-4. Difference in ^{15}N , 1H_N , $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts between $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BSR_{56-63}$ and $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BR_{56-63}$. The grey area corresponds to the region of the N-domain where the bifunctional rhodamine label is attached. The resonances with higher intensity were chosen for twin peaks.

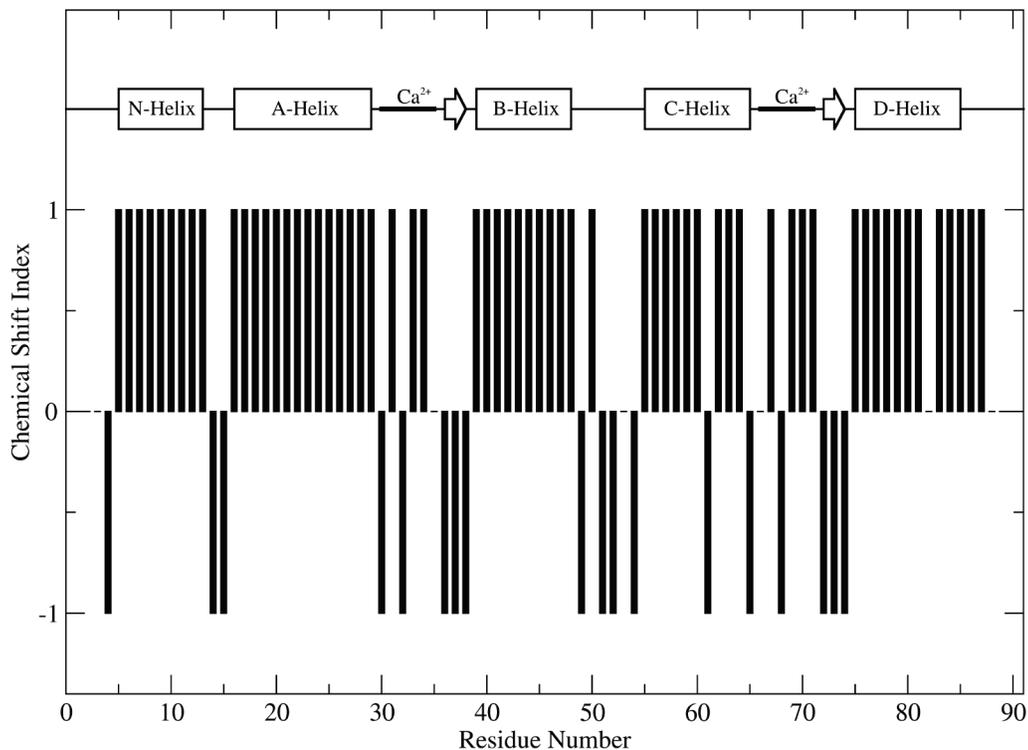


Fig. II-5. Secondary structure prediction of sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃ based on ¹³C α NMR chemical shift homology using the program CSI. The deviation of the C α chemical shift from its random coil value is evaluated by the program: a value of +1 is assigned for the residues having a chemical shift which is characteristic of α -helices, 0 for random coil and -1 for β -sheets. The secondary structure of sNTnC is the following: N helix (D5-L13), A helix (E16-F29), B helix (T39-M48), C helix (K55-V65), D helix (F75-Q85) and small β -sheets (D36-S38 and T72-D74).

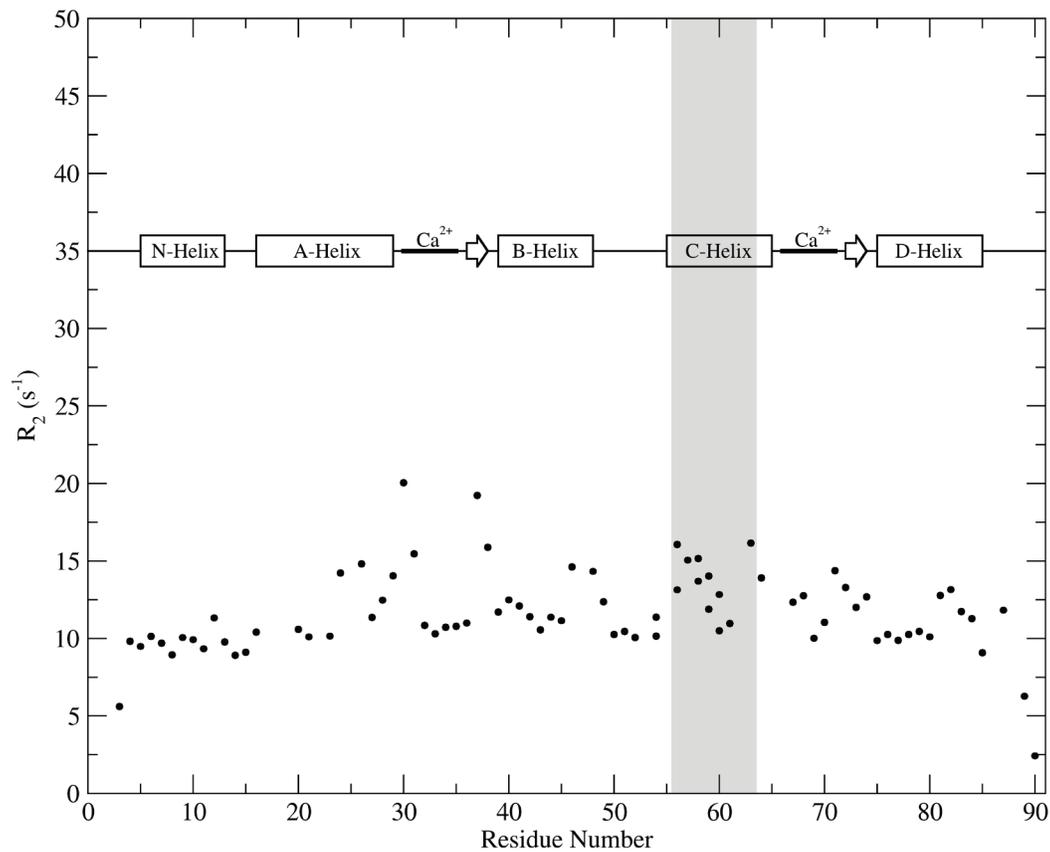


Fig. II-6. Backbone amide ^{15}N - R_2 NMR relaxation rates for each residue of sNTnC•2Ca $^{2+}$ •TnI $_{115-131}$ •BSR $_{56-63}$ at 600 MHz. The grey area corresponds to the region of the N-domain where the bifunctional rhodamine label is attached. Two separate values were calculated for the twin peaks (residues 54, 56, 58, 59 and 60).

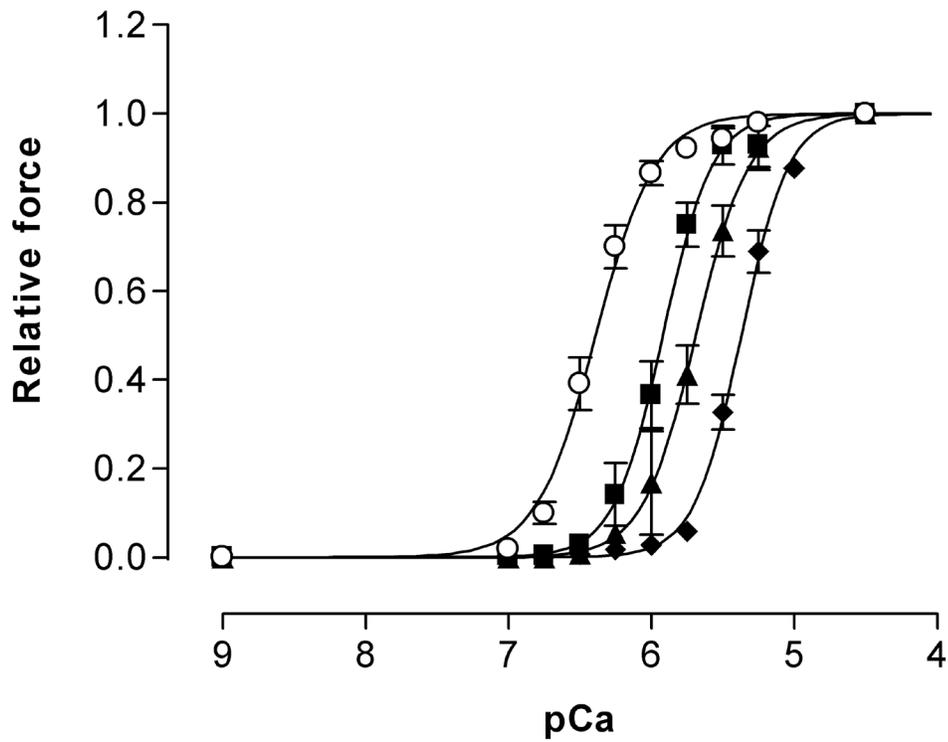


Fig. II-7. The Ca^{2+} dependence of force in control fibers (circles), fibers reconstituted with sTnC•BR (squares) or sTnC•BSR (sTnC•BSR1, triangles; sTnC•BSR2, diamonds). All data are scaled to the value obtained at pCa 4.5 for each labeled sTnC. Each point is the mean \pm SE. Data were fitted to the mean data points with the Hill equation (see Methods).

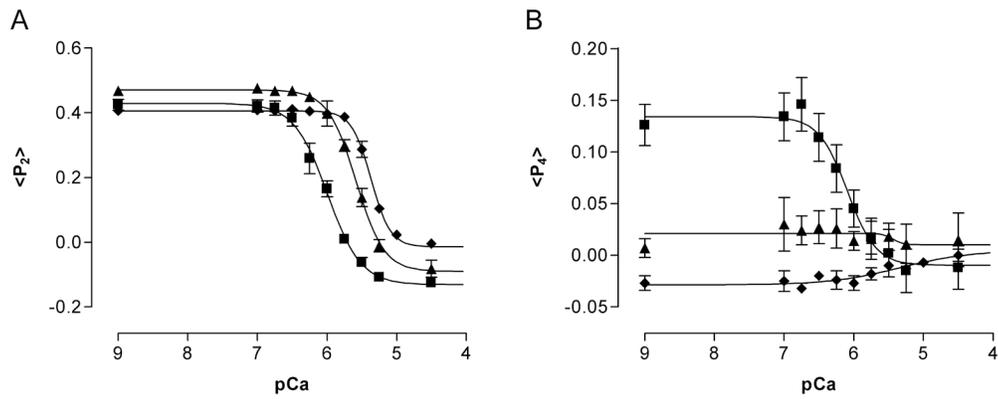


Fig. II-8. Ca^{2+} dependence of dipole order parameters $\langle P_2 \rangle$ and $\langle P_4 \rangle$ for sTnC•BR (squares) and the two diastereoisomers of sTnC•BSR (sTnC•BSR1; triangles, and sTnC•BSR2, diamonds). Values are mean \pm SE. Data were fitted to the mean data points with the Hill equation (see Methods).

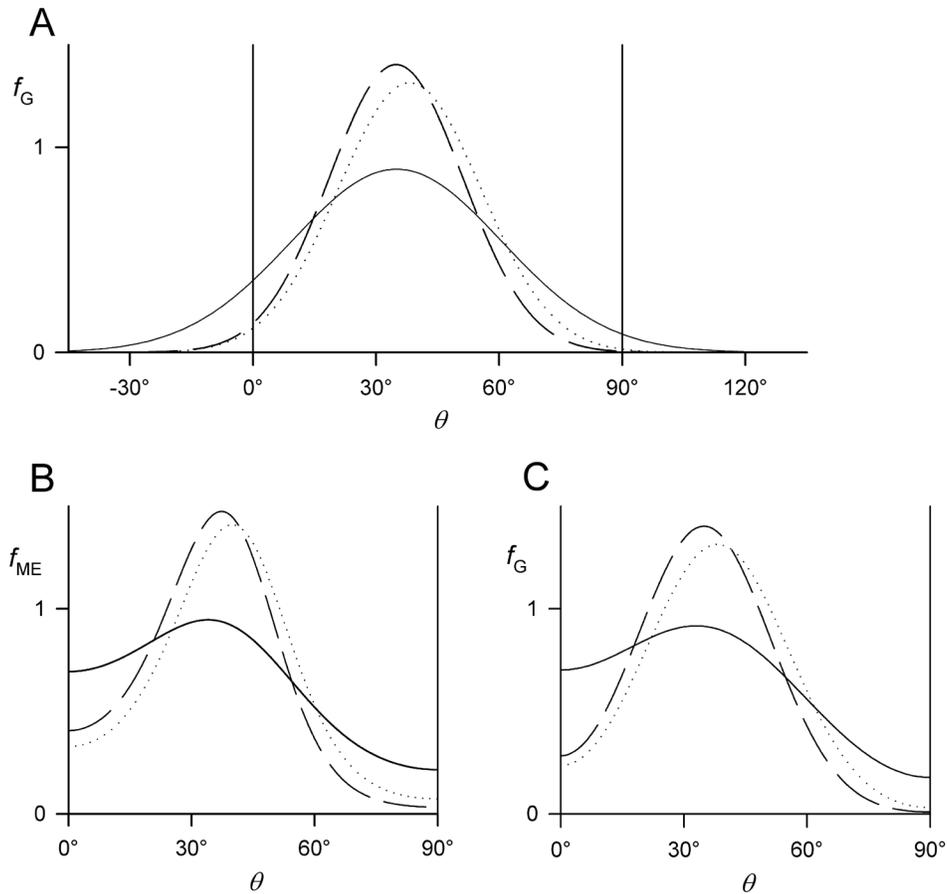


Fig. II-9. Distributions of angles θ between the bifunctional rhodamine dipole and the fiber axis in relaxed muscle, calculated from the mean $\langle P_2 \rangle$ and $\langle P_4 \rangle$ values for pCa 9 in Table 1 for sTnC•BR (continuous line) and the two diastereoisomers of sTnC•BSR (sTnC•BSR1, dashed line; sTnC•BSR2, dotted line). All distributions have been normalized to unit area. A. Gaussian distribution f_G crossing quadrant boundaries at $\theta = 0$ and 90° . B. One-dimensional maximum entropy distribution f_{ME} . C. Folded Gaussian distribution confined to the region $0 < \theta < 90^\circ$.

Table II-1. Fiber force and fluorescence polarization measurements[†]

	pCa	sTnC•BR (n = 5)	sTnC•BSR1 (n = 4)	sTnC•BSR2 (n = 4)
Isometric force				
Peak force (%) [‡]		81.5 ± 5.8	77.3 ± 1.9 ^{ns}	64.2 ± 6.4 ^{ns (ns)}
Half time of force rise (%) [‡]		189 ± 25	293 ± 43 [*]	420 ± 62 ^{** (ns)}
pCa ₅₀ for force		6.00 ± 0.08	5.70 ± 0.04 [*]	5.38 ± 0.03 ^{*** (***)}
n _H for force		2.71 ± 0.18	2.80 ± 0.82 ^{ns}	2.95 ± 0.23 ^{ns (ns)}
Order parameters				
<P _{2d} >	9.0	0.915 ± 0.015	0.926 ± 0.024 ^{ns}	0.922 ± 0.009 ^{ns (ns)}
<P _{2d} >	4.5	0.895 ± 0.013	0.911 ± 0.019 ^{ns}	0.921 ± 0.008 ^{ns (ns)}
<P ₂ >	9.0	0.422 ± 0.021	0.471 ± 0.005 ^{ns}	0.396 ± 0.004 ^{ns (***)}
<P ₂ >	4.5	-0.126 ± 0.006	-0.079 ± 0.019 [*]	-0.004 ± 0.014 ^{*** (*)}
<P ₄ >	9.0	0.128 ± 0.027	0.016 ± 0.011 ^{**}	-0.031 ± 0.005 ^{*** (**)}
<P ₄ >	4.5	-0.012 ± 0.021	0.005 ± 0.021 ^{ns}	0.000 ± 0.006 ^{ns (ns)}
pCa ₅₀ for <P ₂ > [§]		6.02 ± 0.03	5.59 ± 0.07 ^{***}	5.37 ± 0.02 ^{*** (***)}
n _H for <P ₂ > [§]		1.94 ± 0.15	2.21 ± 0.38 ^{ns}	3.32 ± 0.33 ^{** (ns)}
Gaussian fitting				
θ _g (°)	9.0	34.6 ± 1.9	34.9 ± 0.3 ^{ns}	38.2 ± 0.2 ^{ns (***)}
θ _g (°)	4.5	67.2 ± 2.3	64.6 ± 1.4 ^{ns}	62.2 ± 1.1 ^{ns (ns)}
σ(°)	9.0	26.0 ± 1.3	16.4 ± 0.4 ^{***}	17.4 ± 0.2 ^{*** (ns)}
σ(°)	4.5	22.4 ± 2.8	26.0 ± 1.6 ^{ns}	28.0 ± 0.8 ^{ns (ns)}
θ _t (°)	9.0	36.2 ± 0.9	35.0 ± 0.3 ^{ns}	38.3 ± 0.2 ^{ns (***)}
θ _t (°)	4.5	63.3 ± 0.5	61.3 ± 1.0 ^{ns}	57.9 ± 0.4 ^{*** (*)}
Maximum entropy				
θ _{ME} (°)	9.0	36.8 ± 1.1	35.0 ± 0.2 ^{ns}	38.3 ± 0.2 ^{ns (**)}
θ _{ME} (°)	4.5	63.0 ± 0.3	60.9 ± 1.0 ^{ns}	57.2 ± 0.7 ^{*** (*)}

[†]All statistics are given as mean ± SE. The asterisks denote significance level in *t*-tests. ***, P < 0.001; **, P < 0.01; *, P < 0.05; ns, P > 0.05. Comparisons are with the relevant value for sTnC•BR. The superscripts in brackets are for comparisons between sTnC•BSR1 and sTnC•BSR2. [‡] Relative to fibers before TnC exchange for which the half-time of force development was 0.38 ± 0.03 s (n = 19), pCa₅₀ was 6.40 ± 0.04 (n = 5) and n_H was 2.62 ± 0.40 (n = 17). [§] From data in Fig. II-8A. At pCa 4.5 the Gaussian model failed to fit <P₂> and <P₄> for two of the fibers containing sTnC•BR and one of the fibers containing sTnC•BSR1, so fitted parameters are for n = 3 in each case.

REFERENCES

1. Corrie, J. E. T., Brandmeier, B. D., Ferguson, R. E., Trentham, D. R., Kendrick-Jones, J., Hopkins, S. C., Van Der Heide, U. A., Goldman, Y. E., Sabido-David, C., Dale, R. E., Criddle, S., and Irving, M. (1999) Dynamic measurement of myosin light-chain-domain tilt and twist in muscle contraction, *Nature* 400, 425-430.
2. Hopkins, S. C., Sabido-David, C., Van Der Heide, U. A., Ferguson, R. E., Brandmeier, B. D., Dale, R. E., Kendrick-Jones, J., Corrie, J. E. T., Trentham, D. R., Irving, M., and Goldman, Y. E. (2002) Orientation changes of the myosin light chain domain during filament sliding in active and rigor muscle, *J. Mol. Biol.* 318, 1275-1291.
3. Ferguson, R. E., Sun, Y. B., Mercier, P., Brack, A. S., Sykes, B. D., Corrie, J. E. T., Trentham, D. R., and Irving, M. (2003) In situ orientations of protein domains: troponin C in skeletal muscle fibers, *Mol. Cell* 11, 865-874.
4. Brack, A., Brandmeier, B., Ferguson, R., Criddle, S., Dale, R., and Irving, M. (2004) Bifunctional rhodamine probes of Myosin regulatory light chain orientation in relaxed skeletal muscle fibers, *Biophys J* 86, 2329-2341.
5. Forkey, J., Quinlan, M., Shaw, M., Corrie, J., and Goldman, Y. (2003) Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization, *Nature* 422, 399-404.
6. Yildiz, A., Forkey, J., McKinney, S., Ha, T., Goldman, Y., and Selvin, P. (2003) Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization, *Science* 300, 2061-2065.
7. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E. T., Trentham, D. R., and Sykes, B. D. (2003) NMR structure of a bifunctional rhodamine labeled N-domain of troponin C complexed with the regulatory "switch" peptide from troponin I: implications for in situ fluorescence studies in muscle fibers, *Biochemistry* 42, 4333-4348.

8. Corrie, J., Craik, J., and Munasinghe, V. (1998) A homobifunctional rhodamine for labeling proteins with defined orientations of a fluorophore, *Bioconjug Chem* 9, 160-167.
9. Peterman, E., Sosa, H., Goldstein, L., and Moerner, W. (2001) Polarized fluorescence microscopy of individual and many kinesin motors bound to axonemal microtubules, *Biophys J* 81, 2851-2863.
10. Asenjo, A., Krohn, N., and Sosa, H. (2003) Configuration of the two kinesin motor domains during ATP hydrolysis, *Nat Struct Biol* 10, 836-842.
11. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J Biomol NMR* 41, 5-7.
12. Johnson, B. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules, *Methods Mol Biol* 278, 313-352.
13. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
14. Slupsky, C., Boyko, R., Booth, V., and Sykes, B. (2003) Smartnotebook: a semi-automated approach to protein sequential NMR resonance assignments, *J Biomol NMR* 27, 313-321.
15. Wishart, D., and Sykes, B. (1994) The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data, *J Biomol NMR* 4, 171-180.
16. Sabido-David, C., Brandmeier, B., Craik, J., Corrie, J., Trentham, D., and Irving, M. (1998) Steady-state fluorescence polarization studies of the orientation of myosin regulatory light chains in single skeletal muscle fibers using pure isomers of iodoacetamidotetramethylrhodamine, *Biophys J* 74, 3083-3092.
17. Sun, Y., Brandmeier, B., and Irving, M. (2006) Structural changes in troponin in response to Ca^{2+} and myosin binding to thin filaments during activation of skeletal muscle, *Proc Natl Acad Sci USA* 103, 17771-17776.

18. Bell, M., Dale, R., Van Der Heide, U. A., and Goldman, Y. (2002) Polarized fluorescence depletion reports orientation distribution and rotational dynamics of muscle cross-bridges, *Biophys J* 83, 1050-1073.
19. Dale, R. E., Hopkins, S. C., Van Der Heide, U. A., Marszalek, T., Irving, M., and Goldman, Y. E. (1999) Model-independent analysis of the orientation of fluorescent probes with restricted mobility in muscle fibers, *Biophys. J.* 76, 1606-1618.
20. Van Der Heide, U. A., Hopkins, S., and Goldman, Y. (2000) A maximum entropy analysis of protein orientations using fluorescence polarization data from multiple probes, *Biophys J* 78, 2138-2150.
21. Cavallo, L., Moore, M., Corrie, J., and Fraternali, F. (2004) Quantum mechanics calculations on rhodamine dyes require inclusion of solvent water for accurate representation of the structure, *J Phys Chem A* 108, 7744-7751.
22. Mckay, R. T., Tripet, B., Pearlstone, J., Smillie, L., and Sykes, B. (1999) Defining the region of troponin-I that binds to troponin-C, *Biochemistry* 38, 5478-5489.
23. Regnier, M., Rivera, A., Chase, P., Smillie, L., and Sorenson, M. (1999) Regulation of skeletal muscle tension redevelopment by troponin C constructs with different Ca²⁺ affinities, *Biophys J* 76, 2664-2672.
24. Li, M. X., Gagné, S., Tsuda, S., Kay, C., Smillie, L., and Sykes, B. (1995) Calcium binding to the regulatory N-domain of skeletal muscle troponin C occurs in a stepwise manner, *Biochemistry* 34, 8330-8340.
25. Knowles, A. C., Ferguson, R., Brandmeier, B., Kendrick-Jones, J., Corrie, J., Trentham, D., and Irving, M. (2004) Orientation of the essential light chain region of myosin in skeletal muscle fibers determined by polarized fluorescence, *Biophys J* 86, 187A-188A.
26. Borisenko, V., Burns, D., Zhang, Z., and Woolley, G. (2000) Optical switching of ion-dipole interactions in a gramicidin channel analogue, *J Am Chem Soc* 122, 6364-6370.

-
27. Loughheed, T., Borisenko, V., Hennig, T., Ruck-Braun, K., and Woolley, G. (2004) Photomodulation of ionic current through hemithioindigo-modified gramicidin channels, *Org Biomol Chem* 2, 2798-2801.

CHAPTER III – Bifunctional rhodamine dynamics*NMR Studies of the Dynamics of a Bifunctional Rhodamine Probe Attached to Troponin C*

Olivier Julien,¹ Pascal Mercier,¹ Leo Spyropoulos,¹ John E. T. Corrie,²
and Brian D. Sykes¹

¹ *Department of Biochemistry, University of Alberta, Edmonton, Canada.*

² *MRC National Institute for Medical Research, London, United Kingdom.*

Publication: Julien et al. 2008. *J. Am. Chem. Soc.*, 130:2602-2609

Contributions: OJ and BDS designed the NMR experiments. OJ and BDS acquired the NMR spectra, OJ assigned the protein chemical shifts, analyzed the relaxation data, and wrote the manuscript with BDS. PM and LS were involved in the relaxation data analysis. Dave Corson expressed and purified the protein. JETC provided the labeled BR probes and participated in the writing of the manuscript. OJ made all the figures.

OVERVIEW

Fluorescence polarization measurements of bifunctional rhodamine (BR) probes provide a powerful approach to determine the *in situ* orientation of proteins within ordered complexes such as muscle fibers. For accurate interpretation of fluorescence measurements, it is important to understand the probe dynamics relative to the protein to which it is attached. We previously determined the structure of the N-domain of chicken skeletal troponin C, BR-labeled on the C helix, in complex with the switch region of troponin I, and demonstrated that the probe does not perturb the structure or dynamics of the protein. In this study, the motion of the fluorescence label relative to the protein has been characterized using NMR relaxation measurements of ¹³C-labeled methyl groups on the BR

probe and ^{15}N -labeled backbone amides of the protein. Probe dynamics were monitored using off-resonance ^{13}C - $R_{1\rho}$, ^{13}C - R_1 and $\{^1\text{H}\}$ - ^{13}C NOE at magnetic field strengths of 500, 600, and 800 MHz. Relaxation data were interpreted in terms of the overall rotational correlation time of the protein and a two time scale model for internal motion of the BR methyl groups, using a numerical optimization with Monte Carlo parameter error estimation. The analysis yields a 1.5 ± 0.4 ps correlation time for rotation around the three-fold methyl symmetry axis, and a 0.8 ± 0.4 ns rotational correlation time for reorientation of the ^{13}C - ^{14}N bond with an associated S_s^2 of 0.79 ± 0.03 . Order parameters of the backbone NH vectors in the helix to which the probe is attached average $S^2 \sim 0.85$, implying that the amplitude of independent reorientation of the BR probe is small in magnitude, consistent with results from fluorescence anisotropy studies in reconstituted muscle fibers.

INTRODUCTION

Fluorescent labeling of proteins is a widely used approach for studying the properties of biological systems that range from isolated proteins to intact supramolecular assemblies within live cells. Many applications involve imaging or qualitative analysis, whereas others involve techniques like fluorescence resonance energy transfer (FRET) that can yield quantitative intermolecular distances. However, in very few of these applications has there been a detailed characterization of the three dimensional structure of the labeled protein, or the environment and dynamics of the fluorescent label, yet it is evident from many examples that extrinsic fluorescence probes do make interactions with proteins to which they are covalently or non-covalently bound. Information on these interactions is therefore of general interest for better understanding of the use of fluorescence probes. To illustrate the point, an example where structural data are available can be found in a mutant phosphate-binding protein labeled with a particular coumarin fluorophore (*1*). Binding of inorganic phosphate to this labeled protein is accompanied by a large fluorescence enhancement and the

crystallographic structure shows the coumarin specifically bound in a hydrophobic pocket that is created as the protein folds around the phosphate ligand (2). In another recent example, as yet without structural information, we have described two fluorescent sialosides that bind to hemagglutinin (3). Both ligands undergo changes in fluorescence intensity upon binding and the binding affinity increases by three orders of magnitude compared to that for sialic acid itself, implying that there must be additional interactions between the fluorophores and the protein. Many other such examples exist, but without structural and dynamic information on these interactions, progress in understanding the effects is necessarily limited.

The present work relates to the probe dynamics in fluorescence polarization measurements of bifunctional rhodamine-labeled proteins to determine the *in situ* orientation of protein domains whose structures are known *in vitro*. This approach has been used to determine the orientations of domains of myosin regulatory light chain (4,5), and troponin C (6,7) proteins reconstituted into skeletal muscle fibers. BR-labeling involves covalent attachment of the probe to two points of the protein *via* cysteine residues introduced at known positions in the structure (Fig. III-1). The order parameter $\langle P_{2d} \rangle$ extracted from the polarization data gives information on the amplitude of rapid wobble of the probe on a time scale less than the fluorescence lifetime ($\tau \sim 4$ ns) (8). In previous experiments where BR-labeled troponin C was incorporated into the organized system of a single muscle fiber (6,7), the measured values of $\langle P_{2d} \rangle$ were in the range 0.87-0.95. These values are consistent with a restricted "wobble-in-a-cone" motion of the probe with a half-angle in the range 19-23° (8). In the current work, we have made an independent measurement of the dynamics of a bifunctional rhodamine label attached to a mutant (E56,63C) of the N-lobe of chicken skeletal troponin C using NMR relaxation measurements that exploit differential isotope labeling – that is, with ^{13}C on the fluorescence label and ^{15}N -labeling of the protein. In contrast to the fluorescence measurements, these measurements were made for the isolated protein in solution so potential probe flexibility would not be influenced by

possible protein-protein contacts when the labeled protein is in its native environment of the troponin complex reconstituted into a muscle fiber.

Measurements were made on a complex of the N-domain of skeletal troponin C with a BR probe attached to the C helix, saturated with Ca^{2+} , and bound to the switch peptide of TnI ($\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$), for which we have previously determined the structure (Fig. III-1A) and shown that the structure and dynamics of sNTnC are not affected by the presence of the probe (9). The rhodamine was ^{13}C -labeled, as shown in Fig. III-1B and III-1C, and we compared ^{13}C and ^{15}N NMR relaxation measurements at multiple field strengths of the ^{13}C -labeled BR-probe attached to [U - ^{15}N]troponin C to characterize the motion of the label with respect to the protein. Off-resonance ^{13}C - $R_{1\rho}$, ^{13}C - R_1 and $\{^1\text{H}\}$ - ^{13}C NOE were measured at 500, 600 and 800 MHz for the ^{13}C -labeled methyl groups of rhodamine within the $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ complex. These measurements allow us to describe and distinguish the motions of the BR probe attached to sNTnC relative to the mobility of troponin C itself in solution and, in conjunction with fluorescence measurements, to interpret the motions of the bifunctional rhodamine label attached to troponin C reconstituted into muscle fibers. The results suggest that there is relatively little motion of the BR probe on a picosecond to nanosecond time scale, in agreement with the fluorescence results. The distribution of probe orientations in fibers shown on a slower time scale (10) may result from motion of the protein domain to which the probe is attached.

THEORETICAL ASPECTS

^{13}C relaxation measurements of the AX_3 methyl group spin system can provide information about the dynamics of the BR-label (Fig. III-1A), using either direct ^{13}C observed, or indirect ^1H observed NMR experiments. In principle, direct ^{13}C observation with ^1H decoupling is the simplest approach to conducting NMR relaxation experiments for the BR-label. Unfortunately, the inherent low

sensitivity of this method requires a relatively high protein concentration that would lead to protein dimerization in the particular case of sNTnC. To compensate for the low gyromagnetic ratio of ^{13}C nuclei, 1D and 2D NMR experiments using insensitive nuclei enhanced polarization transfer (INEPT) pulse sequence that exploits magnetization transfer from ^1H to ^{13}C have been developed to measure ^{13}C relaxation (11,12). These experiments allow for the indirect measurement of ^{13}C (or ^{15}N) relaxation properties by exploiting the larger ^1H population polarization, and higher sensitivity of ^1H NMR detection. However, care must be taken in the setup of the experiment and the interpretation of the results, because ^{13}C relaxation in AX_3 spin systems is complicated, especially in the slow tumbling limit (11). One has to be aware of the contributions to the relaxation of the ^{13}C nucleus from cross-correlation between ^1H - ^{13}C dipolar interactions, and between ^1H - ^{13}C dipolar interactions and ^{13}C CSA. In the current case, it is possible that there are additional contributions to ^{13}C relaxation due to scalar relaxation involving the attached ^{14}N nucleus (Fig. III-1C). Further contributions to ^{13}C relaxation can potentially occur by chemical exchange processes between rhodamine conformations, and between monomeric and dimeric forms of the protein.

Herein, we employed indirect detection pulse sequences to measure the ^{13}C relaxation rates of the methyl(s) of the BR probe. For these sequences, the magnetization is transferred from the protons of the methyl group to the covalently attached ^{13}C using INEPT, the ^{13}C nuclei are allowed to relax for a variable delay, the magnetization is transferred back to the ^1H using reverse INEPT, and finally the ^1H signal is detected. As mentioned above, significant errors can be introduced in relaxation rates measured using these pulse sequences due to the complex relaxation of both the ^1H and ^{13}C nuclei caused by cross-correlation between relaxation mechanisms. Pulse sequence strategies to minimize these errors have been described elsewhere (11,13,14), and involve adjustment of various timings for standard ^{13}C - R_1 and ^{13}C - $R_{1\rho}$ pulse sequences (BioPack, Varian Inc.). Specifically, the refocusing delay τ was chosen to give $2\pi {}^1J_{\text{CH}}\tau = 54.7^\circ$

(0.955 rad), so magnetization from individual ^{13}C transitions is transferred uniformly to ^1H magnetization. The INEPT delay for polarization transfer was set to less than $1/(4J_{\text{CH}})$ to minimize relaxation losses during delays. Proton 180° pulses were applied during variable relaxation delays to minimize the effect of differential ^1H relaxation. Sensitivity enhancement was not used, in order to reduce the number of delays. The purpose of the experimental modifications noted above is to simplify the interpretation of measured relaxation rates which would otherwise be a more complex mixture of fast and slow relaxation contributions (15). Under the proper experimental conditions, the initial portion of the relaxation curve is a single exponential decay corresponding to the fast component of the relaxation.

The relaxation of the [^{13}C]methyl groups on the rhodamine can potentially include contributions due to scalar relaxation from the covalently attached ^{14}N nucleus or exchange broadening, either from different conformations of the BR-label or from monomer-dimer equilibrium of the protein. In this work, we have used off-resonance $R_{1\rho}$ measurements as a function of the effective spin-locking field in the rotating frame to disperse these contributions to the relaxation. The transverse relaxation rate of a nucleus in the presence of chemical exchange and/or broadening due to scalar relaxation can be calculated by measuring the off-resonance $R_{1\rho}$ and removing the R_1 contribution using the following equations (16):

$$R_2 = \frac{R_{1\rho}^{\text{obs}} - R_1 \cos^2 \theta}{\sin^2 \theta} \quad [1]$$

$$\theta = \arctan\left(\frac{\gamma B_1}{\Delta}\right) \quad [2]$$

where γB_1 is the applied spin-lock field and Δ is the resonance offset from the carrier. The contribution of chemical exchange or scalar interactions to the relaxation can then be determined by measuring R_2 as a function of the strength of the effective spin-locking field in the rotating frame (ω_{eff}):

$$\omega_{\text{eff}} = \sqrt{\Delta^2 + (\gamma B_1)^2} \quad [3]$$

In absence of other relaxation mechanisms, the R_1 , R_2 and NOE values are given by (17,18):

$$R_1 = Nd^2 [J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)] + CJ(\omega_C) \quad [4]$$

$$R_2 = Nd^2 [4J(0) + J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C) + 6J(\omega_H)] + C \left[\frac{2}{3}J(0) + \frac{1}{2}J(\omega_C) \right] + R_{ex} \quad [5]$$

$$\text{NOE} = 1 + \frac{\gamma_H}{\gamma_C} d^2 \left[\frac{N [6J(\omega_H + \omega_C) - J(\omega_H - \omega_C)]}{R_1} \right] \quad [6]$$

where $d^2 = (1/4) (\mu_0/4\pi) [\gamma_H\gamma_C (h/2\pi)]^2 (1/r_{CH}^6)$ and $C = (\omega_C Ds)^2/3$, μ_0 is the permittivity of free space, h is Planck's constant, γ_H and γ_C are the gyromagnetic ratios of ^1H and ^{13}C , r_{CH} is the ^1H - ^{13}C distance (1.11 Å), R_{ex} is the contribution from chemical exchange and is assumed to be proportional to ω^2 , and Ds is the breadth of an axially symmetric ^{13}C CSA. However, the effect of CSA on the calculated relaxation rates was verified to be negligible in the current study (assuming a value of 25 ppm for a methyl ^{13}C). The ^{13}C - ^{14}N dipole-dipole contributions to relaxation rates are also small and were neglected.

Relaxation rates were analyzed using a modified model-free approach, separating the fast and slow internal motions of the CH_3 group (13,19-21):

$$J(\omega) = \frac{2}{5} \left[\frac{S^2 \tau_m}{1 + (\omega \tau_m)^2} + \frac{(1 - S_f^2) \tau_1}{1 + (\omega \tau_1)^2} + \frac{S_f^2 (1 - S_s^2) \tau_2}{1 + (\omega \tau_2)^2} \right] \quad [7]$$

where τ_m is the overall isotropic rotational correlation time of the protein, S_f^2 is the order parameter for the fast rotation of the CH_3 group around the three-fold symmetry axis (equal to 0.111 for ideal tetrahedral geometry (20)), S_s^2 is the order parameter for the slower motion ($\tau_s < \tau_m$) of the ^{13}C - ^{14}N bond, τ_f and τ_s are the lifetimes of these fast and slow motions, respectively, and $S^2 = S_f^2 S_s^2$, $\tau_1^{-1} = \tau_m^{-1} + \tau_f^{-1}$ and $\tau_2^{-1} = \tau_m^{-1} + \tau_s^{-1}$.

MATERIALS AND METHODS

[U - ^{15}N]-labeled and [U - ^{13}C , ^{15}N]-labeled protein was expressed, purified, and BR-labeled as previously described (9). The NMR sample used for relaxation measurements was prepared as follows: the $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ complex was formed in solution using 0.8 mg of [U - ^{15}N]sNTnC•[^{13}C]BR₅₆₋₆₃, 1 μl of 1 M CaCl_2 and 1 mg of sTnI₁₁₅₋₁₃₁. The complex was diluted in 300 μl of NMR buffer (90% H_2O :10% D_2O , 250 mM KCl, and 10 mM imidazole), 6 μl protease inhibitor cocktail (Calbiochem), 5 μl of 0.23 M sodium azide, and 20 μl of 5 mM DSS- d_4 for referencing (Chenomx Inc., Canada). A pH of 6.65 was measured from the chemical shift of the imidazole H2 proton (22). The relatively high salt concentration was necessary to reduce protein dimerization, as previously described (9). The sample was subsequently diluted 3-fold with NMR buffer to further reduce dimerization and transferred to a 5 mm Shigemi tube (Shigemi Ltd., Japan). A higher concentration sample containing 3.0 mg of [U - ^{15}N , ^{13}C]sNTnC•2Ca²⁺•[^{13}C]BR₅₆₋₆₃ with excess sTnI₁₁₅₋₁₃₁ was used to acquire a 3D ^{13}C -NOESY-HSQC NMR spectrum.

The ^1H observed, ^{13}C NMR relaxation experiments were performed on 500, 600 and 800 MHz Varian INOVA spectrometers, the last of which was equipped with a cryogenic probe. The BioPack (Varian Inc.) gChsqc pulse sequence was used to record ^{13}C NMR spectra for the relaxation measurements (R_1 and $R_{1\rho}$). The pulse sequence used for the $\{^1\text{H}\}$ - ^{13}C NOE experiments was written in-house, based on a similar sequence (13) that was modified to employ pulsed field gradients and ^1H spin-lock purge pulses for water suppression. The BioPack (Varian Inc.) gNhsqc pulse sequence was used for measurement of ^{15}N relaxation of [U - ^{15}N]sNTnC, while the gnoesyChsqc pulse sequence was used for the acquisition of the 3D ^{13}C -NOESY-HSQC NMR spectrum. The ^{15}N relaxation rates were measured with 2D ^1H - ^{15}N HSQC spectra, while the ^{13}C relaxation rates were measured from 1D ^1H - ^{13}C HSQC spectra (Table III-1). All spectra were acquired at 303 K. The programs NMRPipe (23) and NMRView 5.2.2 (One Moon Scientific Inc.) were used to process and assign the 2D ^1H - ^{15}N HSQCs and the 3D ^{13}C -NOESY-HSQC. All other spectra were processed and analyzed with VnmrJ v2.1B (Varian Inc.).

The intensities of the resonances used to measure the relaxation rates were fitted to a mono-exponential decay. The errors reported for the relaxation rates were estimated from the non-linear least squares fitting routine within the in-house program xcrvfit (24). Errors on the NOE were calculated from the signal-to-noise of the respective spectra. To measure the ^{13}C - R_2 of the methyl groups, the off-resonance ^{13}C - $R_{1\rho}$ relaxation rates were measured using different ^{13}C carrier positions (35, 56, and 70 ppm) at three different magnetic field strengths. The ^{13}C spin-lock r.f. field strengths (γB_1) were adjusted for magnetic field, relative to the value employed at 600 MHz: $\gamma B_1 = 2000 \text{ Hz} / [({}^1\text{H frequency in MHz}) / 600 \text{ MHz}]$ (Table III-1).

RESULTS

The first step in analyzing the dynamics of the BR-label attached to sNTnC is characterization of the overall rotational motion of the protein. This step is important, since sNTnC is known to dimerize as the concentration increases. The average backbone amide ^{15}N - R_2 is directly proportional to molecular mass and can be used to evaluate the monomer:dimer ratio of sNTnC in the sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BR₅₆₋₆₃ complex in solution. We used the integral of the amide region from 1D CPMG-based ^1H - ^{15}N HSQC NMR spectra as a function of variable ^{15}N relaxation delays to determine the average ^{15}N - R_2 of the protein. The theoretical ^{15}N - R_2 based upon molecular weight is 8.3 s^{-1} for the monomer (9), corresponding to an expected macromolecular rotational correlation time of approximately 7.6 ns (25). The initial concentration of protein (0.3 mM) displayed a ^{15}N - R_2 value of 9.0 s^{-1} , indicating slight dimerization. We then diluted the sample 3-fold with NMR buffer and measured an ^{15}N - R_2 of 8.0 s^{-1} , corresponding to the expected value for monomeric protein. Consequently, this diluted sample was used to perform the ^{13}C relaxation experiments for the BR probe attached to sNTnC.

Fig. III-2A shows a 1D ^1H - ^{13}C HSQC NMR spectrum of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$, acquired in the absence of ^{13}C decoupling, which allows for a longer acquisition time, results in better signal-to-noise ratio and higher resolution in the spectrum, and allows for measurement of the $^1J_{\text{HC}}$ coupling constant for the observed resonances. The two high intensity peaks with ^1H chemical shifts of 3.12 and 3.35 ppm correspond to the two equivalent methyl groups of the BR probe, and are separated by $^1J_{\text{CH}} \sim 139$ Hz. The other narrow peaks in the spectrum arise from natural abundance ^{13}C of excess, unbound $\text{sTnI}_{115-131}$ peptide. Fig. III-2B shows a 1D ^1H - ^{13}C HSQC NMR spectrum in the presence of ^{13}C decoupling, with the methyl resonance(s) at 3.23 ppm. In Fig. III-2C, the 2D ^1H - ^{13}C HSQC spectrum shows two distinct resonances for the methyl groups with ^1H chemical shifts of ~ 3.23 ppm, and ^{13}C chemical shifts of 41.3 and 41.8 ppm. These different cross-peaks reflect presence of the two diastereoisomers that arise because of atropisomerism in the rhodamine structure (see ref. (4) for previous discussion), and are consistent with doubling of backbone amide resonances in 2D ^1H - ^{15}N HSQC NMR spectra of the troponin C for residues situated near the attachment points of the rhodamine probe (9,10). Given that the proton chemical shifts of the methyl resonances are almost identical, these two resonances are not distinguishable in 1D ^1H - ^{13}C HSQC NMR spectra (Fig. III-2B), and are manifested as weak shoulders on both peaks in the higher resolution spectrum shown in Fig. III-2A. Thus, the measured ^{13}C relaxation rates correspond to a combination of two methyl groups from both diastereoisomers. As shown in Fig. III-2C, a third resonance is present near 3.0 ppm, and arises from excess $\text{sTnI}_{115-131}$ peptide, as confirmed by the measurement of $^1J_{\text{CH}} = 142$ Hz, different from the value of 139 Hz for the methyl group of the BR-probe.

Fig. III-2D shows a slice ($^{13}\text{C} = 41.8$ ppm) from a 3D ^{13}C -NOESY-HSQC NMR spectrum. Cross peaks identified by particular marks correspond to NOEs between protons of the [^{13}C]methyl groups of the BR probe and other protons within $\sim 5\text{\AA}$. For example, the peaks marked * correspond to the methylene groups adjacent to the methyl-bearing nitrogen atoms, and those marked (1) and

(2) represent the NOEs with the H2',H7' and H1',H8' protons respectively, of the xanthene ring (see Fig. III-1). The weaker NOE to the H1',H8' protons likely arises from spin diffusion. The peak identified by an arrow corresponds to the diagonal cross-peak of the methyl groups. The same contacts are observable in the ^{13}C plane at 41.3 ppm. The NOE data are consistent with the conformation of the linker arms between rhodamine and the protein as shown in Fig. III-1A and III-1B. Additionally, NOE contacts between the ^{13}C -labeled methyl groups and the protein are not observed, supporting previous results (9).

To characterize the motion of the BR-label, we studied the ^{13}C relaxation rates of the methyl groups. We first measured the longitudinal ^{13}C - R_1 relaxation rate and the $\{^1\text{H}\}$ - ^{13}C NOE of the methyl groups (Table III-2). The NOE values correspond to the intensity ratio of the methyl group peaks measured with and without ^1H saturation, respectively. We obtained an average NOE of 1.5 ± 0.1 for the BR-probe methyl groups at the three magnetic field strengths. By comparison, the NOE measured for natural abundance methyl groups at 0.89 ppm (arising from excess sTnI₁₁₅₋₁₃₁) is 2.2 ± 0.1 (data not shown), the expected value for methyl groups of a peptide. To measure the ^{13}C - R_2 relaxation rate, off-resonance ^{13}C - $R_{1\rho}$ experiments were performed as a function of the effective spin-locking field in the rotating frame to determine whether scalar and/or exchange-broadening contributions to the transverse relaxation of the methyl groups on the BR probe were significant. The contribution from transverse relaxation was subsequently calculated from the measured off-resonance ^{13}C - $R_{1\rho}$ values. The calculated ^{13}C - R_2 values at 500, 600, and 800 MHz are plotted as a function of the effective field in Fig. III-3. No significant relaxation dispersion is observed, indicating that contributions from both scalar relaxation and chemical exchange are insignificant.

We analyzed the relaxation data following methodologies described in previous studies (13,18,26). A numerical optimization approach was used, in conjunction with Monte Carlo parameter error estimation (27). All of the data at three fields

were used for global optimization of the fit by varying the parameters τ_s , τ_f , and S_s^2 , with S_f^2 assumed to be 0.111. The rotational correlation time of the complex was fixed at 8.3 ns, as reported by Mercier et al. (9) and corroborated by our measurement of the average ^{15}N - R_2 backbone amide relaxation rate. We found that the data were better fit at the lower fields, and the fit improved by including a R_{ex} term (0.8 s^{-1} at 500 MHz) to compensate for chemical exchange broadening. Using these conditions, we performed two hundred Monte Carlo simulations yielding $\tau_f = 1.5 \pm 0.4 \text{ ns}$, $\tau_s = 0.8 \pm 0.4 \text{ ns}$, and $S_s^2 = 0.79 \pm 0.03$. To visualize how the individual relaxation measurements constrain the fit, contour plots corresponding to the experimental ^{13}C - R_1 , ^{13}C - R_2 and NOE at 500, 600 and 800 MHz are shown in Fig. III-4 as a function of S_s^2 and τ_f . The intersection where all of the experimental values are satisfied is colored in grey. Comparison of the three panels shows good agreement of the NMR data observed at different fields.

We also looked at the solvent accessibility of the backbone amides of $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$, using hydrogen exchange NMR experiments using a sample similar to the $[\text{U-}^{15}\text{N}]\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ sample but with the complex freshly dissolved in 99.9% D_2O instead of H_2O . A 2D ^1H - ^{15}N HSQC NMR spectrum was acquired for this sample after two hours and compared to $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$ in H_2O (Fig. III-5). The resonances in black correspond to the residues protected from exchange with the deuterated solvent. These residues are located in the middle of α -helical secondary structure elements. This is consistent with slower exchange for backbone amides involved in the hydrogen bond network of secondary structure. The presence of the BR probe does not confer additional protection of the amides of helix C over to those in other helices.

DISCUSSION

A quantitative understanding of the dynamics of the BR-label is important to the proper use of the fluorescence polarization method for determining orientation of

protein domains *in situ*. When this approach was used to determine the orientation of troponin C in muscle fibers (6,7), analysis of the fluorescence polarization measurements indicated that the rhodamine label was relatively immobile on a time scale faster than the fluorescence lifetime (~ 4 ns). This was evidenced by derived order parameters $\langle P_{2d} \rangle$ higher than 0.87 for different attachment sites, including attachment to the C helix (7). These order parameters imply limited motion of the probe and can be interpreted using the wobble-in-a-cone model with a half-angle of 19-23°. The ^{13}C NMR relaxation measurements presented herein provide an independent estimate of the mobility of the BR-label attached to the C helix of troponin C, free of the influence of protein-protein interactions within the thin filament. It is important to note that the fluorescence parameter $\langle P_{2d} \rangle$ is an order parameter, whereas the NMR parameter S^2 , often referred to as an order parameter, is actually the square of the order parameter.

^{13}C NMR relaxation measurements for methyl groups are theoretically and experimentally complicated. The relaxation involves cross-correlation between dipolar and CSA mechanisms, in addition to possible contributions from scalar relaxation and chemical exchange. We therefore followed previous recommendations to circumvent adverse effects of cross-correlated relaxation (11,28) and employed off-resonance $R_{1\rho}$ methods to measure accurate relaxation parameters. In addition, we used low protein concentrations to avoid complications arising from partial protein dimerization. The analysis of the ^{13}C relaxation data for the BR CH_3 groups using the rotational correlation time for the protein derived from the ^{15}N relaxation data ($\tau_m = 8.3$ ns) yielded $\tau_f = 1.5 \pm 0.4$ ps, $\tau_s = 0.8 \pm 0.4$ ns, and $S_s^2 = 0.79 \pm 0.03$. The value for τ_f is quite short when compared with methyl groups within proteins (29), but consistent with reduced steric crowding around the sp^2 hybridized nitrogens of the BR probe. The value for τ_s is not well defined, but it does not have any significant effect on the calculated S_s^2 (18). Of particular interest is the derived order parameter S_s^2 for the amplitude of spatial fluctuations of the N- CH_3 axis of the BR-label. This value is in very close agreement with the squared fluorescence order parameters $\langle P_{2d} \rangle$

($>0.87^2 = 0.76$), indicating that the motion of the CH₃ groups and the xanthen ring are very similar as expected from the partial double bond character of the =N⁺-CH₃ bond. The probe cannot be expected to be more rigid than the helix to which it is attached. The order parameter of the backbone amide NH vectors of the C helix are uniform and ~ 0.85 (30). Consequently, the small difference between the order parameters of the C helix and of the BR probe indicate that the BR probe is virtually immobile relative to the protein.

While this conclusion supports the utility of the overall bifunctional rhodamine approach, it might be considered somewhat surprising. The obvious question is how is the label immobilized, especially considering that it is attached to the protein by several rotatable bonds between the Cys alpha carbon and the rhodamine N-CH₃ moieties. Possibilities are: 1) the fluorophore remains extended from the protein surface and does not move away from its position; 2) the probe is flexible, but the motion is roughly parallel to the N-CH₃ axis and/or the fluorescence excitation dipole moment, and thus the motion is ‘invisible’ to the relaxation and/or fluorescence measurements, or 3) the label is immobilized by contact with the protein surface. The first option seems unreasonable on biophysical grounds, given the flexibility of the linkages to the label. The second option is possible, in whole or in part, given the attachment to two parallel residues on a helix. The third option is supported by molecular dynamics calculations which suggest that the probe forms an interaction with the protein surface (F. Fraternali, personal communication) and these results will be published elsewhere. However, NOE contacts between the methyl groups of the BR-label and the protein are not observed, and we do not observe any additional protection of main chain hydrogen exchange for residues within helix C of troponin C, to which the rhodamine is attached.

A separate maximum entropy analysis of the TnC bifunctional rhodamine fluorescence measurements indicated a distribution of label orientations of approximately 26° on a longer time scale than the fluorescence lifetime (10). The

angular distribution reflected in the polarized fluorescence data from BR-TnC in muscle fibers must come from some other source, such as the movement of the domain to which the fluorophore is attached. This is consistent with NMR relaxation measurements on a troponin core complex (31). In that study, the N-domain of TnC was seen to move independently from the core of the troponin complex when in the apo state. Taking the ratio of the ^{15}N - R_2 relaxation rates between apo and calcium-saturated forms of the troponin complex as an estimate of the order parameter S^2 , the authors obtained $R_2^{\text{apo}} / R_2^{\text{Ca}} = 0.65 - 0.70$. Analysis of these results with a ‘wobble-in-a-cone’ model (20,32), where $S^2_{\text{CONE}} = [(1/2) \cos(\theta_0) (1 + \cos(\theta_0))]^2$, gave an angle of approximately 30° which agrees with the fluorescence analysis. Recently Baber et al. (33) have applied a similar analysis to the motion of the two domains of calmodulin.

CONCLUSION

We have used NMR relaxation measurements to study the dynamics of a [^{13}C]methyl labeled bifunctional rhodamine probe attached to the C helix of the N-domain of troponin C and found no large range independent motion of the BR probe relative to the protein on a time scale shorter than the rotational correlation time of the TnC. This conclusion is in good agreement with fluorescence measurements which gave virtually identical order parameters for the motion of the dipole moment of the xanthene ring. The results presented herein support the utility of the overall bifunctional labeling strategy for determination of *in situ* protein domain orientations.

FIGURES AND TABLES

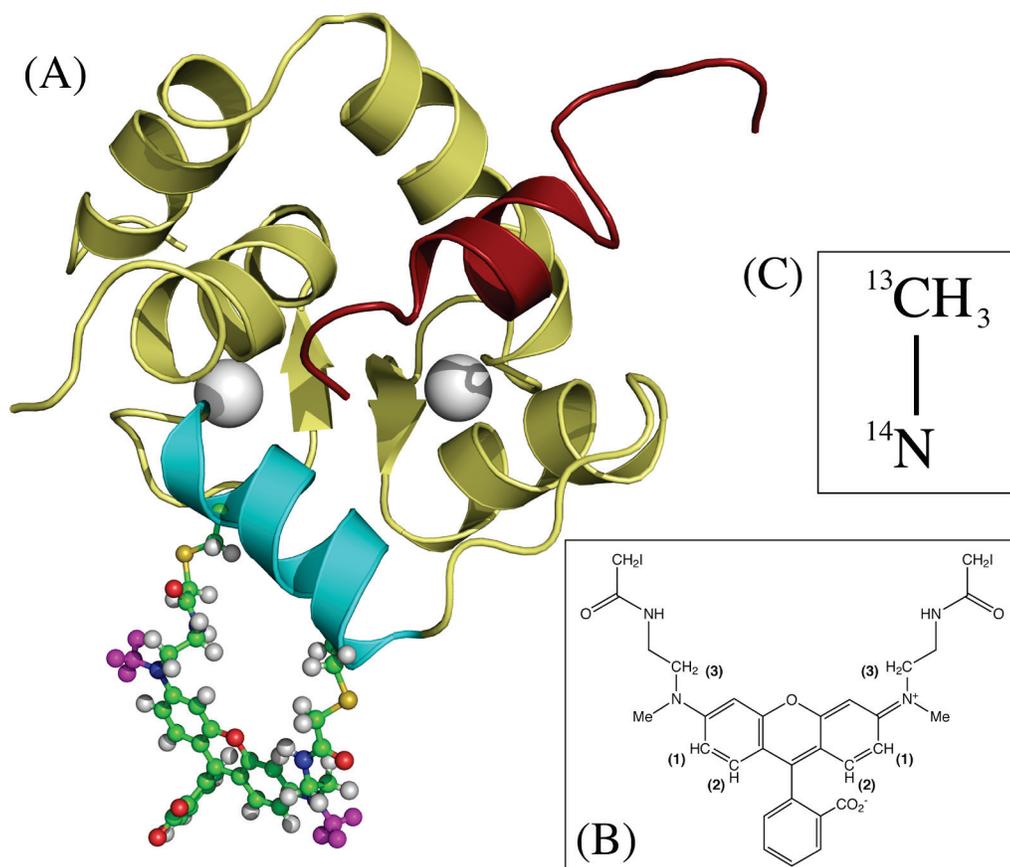


Fig. III-1. (A) Structure of the molecular complex $sNTnC \cdot 2Ca^{2+} \cdot TnI_{115-131} \cdot BR_{56-63}$ with the main chain atoms from $sNTnC_{1-90}$ and $sTnI_{115-131}$ colored yellow and red, respectively and shown in the cartoon representation, with Ca^{2+} ions colored white, and the BR probe (attached to residues E56C and E63C of $sNTnC$) shown in the ball and stick representation with atoms colored as follows: N blue, O red, H white, C green, S yellow, and the ^{13}C -labeled methyl groups in magenta. (B) Chemical structure of the BR probe. (C) Isotopic labeling of the two methyl groups from the BR probe.

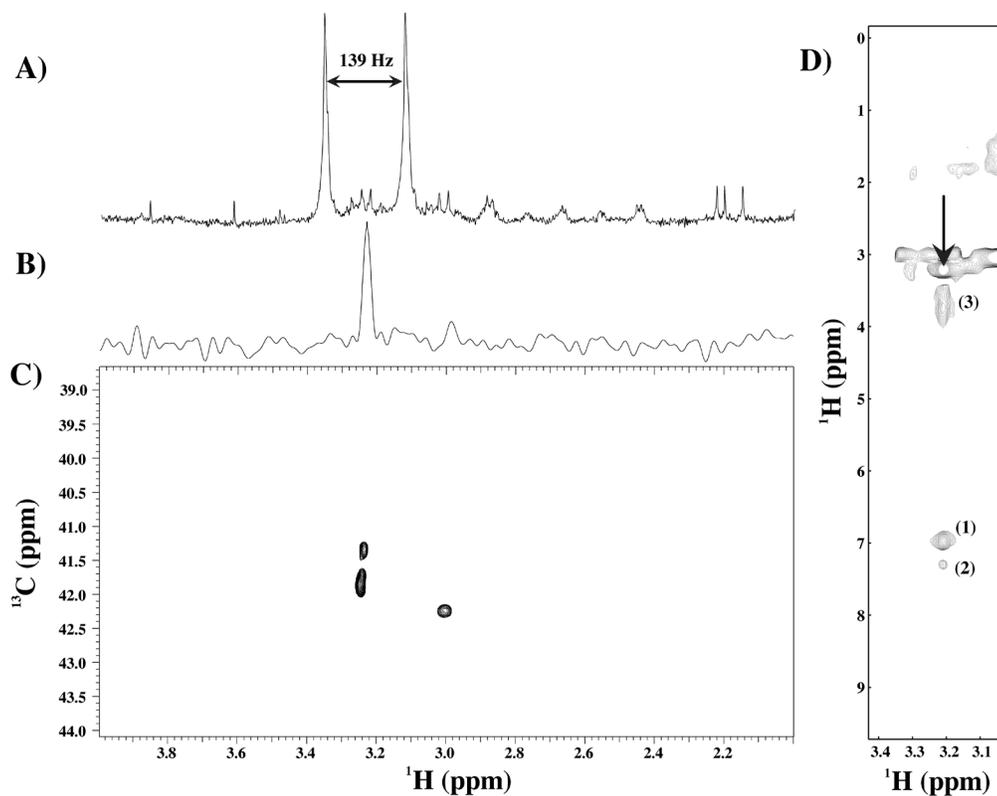


Fig. III-2. NMR spectra for sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BR₅₆₋₆₃. 1D ¹H-¹³C HSQC spectrum for [*U*-¹⁵N]sNTnC•2Ca²⁺•sTnI₍₁₁₅₋₁₃₁₎•[¹³C]BR₅₆₋₆₃ acquired at 600 MHz in the absence (A) and presence (B) of ¹³C decoupling. (C) 2D ¹H-¹³C HSQC of [*U*-¹⁵N]sNTnC•2Ca²⁺•sTnI₍₁₁₅₋₁₃₁₎•[¹³C]BR₅₆₋₆₃. (D) ¹³C plane (41.8 ppm) from a 3D ¹³C NOESY-HSQC for [*U*-¹⁵N,¹³C]sNTnC•2Ca²⁺•sTnI₍₁₁₅₋₁₃₁₎•[¹³C]BR₅₆₋₆₃ acquired at 800 MHz.

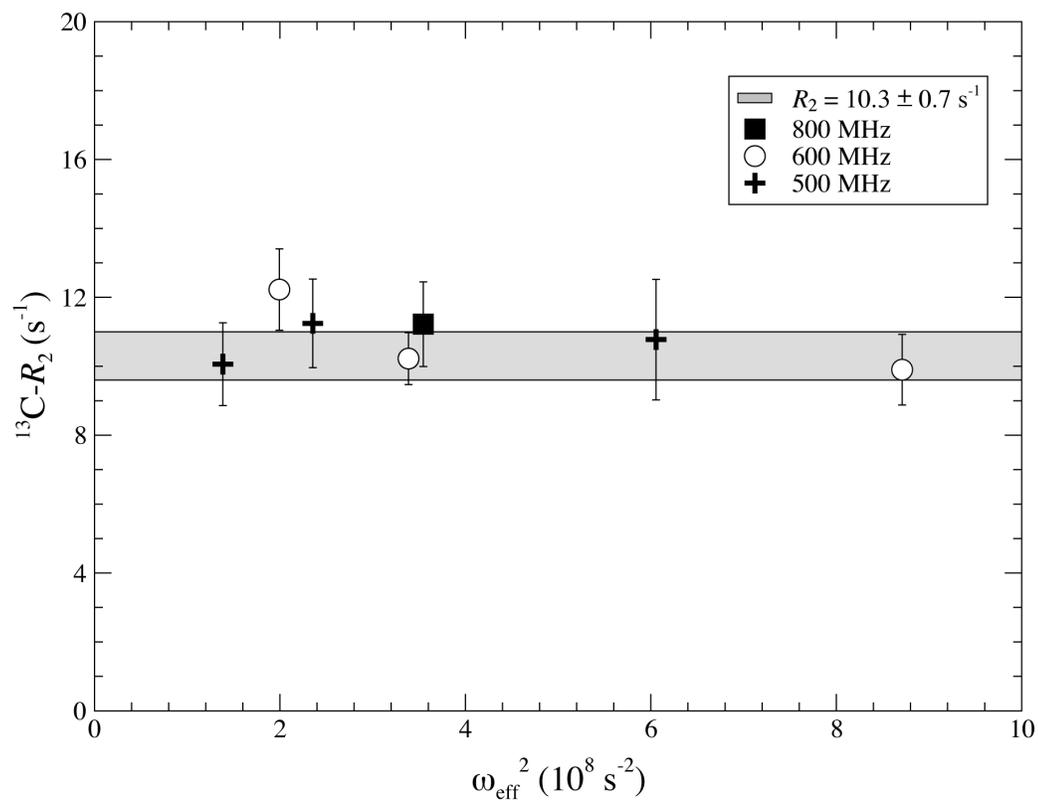


Fig. III-3. ^{13}C - R_2 values calculated from off-resonance $R_{1\rho}$ measurements of the BR probe as a function of the effective field (ω_{eff}^2) at 303 K. The shaded area indicates $\langle R_2 \rangle \pm 1\sigma$.

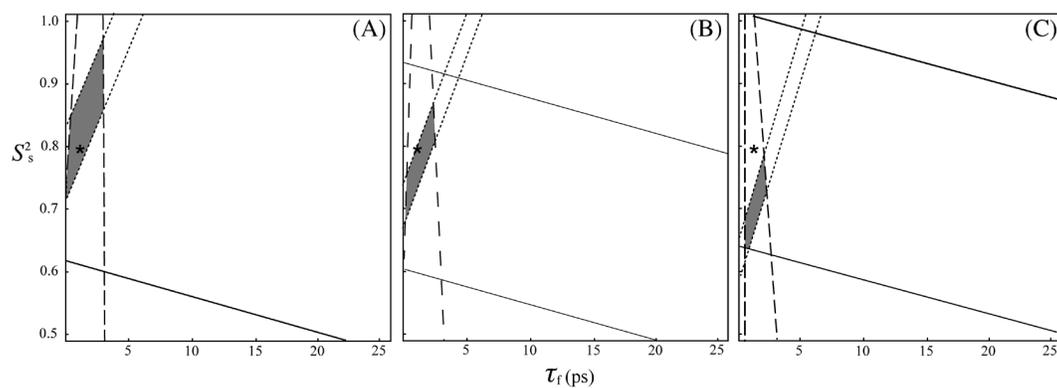


Fig. III-4. Contour plot showing calculated R_1 (---), R_2 (—) and NOE (•••) as a function of S_s^2 and τ_f at (A) 500 MHz (B) 600 MHz and (C) 800 MHz.

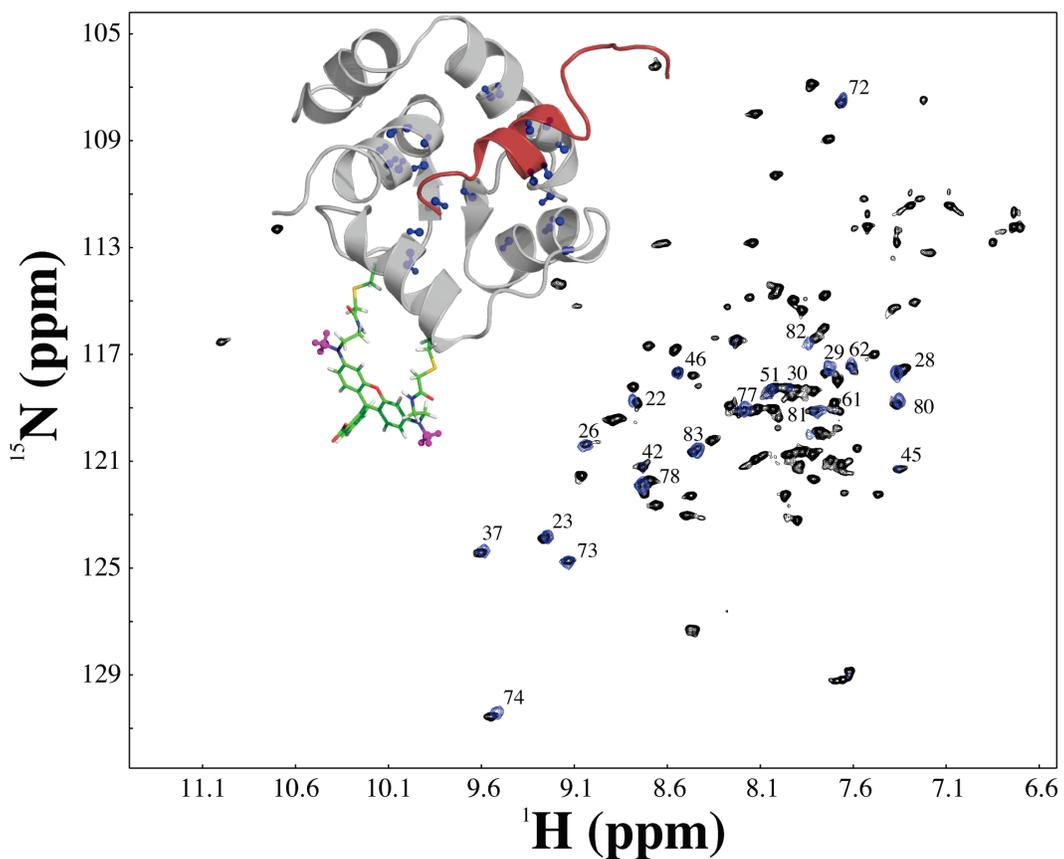


Fig. III-5. Main chain amide hydrogen exchange for $\text{sNTnC}\cdot 2\text{Ca}^{2+}\cdot \text{TnI}_{115-131}\cdot \text{BR}_{56-63}$. For the superimposed 2D ^1H - ^{15}N HSQC spectra, the black peaks correspond to main chain amides in the complex in 90% H_2O :10% D_2O and the blue peaks to slowly-exchanging amides that remain after 2 hours in 99.9% D_2O . The inset shows the structure of the complex, where amides protected from exchange with solvent are highlighted in blue. The spectra were acquired at 30 °C.

Table III-1. NMR Parameters Used to Acquire the Different ^{13}C Relaxation Experiments at 500, 600 and 800 MHz

Experiment (carrier)	Spectrometer (MHz)	γB_1 (Hz)	nt	at (s)	d1 (s)	Relaxation delay [s]; Interval (s)
R_1 (35 ppm)	500	-	1536	0.073	5.0	[0.1-1.7]; (0.2)
	600	-	2048	0.064	5.0	[0.1-1.0]; (0.1)
	800	-	2048	0.085	9.0	[0.1-1.5]; (0.2)
$R_{1\rho}$ (35 ppm)	500	1667	2560	0.073	3.0	[0.1-1.7]; (0.2)
	600	2000	256	0.064	5.0	[0.1-1.0]; (0.1)
	800	2667	2048	0.085	9.0	[0.1-1.5]; (0.2)
$R_{1\rho}$ (56 ppm)	500	1667	2560	0.073	3.0	[0.1-1.7]; (0.2)
	600	2000	2000	0.064	3.0	[0.1-1.0]; (0.1)
$R_{1\rho}$ (70 ppm)	500	1667	2560	0.073	3.0	[0.1-1.7]; (0.2)
	600	2000	3072	0.064	3.0	[0.1-1.0]; (0.1)
$\langle^1\text{H}\rangle\text{-}^{13}\text{C}$	500	-	8192	0.051	> 8.5	-
NOE	600	-	4096	0.051	> 8.5	-
	800	-	4096	0.051	> 8.5	-

* The following parameters were also used: $\lambda = 0.0011$, $t_{\text{CH}} = 0.0011$ and the sensitivity enhancement was turned off.

Table III-2. Longitudinal ^{13}C Methyl Relaxation Rates and $\langle^1\text{H}\rangle$ - ^{13}C NOE

Spectrometer (MHz)	R_1 (s^{-1})	NOE
500	0.65 ± 0.04	1.5 ± 0.1
600	0.57 ± 0.02	1.5 ± 0.1
800	0.46 ± 0.01	1.5 ± 0.1

REFERENCES

1. Brune, M., Hunter, J., Howell, S., Martin, S., Hazlett, T., Corrie, J., and Webb, M. (1998) Mechanism of inorganic phosphate interaction with phosphate binding protein from *Escherichia coli*, *Biochemistry* 37, 10370-10380.
2. Hirshberg, M., Henrick, K., Haire, L. L., Vasisht, N., Brune, M., Corrie, J. E. T., and Webb, M. R. (1998) Crystal structure of phosphate binding protein labeled with a coumarin fluorophore, a probe for inorganic phosphate, *Biochemistry* 37, 10381-10385.
3. Munasinghe, V. R. N., Corrie, J. E. T., Kelly, G., and Martin, S. R. (2007) Fluorescent ligands for the hemagglutinin of influenza A: Synthesis and ligand binding assays, *Bioconjug. Chem.* 18, 231-237.
4. Corrie, J. E. T., Brandmeier, B. D., Ferguson, R. E., Trentham, D. R., Kendrick-Jones, J., Hopkins, S. C., Van Der Heide, U. A., Goldman, Y. E., Sabido-David, C., Dale, R. E., Criddle, S., and Irving, M. (1999) Dynamic measurement of myosin light-chain-domain tilt and twist in muscle contraction, *Nature* 400, 425-430.
5. Hopkins, S. C., Sabido-David, C., Van Der Heide, U. A., Ferguson, R. E., Brandmeier, B. D., Dale, R. E., Kendrick-Jones, J., Corrie, J. E. T., Trentham, D. R., Irving, M., and Goldman, Y. E. (2002) Orientation changes of the myosin light chain domain during filament sliding in active and rigor muscle, *J. Mol. Biol.* 318, 1275-1291.
6. Ferguson, R. E., Sun, Y. B., Mercier, P., Brack, A. S., Sykes, B. D., Corrie, J. E. T., Trentham, D. R., and Irving, M. (2003) In situ orientations of protein domains: troponin C in skeletal muscle fibers, *Mol. Cell* 11, 865-874.
7. Sun, Y. B., Brandmeier, B., and Irving, M. (2006) Structural changes in troponin in response to Ca²⁺ and myosin binding to thin filaments during activation of skeletal muscle, *Proc. Natl. Acad. Sci. U. S. A.* 103, 17771-17776.

8. Dale, R. E., Hopkins, S. C., Van Der Heide, U. A., Marszalek, T., Irving, M., and Goldman, Y. E. (1999) Model-independent analysis of the orientation of fluorescent probes with restricted mobility in muscle fibers, *Biophys. J.* 76, 1606-1618.
9. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E. T., Trentham, D. R., and Sykes, B. D. (2003) NMR structure of a bifunctional rhodamine labeled N-domain of troponin C complexed with the regulatory "switch" peptide from troponin I: implications for in situ fluorescence studies in muscle fibers, *Biochemistry* 42, 4333-4348.
10. Julien, O., Sun, Y. B., Knowles, A. C., Brandmeier, B. D., Dale, R. E., Trentham, D. R., Corrie, J. E. T., Sykes, B. D., and Irving, M. (2007) Toward protein structure in situ: comparison of two bifunctional rhodamine adducts of troponin C, *Biophys. J.* 93, 1008-1020.
11. Kay, L. E., Bull, T. E., Nicholson, L. K., Griesinger, C., Schwalbe, H., Bax, A., and Torchia, D. A. (1992) The Measurement of Heteronuclear Transverse Relaxation-Times in Ax³ Spin Systems Via Polarization-Transfer Techniques, *J. Magn. Reson.* 100, 538-558.
12. Sklenar, V., Torchia, D., and Bax, A. (1987) Measurement of C-13 Longitudinal Relaxation Using H-1 Detection, *J. Magn. Reson.* 73, 375-379.
13. Nicholson, L. K., Kay, L. E., Baldisseri, D. M., Arango, J., Young, P. E., Bax, A., and Torchia, D. A. (1992) Dynamics of Methyl-Groups in Proteins as Studied by Proton-Detected C-13 Nmr-Spectroscopy - Application to the Leucine Residues of Staphylococcal Nuclease, *Biochemistry* 31, 5253-5263.
14. Palmer, A. G., Wright, P. E., and Rance, M. (1991) Measurement of Relaxation-Time Constants for Methyl-Groups by Proton-Detected Heteronuclear NMR-Spectroscopy, *Chem. Phys. Lett.* 185, 41-46.
15. 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)N transverse relaxation, *J. Am. Chem. Soc.* 129, 2971-2980.

16. Deverell, C., Morgan, R. E., and Strange, J. H. (1970) Studies of Chemical Exchange by Nuclear Magnetic Relaxation in Rotating Frame, *Mol. Phys.* 18, 553-559.
17. Abragam, A. (1961) *The Principles of Nuclear Magnetism*, pp 289-305, Clarendon Press, Oxford.
18. Lee, A. L., Flynn, P. F., and Wand, A. J. (1999) Comparison of H-2 and C-13 NMR relaxation techniques for the study of protein methyl group dynamics in solution, *J. Am. Chem. Soc.* 121, 2891-2902.
19. 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.
20. 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.
21. Lipari, G., and Szabo, A. (1982) Model-Free Approach to the Interpretation of Nuclear Magnetic-Resonance Relaxation in Macromolecules.2. Analysis of Experimental Results, *J. Am. Chem. Soc.* 104, 4559-4570.
22. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J Biomol NMR* 41, 5-7.
23. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
24. <http://www.bionmr.ualberta.ca/bds/software/xcrvfit>.
25. Spyropoulos, L., Gagné, S. M., and Sykes, B. D. (2001) Know your protein: effects of weak dimerization, in *Dynamics, Structure, and Function of Biological Macromolecules* pp 37-44, Jardetzky O., Finucane, M. D., Eds.; IOS Press.
26. Ishima, R., Petkova, A. P., Louis, J. M., and Torchia, D. A. (2001) Comparison of methyl rotation axis order parameters derived from model-

- free analyses of H-2 and C-13 longitudinal and transverse relaxation rates measured in the same protein sample, *J. Am. Chem. Soc.* 123, 6164-6171.
27. Spyropoulos, L. (2006) A suite of Mathematica notebooks for the analysis of protein main chain ¹⁵N NMR relaxation data, *J. Biomol. NMR* 36, 215-224.
 28. Massi, F., Grey, M. J., and Palmer, A. G. (2005) Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1rho relaxation experiments, *Protein Sci.* 14, 735-742.
 29. Xue, Y., Pavlova, M. S., Ryabov, Y. E., Reif, B., and Skrynnikov, N. R. (2007) Methyl rotation barriers in proteins from H-2 relaxation data. Implications for protein structure, *J. Am. Chem. Soc.* 129, 6827-6838.
 30. Gagné, S. M., Tsuda, S., Spyropoulos, 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, 667-686.
 31. 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, 21924-21932.
 32. Woessner, D. E. (1962) Nuclear Spin Relaxation in Ellipsoids Undergoing Rotational Brownian Motion, *J. Chem. Phys.* 37, 647-654.
 33. Baber, J. L., Szabo, A., and Tjandra, N. (2001) Analysis of slow interdomain motion of macromolecules using NMR relaxation data, *J. Am. Chem. Soc.* 123, 3953-3959.

CHAPTER IV – Trp mutants of TnC*Tryptophan Mutants of Cardiac Troponin C:
3D Structure, Troponin I Affinity, and in situ Activity*

Olivier Julien,¹ Yin-Biao Sun,² Xu Wang,¹ Darrin A. Lindhout,¹
Angela Thiessen,¹ Malcolm Irving,² and Brian D. Sykes¹

¹ *Department of Biochemistry, University of Alberta, Edmonton, Canada.*

² *Randall Division of Cell and Molecular Biophysics, King's College London, London, United Kingdom.*

Publication: Julien et al. 2008. *Biochemistry*, 47:597-606.

Contributions: OJ and BDS designed the NMR experiments. OJ acquired the NMR spectra, assigned the protein chemical shifts, analyzed the relaxation data, solved the NMR and wrote the manuscript with BDS. OJ made all the figures. YBS and MI performed the physiological experiments. DAL and AT cloned the mutants. AT and MC expressed and purified the protein.

OVERVIEW

In situ fluorescence/NMR spectroscopic approaches have been used to elucidate the structure, mobility, and domain orientations of troponin C in striated muscle. This led us to consider complementary approaches such as solid-state NMR spectroscopy. The biophysical properties of tryptophan and Trp-analogues, like fluorotryptophan or hydroxytryptophan, are often exploited to probe protein structure and dynamics using solid-state NMR or fluorescence spectroscopy. We have characterized Phe-to-Trp mutants in the ‘structural’ C-domain of cardiac troponin C, designed to immobilize the indole ring in the hydrophobic core of the domain. The mutations and their fluorinated analogues (F104W, F104(5fW), F153W, and F153(5fW)) were shown not to perturb the structural properties of the

protein. In this paper, we characterize the mutations F77W and F77W-V82A in the ‘regulatory’ N-domain of cardiac troponin C. We used NMR to determine the structure and dynamics of the mutant F77W-V82A-cNTnC, which shows a unique orientation of the indole ring. We observed a decrease in calcium binding affinity and a weaker affinity for the switch region of TnI for both mutants. We present force recovery measurements for all of the N- and C-domain mutants reconstituted into skeletal muscle fibers. The F77W mutation leads to a reduction of the *in situ* force recovery, whereas the C-domain mutants have the same activity as the wild type. These results suggest that the perturbations of the N-domain caused by the Trp mutation disturb the interaction between TnC and TnI, which in turn diminishes the activity in fibers, providing a clear example of the correlation between *in vitro* protein structures, their interactions, and the resulting *in situ* physiological activity.

INTRODUCTION

Many, if not most, proteins exist as parts of large, complex bio-macromolecular assemblies such as fibers, viruses, cell walls, nano-motors, and membrane surface receptors. One of the most studied of these systems is striated muscle, where structures of many important isolated components involved in force generation and calcium regulation have been determined. The wealth of troponin structures includes X-ray and NMR structures of the apo and calcium saturated versions of troponin C from skeletal and cardiac muscle (1-6) of troponin C in complex with fragments of the inhibitory target protein TnI (7,8); and X-ray structures of the core regions of the skeletal and cardiac troponin complexes (9,10). These *in vitro* structures provide a detailed molecular description of many of the important interactions amongst the regulatory proteins. It is, however, very important but typically much more difficult to determine the structures of proteins in their native environments in order to elucidate the molecular details of their mechanism of action. One approach which has been exploited quite successfully in the motility area uses fluorescence spectroscopy of bifunctional rhodamine labeled proteins to

determine the *in situ* orientation of domains of proteins whose structures are known *in vitro*. This approach has been used to determine the orientations of the domains of the bi-lobal myosin regulatory light chain and troponin C proteins reconstituted into skeletal muscle fibers (11,12). This data has subsequently been used in combination with other studies such as cryoelectron microscopy to build detailed model structures positioning the major protein components in the intact muscle fiber (13,14).

The bifunctional rhodamine label is, however, a large, hydrophobic, and potentially mobile probe. We have sought to develop an intrinsic, more rigid, non-perturbing probe that might be used with either fluorescence or solid-state NMR spectroscopies. The spectroscopic properties of tryptophan as a probe are appropriate, and tryptophan is relatively scarce in proteins, so that single tryptophan mutants have been widely used in biophysical studies. These include the Trp F78W and F154W mutants of skeletal TnC, which are analogous to the mutants described herein (15), and many others (15-17) including the widely used F29W (18). In many cases, however, the design strategy was the opposite to ours; to have the largest spectroscopic change resulting from the large conformational change concomitant with calcium binding. Tryptophan analogues such as fluorotryptophan and hydroxytryptophan also have special spectroscopic properties that can be exploited using solid-state NMR (19) and fluorescence (20) spectroscopies, respectively. We have previously characterized two Phe-to-Trp mutants in the C-terminal domain of cTnC (21). The mutations were designed to immobilize the indole ring within the hydrophobic core of the EF-hand calcium binding domain, by replacing the phenylalanine residue anchoring the calcium binding loops of sites III and IV. These single mutants of cTnC and their fluorinated analogues (F104W, F104(5fW), F153W, and F153(5fW)) were shown to have very similar overall structures compared to the wild type protein with their indole ring immobilized (21), making them potentially useful minimally perturbing probes of *in situ* domain orientation. No structure exists for a similar mutant in the regulatory domain of TnC, nor have any of these mutations been

characterized in terms of their interactions with the appropriate target regions of TnI. At the highest level, it is important to characterize the physiological activity (i.e., force development) of these proteins when reconstituted into muscle fibers.

The focus of this paper is the biophysical characterization and structure determination of the calcium binding site II mutant F77W in the regulatory domain of cTnC. In the course of this work, we discovered by NMR that we had unintentionally expressed the double mutant F77W-V82A-cNTnC. Consequently, the calcium and cTnI switch peptide (cSp) affinities were studied for both F77W and F77W-V82A-cNTnC in this paper, as well as the determination of the three dimensional solution structure of F77W-V82A-cNTnC•Ca²⁺ using multinuclear, multidimensional NMR spectroscopy. Finally, we present the *in situ* force recovery (activity) of all of the Phe-to-Trp mutants reconstituted in skeletal muscle fibers, including 4fW and 5fW analogues. The results revealed a correlation between the structural changes introduced by the tryptophan mutation in the regulatory N-domain, the affinity changes for the binding of the ‘switch’ region of cTnI, and the concomitant changes in the maximal force developed *in situ* in muscle fibers. No such perturbation is observed for the mutations made in the C-terminal domain of cTnC, consistent with its stronger interactions with cTnI and its structural role in the mechanism.

MATERIALS AND METHODS

Protein and Peptide Nomenclature

Full length human cardiac troponin C is referred as cTnC (residues 1-161). The N-domain and C-domain of cTnC are referred as cNTnC (residues 1-89) and cCTnC (residues 91-161). The Phe-to-Trp mutation at position 77 in cTnC is referred as F77W-cTnC for the full length and F77W-cNTnC for the N-domain. cTnC containing the mutations Phe 77 to Trp and Val 82 to Ala is referred as F77W-V82A-cTnC for the full length protein and F77W-V82A-cNTnC for the N-domain. The human cardiac troponin I is referred as cTnI. The specific region of

cTnI containing residues 147-163 is referred as cTnI₁₄₇₋₁₆₃ or cSp (cardiac switch peptide): (Ac)-RISADAMMQALLGARAK-(amide).

Protein Expression and Purification of TnC and cNTnC Mutants

The plasmid DNA pET3a-F77W-cNTnC (1-89), pET3a-F77W-V82A-cNTnC (1-89), pET3a-F77W-cTnC (1-161) and pET3a-F77W-V82A-cTnC (1-161) were used to transform BL21(DE3)pLysS cells. The expressions of unlabeled, [¹⁵N]-labeled and [¹⁵N, ¹³C]-labeled protein in *E. coli* were as described by Gagné et al. (22) and Li et al. (23). Purification of the proteins followed the previously published protocol for TnC. Decalcification was done as described by Li et al. (23).

Calcium Titrations

The calcium binding affinity of cNTnC, F77W-cNTnC and F77W-V82A-cNTnC were measured via competition with a chromophoric chelator. Ca²⁺-free buffer (100mM KCl, 50mM MOPS, pH=7.5) was prepared as described by Linse (24). The calcium binding affinity of the chromophoric chelator (BAPTA) was measured at $\lambda = 238.5$ nm ($K_d = 0.5 \pm 0.3$ μ M, $n = 5$). The binding constant of each protein was then determined by calcium titration of the protein in presence of BAPTA. The absorbance changes were verified to come from the chelator only at the wavelength specified above. The protein concentrations (from amino acid analysis) were kept as close as possible to the chelator concentration used: BAPTA (44 μ M), cNTnC (48 μ M), F77W-cNTnC (45 μ M) and F77W-V82A-cNTnC (50 μ M). The calcium binding constants of the proteins were extracted from the fitting of the normalized absorbance at each point of titration plotted as a function of the calcium concentration, using an in-house script, following the procedures outlined by Linse (24).

cSp Titrations

1.0 mg of ¹⁵N-labeled protein were dissolved in 550 μ l of NMR buffer (90% H₂O:10% D₂O and 100 mM KCl) for a final protein concentration of 0.1 mM

(from amino acid analysis). To the sample were added: CaCl₂ (5.0 mM), DTT (12.5 mM), 0.03 % sodium azide and protease inhibitors to prevent sample degradation. DSS was added to reference the spectra. The pH was adjusted to 6.8 according to the imidazole signal from the 1D spectrum. A 2D ¹H-¹⁵N HSQC, with 512 (t₂) x 128 (t₁) complex points, was acquired for every of step of the cSp titration. Solid peptide was weighed and added directly into the NMR sample at every step; cSp has insufficient water solubility to allow the preparation of a solution stock concentrated. The following cSp concentrations were used for F77W-V82A-cNTnC: 0.00, 0.10, 0.20, 0.31, 0.42, 0.51, 0.60, 0.77, 1.01, 1.41, 2.29 and 3.70 mM of cSp. For F77W-cNTnC: 0.00, 0.10, 0.23, 0.37, 0.52, 0.69, 1.03, 1.51, 3.11, 5.77 mM of cSp. The spectra were acquired on a Varian INOVA 600MHz spectrometer.

¹⁵N-Relaxation Measurements

1 mg of [¹⁵N]-labeled protein was diluted in 550 µl of NMR buffer. Deuterated-DSS, DTT (12.5 mM), CaCl₂ (5 mM), 0.03 % of sodium azide and protease inhibitors were also added. The pH was adjusted to 6.84 according to the imidazole signal. The protein concentration was determined to be 0.2 mM from an amino acid analysis. For both samples, the ¹⁵N-transverse relaxation rates were measured by acquiring six 2D ¹H-¹⁵N HSQC NMR spectra using the BioPack pulse sequence with different T₂ relaxation time delays (τ = 10, 30, 50, 70, 90, 110 ms). To measure the ¹⁵N-longitudinal relaxation rates, seven ¹H-¹⁵N HSQC spectra were acquired with the following time delays: τ = 10, 50, 100, 200, 300, 500 and 800 ms. The ¹H-¹⁵N HSQC spectra were all collected with 512 (t₂) x 128 (t₁) complex points and the relaxation delay between pulse sequences was set to 3.0 s. The spectra were acquired on a Varian INOVA 600MHz NMR spectrometer.

Data acquisition, Chemical Shift Assignments and NMR restraints

A [¹⁵N-¹³C]-F77W-V82A-cNTnC sample was prepared similarly to the one made for the acquisition of the ¹⁵N- relaxation data, but a higher protein concentration

(~1.5 mM) was used in addition to 19% TFE (v/v). This sample was used to acquire the spectra needed for the chemical shift assignments and the structure determination of F77W-V82A-cNTnC•Ca²⁺. The spectra were all processed using NMRPipe (25). The chemical shifts were assigned using smartnotebook v5.1.3 (26) and NMRView 5.2.2 (One Moon Scientific, Inc.). The following spectra were used for the chemical shift assignments: 2D ¹H-¹⁵N HSQC, 3D ¹H-¹⁵N TOCSY-HSQC, 3D CBCACONNH, 3D HNCACB for the backbone and 3D CCONH, 3D HCCONH, 3D HCCH-TOCSY and two 2D ¹H-¹³C HSQC (aliphatic and aromatic) for the side chains. The structural restraints were obtained from a 3D ¹⁵N NOESY-HSQC and two 3D ¹H-¹³C NOESY-HSQC (aliphatic and aromatic) for the NOEs, and a 3D-HNHA to measure the ³J_{HNH α} coupling constants (phi angles).

Structure Calculation

The structure calculations were performed using CYANA 2.1 (27). Most of the NOEs were assigned manually and calibrated within CYANA. The noeassign procedure in cyana assigned 235 out of 1239 NOEs. The minimum and maximum NOE calibration values were set to 1.8 and 6.0 Å respectively. The program was forced to keep all of the manually assigned NOEs for the seven runs (8000 steps) of calculation. In addition, 8 calcium binding restraints (based on calcium binding site homology), 103 dihedral restraints from TALOS (28), and 68 ³J_{HNH α} were used to obtain the NMR ensemble (see Table IV-1). The structure validation was performed using procheck v.3.5.4. (29). No further refinement was necessary in order to obtain this well-resolved NMR ensemble.

Muscle fibers and solutions

Adult New Zealand white rabbits were killed by sodium pentobarbitone injection (200 mg kg⁻¹). Small fiber bundles were dissected from the psoas muscle, demembrated, and stored for up to four weeks in relaxing solution containing 50 % (v/v) glycerol at -20°C (30). Single fiber segments 2.5–3.5 mm long were dissected in the above storage solution on a cooled microscope stage and

mounted, via aluminum T-clips, at sarcomere length 2.4 μm between a force transducer (AE801, Memscap, Bernin, France) and a fixed hook, in a 60- μl glass trough containing relaxing solution. The experimental temperature was $10.0 \pm 0.5^\circ\text{C}$. All experimental solutions used in muscle fiber experiments contained 5 mM Mg-ATP, 1 mM free Mg^{2+} , and 25 mM imidazole, except where noted. Ionic strength was adjusted to 150 mM by addition of potassium propionate (KPr), and the pH was 7.1 at 10°C . The relaxing solution also contained 10 mM EGTA. The activating solution contained 10 mM Ca-EGTA (pCa 4.5, where $\text{pCa} = -\log[\text{Ca}^{2+}]$). The pre-activating solution contained 0.2 mM EGTA. TnC-extraction solution contained 0.5 mM trifluoperazine (TFP), 20 mM MOPS, 5 mM EDTA and 130 mM KPr, pH 7.1 at 10°C .

TnC extraction and reconstitution

Ca-activated force prior to TnC extraction (T_0) was determined at pCa 4.5. Endogenous skeletal TnC was selectively extracted from single glycerinated rabbit psoas muscle fibers by 10 cycles of 30-s incubations in TnC-extraction solution followed by 30 s in relaxing solution (14). Fibers were reconstituted with cTnC by bathing them in relaxing solution containing 1 - 2 mg/ml cTnC for up to 20 min at 10°C . No further increase in active force could be detected with subsequent incubations in cTnC.

RESULTS

The focus of this paper is the biochemical and structural characterization of the calcium binding site II mutant F77W in the regulatory N-domain of cTnC, and the physiological characterization of a series of Trp mutants in both the structural and regulatory domains. Calcium binding site I in cTnC is naturally defunct and flexible, and therefore the mutant F27W was not studied. Midstream in our work, we discovered by NMR that we had unintentionally expressed the double mutant F77W-V82A-cTnC. Consequently, the F77W single mutant was re-cloned, expressed and labeled, and the NMR and biochemical measurements including

calcium and cTnI switch peptide affinities were studied for both F77W- and F77W-V82A-cTnC. The results show that the V82A mutation has no consequence on any of the biophysical properties of the protein, which is supported by the *in situ* fiber force recovery measurements conducted on both F77W mutants. The 2D ^1H - ^{15}N HSQC NMR spectra at 600 MHz of F77W- and F77W-V82A-cTnC• Ca^{2+} are presented in Fig. IV-1. These NMR spectra were run at 100 μM protein where dimerization (see below) is negligible. Residue specific assignments are indicated. Since amide ^1H and ^{15}N NMR chemical shifts are very sensitive to minute changes in structure, chemical environment and dynamics, the comparison of the spectra for F77W- and F77W-V82A-cTnC• Ca^{2+} shows in exquisite detail on a residue specific basis that the two structures are virtually identical. Consequently the NMR structure and dynamics are presented only for F77W-V82A-cTnC• Ca^{2+} since replication would provide little additional information.

Dimerization of the N-domain of Troponin C

It is known that the N-domain of TnC dimerizes weakly in solution at typical NMR concentrations, and that the tendency to dimerize is greater for skeletal than cardiac NTnC ($K_{\text{dimer}} = 1.3$ vs 7.3 mM, respectively) (31). To evaluate the monomer:dimer equilibria for F77W- and F77W-V82A-cTnC• Ca^{2+} , we measured backbone amide ^{15}N - R_2 relaxation rates as a function of protein concentration ranging from 0.11 to 3.4 mM. Several peaks were impossible to identify at higher concentrations, but appeared as unambiguous resonances at low concentration. This is the result of extensive exchange broadening of the NMR cross peaks. Analysis of the data gave a K_{dimer} of approximately 5 mM for both proteins, indicating that the mutation V82A is not responsible for the increase of dimerization. The increased dimerization reflects a more open structure and exposed hydrophobic domain (see below). Since the weak dimerization affects the observed calcium and ligand affinities, all interactions for both proteins were studied at low protein concentrations. Multinuclear 3D NMR structural data was collected at higher concentrations in the presence of the co-solvent TFE, used as a denaturant of quaternary structure (5,32).

Calcium and cSp affinities of F77W mutants

Calcium and cSp affinities of both mutants were measured to determine whether the presence of the Trp disturbs the interaction with calcium, the coupled ‘closed’ to ‘open’ conformational equilibrium, or the subsequent binding of cTnI in the hydrophobic pocket of the calcium saturated cNTnC. The calcium dissociation constants were measured using an optical method in which the apparent calcium binding affinity of the chromophoric chelator BAPTA is measured in the absence and presence of the calcium binding protein, and the difference attributed to competition from binding to the protein. These measurements are performed at protein concentrations $< 50 \mu\text{M}$ where dimerization is negligible (data not shown). A K_d of $4 \pm 3 \mu\text{M}$ was obtained for wild type cNTnC, and K_d 's of 15 ± 5 and $18 \pm 5 \mu\text{M}$ for the mutants F77W and F77W-V82A, respectively.

The cSp affinity of both mutants were measured by following the chemical shift changes of the amide resonances in 2D ^1H - ^{15}N 2D HSQC NMR spectra of cNTnC•Ca²⁺ acquired during titration with cSp. These NMR spectra were run at $100 \mu\text{M}$ protein concentration where dimerization is negligible. This also allows for the asymptote of the titration to be well defined, resulting in an accurate determination of the dissociation constants. The cSp titrations for F77W- and F77W-V82A-cNTnC•Ca²⁺ are shown as overlay NMR plots in Fig. IV-2. The comparison of the spectra at the beginning and end of both titrations shows that nearly identical structural changes occur in both proteins upon cSp binding. Seven residues showing unambiguous chemical shifts in both cSp titrations were followed: Ala31, Glu66, Gly30, Gly42, Leu29, Ser37 and Thr71. The average chemical shift changes were plotted as a function of the ratio of the total cSp concentration over the protein concentration (see inserts Fig. IV-2). A K_d of $456 \pm 35 \mu\text{M}$ was measured for the mutant F77W-V82A-cNTnC•Ca²⁺ and a K_d of $470 \pm 35 \mu\text{M}$ for F77W-cNTnC•Ca²⁺. Those values are three times higher than the K_d obtained by Li et al. (33) for cNTnC ($154 \pm 10 \mu\text{M}$), demonstrating the weaker affinity of the Phe-to-Trp mutants for cTnI.

Characterization of the dynamics of F77W-V82A-cNTnC•Ca²⁺

The backbone amide ¹⁵N-relaxation rates (¹⁵N-R₁ and ¹⁵N-R₂) of F77W-V82A-cNTnC•Ca²⁺ were determined on a per residue basis at a protein concentration of 200 μM (Fig. IV-3). The comparison of those relaxation rates with the ones observed by Spyropoulos et al. (34) for the wild type protein does not reveal any major differences in the profile of R₁ and R₂ values: the N- and C-termi are more flexible than the rest of the protein (low R₁ and R₂), the defunct calcium binding site I is also more flexible than the functional binding site II, and the secondary structures in general are more stable than the loop regions (as judged by the R₁/R₂ ratios). We obtained an average ¹⁵N-R₁ of 2.5 ± 0.4 s⁻¹ (cNTnC: 2.3 ± 0.6 s⁻¹) and ¹⁵N-R₂ of 9.4 ± 0.8 s⁻¹ (cNTnC: 6.5 ± 1.3 s⁻¹) for all characterized residues (n = 79). The average ¹⁵N-R₂ drops to 7.9 ± 0.4 s⁻¹ if the seven ¹⁵N-R₂ values higher than 15 s⁻¹ are excluded. We measured similar relaxation rates for mutant F77W-cNTnC (data not shown). The lower average transverse relaxation rate measured for the Trp mutants is due to residual dimerization in solution.

Structure of F77W-V82A-cNTnC•Ca²⁺ in TFE

The structure of the mutant F77W-V82A-cNTnC•Ca²⁺ was determined by NMR in the presence of 19% TFE (v/v). The co-solvent TFE is used here as a denaturant of quaternary structure to reduce dimerization (5,32) and was shown to not modify the structure of cNTnC•Ca²⁺ by comparison of the NMR structures determined without (34) and with TFE (32). The NMR ensemble of F77W-V82A-cNTnC•Ca²⁺ containing the 30 structures with lowest RMSD (out of 50) is presented on Fig. IV-4A. All of the structures in the ensemble have a good protein geometry with 91.6 % of the dihedral angles in the most favourable region of the Ramachandran plot, and 8.4 % in the additionally allowed region (Table IV-1). The backbone RMSD is 0.67 ± 0.16 Å, and 1.16 ± 0.12 Å for the heavy atoms. There are no NOE violations more than 0.4 Å. Overall, the mutations F77W and V82A do not change the structure of cNTnC. The protein shows two EF-hand motifs. The calcium binding site I is still structured and its overall structure is

unchanged. The length and sequential position of the secondary structures are identical: N-helix (5-11), A-helix (14-27), B-helix (41-48), C-helix (54-62), D-helix (74-84) and the small β -sheet (35-37, 71-73). The superimposition of F77W-V82A-cNTnC onto cNTnC (1AP4) shows a RMSD of 2.4 Å (Fig. IV-4B) for the backbone atoms of the secondary structures (N, C α , C'). Residue 77 of the two structures is represented in sticks, confirming that both aromatic rings share the same position in the hydrophobic core of the N-domain of TnC.

Even though the overall fold of the protein is conserved, the structure of F77W-V82A-cNTnC•Ca²⁺ reveals two significant changes. Firstly, a comparison of this Trp mutation in cNTnC with other Trp mutation in homologous proteins revealed that the indole ring of F77W is in the opposite/anti-parallel orientation (see Discussion section). Secondly, the F77W mutation, replacing a phenyl side chain by a larger indole ring, has created a more open conformation than for cNTnC. We have measured the interhelical angles between helix A and B, and between C and D, for F77W-V82A-cNTnC and cNTnC in the presence or absence of the cSp (PDB 1ap4 and 1mxl) using interhlx (K. Yap, University of Toronto). The interhelical angles A/B and C/D are a good indication of the opening of the N-domain of cTnC. A decrease in the interhelical angle indicates an opening of the structure. Interestingly, interhelical angle values from F77W-V82A-cNTnC•Ca²⁺ (A/B = 112°, B/C = 93°) are similar to those found in cNTnC•Ca²⁺•TnI₁₄₇₋₁₆₃ (A/B = 105°; C/D = 89°), but different from the wild type cNTnC•Ca²⁺ (A/B = 132°; C/D = 122°). Furthermore, the structural superimposition of F77W-V82A-cNTnC•Ca²⁺ with cNTnC•Ca²⁺•TnI₁₄₇₋₁₆₃ (Fig. IV-5A) has a lower RMSD (2.1 Å) than when the structure is superimposed on cNTnC•Ca²⁺ (2.4 Å) using the backbone atoms of the secondary structures (N, C α , C'). All these comparisons demonstrate that the insertion of a Trp in the core of the N-domain of cTnC generates a more open structure, similar to wild type cNTnC•Ca²⁺, but even more to cNTnC•Ca²⁺•TnI₁₄₇₋₁₆₃.

Regulation of active force by Trp mutants in skeletal muscle fibers

Extraction of the native skeletal TnC from single demembrated fibers of rabbit psoas muscle using trifluoperazine (see Materials and Methods) reduced the active isometric force measured in standard activating solution (pCa 4.5) to 8.7 ± 1.0 % (mean \pm SE, $n=29$) that observed prior to TnC extraction (T_0). After reconstitution with recombinant skeletal TnC the active force recovered to 92.2 ± 4.3 % of T_0 ($n=3$). Force recovery after reconstitution with wild type cardiac TnC (cTnC) was only 65.4 ± 3.4 % ($n=7$) of T_0 , similar to previously published values (65.4 %, (35); 71 %, (36)).

Active force recovery in fibers reconstituted with the F104(4fW), F153(4fW) and F153(5fW) mutants of cTnC was 75.0 ± 6.1 % ($n=4$), 61.3 ± 5.8 % ($n=4$) and 68.2 ± 4.2 % ($n=3$) of T_0 , respectively (Table IV-2), and these values are not significantly different from that reported above for wild type cTnC ($P > 0.05$). Fibers were fully relaxed at pCa 9.0 after TnC extraction and after reconstitution with the F104(4fW), F153(4fW) and F153(5fW) mutants of cTnC. Thus the introduction of 4fW or 5fW at either of these sites in the C-terminal lobe of cTnC has no effect on Ca^{2+} -regulation in muscle fibers beyond that associated with replacement of TnC by cTnC. In contrast, active force recovery in fibers reconstituted with the F77W and F77W-V82A mutants of cTnC was only 40.0 ± 5.5 % ($n=4$) and 39.6 ± 4.5 % ($n=5$) respectively, significantly lower than that observed for wild type cTnC ($P < 0.01$). Moreover the fibers containing F77W- and F77W-V82A-cTnC were not completely relaxed at pCa 9.0, but generated a force of 2.8 ± 0.3 % of T_0 . Thus introduction of Trp at these sites in the N-terminal lobe of cTnC does affect Ca^{2+} regulation *in situ*. The decrease in active force in fibers reconstituted with cTnCs is unlikely to be due to incomplete occupation of the TnC binding sites in the fiber, because an additional 5 min incubation with recombinant skeletal TnC produced no further recovery of active force (-1.1 ± 0.9 %, $n=7$).

DISCUSSION

Ca²⁺ and cSp affinity measurements

We have measured the calcium binding affinities of both F77W mutants in comparison to wild type cNTnC. We obtained K_d 's of 5, 15 and 18 μM for cNTnC, F77W and F77W-V82A, respectively. The K_d for cNTnC is consistent with the previous published values (37,38). The results suggest that the mutation F77W reduces the calcium affinity of cNTnC by ~ 3 fold. Moncrieffe et al. (15) observed a two-fold reduction in Ca^{2+} affinity for the skeletal F154W-TnC mutant (see also (39)), but only a small reduction in Ca^{2+} affinity for the F78W mutant, which is the skeletal analog to F77W. The cSp binding affinity of mutants F77W and F77W-V82A was also monitored by NMR at low protein concentrations. The ^1H - ^{15}N HSQC NMR spectra at high protein concentration were of poor quality, while the high quality spectra obtained during the titrations at a lower protein concentration are presented in Fig. IV-2. Another advantage of using low protein concentrations in our binding affinity measurements is that as high as 40- to 60-fold molar excess of [cSp] over [protein] is possible. This is not possible at higher protein concentrations because of limited peptide solubility. This allows one to reach a plateau in the chemical shift changes, which is reflected in a much more accurate and precise determination of the dissociation constant. The cSp binding affinity of both mutants F77W and F77W-V82A are decreased by a factor 3 in comparison to wild type cNTnC. The fact that the results were similar for both mutants indicates that the mutation V82A does not affect the binding affinity of the mutants to cTnI.

Structure of F77W-V82A-cNTnC•Ca²⁺ and Trp orientation

The solution structure of mutant F77W-V82A-cNTnC in the calcium state has been determined in the presence of TFE. The overall structure of the double mutant is very similar to that of wild type cNTnC. The sidechain of Trp77 is located at the same position as Phe77 in the wild type (Fig. IV-4B) with similar χ_1 and χ_2 values. The indole ring is well defined in the core of the protein (Fig. IV-4A). However, a comparison of this Trp mutant in cNTnC and other Trp mutants in homologous proteins reveals an important change: the indole ring in the mutant

F77W-V82A is in the opposite orientation compared to some homologous proteins. For example, the NMR structure of the F153W mutant has an indole ring orientation with the HN-bond pointing up (Fig. IV-4D) (21). The mutant F153W in the C-domain shares the homologous position as F77W in the N-domain, both sharing similar χ_1 angles (F77W = 177° and F153W = 177°), but having different χ_2 angles (F77W = 90° and F153W = -96°). Silver Hake Parvalbumin has a single Trp (W102) buried in the core of its domain and in an analogous position to F77W-V82A-cNTnC and F153W-cCTnC (Fig. IV-4C). In the X-ray structure the indole orientation is also opposite with a similar χ_1 value (W102 = 170°), but again a different χ_2 value (W102 = -106°) (40). The same orientation is observed in the structure of carp parvalbumin mutant F102W (41). Moncrieffe et al. discuss possible heterogeneity of the indole ring orientation in the context of Trp fluorescence lifetime experiments, but conclude that they only observe evidence for one major conformation (> 97 %). Their minimum perturbation maps suggest two possible minima having similar χ_1 angles of $\sim 180^\circ$ but different χ_2 values of approximately 100° and -100° . A similar computer simulation was performed for F78W-sNTnC, the skeletal homolog of F77W, and two energy minima were also obtained, the more stable of which corresponding to the orientation observed in F77W-V82A-cNTnC (15). We have also generated a model for F77W-V82A-cNTnC based on the structure of cNTnC (PDB 1AP4) using Modeller 9v1 (42). The ensemble of structures obtained after minimization revealed a Trp orientation that is in accordance with our NMR ensemble (data not shown).

Structural Consequences

There is no significant difference in the binding affinity of the two mutants for the cSp or Ca^{2+} . This suggests that the mutation V82A does not influence the properties of the F77W mutants. The structural characteristics expected from the Trp mutants were successfully attained: the same overall structure and an immobilized Trp ring in the core of the N-domain of cTnC. One structural difference between wild type cNTnC and mutant F77W-V82A is the larger extent of opening of the mutant. The comparison of the structures of F77W-V82A-

cNTnC•Ca²⁺ with cNTnC•Ca²⁺•cSp shows a similar opening (Fig. IV-5A). One would expect that a more open N-domain would lead to an increase in the cSp affinity; i.e., sNTnC (more open) has a higher affinity for sSp than cNTnC (less open) has for cSp. However, the cSp binding is weaker even though the structure of the Phe-to-Trp mutant is more open. One possibility is that the Phe-to-Trp substitution makes unfavourable contacts with cTnI, as the indole ring appears at the surface of the F77W-V82A-cNTnC•Ca²⁺ structure (Fig. IV-5B). Interestingly, the HN resonance of the Trp indole disappears from the ¹H-¹⁵N HSQC NMR spectrum after the first step of the cSp titration, and does not reappear even after an addition of more than a 40-fold molar excess of peptide. This indicates that the chemical environment of the Trp is highly modified by the presence of the cSp, supporting the interpretation of a direct contact between Trp77 and cTnI.

Force Recovery of different Phe-to-Trp mutants

Similarly to the cSp and Ca²⁺ binding affinities discussed above, the recovery of force in reconstituted skeletal muscle fibers is about the same for the F77W- and F77W-V82A mutants, indicating once again that the V82A mutation is benign. The force recovery profiles for the F104W (site III) and F153W (site IV) are similar, and not significantly different from wild type cTnC, which is used as a control in the skeletal muscle fibers. However, the force recovery for the F77W and F77W V82A (site II) and F153W mutants are much less than that of the wild type cTnC. This is consistent with the expected structural role of the C-domain, and the expected regulatory role for the N-domain derived from the large body of existing studies. At a molecular level, the insensitivities of muscle function to C-domain Phe-to-Trp mutations indicate that either the Phe-to-Trp mutations do not affect TnC-TnI interaction, or the change in affinity of cCTnC for cTnI is not enough to compromise the anchoring of TnI to TnC in the troponin complex. For F77W, the mutation reduced cTnC's affinity for cSp, thus attenuating the critical Ca²⁺-dependent interaction responsible for triggering muscle contraction. This is directly reflected in the change in recovery of force. One qualification of these studies would be that the interaction in the reconstituted fibers is with sSp and not

cSp, but their sequences are very similar and critical residues like M154 are identical between the two. The sequence of cSp₁₄₇₋₁₆₃ is RISADAMMQALLGARAK, and that of sSp₁₁₅₋₁₃₁ is RMSADAMLRALLGSKHK. In some fibers, chicken skeletal TnC was also used after reconstitution with human cardiac TnC and no further force recovery was observed, indicating that the incomplete reconstitution of cTnC could be ruled out as the reason for the incomplete force recovery after reconstitution with human cardiac TnC.

The overall conclusions of this study are that while the Phe-to-Trp mutations in cNTnC studied cause only relatively small changes at the level of molecular structure, they perturb the interactions with target proteins and this is reflected at the level of force generation in reconstituted fibers. Further, the effects of the Phe-to-Trp mutations are differential: those in the C-lobe do not result in any change in activity, consistent with its role as a structural anchor; whereas those in the N-lobe, while small, are critical to activity since they directly regulate the Ca²⁺ signaling. A reduction in affinity between regions of cTnC and cTnI of only 3-fold leads to a 40% reduction in physiological activity. Thus we can begin to correlate interactions now understood at a 3D structural level with the resulting downstream effects on function. At present considerable effort is being spent investigating the influence of familial hypertrophic cardiomyopathy mutations on cardiac muscle performance. Our study indicates that even the very small changes in biophysical properties that are observed can have a significant effect on long term function.

FIGURES AND TABLES

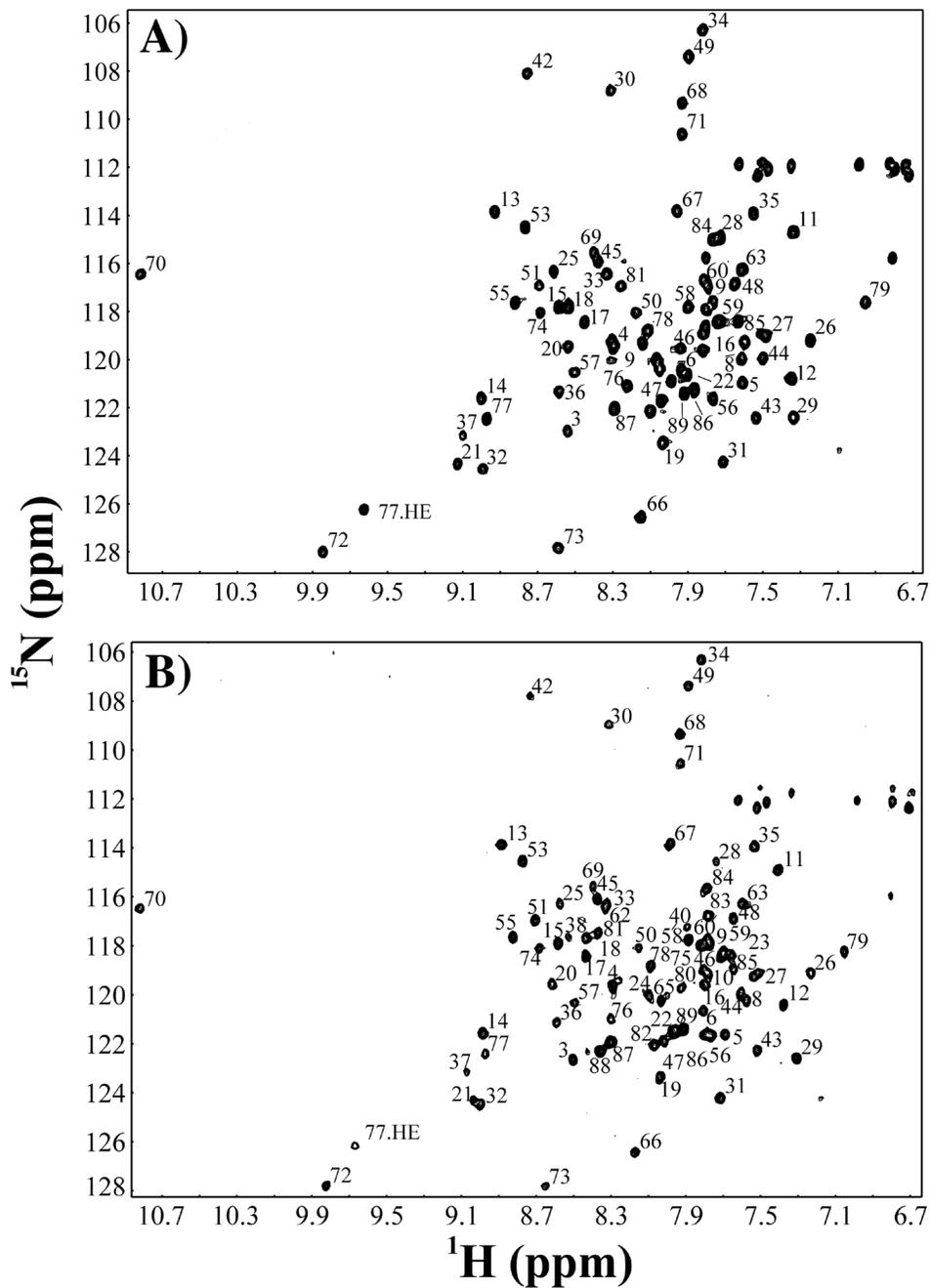


Fig. IV-1. 2D ^1H - ^{15}N HSQC NMR spectra at 600 MHz of (A) F77W-cNTnC•Ca $^{2+}$ and (B) F77W-V82A-cNTnC•Ca $^{2+}$ showing the very close similarity between both F77W mutants. The numeric labels indicate the residue specific assignments.

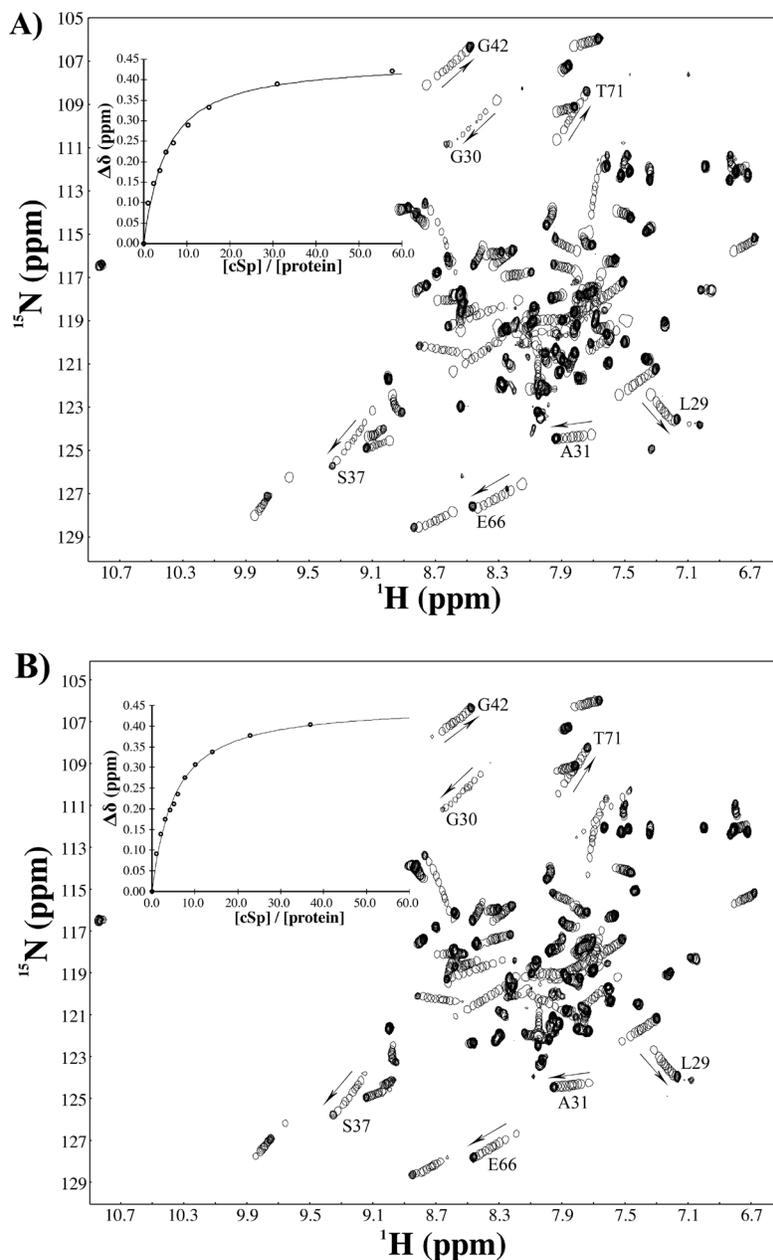


Fig. IV-2. cSp titrations of F77W and F77W-V82A-cNTnC•Ca²⁺ performed at a low concentration of protein (~0.1 mM). The seven residues identified were used to measure an average chemical shift change for each step of the titrations. Inset: the chemical shift changes are plotted as a function of the $[\text{cSp}]/[\text{protein}]$ ratio and the best fit is represented by a continuous line. A K_d of $456 \pm 35 \mu\text{M}$ was measured for mutant F77W-V82A (A) and $470 \pm 47 \mu\text{M}$ for mutant F77W (B).

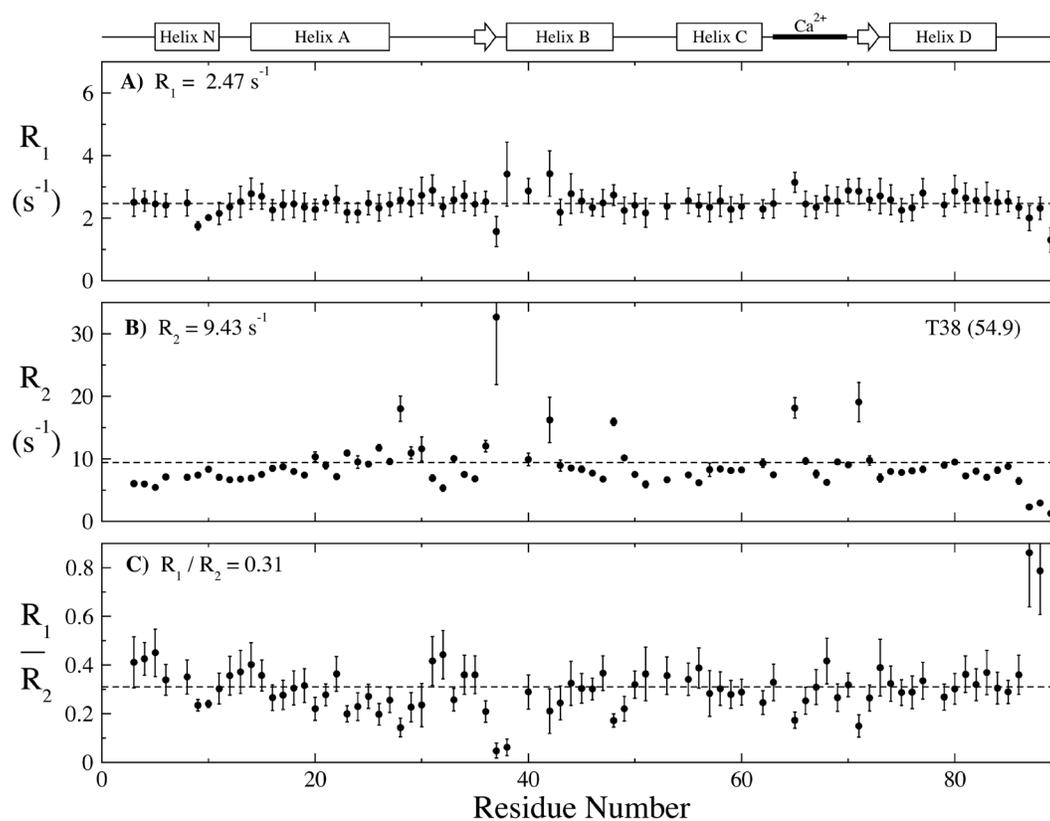


Fig. IV-3. ^{15}N -Relaxation of F77W-V82A-cNTnC•Ca²⁺ at ~0.2 mM. The ^{15}N - R_1 , ^{15}N - R_2 , and R_1/R_2 ratio are plotted as a function of the primary sequence on a per residue basis with their respective error bars. The averages for the three relaxation values are shown with dashed lines.

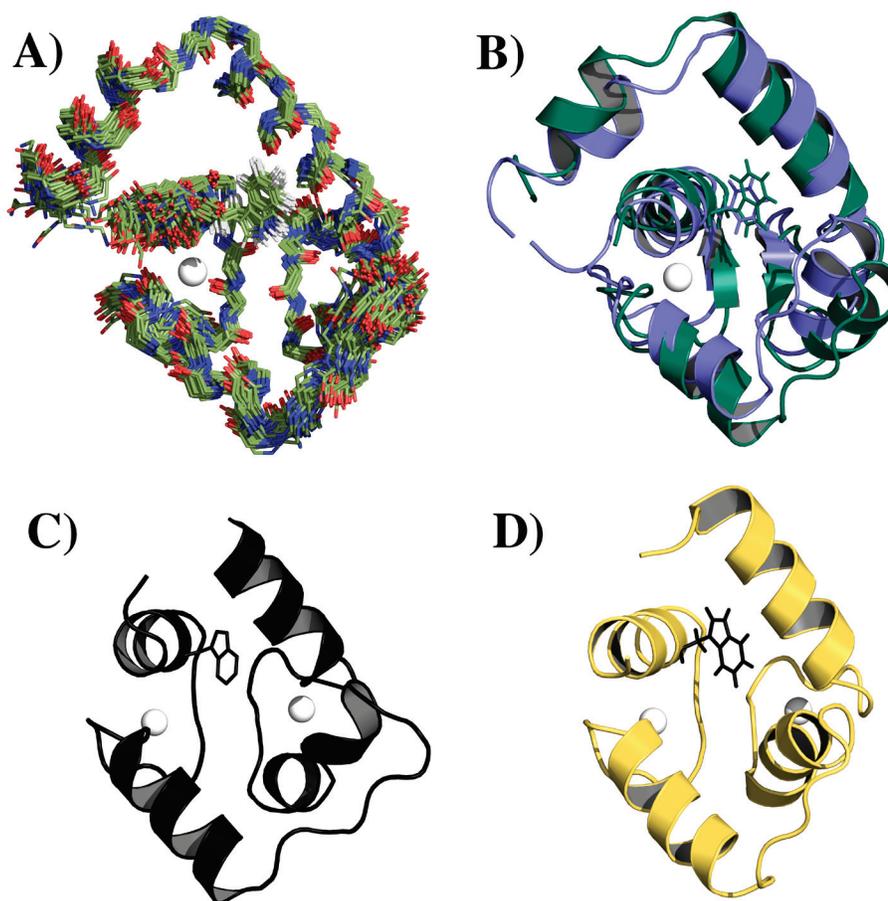


Fig. IV-4. A) NMR ensemble (30 structures) of F77W-V82A-cNTnC•Ca²⁺ in TFE. B) Superposition of F77W-V82A-cNTnC and cNTnC and a comparison of the orientation of both residues 77 (W77 vs F77). Structural comparison with C) silver hake parvalbumin (x-ray) and D) cCTnC mutant F153W (NMR) with their respective Trp residue represented with sticks.

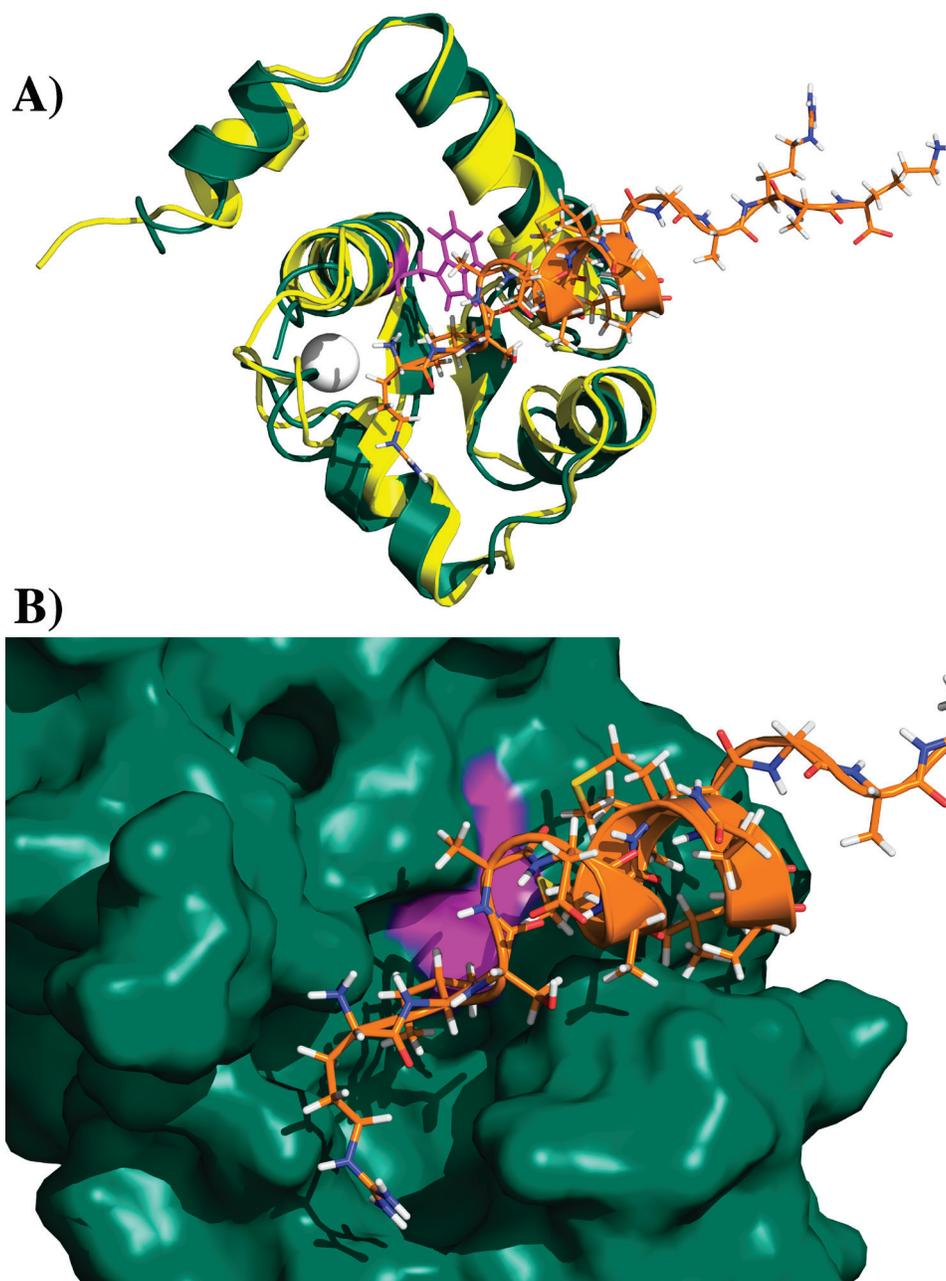


Fig. IV-5. A) Superposition of F77W-V82A-cNTnC•Ca²⁺ (green) and cNTnC•Ca²⁺•cSp (yellow) structures both determined by NMR. The Trp77 of the mutant is colored in pink and the cSp bound to the wild type is colored in orange. B) Potential interaction of cSp with F77W-V82A-cNTnC•Ca²⁺. The Fig. IV-was created using the same orientation as in A). The green surface corresponds to the surface area of F77W-V82A-cNTnC with residue Trp77 colored in pink.

Table IV-1. Structural Statistics of F77W-V82A-cNTnC•Ca²⁺ (30 structures)

Number of experimental restraints	
NOE restraints	1250
Sequential ($ i - j = 1$)	734
Medium range ($1 < i - j < 5$)	333
Long range ($ i - j \geq 5$)	183
Calcium binding restraints	8
Dihedral angle restraints from TALOS	152
Φ	76
Ψ	76
Coupling constants	58
Distances violations (n°) ^a	12
Ramachandran plot for residues 1-89 ^b	
Φ/Ψ in most favoured region	91.6 %
Φ/Ψ in additionally allowed region	8.4 %
Φ/Ψ in generously allowed region	0.06 %
Φ/Ψ in disallowed region	0.0 %
Atomic RMSD (Å) ^c	
Backbone atoms (N,C α ,C')	0.67 ± 0.16
Heavy atoms	1.16 ± 0.12

^a Violated in > 10 structures. No violations > 0.4 Å.

^b Calculated with PROCHECK_NMR v.3.5.4, excluding GLY and PRO.

^c Calculated with CYANA over residues 5-85

Table IV-2. Active force at pCa 4.5 after reconstitution with cTnC (% of the maximal force observed prior to extraction).

Human cTnC	Active force (% T_0)	No. of experiments
F104 (4F-Trp)	75.0 ± 6.1	4
F153 (4F-Trp)	61.3 ± 5.8	4
F153 (5F-Trp)	68.2 ± 4.2	3
F77W	40.0 ± 5.5*	4
F77W-V82A	39.6 ± 4.5*	5
Wild type	65.4 ± 4.3	7

Mean ± SEM; * P<0.01 when compared with wild type.

REFERENCES

1. Gagné, 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, 784-789.
2. Herzberg, O., Hayakawa, K., and James, M. N. (1984) Crystallographic data for troponin C from turkey skeletal muscle, *J. Mol. Biol.* 172, 345-346.
3. Houdusse, A., Love, M. L., Dominguez, R., Grabarek, Z., and Cohen, C. (1997) Structures of four Ca²⁺-bound troponin C at 2.0 Å resolution: further insights into the Ca²⁺-switch in the calmodulin superfamily, *Structure* 5, 1695-1711.
4. Sia, S. K., Li, M. X., Spyropoulos, L., Gagné, S. M., Liu, W., Putkey, J. A., and Sykes, B. D. (1997) Structure of cardiac muscle troponin C unexpectedly reveals a closed regulatory domain, *J. Biol. Chem.* 272, 18216-18221.
5. Slupsky, C. M., Kay, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) Calcium-induced dimerization of troponin C: mode of interaction and use of trifluoroethanol as a denaturant of quaternary structure, *Biochemistry* 34, 7365-7375.
6. Sundaralingam, M., Bergstrom, R., Strasburg, G., Rao, S. T., Roychowdhury, P., Greaser, M., and Wang, B. C. (1985) Molecular structure of troponin C from chicken skeletal muscle at 3-angstrom resolution, *Science* 227, 945-948.
7. 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, 559-579.
8. Sykes, B. D. (2003) Pulling the calcium trigger, *Nat. Struct. Biol.* 10, 588-589.
9. 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.

10. Vinogradova, M. V., Stone, D. B., Malanina, G. G., Karatzaferi, C., Cooke, R., Mendelson, R. A., and Fletterick, R. J. (2005) Ca²⁺-regulated structural changes in troponin, *Proc. Natl. Acad. Sci. U.S.A.* 102, 5038-5043.
11. Corrie, J. E. T., Brandmeier, B. D., Ferguson, R. E., Trentham, D. R., Kendrick-Jones, J., Hopkins, S. C., Van Der Heide, U. A., Goldman, Y. E., Sabido-David, C., Dale, R. E., Criddle, S., and Irving, M. (1999) Dynamic measurement of myosin light-chain-domain tilt and twist in muscle contraction, *Nature* 400, 425-430.
12. 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, 865-874.
13. Pirani, A., Vinogradova, M. V., Curmi, P. M., King, W. A., Fletterick, R. J., Craig, R., Tobacman, L. S., Xu, C., Hatch, V., and Lehman, W. (2006) An atomic model of the thin filament in the relaxed and Ca²⁺-activated states, *J. Mol. Biol.* 357, 707-717.
14. Sun, Y. B., Brandmeier, B., and Irving, M. (2006) Structural changes in troponin in response to Ca²⁺ and myosin binding to thin filaments during activation of skeletal muscle, *Proc. Natl. Acad. Sci. U.S.A.* 103, 17771-17776.
15. Moncrieffe, M. C., Eaton, S., Bajzer, Z., Haydock, C., Potter, J. D., Laue, T. M., and Prendergast, F. G. (1999) Rotational and translational motion of troponin C, *J. Biol. Chem.* 274, 17464-17470.
16. Gryczynski, I., Malak, H., Lakowicz, J. R., Cheung, H. C., Robinson, J., and Umeda, P. K. (1996) Fluorescence spectral properties of troponin C mutant F22W with one-, two-, and three-photon excitation, *Biophys. J.* 71, 3448-3453.
17. She, M., Dong, W. J., Umeda, P. K., and Cheung, H. C. (1998) Tryptophan mutants of troponin C from skeletal muscle--an optical probe of the regulatory domain, *Eur. J. Biochem.* 252, 600-607.
18. Pearlstone, J. R., Borgford, T., Chandra, M., Oikawa, K., Kay, C. M., Herzberg, O., Moulton, J., Herklotz, A., Reinach, F. C., and Smillie, L. B.

- (1992) Construction and characterization of a spectral probe mutant of troponin C: application to analyses of mutants with increased Ca²⁺ affinity, *Biochemistry* 31, 6545-6553.
19. Grage, S. L., Wang, J., Cross, T. A., and Ulrich, A. S. (2002) Solid-state ¹⁹F-NMR analysis of ¹⁹F-labeled tryptophan in gramicidin A in oriented membranes, *Biophys. J.* 83, 3336-3350.
 20. Correa, F., and Farah, C. (2005) Using 5-hydroxytryptophan as a probe to follow protein-protein interactions and protein folding transitions, *Protein Pept Lett* 12, 241-244.
 21. Wang, X., Mercier, P., Letourneau, P. J., and Sykes, B. D. (2005) Effects of Phe-to-Trp mutation and fluorotryptophan incorporation on the solution structure of cardiac troponin C, and analysis of its suitability as a potential probe for in situ NMR studies, *Protein Sci.* 14, 2447-2460.
 22. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C, *Protein Sci.* 3, 1961-1974.
 23. Li, M. X., Gagné, 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, 8330-8340.
 24. Linse, S. (2002) Calcium binding to proteins studied via competition with chromophoric chelators, *Methods Mol. Biol.* 173, 15-24.
 25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
 26. 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, 313-321.
 27. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol. Biol.* 278, 353-378.

28. 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.
29. Laskowski, R. A., Rullmann, J. A., Macarthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR* 8, 477-486.
30. Sabido-David, C., Brandmeier, B., Craik, J. S., Corrie, J. E. T., Trentham, D. R., and Irving, M. (1998) Steady-state fluorescence polarization studies of the orientation of myosin regulatory light chains in single skeletal muscle fibers using pure isomers of iodoacetamidotetramethylrhodamine, *Biophys. J.* 74, 3083-3092.
31. Spyropoulos, L. (2001) Know your protein: effects of weak dimerization, Edited by Oleg Jardetzky and Michael D. Finucane, pp 37-44, IOS Press.
32. Paakkonen, K., Annala, A., Sorsa, T., Pollesello, P., Tilgmann, C., Kilpelainen, I., Karisola, P., Ulmanen, I., and Drakenberg, T. (1998) Solution structure and main chain dynamics of the regulatory domain (Residues 1-91) of human cardiac troponin C, *J. Biol. Chem.* 273, 15633-15638.
33. 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, 8289-8298.
34. Spyropoulos, L., Gagné, S. M., Li, M. X., and Sykes, B. D. (1998) Dynamics and thermodynamics of the regulatory domain of human cardiac troponin C in the apo- and calcium-saturated states, *Biochemistry* 37, 18032-18044.
35. Morris, C. A., Tobacman, L. S., and Homsher, E. (2001) Modulation of contractile activation in skeletal muscle by a calcium-insensitive troponin C mutant, *J. Biol. Chem.* 276, 20245-20251.
36. Moreno-Gonzalez, A., Fredlund, J., and Regnier, M. (2005) Cardiac troponin C (TnC) and a site I skeletal TnC mutant alter Ca²⁺ versus

- crossbridge contribution to force in rabbit skeletal fibres, *J. Physiol.* 562, 873-884.
37. Hazard, A. L., Kohout, S. C., Stricker, N. L., Putkey, J. A., and Falke, J. J. (1998) The kinetic cycle of cardiac troponin C: calcium binding and dissociation at site II trigger slow conformational rearrangements, *Protein Sci.* 7, 2451-2459.
 38. Li, M. X., Gagné, S. M., Spyropoulos, L., Kloks, C. P., Audette, G., Chandra, M., Solaro, R. J., Smillie, L. B., and Sykes, B. D. (1997) NMR studies of Ca²⁺ 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, 12519-12525.
 39. Chandra, M., Mccubbin, W., Oikawa, K., Kay, C., and Smillie, L. (1994) Ca²⁺, Mg²⁺, and troponin I inhibitory peptide binding to a Phe-154 to Trp mutant of chicken skeletal muscle troponin C, *Biochemistry* 33, 2961-2969.
 40. Richardson, R. C., King, N. M., Harrington, D. J., Sun, H., Royer, W. E., and Nelson, D. J. (2000) X-Ray crystal structure and molecular dynamics simulations of silver hake parvalbumin (Isoform B), *Protein Sci.* 9, 73-82.
 41. Moncrieffe, M. C., Juranic, N., Kemple, M. D., Potter, J. D., Macura, S., and Prendergast, F. G. (2000) Structure-fluorescence correlations in a single tryptophan mutant of carp parvalbumin: solution structure, backbone and side-chain dynamics, *J. Mol. Biol.* 297, 147-163.
 42. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234, 779-815.

CHAPTER V – TFE*The effect of the co-solvent trifluoroethanol on a tryptophan side chain orientation in the hydrophobic core of troponin C*Olivier Julien,¹ Pascal Mercier,¹ Melissa L. Crane,¹ and Brian D. Sykes¹¹ *Department of Biochemistry, University of Alberta, Edmonton, Canada.*Publication: Julien et al. 2009. *Protein Science*, 18:1165-74.

Contributions: OJ and BDS designed the NMR experiments. OJ acquired the NMR spectra, assigned the protein chemical shifts, solved the NMR structure (with PM) and wrote the manuscript with BDS. OJ made all of the figures. MLC cloned, expressed and purified the mutant proteins.

OVERVIEW

The unique biophysical properties of tryptophan residues have been exploited for decades to monitor protein structure and dynamics using a variety of spectroscopic techniques, such as fluorescence and nuclear magnetic resonance (NMR). We recently designed a tryptophan mutant in the regulatory N-domain of cardiac troponin C (F77W-cNTnC) to study the domain orientation of troponin C in muscle fibers using solid-state NMR. In our previous study, we determined the NMR structure of calcium-saturated mutant F77W-V82A-cNTnC in the presence of 19% 2,2,2-trifluoroethanol (TFE). TFE is a widely used co-solvent in the biophysical characterization of the solution structures of peptides and proteins. It is generally assumed that the structures are unchanged in the presence of co-solvents at relatively low concentrations, and this has been verified for TFE at the level of the overall secondary and tertiary structure for several calcium regulatory proteins. Here, we present the NMR solution structure of the calcium saturated

F77W-cNTnC in presence of its biological binding partner troponin I peptide (cTnI₁₄₄₋₁₆₃) and in the absence of TFE. We have also characterized a panel of six F77W-cNTnC structures in the presence and absence TFE, cTnI₁₄₄₋₁₆₃, and the extra mutation V82A, and used ¹⁹F NMR to characterize the effect of TFE on the F77(5fW) analog. Our results show that although TFE did not perturb the overall protein structure, TFE did induce a change in the orientation of the indole ring of the buried tryptophan side chain from the anticipated position based upon homology with other proteins, highlighting the potential dangers of the use of co-solvents.

INTRODUCTION

The use of co-solvents in studies of biological macromolecules is a common practice in biochemistry and structural biology. Their use facilitates the study of small polypeptides as well as large molecular complexes of protein and nucleic acids. Co-solvents such as alcohols, glycerol and others have been used in structure determination by x-ray crystallography and solution studies using circular dichroism and nuclear magnetic resonance (NMR) spectroscopy. In particular, 2,2,2-trifluoroethanol (TFE) is a widely used co-solvent in the determination of NMR solution structures of peptides and proteins. TFE is often used to induce secondary structure of small amino acid chains by favoring the formation of α -helices or even β -sheet content (1,2). TFE has also been shown to stabilize β -hairpin structures (3) and induce amyloid-like fibril formation (4,5). Furthermore, TFE is used as a membrane-mimetic solvent (6) to solubilize and stabilize the structure of membrane peptides and proteins, and also to prevent their aggregation. TFE also decreases the tendency of proto-filaments to form clusters (7). Studies of the NMR solution structures of the calcium regulatory protein troponin C from skeletal and cardiac muscle (8,9) have also demonstrated that TFE can be used as a denaturant of quaternary structure, weakening hydrophobic interactions within or between molecules in solution. In fact, the NMR solution structure of several troponin C constructs have been determined in the presence of

TFE by ourselves and other groups (9-11), and more than 60 NMR structures in the Protein Data Bank (PDB) were determined in a mixture of water and TFE. For a complete review on the use of TFE and co-solvents, see Ref. (12). The extensive use of co-solvents such as TFE makes the understanding of their mechanism of action desirable, but the widely different examples suggest that the mechanism may be complex and unpredictable -which is the case in our present study.

The long-term goal of this research project is to obtain *in situ* structural information regarding the molecular changes of sarcomere proteins during muscle contraction. Many techniques and spectroscopic probes have been used to monitor the structure, function, and dynamics of the muscle protein machinery: for example, fluorescence resonance energy transfer measurements (13), electron paramagnetic resonance spectroscopy using nitroxide probes (14), electron microscopy (15), NMR spectroscopy (16), and x-ray crystallography (17,18). One recent and particularly powerful approach involves the use of bifunctional rhodamine fluorescence labels to monitor protein domain orientation *in situ*. Examples include the orientation of the domains of myosin light chains and troponin C (TnC) in active muscle fibers (19,20). However, these bifunctional rhodamine probes are bulky and hydrophobic with a proclivity for complex dynamics (21,22). The use of solid-state NMR methods is rapidly becoming important in many immobilized biological systems such as membrane proteins in phospholipid bilayers. In NMR, the probe is a nuclear spin (either naturally occurring or isotopically labeled). Even when a non-natural nuclear spin label such as ^{19}F is used, it is expected to be much less perturbing than the larger aromatic optical or electron spin resonance probes.

For more than half a century, tryptophan fluorescence has been a popular method to study protein-protein interactions, protein-drug bindings or to monitor other conformational changes affecting protein structures (see Ref (23) for a review on protein fluorescence). For example, tryptophan fluorescence has been extensively used to study the effect of calcium binding to chicken skeletal troponin C with the

engineering of several tryptophan mutants such as F29W (24), and F78W and F154W (25,26). Similar studies also exploited the mutation F102W in parvalbumin and oncomodulin (27,28). Over the past few years, we have designed several tryptophan mutants of cardiac troponin C in order to incorporate ^{19}F -labeled tryptophan and then to use solid-state ^{19}F NMR spectroscopy to determine the orientation and dynamics of TnC during calcium triggered muscle contraction (10,29,30). The design criterion was to incorporate a single tryptophan immobilized within the core of protein in a presumed well-defined position. TnC is a calcium-binding protein containing several EF hand motifs, which are formed by a contiguous stretch of 12 residues with 6 of them involved in calcium binding, denoted X, Y, Z, -Y, -X and -Z at positions 1, 3, 5, 7, 9 and 12. The immediately preceding and following residues are always hydrophobic, with the subsequent residue virtually always being a phenylalanine. The EF hands are virtually always two site domains, even if sometimes the individual sites no longer bind metals. The phenylalanine residues following an EF hand are buried in middle of the domain and serve to anchor the Ca^{2+} binding site. In some cases, the phenylalanine residue is replaced by a naturally occurring tryptophan in this location, as found in silver hake Parvalbumin and scallop TnC (31,32). We have previously determined the structures of F153W and F153(5fW) cardiac troponin C (29), and demonstrated that the indole side chain in this position is immobilized and in an expected orientation compared to silver hake Parvalbumin W102. Recently we described the effect of the F77W substitution on the 3D structure, troponin I affinity, and *in situ* activity of troponin C (10). Interestingly, the NMR structure of the F77W-V82A mutant of the N-domain of troponin C in the presence of 19% TFE revealed that the W77 side chain was in a different orientation compared to the one found in F153W-cCTnC and silver hake parvalbumin, suggesting a rationalization for the altered biological activity. The presence of TFE was used there to counteract the higher tendency of the mutant to dimerize compared to the wild type protein. The importance of tryptophan side chain orientations in biology has been recently highlighted by two studies. Ozkirimli et al. have shown using enzyme kinetics and molecular dynamics

simulations that the activation of Src kinase is coupled with the side chain conformer of W260 (33). Using ^{19}F solid-state NMR spectroscopy, Witter et al. have shown that the activation/inactivation of protein M2 of influenza A virus is regulated by the side chain rotamer of the W41 (34).

We have determined the NMR solution structure of the calcium saturated mutant F77W-cNTnC in the absence of TFE and in complex with its binding partner, residues 144 to 163 of troponin I (cTnI₁₄₄₋₁₆₃). We have structurally characterized a panel of structures of F77W-cNTnC in the presence and absence of TFE, cTnI₁₄₄₋₁₆₃, and V82A. Interestingly, the tryptophan orientation of W77 in the F77W-cNTnC•cTnI₁₄₄₋₁₆₃ complex was observed to be in the canonical orientation observed for F153W-cCTnC and W102 in silver hake parvalbumin, and opposite to the orientation determined for F77W-V82A-cNTnC in the presence of 19% TFE. Of the three possibilities that could cause this conformational difference, the panel of structures presented herein revealed that the TFE is responsible for the different orientation of the indole side chain of F77W. We have also used ^{19}F NMR to characterize the effect of TFE on the F77(5fW) analog.

MATERIAL AND METHODS

Protein Expression and Purification

pET3A plasmids containing the sequence for F77W-cNTnC (1-89) or F77W, V82A-cNTnC (1-89) were transformed into BL21(DE3)pLysS host cells. The expression and purification of [U - ^{13}C , ^{15}N]-labeled protein in *E. coli* has previously been described (35,36). Decalcification was performed as described (36). Expression of F77(5fW)-cNTnC was performed as described (29) and purified as described (36). The human cardiac troponin I (cTnI) peptide containing residues 144 to 163 (Ac-RRVRISADAMMQALLGARAK-NH₂) was bought from GL Biochem ltd. (Shanghai, China).

NMR sample preparation

The contents of the different NMR samples used in this study are described below. A cocktail of protease inhibitors (Calbiochem) and NaN_3 were used to prevent sample degradation. The NMR buffer consists of 100 mM KCl and 10 mM imidazole in H_2O . Dithiothreitol (DTT) was used to prevent the oxidation of cysteines and CaCl_2 to saturate calcium binding site I. The pH of the sample was calculated from the measurement of the chemical shift of the imidazole H2 proton.⁽³⁷⁾ Every ^1H spectrum was referenced according to the DSS- d_6 signal (Chenomx Inc., Canada). Standard 5mm NMR tubes with 500 μL of solution were used unless specified otherwise.

The following samples were prepared to obtain NMR data of the different proteins: **1)** F77W-V82A-cNTnC: 4.2 mg [U - ^{13}C , ^{15}N]-labeled protein, 500 μL NMR buffer, 0.3 mM DSS- d_6 , 0.03% NaN_3 , 6 mM CaCl_2 , 15 mM DTT, 6.0 μL protease inhibitors. The pH was adjusted to 6.40. **2)** F77W-cNTnC: 4.0 mg [U - ^{13}C , ^{15}N]-labeled protein, 500 μL NMR buffer, 0.3 mM DSS- d_6 , 0.015 % NaN_3 , 6 mM CaCl_2 , 15 mM DTT, 6.0 μL protease inhibitors. The pH was adjusted to 6.76. **3)** F77W-V82A-cNTnC•TFE: 5.3 mg [U - ^{13}C , ^{15}N]-labeled protein, 215 μL NMR buffer, 19% TFE, 0.025% NaN_3 , 0.2 mM DSS- d_6 , 6 mM CaCl_2 , 10 mM DTT, 3.0 μL protease inhibitors. The pH was adjusted to 6.80 and the sample transferred into a Shigemi tube (Shigemi Ltd., Japan). **4)** F77W-cNTnC•TFE: 5.2 mg [U - ^{13}C , ^{15}N]-labeled protein, 245 μL NMR buffer, 60 μL TFE, 0.3 mM DSS- d_6 , 0.025% NaN_3 , 6 mM CaCl_2 , 16 mM DTT, 3.0 μL protease inhibitors. The pH was adjusted to 6.71 and the sample transferred to a Shigemi tube (Shigemi Ltd., Japan). **5)** F77W-cNTnC•cTnI₁₄₄₋₁₆₃: 3.7 mg [U - ^{13}C , ^{15}N]-labeled protein, 500 μL NMR buffer, 0.3 mM DSS- d_6 , 0.03% NaN_3 , 6 mM CaCl_2 , 15 mM DTT, 6.0 μL protease inhibitors and 1.7 mg cTnI₁₄₄₋₁₆₃. The pH was adjusted to 6.79. **6)** TFE Titration of F77W-cNTnC: 6.1 mg [U - ^{13}C , ^{15}N]-labeled protein, 500 μL NMR buffer, 0.3 mM DSS- d_6 , 6 mM CaCl_2 , 15 mM DTT, 0.03 % NaN_3 and 6 μL protease inhibitors. The pH was adjusted to 6.69. An identical sample was

prepared to titrate F77(5fW)-cNTnC with TFE, and another one to examine the effect of cTnI₁₄₄₋₁₆₃.

NMR data acquisition and processing

The ¹H, ¹³C, ¹⁵N NMR spectra presented in this study were acquired on Varian INOVA 500 MHz, 600MHz and 800 MHz NMR spectrometers (the latter equipped with a cryogenic probe) using Biopack pulse sequences (version 2008-01-18, VnmrJ 2.1B, Varian Inc.) Each spectrum presented in Fig. V-2 was obtained with the gnoesyChsqc pulse sequence(38) and acquired with 16 scans at a resolution of 512 (*t3*) x 64 (*t2*) x 32 (*t1*) or more complex points. The spectral widths were 14 or 10 ppm for the ¹H-dimensions, and 20 or 30 ppm for the ¹³C-dimension with a carrier position at 125 ppm specific for the aromatic region. For the TFE titration (see Fig. V-3), a 2D ¹H-¹³C HSQC with 32 scans was acquired with a resolution of 512 (*t2*) x 128 (*t1*) complex points for every of step of the titration: 0, 2, 4, 6, 8, 10, 14, 19 and 25% TFE (v/v). All of the NMR spectra acquired in this study were processed with NMRPipe (39) or vnmrJ 2.1B (Varian inc.). Prior to the Fourier transform, a sinebell function of 60 to 90° offsets was applied, and then zero filling was used to extend a maximum of twice the number of real points. Linear prediction was only used for the *t3* dimension of 3D spectra with a maximum of half the number of acquired points predicted. The chemical shift assignments were completed using with smartnotebook v5.1.3 (40) and NMRViewJ (One Moon Scientific inc.).

The ¹⁹F NMR spectroscopy was performed on the Varian INOVA 600MHz, by replacing the gradient probe with our home-built ¹⁹F/¹H probe.(41) Each ¹⁹F NMR spectrum presented in Fig. V-4 was acquired with 1000 scans, a spectral width of 50,000 Hz and a time delay of 1 s. The frequency of the trifluoroacetate (TFA) was used as an internal standard to reference the spectra at 0 ppm. The spectra were processed with VnmrJ 2.1B (Varian inc.) using a line broadening of 40 Hz and back linear prediction for 6 points.

NMR Structure determination of F77W-cNTnC•cTnI₁₄₄₋₁₆₃

The following spectra were used for the chemical shift assignments: 2D ¹H-¹⁵N HSQC, 2D ¹H-¹³C HSQC, 3D ¹H-¹⁵N TOCSY-HSQC, 3D CBCACONNH, 3D HNCACB for the backbone and 3D CCONH, 3D HCCONH, 3D HCCH-TOCSY. The intra-molecular NOE restraints for F77W-cNTnC (1284 total) were obtained from a 3D ¹⁵N NOESY-HSQC and two 3D ¹³C NOESY-HSQC (aliphatic and aromatic) acquired at 800 MHz. The 39 intra-molecular NOE distances for cTnI₁₄₄₋₁₆₃ were taken from a previously determined NMR structure involving cNTnC and cTnI₁₄₇₋₁₆₃ (PDB ID 1lxf), as well as the 30 inter-molecular distances used to position the peptide on the protein. The structure calculations were performed using CYANA 2.1 (42). The NOEs were calibrated within CYANA. The minimum and maximum NOE calibration values were set to 1.8 and 6.0 Å, respectively. In addition, 8 calcium-binding restraints (based on calcium binding site homology) and 134 dihedral restraints from TALOS (43) were used to obtain the NMR ensemble. The option autoaco was also used in cyana to generate additional angle constraints. The structure validation was performed using procheck v.3.5.4 (44). The favorable dihedral angles, restraint violations and cyana target function are presented in Table V-1.

RESULTS

The focus of this paper is to evaluate the effect of TFE on the structure of cNTnC by comparing NMR data obtained for the mutant F77W in the presence and absence of the mutation V82A, the co-solvent TFE, and the binding partner, residues 144 to 163 of troponin I (cTnI₁₄₄₋₁₆₃). All of the calcium-binding proteins described herein were all studied in their calcium-saturated states. The previously determined NMR solution structure of F77W-V82A-cNTnC in 19% TFE is shown in Fig. V-1A. In this structure, the orientation of the indole ring is inconsistent with that observed in the calcium saturated F153W-cCTnC (29) and silver hake parvalbumin W102 (31) (see Fig. V-1A insert). In this paper, we describe the new NMR solution structure of the F77W-cNTnC•cTnI₁₄₄₋₁₆₃

complex in the absence of TFE (see Fig. V-1B). In this new structure, the tryptophan indole ring is in the canonical orientation with a NH indole pointing towards helices N and A (see Fig. V-1B inset), as observed for F153W-cTnC and parvalbumin. To further understand the effect of the trifluoroethanol on the structure of cardiac troponin C, we have structurally characterized a panel of structures in the presence and absence of TFE, V82A, and cTnI₁₄₄₋₁₆₃ peptide; and we have used ¹⁹F NMR to characterize the effect of TFE on the F77(5fW) analog.

The overall fold of the newly determined NMR solution structure of F77W-cTnC•cTnI₁₄₄₋₁₆₃ in the absence of TFE is very similar to the previously determined structure of F77W-V82A-cTnC in 19% TFE (PDB ID 2jxl) and that of the wild type calcium saturated cTnC•cTnI₁₄₇₋₁₆₃ complex (PDB ID 1mxl). The structural statistics for the structure are presented in Table V-1, and the ensemble of NMR structures deposited to the PDB with access code 2kgb. The RMSD for the superimposition of the backbone atoms for residues 5-85 of F77W-cTnC•cTnI₁₄₄₋₁₆₃ with F77W-V82A-cTnC in 19% TFE and cTnC•cTnI₁₄₇₋₁₆₃ is 1.6 Å and 2.2 Å, respectively. These relatively low RMSD values indicate that the backbones of the three proteins are very similar to each other. Of particular interest is the orientation of the side chain of residue 77. Amino acid side chains are usually described using two dihedral angles χ_1 and χ_2 , defined by the N-C α -C β -C γ and C α -C β -C γ -C δ_1 bonds, respectively. The tryptophan in F77W-cTnC•cTnI₁₄₄₋₁₆₃ has χ_1 and χ_2 values of -171° and -110° , respectively. In general, residues with two H β preferentially adopt rotamers that have one of the three standard χ_1 values of $-60^\circ \pm 30^\circ$, $60^\circ \pm 30^\circ$ and $180^\circ \pm 30^\circ$, often referred to in the literature as gauche negative (g- or m), gauche positive (g+ or p) and trans (t), respectively (45-47). For a tryptophan like W77, the combination of the three standard χ_1 and two possible χ_2 values leads to six favorable rotamers (48,49). Accordingly, all of the tryptophan side chains discussed herein for troponin C (with and without TFE) and parvalbumin (W102) possess a trans orientation with similar χ_1 dihedral angles. This leads to virtually coplanar aromatic rings in all

structures. However, the orientation of W77 in the structure in 19% TFE is in the opposite direction with respect to W77 in F77W-cNTnC•cTnI₁₄₄₋₁₆₃ and W153 in F153W-cCTnC (both in the absence of TFE). This difference can be explained by their opposite χ_2 values. Using the nomenclature described by Lovell et al., the tryptophan orientation of F77W-V82A-cNTnC in 19% TFE can be categorized as t90° ($\chi_1 = \text{trans}$ and $\chi_2 \sim 90^\circ$) in comparison to t-105° ($\chi_1 = \text{trans}$ and $\chi_2 \sim -105^\circ$) for F77W-cNTnC•cTnI₁₄₄₋₁₆₃, F153W-cCTnC and silver hake parvalbumin W102.

To determine the cause(s) of the re-orientation of this buried tryptophan side chain, we collected structural data for the panel of F77W samples in the presence and absence of cTnI₁₄₄₋₁₆₃, V82A or TFE. We acquired 3D ¹³C NOESY-HSQC NMR spectra for all of the samples to look at the differential NOE contacts of the W77 indole ring. Fig. V-2 shows the NOE contacts between the W77.HD1 (C2') and its surrounding protons for different sample conditions. By comparing Fig. V-2A and 2E, one can clearly see the different NOE patterns for F77W-V82A-cNTnC in 19% TFE (t90°) and F77W-cNTnC•cTnI₁₄₄₋₁₆₃ (t-105°). A complete set of the unambiguous NOEs observable between W77 in a t90° and a t-105° orientation is presented in Table V-2. Using these NOE patterns for the different aromatic protons of the tryptophan, one can clearly identify the indole ring orientation of residue 77 in a specific sample condition or within a given protein complex. Fig. V-2D clearly shows that the indole orientation of F77W-cNTnC in water and in the absence of cTnI results in NOE contacts similar to Fig. V-2E, where cTnI₁₄₄₋₁₆₃ is present. This is typical of an indole t-105° situation, meaning that the presence of cTnI₁₄₄₋₁₆₃ is not the key for the W77 orientation change. Focusing on the V82A mutation, we looked at the aromatic NOE contacts of F77W-V82A-cNTnC in water without TFE or peptide (Fig. V-2C). Again, the NOE contacts were typical of a W77 indole having a χ_2 of -105°, ruling out the V82A mutation as being responsible for the aromatic flip. To confirm that the presence of TFE is the cause of the W77 orientation change, we compared the NOE contacts of both F77W and F77W-V82A in 19% TFE (Fig. V-2A and 2B) and observed similar NOE contacts, both characteristic of a W77 in a t90° rotamer

conformation. These results demonstrated that TFE is responsible for the tryptophan indole orientation change.

To investigate the effect of the concentration of TFE on the W77 aromatic ring orientation, we titrated the protein with TFE from 0 to 25% (v/v) and monitored the structural changes on F77W-cNTnC. For each step in the titration, a 2D ^{15}N HSQC NMR spectrum and a 2D ^{13}C HSQC NMR spectrum of the aromatic region were acquired to look at the effect of TFE on the chemical environment of W77 (Fig. V-3A). The results suggest that early in the titration (i.e. ~2% TFE) a conformational change occurs; as evidenced by the absence of some of the aromatic proton signals like W77.HZ3 (Fig. V-2C) and W77.HE1 (from ^{15}N HSQC data not shown). Both protons do not reappear until late in the titration around 19% TFE. Similarly, the W77.HZ2 signal broadens from 0 to about 10%, and starts to sharpen again as the titration continues. Other aromatic positions show different patterns, but most seem to present at two-step transition. For example, the proton W77.HD1 shows similar broadening to HZ2, but the chemical shift changes direction at 10% TFE and reverts toward its original position (that of 0% TFE) at the end of the titration.

We also looked at the effect of TFE on the methyl group of the methionine side chains by following the 2D ^{13}C HSQC NMR spectra of the aliphatic region during the TFE titration (Fig. V-3E). The spectral changes reveal that the side chain methyl resonances of M80 and M81 are by far the most perturbed by the increased concentration from 0 to 25% TFE. The total magnitude of the ^1H and ^{13}C chemical shift changes ($\Delta\delta_{total} = \sqrt{((\Delta\Delta\delta_{^1\text{H}})^2 + (\Delta\Delta\delta_{^{13}\text{C}})^2)}$) of M80 and M81 are 149 Hz and 448 Hz, respectively, on a 500 MHz NMR spectrometer. These two residues are located in the hydrophobic pocket of cNTnC in close proximity to the indole ring of W77. Interestingly, the resonance for the methyl group of M81 is too broad to be observed between 2 and 19% TFE, as was observed for W77.HZ3.

To support these results, we have substituted the proton at position 5' of W77 (W77.HZ3) with a ^{19}F atom by the incorporation of 5-fluorotryptophan (5fW) at position 77 in order to monitor the effect of TFE on the tryptophan using ^{19}F NMR. The 1D ^{19}F NMR spectra obtained during the TFE titration of F77(5fW)-cNTnC are presented on Fig. V-4. At 0% TFE, the 5fW resonance has a chemical shift of -47.14 ppm. The peak is very broad with a linewidth ($\Delta\nu$) of 465 Hz. As the TFE concentration is increased, the ^{19}F resonance moves up-field and gets sharper, ending up at -47.87 ppm in 25% TFE with a $\Delta\nu$ of 78 Hz. This is roughly the linewidth expected for a protein the size of F77(5fW)-cNTnC. When the cTnI₁₄₄₋₁₆₃ peptide is bound to the protein (without TFE), the 5fW resonance is also sharper, with a linewidth of 138 Hz. These results are consistent with exchange broadening of the ^{19}F NMR linewidth of the 5fW residue in the absence of either TFE or cTnI₁₄₄₋₁₆₃. This broadening is caused by an equilibrium between the t90° and t-105° conformations such that both TFE and cTnI₁₄₄₋₁₆₃ force the equilibrium to one side or the other.

DISCUSSION

More than 60 NMR structures in the PDB were determined in a mixture of water and TFE. The implicit assumption in these structures is that the structures are not changed by this co-solvent. For example, in our previous description of the effect of the F77W substitution on the 3D structure, troponin I affinity, and *in situ* activity of troponin C (10), we modeled the TnC-TnI complex without questioning the unusual orientation of the tryptophan side chain observed in TFE. In this manuscript we have demonstrated that TFE can change the orientation of a presumably buried and immobile tryptophan side chain by comparing the NMR solution structures of F77W-cNTnC•cTnI₁₄₄₋₁₆₃ and F77W-cNTnC in the presence of 19% TFE. Using NOESY experiments, we ruled out the effect of the V82A mutation and the presence of the bound peptide in the hydrophobic pocket, and

identified the co-solvent TFE to be responsible for the change in the indole ring orientation.

How does TFE change the orientation of W77? One possibility is that the hydrophobic TFE molecules bind to and disturb the hydrophobic pocket of cNTnC. This could involve a general non-specific binding or a binding to a specific site(s). In the former case the binding of TFE could lead to a change in the equilibrium between the open and closed states of cNTnC existing in solution (16). Inspection of the structure of F77W-V82A-cNTnC in 19% TFE reveals that the N-domain has a slightly more open structure compared to wild type, which might allow more space for the W77 and facilitate the rotation of aromatic side chain. In the absence of TFE, the tryptophan indole NH is solvent-exposed and forms hydrogen bonds, or at least favorable electrostatic interactions, with the water molecules, whilst the bulky hydrophobic ring is buried inside the protein core making favorable contact with I36 and V72 (Fig. V-1B). However, when TFE is present, the hydrophobic aromatic ring of W77 could make favorable interactions with the TFE molecules.

In the case of a specific binding site, TFE may bind to the methionine side chains as previously suggested for a similar molecule, 2-bromo-2-chloro-1,1,1-trifluoroethane (halothane) (50). This study suggested that the anesthetic molecule was bound between the side chains of a methionine and a tryptophan; the $-\text{CF}_3$ of the drug making favorable contacts with the sulfur atom of the methionine side chain in the core of a four α -helix bundle. In our case, there is a surface pocket in the structure formed by M80, M81 and W77. The large chemical shift changes observed for the methyl groups of M80 and M81 support this possibility. However, the changes in chemical shifts for these two residues could also be a consequence of a change in the orientation of an aromatic ring, which would occur with a W77 indole flip in close proximity. We have, however, been unable to observe any clear heteronuclear ^1H - ^{19}F NOE contacts between the CF_3 group of

TFE and either the M80 and M81 methyl groups or the W77 ring protons (data not shown).

Previous studies using tryptophan fluorescence measurements and molecular dynamics simulations on the F78W mutant of skeletal TnC reported the possibility of two or more side chain orientations for the homologous tryptophan (skeletal F78W mutation is equivalent to cardiac F77W) (26). The minimum perturbation map of W78 χ_1 and χ_2 presented by Moncrieffe et al. suggests two minimum wells of $\chi_1 \times \chi_2$ corresponding to the two tryptophan indole orientations described in this paper. Fig. V-5 shows the $\chi_1 \times \chi_2$ isomerization map of the different tryptophans naturally occurring and engineered into troponin C and parvalbumin. Of course, the conformations of all protein side chains are dynamic in solution, so that an equilibrium between different indole conformations for W77 presumably exists in solution. In water, the main conformation of the W77 indole ring is oriented towards helices B and C as observed in the F77W-cNTnC•cTnI₁₄₄₋₁₆₃ structure. With the addition of a small amount of TFE (i.e. 2% TFE), the equilibrium of the aromatic moiety is perturbed as reflected by the exchange broadening of some aromatic proton resonances in the ¹³C and ¹⁵N HSQC NMR spectra. By adding more TFE (i.e. ~19% TFE), the equilibrium is then shifted towards the opposite indole orientation.

Obviously, if TFE changes the orientation and dynamics of a buried tryptophan side chain, it will most likely affect a number of other residues. However, this usually does not have a major impact on the overall structure of the protein, since the side chains that are solvent accessible are most often flexible and poorly defined in solution. The caution is to be very careful when making mechanistic conclusions based upon the orientation of surface residues. Our results reemphasize that proteins are highly dynamic, and support a multitude of previous fluorescence studies that indicated that tryptophan rings are not restrained to only one conformation even if well packed in a protein.

FIGURES AND TABLES

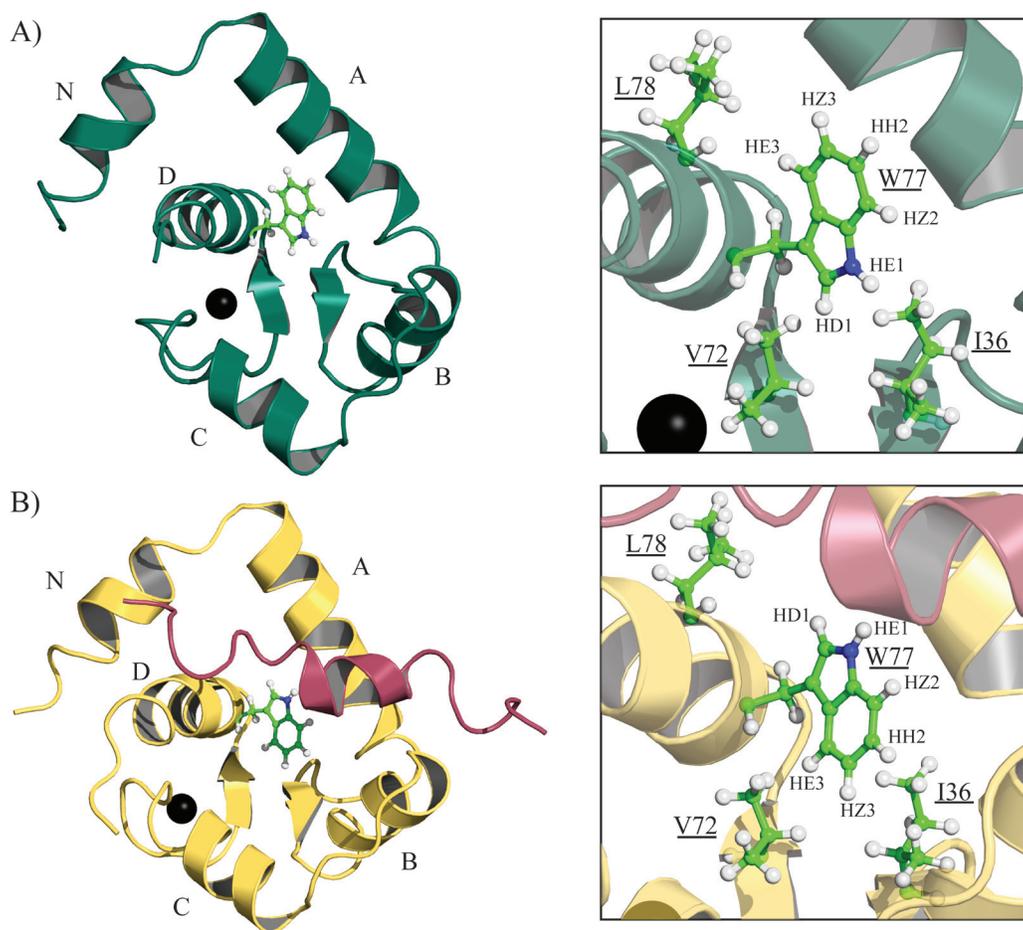


Fig. V-1. Opposite side chain orientation of W77 in water and TFE (A) Solution structure of calcium saturated mutant F77W-V82A of N-domain cardiac troponin C (cNTnC) determined in 19% TFE (v/v) (10). The side chain of W77 and residues at close proximity are shown in ball-and-sticks. The carbon atoms are colored in green, the hydrogens in white, the nitrogens in blue and the calcium ion in black. The W77 side chain orientation revealed an indole ring pointing towards helices N and A ($t90^\circ$). (B) Solution structure of calcium saturated F77W-cNTnC determined in H_2O and bound to its protein binding partner cTnI₁₄₄₋₁₆₃ (shown in red). The W77 side chain orientation revealed an indole ring pointing towards helices B and C ($t-105^\circ$). In comparison to A), the W77 indole ring is rotated close to 180° around its C β -C γ bond. PyMOL (DeLano Scientific LLC) was used to make these figures.

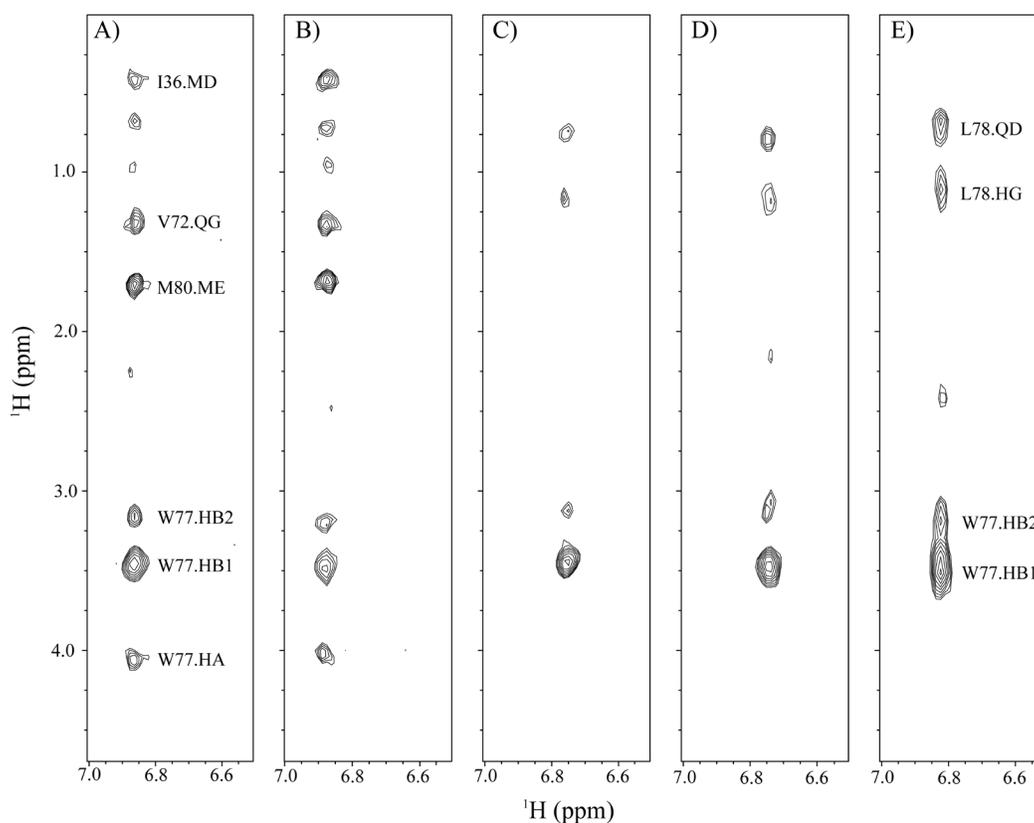


Fig. V-2. W77 CD1 planes of 3D ^{13}C NOESYHSQCs acquired at 600 MHz. The observable NOEs in the presence of TFE are shown in panels A and B, for F77W-V82A-cNTnC and F77W-cNTnC, respectively, revealing a W77 rotamer in the $t90^\circ$ position ($\chi_1 = \text{trans}$ and $\chi_2 \sim 90^\circ$). In opposition, the spectra acquired in H_2O are presented in panels C, D and E, for F77W-V82A-cNTnC, F77W-cNTnC and F77W-cNTnC•cTnI₁₄₄₋₁₆₃, respectively. These are characteristic of an tryptophan indole occupying a $t-105^\circ$ rotamer position ($\chi_1 = \text{trans}$ and $\chi_2 \sim -105^\circ$).

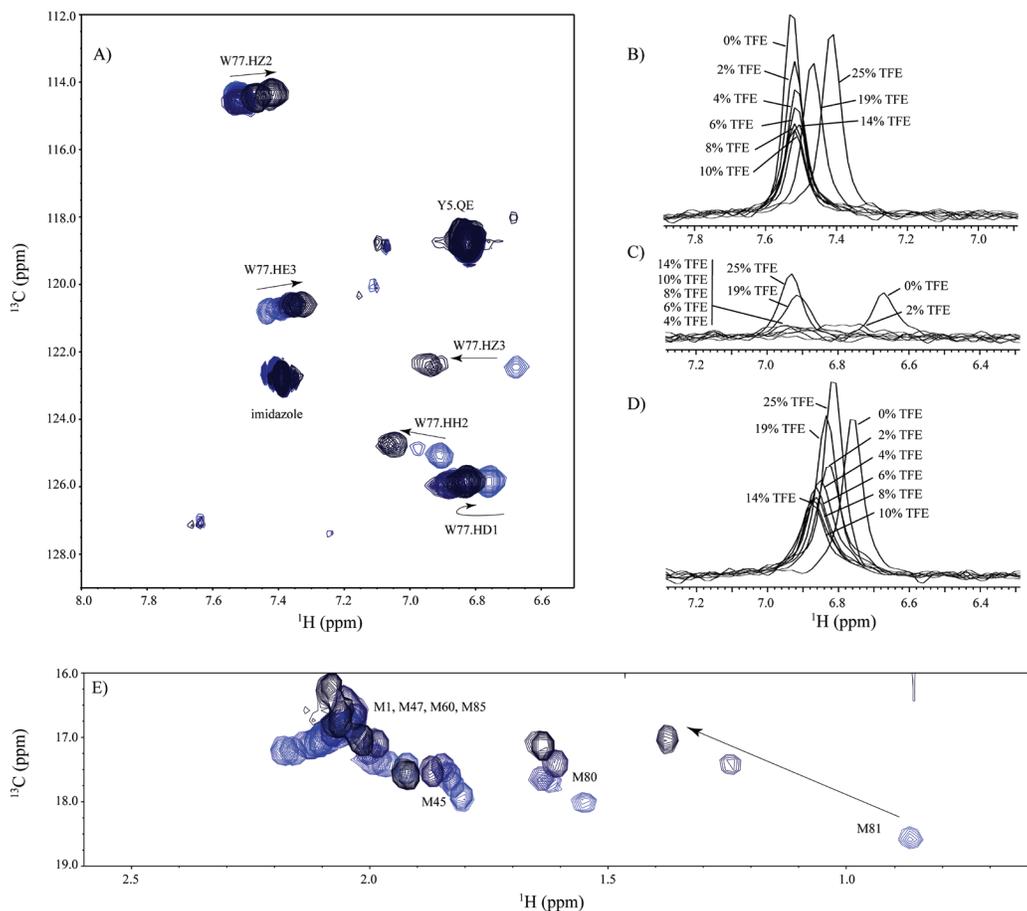


Fig. V-3. Titration of F77W-cNTnC with the co-solvent TFE monitored on a 500 MHz NMR spectrometer. A) and E) show the superimposition of the ^{13}C -HSQCs for the aromatic region and methionine methyl regions, respectively. From the light to dark blue correspond to 0 to 25% TFE (v/v). B-D) 1D trace of W77.HZ2, W77.HZ3 and W77.HD1 cross-peaks, respectively. The peak intensities were adjusted to take into account the dilution factors caused by the addition of TFE.

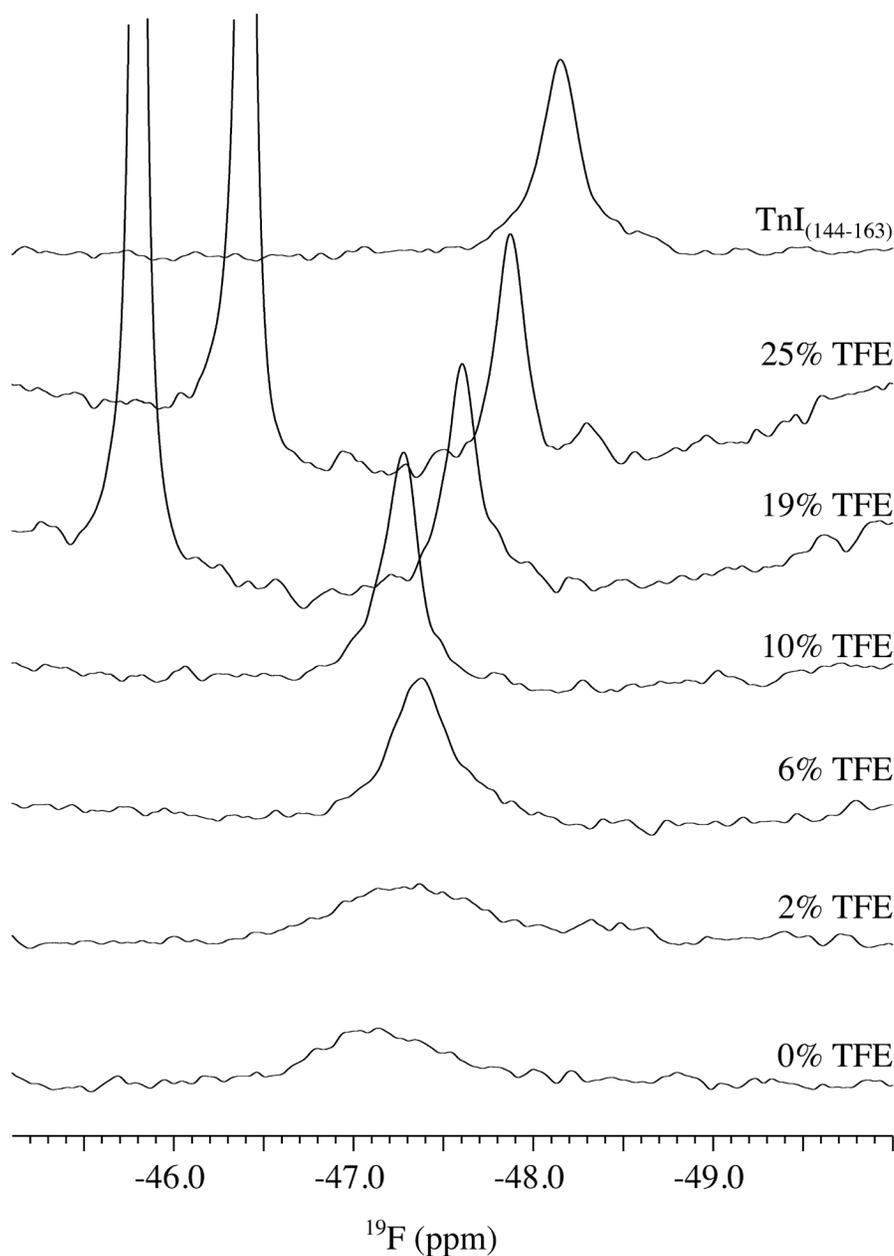


Fig. V-4. ^{19}F NMR spectra of F77(5fW)-cNTnC acquired during the titration with the co-solvent TFE (from 0 to 25% TFE). The spectra were acquired on a 600 MHz spectrometer with a $^1\text{H}/^{19}\text{F}$ probe. The spectrum at the top does not contain TFE, but is saturated with cTnI₁₄₄₋₁₆₃. The peak intensities were adjusted to take into account the dilution factors caused by the addition of TFE. Chemical shifts are with respect to TFA at 0 ppm. The resonance near -46 ppm arises from the AEBSF included in the cocktail of protease inhibitors.

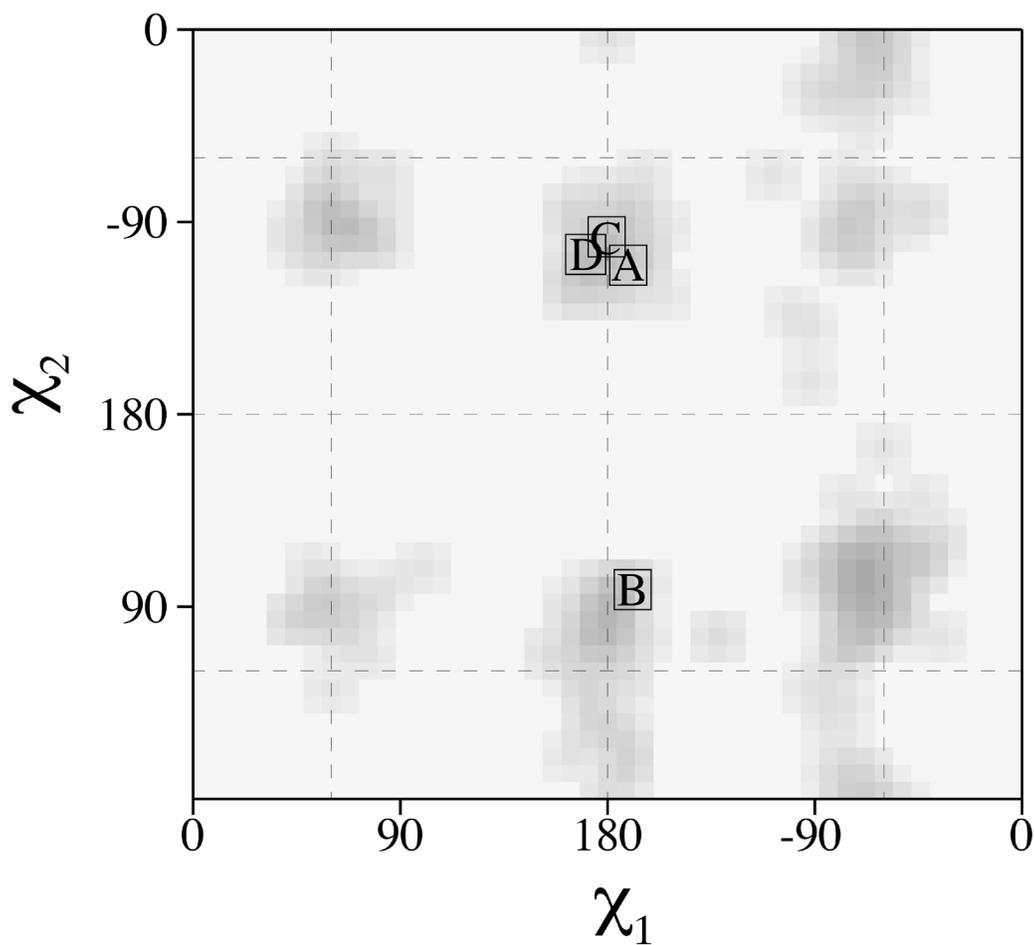


Fig. V-5. Tryptophan side chain $\chi_1 \times \chi_2$ conformations in A) F77W-cNTnC•cTnI₁₄₄₋₁₆₃ B) F77W-V82A-cNTnC in 19% TFE C) F153W-cCTnC and D) silver hake parvalbumin W102. The dark regions correspond to the most favorable conformations. The figure was adapted from the figure obtained using procheck nmr.⁵⁰

Table V-1. Structural statistics for F77W-cNTnC in the calcium saturated F77W-cNTnC•cTnI₁₄₄₋₁₆₃ complex for the 20 best target-function structures.

Intra-molecular NOE restraints	
Total	1280
Short range ($ i-j \leq 1$)	747
Medium range ($1 \leq i-j < 5$)	342
Long range ($ i-j \geq 5$)	191
Dihedral Restraints	
Total	134
ϕ	67
ψ	67
χ_1	0
Restraint violations	
Distance > 0.2 Å	4
Dihedral > 1°	0
RMSD to averaged coordinates (Å)	
Well defined regions (N, C _α , C) ^a	0.77
All regions (N, C _α , C)	0.95
N helix (5-11)	0.14
A helix (14-27)	0.36
B helix (41-48)	0.14
C helix (54-62)	0.23
D helix (74-85)	0.30
Average cyana target function	0.55
ϕ, ψ in core or allowed regions ^a	
Residues in most favoured regions	91.8%
Residues in additional allowed regions	8.2 %
Residues in generously allowed regions	0.0 %
Residues in disallowed regions	0.0 %

^a Using residues 5-85

Table V-2. Comparison of the W77 aromatic protons inter-residual NOEs observed in TFE and H₂O

NOEs	t90° rotamer (TFE)	t-105° rotamer (H ₂ O)
W77.HD1	I36.MD, V72.QG, M80.ME	L78.HG, L78.QD
W77.HZ2	A23.MB, M81.ME	A23.MB, L41.QD
W77.HH2	A23.MB	L41.QD
W77.HZ3	F20.MB F20.QD, F24.QD, W77.HA, L78.HN,	
W77.HE3	L78.HA, L78.QB, L78.QD, M80.ME, M81.ME	I36.MD, V72.QG

REFERENCES

1. Khemtemourian, L., Sani, M., Bathany, K., Grobner, G., and Dufourc, E. (2006) Synthesis and secondary structure in membranes of the Bcl-2 anti-apoptotic domain BH4, *J. Pept. Sci.* 12, 58-64.
2. Sonnichsen, F., Van Eyk, J. E., Hodges, R., and Sykes, B. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide, *Biochemistry* 31, 8790-8798.
3. Santiveri, C., Pantoja-Uceda, D., Rico, M., and Jimenez, M. (2005) Beta-hairpin formation in aqueous solution and in the presence of trifluoroethanol: a ¹H and ¹³C nuclear magnetic resonance conformational study of designed peptides, *Biopolymers* 79, 150-162.
4. Calamai, M., Chiti, F., and Dobson, C. (2005) Amyloid fibril formation can proceed from different conformations of a partially unfolded protein, *Biophys. J.* 89, 4201-4210.
5. Pallares, I., Vendrell, J., Aviles, F., and Ventura, S. (2004) Amyloid fibril formation by a partially structured intermediate state of alpha-chymotrypsin, *J. Mol. Biol.* 342, 321-331.
6. Shanmugam, G., Polavarapu, P., Gopinath, D., and Jayakumar, R. (2005) The structure of antimicrobial pexiganan peptide in solution probed by Fourier transform infrared absorption, vibrational circular dichroism, and electronic circular dichroism spectroscopy, *Biopolymers* 80, 636-642.
7. Kanno, T., Yamaguchi, K., Naiki, H., Goto, Y., and Kawai, T. (2005) Association of thin filaments into thick filaments revealing the structural hierarchy of amyloid fibrils, *J. Struct. Biol.* 149, 213-218.
8. Slupsky, C. M., Kay, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) Calcium-induced dimerization of troponin C: mode of interaction and use of trifluoroethanol as a denaturant of quaternary structure, *Biochemistry* 34, 7365-7375.
9. Pääkkönen, K., Annala, A., Sorsa, T., Pollesello, P., Tilgmann, C., Kilpeläinen, I., Karisola, P., Ulmanen, I., and Drakenberg, T. (1998)

- Solution structure and main chain dynamics of the regulatory domain (Residues 1-91) of human cardiac troponin C, *J. Biol. Chem.* 273, 15633-15638.
10. Julien, O., Sun, Y., Wang, X., Lindhout, D., Thiessen, A., Irving, M., and Sykes, B. (2008) Tryptophan mutants of cardiac troponin C: 3D structure, troponin I affinity, and in situ activity, *Biochemistry* 47, 597-606.
 11. Slupsky, C. M., and Sykes, B. D. (1995) NMR solution structure of calcium-saturated skeletal muscle troponin C, *Biochemistry* 34, 15953-15964.
 12. Buck, M. (1998) Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins, *Q. Rev. Biophys.* 31, 297-355.
 13. Dong, W., Xing, J., Villain, M., Hellinger, M., Robinson, J., Chandra, M., Solaro, R., Umeda, P., and Cheung, H. (1999) Conformation of the regulatory domain of cardiac muscle troponin C in its complex with cardiac troponin I, *J. Biol. Chem.* 274, 31382-31390.
 14. Cooke, R., Crowder, M., and Thomas, D. (1982) Orientation of spin labels attached to cross-bridges in contracting muscle fibres, *Nature* 300, 776-778.
 15. Pirani, A., Vinogradova, M. V., Curmi, P. M., King, W. A., Fletterick, R. J., Craig, R., Tobacman, L. S., Xu, C., Hatch, V., and Lehman, W. (2006) An atomic model of the thin filament in the relaxed and Ca²⁺-activated states, *J. Mol. Biol.* 357, 707-717.
 16. Gagné, 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, 784-789.
 17. Takeda, S., Yamashita, A., Maeda, K., and Maeda, Y. (2003) Structure of the core domain of human cardiac troponin in the Ca²⁺-saturated form, *Nature* 424, 35-41.
 18. Herzberg, O., and James, M. (1985) Structure of the calcium regulatory muscle protein troponin-C at 2.8 Å resolution, *Nature* 313, 653-659.
 19. Corrie, J. E. T., Brandmeier, B. D., Ferguson, R. E., Trentham, D. R., Kendrick-Jones, J., Hopkins, S. C., Van Der Heide, U. A., Goldman, Y. E., Sabido-David, C., Dale, R. E., Criddle, S., and Irving, M. (1999) Dynamic

- measurement of myosin light-chain-domain tilt and twist in muscle contraction, *Nature* 400, 425-430.
20. 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, 865-874.
 21. De Simone, A., Corrie, J. E. T., Dale, R., Irving, M., and Fraternali, F. (2008) Conformation and dynamics of a rhodamine probe attached at two sites on a protein: implications for molecular structure determination in situ, *J. Am. Chem. Soc.* 130, 17120-17128.
 22. Julien, O., Mercier, P., Spyropoulos, L., Corrie, J. E. T., and Sykes, B. D. (2008) NMR studies of the dynamics of a bifunctional rhodamine probe attached to troponin C, *J. Am. Chem. Soc.* 130, 2602-2609.
 23. Lakowicz, J. (2000) On spectral relaxation in proteins, *Photochem. Photobiol.* 72, 421-437.
 24. Li, M. X., Chandra, M., Pearlstone, J. R., Racher, K. I., Trigo-Gonzalez, G., Borgford, T., Kay, C. M., and Smillie, L. B. (1994) Properties of isolated recombinant N and C domains of chicken troponin C, *Biochemistry* 33, 917-925.
 25. Moncrieffe, M. C., Venyaminov, S. Y., Miller, T. E., Guzman, G., Potter, J. D., and Prendergast, F. G. (1999) Optical spectroscopic characterization of single tryptophan mutants of chicken skeletal troponin C: evidence for interdomain interaction, *Biochemistry* 38, 11973-11983.
 26. Moncrieffe, M. C., Eaton, S., Bajzer, Z., Haydock, C., Potter, J. D., Laue, T. M., and Prendergast, F. G. (1999) Rotational and translational motion of troponin C, *J. Biol. Chem.* 274, 17464-17470.
 27. Hutnik, C., Macmanus, J., Banville, D., and Szabo, A. (1990) Comparison of metal ion-induced conformational changes in parvalbumin and oncomodulin as probed by the intrinsic fluorescence of tryptophan 102, *J. Biol. Chem.* 265, 11456-11464.
 28. Moncrieffe, M. C., Juranic, N., Kemple, M. D., Potter, J. D., Macura, S., and Prendergast, F. G. (2000) Structure-fluorescence correlations in a single

- tryptophan mutant of carp parvalbumin: solution structure, backbone and side-chain dynamics, *J. Mol. Biol.* 297, 147-163.
29. Wang, X., Mercier, P., Letourneau, P. J., and Sykes, B. D. (2005) Effects of Phe-to-Trp mutation and fluorotryptophan incorporation on the solution structure of cardiac troponin C, and analysis of its suitability as a potential probe for in situ NMR studies, *Protein Sci.* 14, 2447-2460.
 30. Zhao, X., Devries, J., McDonald, R., and Sykes, B. (2007) Determination of the ¹⁹F NMR chemical shielding tensor and crystal structure of 5-fluoro-dl-tryptophan, *J. Magn. Reson.* 187, 88-96.
 31. Richardson, R. C., King, N. M., Harrington, D. J., Sun, H., Royer, W. E., and Nelson, D. J. (2000) X-Ray crystal structure and molecular dynamics simulations of silver hake parvalbumin (Isoform B), *Protein Sci.* 9, 73-82.
 32. Nishita, K., Tanaka, H., and Ojima, T. (1994) Amino acid sequence of troponin C from scallop striated adductor muscle, *J. Biol. Chem.* 269, 3464-3468.
 33. Ozkirimli, E., Yadav, S., Miller, W., and Post, C. (2008) An electrostatic network and long-range regulation of Src kinases, *Protein Sci.* 17, 1871-1880.
 34. Witter, R., Nozirov, F., Sternberg, U., Cross, T., Ulrich, A., and Fu, R. (2008) Solid-state ¹⁹F NMR spectroscopy reveals that Trp41 participates in the gating mechanism of the M2 proton channel of influenza A virus, *J. Am. Chem. Soc.* 130, 918-924.
 35. Gagné, S. M., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C, *Protein Sci.* 3, 1961-1974.
 36. Li, M. X., Gagné, 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, 8330-8340.
 37. Baryshnikova, O., Williams, T., and Sykes, B. D. (2008) Internal pH indicators for biomolecular NMR, *J. Biomol. NMR* 41, 5-7.

38. Zhang, O. W., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) Backbone H-1 and N-15 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, 845-858.
39. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
40. 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, 313-321.
41. Graether, S., Devries, J., McDonald, R., Rakovszky, M., and Sykes, B. (2006) A 1H/19F minicoil NMR probe for solid-state NMR: application to 5-fluoroindoles, *J. Magn. Reson.* 178, 65-71.
42. Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods Mol. Biol.* 278, 353-378.
43. 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.
44. Laskowski, R. A., Rullmann, J. A., Macarthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR* 8, 477-486.
45. Willis, K., Neugebauer, W., Sikorska, M., and Szabo, A. (1994) Probing alpha-helical secondary structure at a specific site in model peptides via restriction of tryptophan side-chain rotamer conformation, *Biophys. J.* 66, 1623-1630.
46. Markley, J., Bax, A., Arata, Y., Hilbers, C., Kaptein, R., Sykes, B., Wright, P. E., and Wuthrich, K. (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids--IUPAC-IUBMB-IUPAB Inter-Union Task Group on the standardization of data bases of protein and

-
- nucleic acid structures determined by NMR spectroscopy, *Eur. J. Biochem.* 256, 1-15.
47. Lovell, S. C., Word, J., Richardson, J., and Richardson, D. (2000) The penultimate rotamer library, *Proteins* 40, 389-408.
 48. Dezube, B., Dobson, C., and Teague, C. (1981) Conformational-analysis of tryptophan in solution using nuclear magnetic-resonance methods, *J. Chem. Soc. Perk. Trans. 2.* 730-735.
 49. Ross, J. B., Wyssbrod, H., Porter, R., Schwartz, G., Michaels, C., and Laws, W. (1992) Correlation of tryptophan fluorescence intensity decay parameters with ¹H NMR-determined rotamer conformations: [tryptophan₂]oxytocin, *Biochemistry* 31, 1585-1594.
 50. Johansson, J., Scharf, D., Davies, L., Reddy, K., and Eckenhoff, R. (2000) A designed four-alpha-helix bundle that binds the volatile general anesthetic halothane with high affinity, *Biophys. J.* 78, 982-993.

CHAPTER VI – Chimera*Is There Nascent Structure in the Intrinsically Disordered Region of Troponin I?*

Olivier Julien,¹ Pascal Mercier,¹ Claire N. Allen,² Olivier Fiset,³ Carlos H. I. Ramos,⁴ Patrick Lagüe,³ Tharin M. A. Blumenschein,² and Brian D. Sykes^{1,*}

¹ *Department of Biochemistry, University of Alberta, Edmonton, AB, Canada*

² *School of Chemistry, University of East Anglia, Norwich, UK*

³ *Département de biochimie et de microbiologie, Université Laval, Québec, QC, Canada*

⁴ *Instituto de Química, Universidade Estadual de Campinas, SP, Brazil*

Publication: Julien et al. 2011. *Proteins: Structure, Function, and Bioinformatics*, (in press).

Contributions: OJ created the four models. OF and PL ran the MD simulations. CNA expressed the protein. TMB acquired the NMR data. OJ and PM analysed the MD trajectories, assigned the chemical shifts, analysed the NMR relaxation data. OJ, TMB and BDS wrote the paper. OJ made all of the figures.

OVERVIEW

In striated muscle, the binding of calcium to troponin C (TnC) results in the removal of the C-terminal region of the inhibitory protein troponin I (TnI) from actin. While structural studies of the muscle system have been successful in determining the overall organization of most of the components involved in force generation at the atomic level, the structure and dynamics of the C-terminal region of TnI remains controversial. This domain of TnI is highly flexible, and it has been proposed that this intrinsically disordered region (IDR) regulates contraction via a ‘fly-casting’ mechanism. Different structures have been presented for this region using different methodologies: a single α -helix, a “mobile domain” containing a small β -sheet, an unstructured region, and a two helix segment. To

investigate whether this IDR has in fact any nascent structure, we have constructed a skeletal TnC-TnI chimera that contains the N-domain of TnC (1-90), a short linker (GGAGG), and the C-terminal region of TnI (97-182) and have acquired ^{15}N NMR relaxation data for this chimera. We compare the experimental relaxation parameters with those calculated from molecular dynamic simulations using four models based upon the structural studies. Our experimental results suggest that the C-terminal region of TnI does not contain any defined secondary structure, supporting the ‘fly-casting’ mechanism. We interpret the presence of a “plateau” in the ^{15}N NMR relaxation data as being an intrinsic property of IDRs. We also identified a more rigid adjacent region of TnI that has implications for muscle performance under ischemic conditions.

INTRODUCTION

Human skeletal, cardiac and smooth muscles all produce force in a similar manner and all are calcium regulated. For the striated muscle found in skeletal and cardiac muscle tissues, the sarcoplasmic calcium concentration dictates the activation or inhibition of the contractile machinery through the troponin complex. Troponin is composed of three structural subunits found in an equimolar ratio. Troponin C (TnC) is the calcium-binding subunit, which senses changes in the cellular calcium concentration. Troponin T (TnT) is the tropomyosin-interacting subunit that serves as an anchor to position the troponin complex in place on the thin filament. Troponin I (TnI) is the inhibitory subunit, which modulates the actomyosin ATPase activity by binding to actin (for reviews, see (1-4)).

When calcium binds to TnC, a structural conformational change occurs in the N-terminal domain of TnC, which leads to the binding of the ‘switch’ region of TnI (sSp) and subsequent release of TnI from actin. The resulting movement of tropomyosin exposes myosin-binding sites on actin, resulting in the power stroke. While many structural studies of the muscle system have been successful in determining the overall organization of most of the components involved in force

generation at the atomic level, the structure and dynamics of TnI remains controversial. Specifically, the C-terminal region of TnI (residues 97-182) has been the focus of many studies (discussed below), but a clear picture of the structure, dynamics and function of this region remains unclear. This domain of TnI is highly flexible, and it has been proposed that this intrinsically disordered region (IDR) regulates contraction via a ‘fly-casting’ mechanism (5-7).

The focus of this article is the skeletal C-terminal region of TnI (sTnI), because of its important role in the regulation of muscle contraction. For example, many familial hypertrophic cardiomyopathy mutations occur in this region (8). In a previous study of the 52 kDa troponin complex by nuclear magnetic resonance (NMR) spectroscopy, we recognized the C-terminal region of sTnI as having the characteristics of an IDR and probed the dynamics of this region in presence and absence of calcium (6). Our results indicated that this region is highly flexible based on NMR relaxation data and chemical shifts. Similar results have also been observed for cardiac TnI (9). Interestingly, we observed a "plateau" in the relaxation data in the middle of the IDR of TnI, which we interpreted as a possible indication of some nascent structure.

The question remains: is there a nascent structure in the C-terminal region of TnI? Different structures have been presented over the years for this region using different methodologies: a long single α -helix has been observed by X-ray crystallography in the structure of the cardiac troponin complex (10), a “mobile domain” containing a small β -sheet flanked by two short α -helices was determined using solution NMR (11), two anti-parallel helical segments were modelled to fit SANS and SAXS data (12), and a mainly unstructured region proposed from NMR relaxation and chemical shifts (6), further supported by the absence of electronic density for residues 144-182 of TnI in the X-ray structure of skeletal troponin (13). It is important to mention that all these studies were done in the off-state, when TnI is bound to TnC and removed from actin. The dynamics has also been probed bound to actin using SDSL-EPR spectroscopy (14). The

question of whether there is in fact any structure, nascent or otherwise, in this region of TnI is the focus of this paper.

To avoid the experimental constraints created by the size of the troponin complex, we have constructed a skeletal TnC-TnI chimera to study the structure and dynamics of the regulatory region of troponin, combining the skeletal N-terminal domain of TnC (sNTnC) and the C-terminal region of TnI (sTnI) in a single polypeptidic chain (15). The sequence of the skeletal TnC-TnI chimera contains 181 amino acids; it starts with a methionine due to the expression system (omitted in the numbering), followed by residues 1 to 90 of sTnC linked by a short GGAGG flexible linker to residues 98 to 182 of sTnI. When the TnI portion of the chimera is discussed in the text, it is referred to as residues 98-182 in order to agree with the human skeletal TnI sequence numbering. The use of this 20 kDa construct has many advantages over the 52 kDa ternary troponin complex, avoiding problems with limited stability, solubility, aggregation, and the rigid requirement of the equimolar ratios of the different subunits for the proper assembly of the ternary complex in solution. The use of high salt concentrations (250 mM) in previous studies could possibly have lead to apparent IDR, and the use of ^2H labelling might have altered the stability of the complex. The increased stability of the chimera makes it easier to study under a variety of conditions. The chimera has also been previously shown to be functional, capable of regulating actomyosin ATPase (15).

We have assigned the backbone chemical shifts of the chimera (142 out of 180 residues) and measured backbone amide ^{15}N NMR relaxation data. Experimental relaxation parameters were compared with those calculated from molecular dynamic simulations using four models for the structure of the chimera (Fig. VI-1), based upon the previous structural studies mentioned above. We conclude that none of the secondary structure models proposed so far for the C-terminal region of TnI is correct in solution. This region is likely an IDR, although the molecular dynamic simulations were unable to accurately reproduce the large motions of an

IDR. The chimera is an authentic mimic for the study of the ‘structure’ and dynamics of the C-terminal region of TnI, since the chemical shifts and relaxation data for both the TnC and TnI regions are unperturbed compared to the same regions in the core troponin complex.

MATERIALS AND METHODS

Sample preparation

Uniformly [^{15}N]- or [^{15}N , ^{13}C]-labeled skeletal TnC- skeletal TnI chimera was expressed as described (15) from the plasmid pET3aTnC(1–91)–TnI(98–182), but with the addition of 0.8 mM IPTG instead of lactose. Purification procedures were adapted from Tiroli et al. (15). Cells were lysed by sonication, followed by centrifugation for 3h20 min at 18,500 rpm at 4 °C in a Beckman Avanti centrifuge, rotor JA20 (Palo Alto, CA, USA). The supernatant was dialysed against 100 mM acetate buffer, pH 5.0, 1 mM CaCl_2 , 50 mM NaCl and centrifuged for 20 min at 18,500 rpm at 4 °C in a Beckman Avanti centrifuge, rotor JA20. The soluble fraction was loaded into a HiTrap Q column (GE Healthcare) equilibrated with the same buffer. The chimera was eluted in the flow-through. NMR samples were prepared in 90% H_2O /10% D_2O , 100 mM KCl, 10 mM Imidazole, 1 mM CaCl_2 , 0.03% sodium azide, and 0.2 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid. The pH was adjusted to 6.8 according to the frequency of the imidazole signal (16).

NMR spectroscopy

NMR spectra were collected at 30 °C (unless specified otherwise). Varian Inova 600 and 800 MHz, and Bruker Avance III 800 MHz NMR spectrometers equipped with triple resonance probes with Z-pulsed field gradients and a computer-controlled variable temperature (VT) module to regulate the temperature were used in this study. The 1D ^1H spectra were processed and plotted with VnmrJ 2.1B. The 2D and 3D spectra were all processed with NMRPipe (17) and analyzed with NMRViewJ (One Moon Scientific, inc.). The

relaxation data at 600 and 800 MHz were analyzed with the rate analysis module of NMRViewJ (v8.0-rc17).

A temperature series (from 5 °C to 50 °C) was performed on a [¹⁵N]-labeled sample at 600 MHz. One-dimensional ¹H and 2D ¹H-¹⁵N HSQC NMR spectra were acquired at each temperature. The spectra were acquired with the water and gNhsqc pulse sequences of Biopack (Varian Inc.), respectively. The 1D spectra were acquired with a spectral width of 14 ppm, 128 transients, a relaxation delay of 1 s and an acquisition delay of 2 s. The ¹H-¹⁵N HSQC were acquired with 16 transients, a relaxation delay of 1 s and spectral widths of 14 ppm (ω_2) and 40 ppm (ω_1) with 2048 and 192 complex points, respectively. The 3D CBCACONH and HNCACB were acquired at 800 MHz with a relaxation delay of 1 s, spectral widths of 15 ppm (ω_3), 30 ppm (ω_2) and 75 ppm (ω_1) with 2048, 64 (or 42) and 164 (or 128) complex points, respectively. The 3D HNCACO and HNCO were acquired at 800 MHz with a relaxation delay of 1 s, spectral widths of 15 ppm (ω_3), 30 or 36 ppm (ω_2) and 36 ppm (ω_1) with 2048, 64 and 164 (or 192) complex points, respectively. In order to complete the chemical shifts assignment, a 3D ¹⁵N NOESYHSQC was also acquired at 800 MHz, with a relaxation delay of 1.2 s, a mixing time of 0.1 s, spectral widths of 15 ppm (ω_3), 10 ppm (ω_2) and 40 ppm (ω_1) with 2048, 184 and 116 complex points, respectively.

Backbone amide relaxation data

All backbone amide relaxation data (¹⁵N-T₁, ¹⁵N-T₂ and {¹H}-¹⁵N-NOE) were measured from a series of 2D ¹H-¹⁵N HSQC spectra acquired at both 600 and 800 MHz at 30 °C. The relaxation experiments were all acquired with spectral widths of 15 ppm (ω_2) and 40 ppm (ω_1), and with 2048 (t_1) x 192 (t_2) complex points and 2048 (t_1) x 128 (t_2) complex points at 600 and 800 MHz, respectively. The T₁ relaxation delays were (10, 100, 200, 500, 750, 1000, 2000, 3000 ms) at 600 MHz and (20, 100, 200, 500, 750, 1000, 2000, 3000, 5000 ms) at 800 MHz. For ¹⁵N-T₂ measurements, the relaxation delays were set to (10, 50, 90, 130, 170, 210, 250 ms) at 600 MHz and (17, 51, 85, 136, 170, 214, 255 ms) at 800 MHz. At 600

MHz, the delay between repetitions of the pulse sequence was 4 s for both $^{15}\text{N-T}_1$ and $^{15}\text{N-T}_2$. The NOE were measured in the absence and presence of proton saturation (5 s at 600 MHz, 6 s at 800 MHz).

Computational methods

The *CHARMM* program (18), version c35b1, was used to build all four troponin hybrid TnC-TnI systems and perform simulations. Initial coordinates for each model were obtained by modeling residues 1-180 of the chimera using different published and deposited protein structures of the troponin system, or structural information of the Tn complex. The program modeller was used to generate starting coordinates (19). For the four models presented in Fig. VI-1, the sTnC portion of the chimera was modeled using both the structure of sTnC in the X-ray structure of the skeletal Tn complex (PDB ID 1YTZ), as well as the NMR structure of sTnC in complex with residues 115-131 of sTnI with a bifunctional rhodamine probe attached on the C helix of sTnC (PDB ID 1NPQ). For the sTnI portion of the chimera, the starting structures were built as follow. Model 1 (Fig. VI-1A) is based on a solution NMR structure of the C-terminal region of TnI (11). This structure, known as the mobile domain, contains a two-stranded anti-parallel β -sheet flanked by two α -helices (PDB ID 1VJD). Model 2 (Fig. VI-1B) is derived from the crystallographic X-ray structure of the cardiac troponin complex (10), and exhibits a long α -helix followed by a short disordered region as seen in one of the two presented conformers presented (PDB ID 1JLE). Model 3 (Fig. VI-1C) contains no restraints for sTnI portion of the chimera, based on the chemical shifts and NMR relaxation data for the skeletal Tn complex (6). Model 4 (Fig. VI-1D) is derived from the crystal structure of the skeletal troponin complex where residues 126-143 of sTnI in the structure of the sTn complex are disordered, but electronic density is observable in one of the two conformers presented (PDB ID 1YTZ) (13). The structure of this region highlights the interaction between residues 125-132 of sTnI with α -helix A of sTnC. In this model, we have included the H2 helix of the "mobile domain" (Model 1) to assess the possibility of a short α -helix as suggested by our chemical shift data, and ESR data for the homologue

region of TnI in the cardiac system that becomes more ordered when binding to actin in the ‘on’ state (14).

Crystallographic water and ions were removed from all models. Hydrogen positions were computed using *CHARMM*'s hbuild routines. The initial protonation state of acid-base residues was determined using Poisson-Boltzman calculations and the *CHARMM* pbeq module (20) with the recommended protocol. The protein dielectric constant was set to 5, and a value of 80 represented the bulk solvent. All residues showed standard protonation states at a pH of 7.0. Calcium atoms, two per system, were placed in each calcium-binding domain, for a total system size of 2817 atoms.

Langevin dynamics were carried out using the CHARMM22 forcefield (21) with ϕ , ψ cross-term map correction (CMAP) (22), a leapfrog integrator, and one femtosecond timestep. The Generalized Born with a simple switching (GBSW) implicit solvent model (23) was used with the recommended settings: 0.6 Å smoothing length, 0.03 kcal/mol/Å² nonpolar surface tension coefficients, 1.5 Å grid spacing and the optimized radii from the gbsw module. The non-bonded interaction cut-off for both Van der Waals and electrostatics was set to 20 Å and smoothed up to 16 Å using a switching function. The length of bonds involving hydrogen atoms was constrained with the SHAKE algorithm (23). The Langevin temperature bath was set to 25 °C, with a 5.0 °C allowed temperature deviation. Prior to dynamics, proper initial velocities were assigned randomly at -70 °C; the system was then heated progressively to 25 °C over 100 ps, and equilibrated for 4900 ps. Production runs were then recorded, one 20 ns trajectory (or more) for each model, with coordinates saved every picosecond.

The stability of the simulation was assessed from total system energy and backbone RMSD excluding the flexible domain, both of which were stable after about 1 ns, well under the equilibration time. The CHARMM NMR module was used to compute the NMR relaxation parameters (¹⁵N-T₁, ¹⁵N-T₂ and {¹H}-¹⁵N-

NOE). Prior to analysis, global rotation and translation were removed by an RMSD-minimizing superposition, taking into account only residues located in α -helices and β -strands in the TnC moiety. Therefore, only local TnI motions are considered. Relaxation times were calculated for a spectrometer frequency of 600 MHz using a 10 ns global tumbling time (estimated from molecular weight, $\tau_c(\text{ns}) \sim \text{MW}(\text{kDa})/2$) and a chemical shift anisotropy of 160 ppm for the ^{15}N nucleus. The correlation functions for the molecular motions (24) were computed over 75% of the time domain of the trajectory length. Order parameters (S^2) were computed as the value of the autocorrelation functions at 15 ns.

RESULTS

The residues 1-90 of the chimera correspond to the regulatory domain of sTnC (methionine from the expression system omitted), residues 91-95 are the GGAGG linker, and residues 96-180 correspond to residues 98-182 of the C-terminal portion of sTnI. When the sTnI portion of the chimera is discussed in the text, it is referred to as residues 98-182 in order to agree with the human skeletal TnI sequence numbering.

Chemical Shift Assignment

Temperature series were performed to determine the optimal temperature that should be used for the acquisition of the 3D spectra. On the one hand, higher temperatures allow for faster tumbling of the chimera in solution, which translate into narrower line widths. On the other hand, any nascent structure might get denatured at higher temperatures. A stacked plot of the 1D ^1H - ^{15}N HSQC spectra is in Fig. VI-2. We chose to acquire our experiments at 30 °C where virtually all of the previous NMR data has been acquired. The 2D ^1H - ^{15}N HSQC NMR spectrum acquired at 30 °C is presented in Fig. VI-3A and shows the assignments of the amide cross-peaks. The spectrum is well resolved, and virtually all expected resonances are present. Fig. VI-3B shows a superimposition of reconstructed ^1H - ^{15}N HSQCs based on the chemical shifts of the chimera and the

respective domains of TnC (25) and TnI (11) in the Tn complex. The chemical shifts were globally adjusted to compensate for the TROSY acquisition mode and referencing methods. This highlights that the ^1H - ^{15}N HSQC spectrum of the chimera resembles a combination of the sTnI and sNTnC spectra as seen in the troponin complex.

Backbone chemical shift assignments were obtained using 3D HNCACB, 3D HNCACO, 3D HNCO and 3D ^{15}N NOESYHSQC NMR spectra. The assignments were completed to 80%, 80%, 82%, 80% and 79% for the HN, N, CO, C α and C β nuclei, respectively. We have compared the chemical shifts of the sTnC and sTnI portions of the chimera with those of the respective proteins in the Tn complex (Fig. VI-4). The HN, N and C α chemical shifts of sTnC in the chimera revealed only minimum changes compared to those obtained previously for sNTnC in the Tn complex (25), suggesting a high level of structural similarity. The small discrepancies for the C α nucleus originate from the deuterium isotope shift in the sTn complex. The chemical shifts of the sTnI domain of the chimera were also similar to those of the C-terminal region of TnI in the ternary Tn complex (BMRB 9500) (11). These comparisons indicate that the structural characteristics of the IDR of TnI are virtually identical in the skeletal Tn complex and in the TnC-TnI chimera used in this study. The ^{13}C chemical shifts (CO, C α and C β) of the chimera were used to calculate the chemical shift index (CSI) (26) to predict the secondary structure of the chimera in Fig. VI-4B. These results confirm the expected secondary structures for the sTnC domain, but do not indicate any structure in the IDR of sTnI other than a short helix of five residues towards the end of the C-terminus.

NMR relaxation data

To assess the dynamic properties of the different regions of the chimera, we acquired the three standard backbone amide NMR relaxation experiments ^{15}N -T $_1$, ^{15}N -T $_2$ and $\{^1\text{H}\}$ - ^{15}N -NOE at 600 and 800 MHz (Fig. VI-5). The ^{15}N -T $_1$ data for the sTnC region (residues 4-85) were 0.3 s higher at 800 MHz than at 600 MHz,

with averages (\pm average errors) of 1.00 ± 0.03 s and 0.70 ± 0.04 s, respectively, as expected for a rigid globular domain. The ^{15}N - T_1 data for the C-terminal region of sTnI (residues 137-181), however, were similar at both fields (average of 0.57 ± 0.02 at 600 MHz, and 0.65 ± 0.02 at 800 MHz), characteristic of an IDR. The average ^{15}N - T_2 values at 600 MHz [800 MHz] for the sTnC and sTnI regions were 0.071 ± 0.004 [0.061 ± 0.003] and 0.258 ± 0.009 [0.215 ± 0.003], respectively. These values are typical for a globular domain for sTnC and a disordered region for the C-terminal region of sTnI. The ^{15}N - T_2 data show the presence of a "plateau" at approximately $T_2 = 0.25$ s for residues 140 to 170 of the IDR of sTnI, as observed for sTnI in the Tn complex (6). Of particular interest, the average ^{15}N - T_2 at 600 MHz [800 MHz] for residues 124 to 130 of sTnI (sTnI₁₂₄₋₁₃₀) is 0.062 ± 0.005 s [0.049 ± 0.003 s]. These values for the switch region of sTnI are more similar to sTnC as opposed to sTnI (see discussion section for more details). The $\{^1\text{H}\}$ - ^{15}N -NOE data reflects a well-folded protein for the sTnC region, with NOE values approaching 1.0, but dropping rapidly to lower values starting at residue 137 in the sTnI domain, suggesting a much more flexible region. Again, the NOE values for residues 124-130 of sTnI are more similar to the sTnC region, as opposed to the rest of the C-terminal region of sTnI. Overall, the ^{15}N - T_1 , ^{15}N - T_2 and $\{^1\text{H}\}$ - ^{15}N -NOE data obtained for the chimera are very similar to the experimental relaxation data for sTn (6,25), justifying the use of the chimera to probe the motions of both regulatory domains when interacting to each other.

Molecular dynamics simulations

The validity of the different proposed models was assessed by molecular dynamic simulation using CHARMM. Four different trajectories using implicit solvent were acquired and 15 ns of each trajectory was analysed to probe the motions of the chimera *in silico*. The correlation functions $C(t)$ of each residue in each model as a function of time are presented in the top panel of Fig. VI-6A-D. The correlation functions of the backbone amide NH vector reported are a measure of the mobility of a particular region; if the value of $C(t)$ remains high this indicates

that the motions of these residues are restricted. On the other hand, if the value of $C(t)$ decreases as a function of time, this indicates that the motions of a given residue are large on the time scale studied, and the conformation diverges over time. Of course, since the rotational reorientation of the TnC domain was removed prior to analysis (see Methods), $C(t)$ is close to one for residues in this domain that do not experience internal motions. While the decays of the $C(t)$ s for the individual residues are not labelled, the black band at the top of each panel corresponds to residues in the sTnC domain, and the ever more rapidly decaying curves correspond to residues in the C-terminus of the sTnI domain. To illustrate this, $C(t)$ is plotted on a per-residue basis at 15 ns in lower graphs for each model. These values of $C(t)$ reflect the order parameter of the internal motions of the chimera. The values close to 1.0 correspond to residue with limited backbone amide dynamics, observed for the rigid residues of the sTnC region in each model. For the sTnI region, the results show low order reflecting considerable motion. Only model 4 predicts the correct position where the motion begins as seen in the experimental relaxation data (see below).

DISCUSSION

The chemical shifts of residues in the sTnC and sTnI regions of the chimera were virtually identical to those observed for both proteins when part of the full-length troponin complex (6,25). Further, the dynamics of the chimera probed by ^{15}N NMR relaxation data also showed the same pattern as previously published. This supports the use of this chimera to characterize the structure and dynamics of the C-terminal region of TnI. The skeletal TnC-TnI chimera is also superior for all the technical reasons explained above, including enhanced stability (robustness to degradation over time) and NMR simplicity. We have used the ^{13}C chemical shifts as an indicator of secondary structure. Overall, the chemical shifts predict all expected secondary structure elements for the sTnC region, and point towards an IDR for the sTnI region, except for a short but possible nascent H2 helix as reported in the mobile domain structure (11).

Two regions of the chimera revealed interesting relaxation data pattern. First, the "plateau" observed in the ^{15}N - T_1 , ^{15}N - T_2 and $\{^1\text{H}\}$ - ^{15}N -NOE NMR relaxation data for residues 140 to 170 of sTnI in the Tn complex (6) was also observed in the chimera. Although the relaxation data clearly show that this region is flexible, one possibility to explain the presence of the plateau is the presence of some sort of nascent structure. A second possibility is that this profile is an intrinsic property of disordered regions in general. We found that similar results can be recognized in other IDRs attached to globular domains using ^{15}N relaxation data, such as the data of Zhukov et al. who reported the ^{15}N NMR relaxation study of the ribosome-associated cold shock response protein Yfia (27). This 90 amino acid protein structure corresponds to an anti-parallel β -sheet with two α -helices positioned on the same side, with a 25 residues IDR at the C-terminus. The relaxation data measured at three different fields reveals a plateau of about 12 residues that resembles the one found in our relaxation data. In another case, Liu et al. report backbone amide ^{15}N relaxation data for human PDCD5 (28). The globular domain of 100 amino acids is mainly α -helical, while the C-terminus corresponds to an IDR of 30 residues that shows a clear plateau in all three NMR relaxation parameters.

The second region of interest consists of residues 124 to 130 of sTnI. The backbone dynamics of this region monitored by NMR relaxation show no evidence of mobility, with values of ^{15}N - T_1 , ^{15}N - T_2 and $\{^1\text{H}\}$ - ^{15}N -NOE much more comparable to sTnC than to the flexible C-terminal region of sTnI. For example, the average ^{15}N - T_2 at 800 MHz for residues 124-130 of sTnI₁₂₄₋₁₃₀ (0.049) is similar to the sTnC region (0.061) as opposed to the C-terminal region of sTnI (0.215). This indicates that residues 124-130 of sTnI are as rigid as the globular domain of sTnC. The MD simulation for model 4 also reflects this experimental result. This observation was surprising, since these residues were not shown to be rigid in the NMR solution structure of the sNTnC:sTnI₁₁₅₋₁₃₁ (29,30) as well as in the homologous cardiac structure (31). This region of sTnI

makes contact with α -helix A of sTnC in the X-ray structure of the sTn complex, but with electronic density observable only in one of the two conformers seen in the crystal (chain I) (13). Importantly, the proximity of H130 of sTnI with E20 of sTnC was proposed to be a pH-dependent interaction known as a "histidine button" responsible for the pH sensitivity difference between skeletal and cardiac muscle, and therefore sensitivity to lactic acidosis generated in ischemic conditions (32,33). Our experimental relaxation data for this region support this theory, and reveal that residues 124-130 are not flexible as previously interpreted.

All of the MD simulations show a gradient of correlations times for residues in the C-terminus but none of the simulations fit the details of the experimental NMR data. This can be seen in Fig. VI-7 where the calculated relaxation parameters are compared with the experimental data. The relaxation data calculated for the N-terminal globular sTnC domain fit the experimental NMR relaxation parameters well, including the details of flexibility of the linker between the two calcium-binding sites (34). However, the amplitude of the motions sampled in the MD simulations is never large enough to result in calculated data that reach the experimental data (see $^{15}\text{N-T}_2$ in particular), nor is the plateau reproduced. Even the flexible model 3 featuring an absence of nascent secondary structure was not able to reproduce the mobility of the C-terminal region of TnI as seen in the experimental $^{15}\text{N-T}_2$ NMR data. While this could suggest that none of the models were appropriate, it more likely suggests that the molecular dynamic simulations used in this study were unable to reproduce the larger conformational sampling required to model the intrinsically disordered region of TnI.

CONCLUSION

Overall, our results indicate that the use of the sTnC-sTnI chimera is a good strategy to probe the dynamics of this key regulatory interaction within the sTn complex. Our relaxation data showed that the N-domain of sTnC is well folded

and interacts with its sTnI binding partner in the chimera as expected. The NMR relaxation and chemical shift data of the C-terminal region of sTnI show the presence of a flexible and disordered region. While standard MD simulations were successful in characterizing the internal motions within a globular domain of our construct, we were unable to reproduce the conformational sampling required to describe the large-scale motions of the intrinsically disordered region in detail. On balance, our experimental and MD data argues against any stable secondary structure significantly populated in the IDR of sTnI. This conclusion supports the role of a ‘fly-casting’ mechanism for regulation of muscle contraction. The fact that a similar plateau in the relaxation data can be observed for other proteins suggests that this is an intrinsic property of extended mobile IDRs. We also observed an interaction of the region of sTnI immediately adjacent to the IDR with sTnC, which has important implications in the performance of skeletal muscle under ischemic conditions.

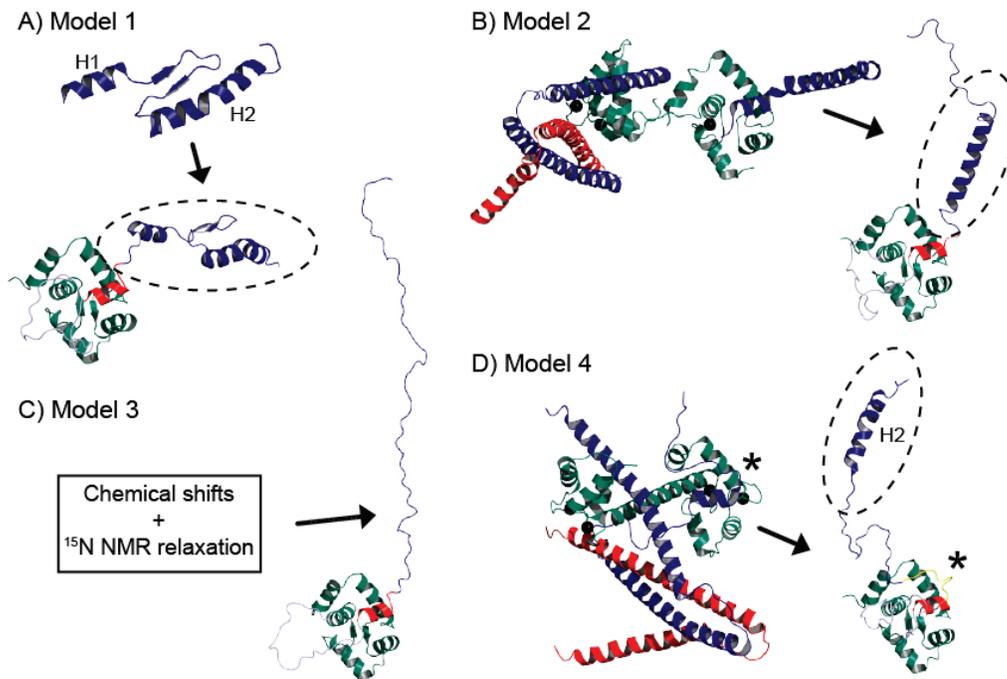
FIGURES


Fig. VI-1. Four plausible models for the skeletal TnC-TnI chimera based on different TnI conformations. A) Model 1, based on the reported NMR structure of the C-terminal region of TnI determined by Murakami et al. (11). B) Model 2, based on the X-ray structure of the cardiac troponin complex determined by Takeda et al. (10). C) Model 3, based on the NMR relaxation data obtained by Blumenschein et al. (6). D) Model 4, based on the X-ray structure of the skeletal troponin complex (13), featuring an interaction between residue H130 of sTnI and E20 of sTnC known as "histidine button" (32,33).

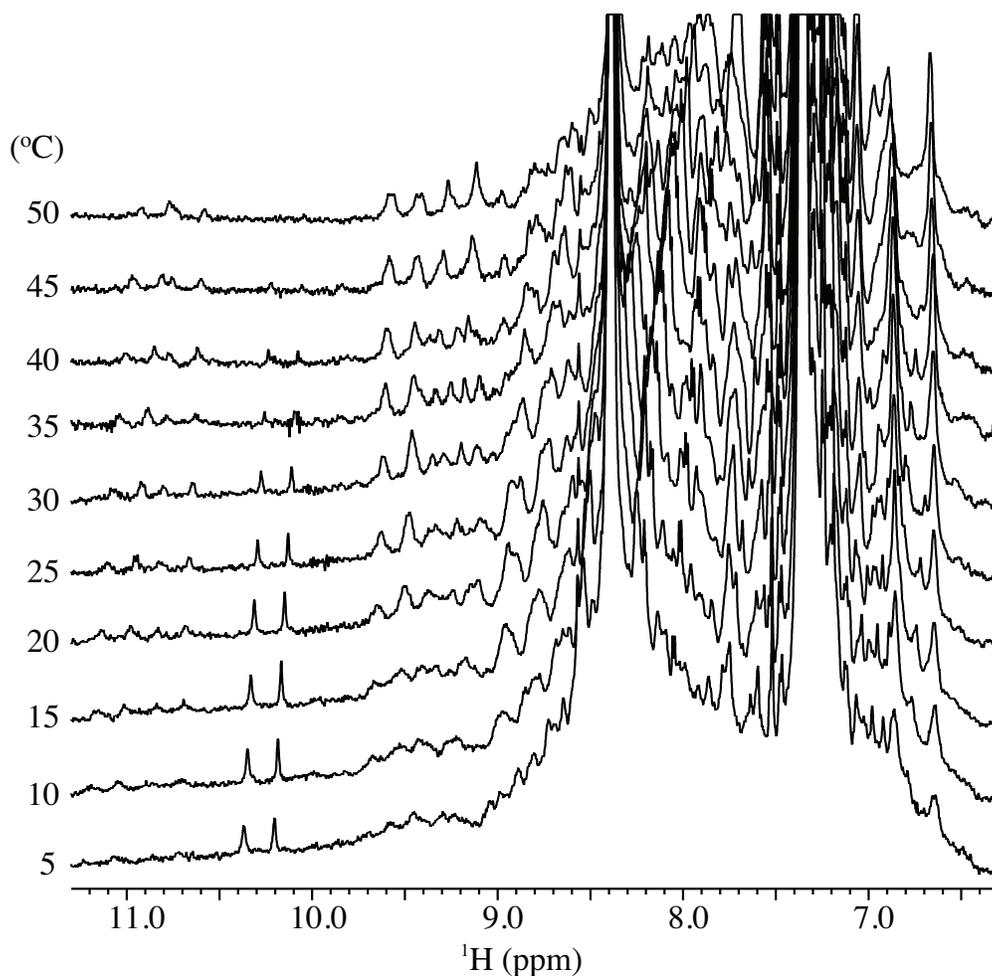


Fig. VI-2. 1D NMR spectra of the skeletal TnC-TnI chimera acquired at different temperatures. The stack plot shows the line width dependency of the aromatic and amide regions. The peaks get narrower around 30-35 °C. As the temperature increases, one can notice the disappearance of the Trp ^1H - ^{15}N indole signal (doublet at 10.3 ppm). The spectra were recorded on a 600 MHz NMR spectrometer.

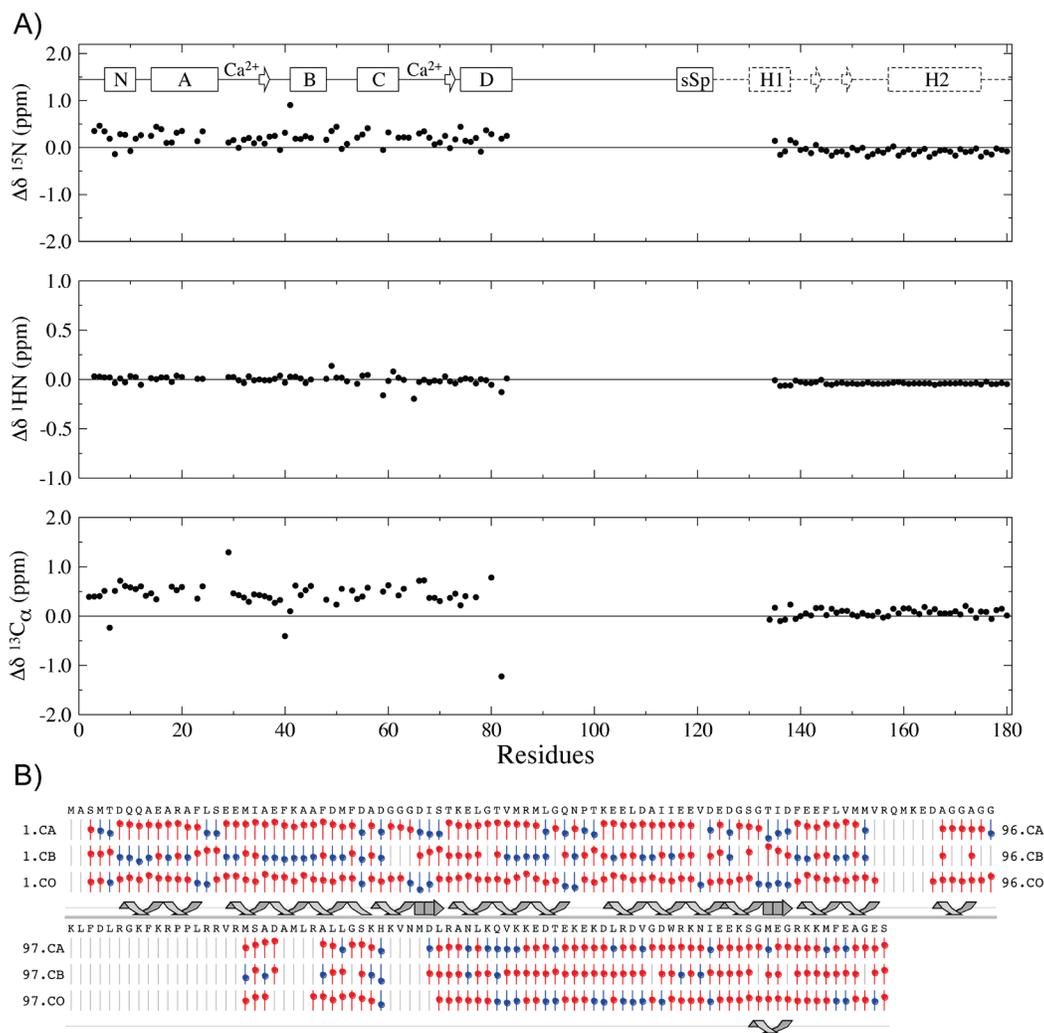


Fig. VI-4. Chemical shifts of the skeletal TnC-TnI chimera. A) Chemical shift difference between the chimera and sTnC and sTnI in the Tn complex. B) Chemical Shift Index (CSI) of the TnC-TnI chimera; the residues colored in red have a high likelihood of being located in an α -helix region, while residues colored in blue suggest a β -sheet region. The chemical shifts of the C_α , C_β and CO were used to calculate the secondary structure likelihood of each assigned residue, with the consensus of the three nuclei shown with a cartoon representation. The chemical shifts were obtained from 3D HNCACB, CBCACONH, HNCO and HNCACO experiments.

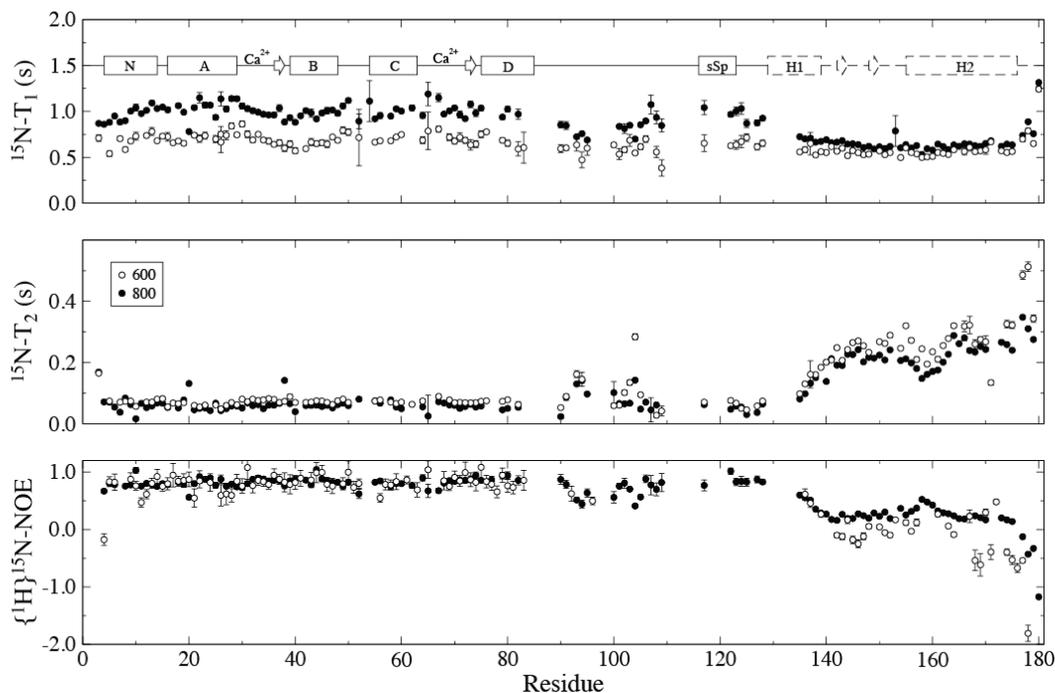


Fig. VI-5. Backbone amide NMR relaxation parameters $^{15}\text{N-T}_1$, $^{15}\text{N-T}_2$ and $\{^1\text{H}\}$ - $^{15}\text{N-NOE}$ of the skeletal TnC-TnI chimera. The data were measured at 600 MHz (white circles) and 800 MHz (black circles). The secondary structure elements of the chimera based on previous structures and chemical shift measurements presented in this paper are shown with solid lines, while the previously proposed secondary structures proposed for the mobile domain are shown with dashed lines. Relaxation parameters for ten residues visible in the ^1H - ^{15}N HSQC but not unambiguously assigned were arbitrary attributed to residues 100-109 for graphing purposes.

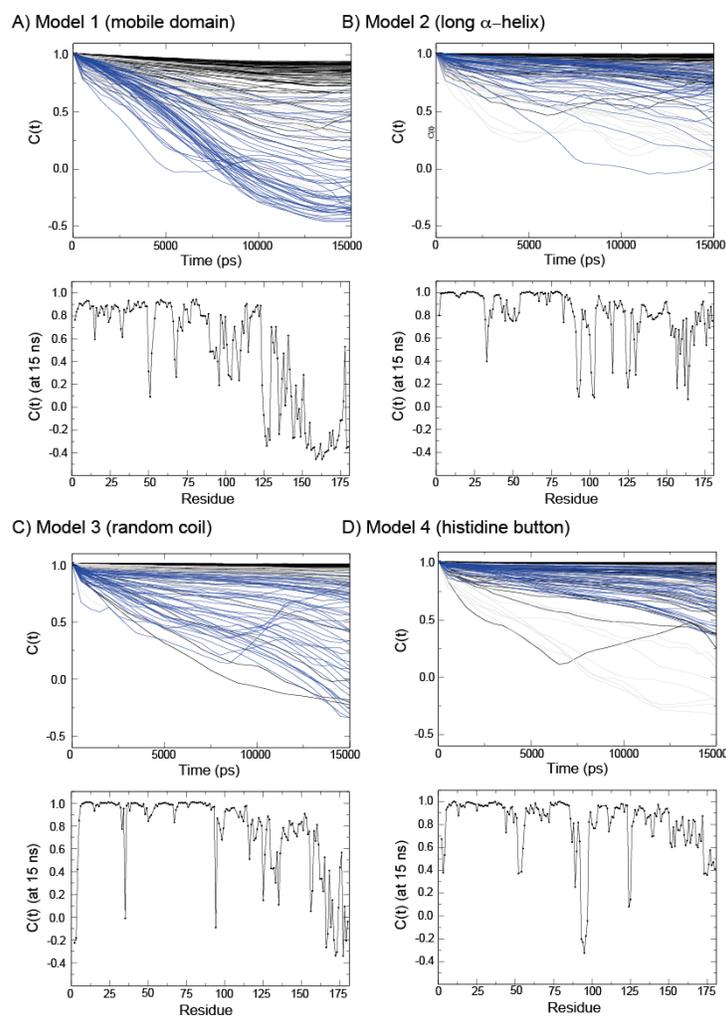


Fig. VI-6. Correlation functions extracted from molecular dynamic simulations. A) Model 1, based of the mobile domain structure B) Model 2, the helical model C) Model 3, the flexible model D) Model 4, the histidine button model. The figures at the top show the correlation functions as a function of correlation time, with residues 1-90 of sTnC colored in black, the linker and switch region of TnI colored in grey, and the C-terminal region of sTnI colored in blue. The graphs below show the amplitude of the correlation functions at 15 ns (estimated order parameter) as a function of residue number. The MD simulations were carried out using a Generalized Born with a simple switching (GBSW) implicit solvent model in CHARMM. The structures were first minimized, heated to 298K, equilibrated, and followed by a 20 ns (or more) simulation for the analysis.

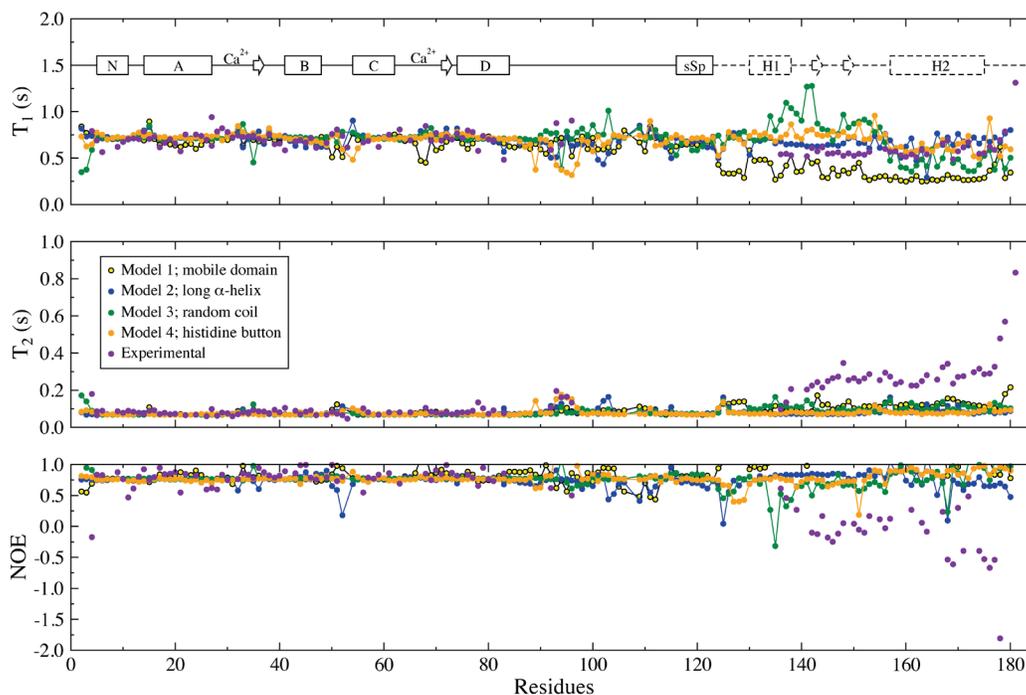


Fig. VI-7. Comparison of the experimental NMR relaxation parameters ^{15}N - T_1 , ^{15}N - T_2 , $\{^1\text{H}\}$ - ^{15}N -NOE with the one calculated from the MD trajectories. The models 1 to 4 are colored in yellow, blue, green and orange, respectively. Experimental data acquired at 600 MHz are shown in violet. An isotropic overall tumbling time of 10 ns for the chimera was assumed for the analysis.

REFERENCES

1. Sykes, B. (2003) Pulling the calcium trigger, *Nat. Struct. Biol.* 10, 588-589.
2. Gordon, A. M., Homsher, E., and Regnier, M. (2000) Regulation of contraction in striated muscle, *Physiol. Rev.* 80, 853-924.
3. Tobacman, L. (1996) Thin filament-mediated regulation of cardiac contraction, *Annu. Rev. Physiol.* 58, 447-481.
4. Farah, C., and Reinach, F. (1995) The troponin complex and regulation of muscle contraction, *FASEB J.* 9, 755-767.
5. Shoemaker, B., Portman, J., and Wolynes, P. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism, *Proc. Natl. Acad. Sci. U.S.A.* 97, 8868-8873.
6. Blumenschein, T., Stone, D., Fletterick, R., Mendelson, R., and Sykes, B. (2006) Dynamics of the C-terminal region of TnI in the troponin complex in solution, *Biophys. J.* 90, 2436-2444.
7. Hoffman, R., Blumenschein, T., and Sykes, B. (2006) An interplay between protein disorder and structure confers the Ca^{2+} regulation of striated muscle, *J. Mol. Biol.* 361, 625-633.
8. Gomes, A., and Potter, J. (2004) Cellular and molecular aspects of familial hypertrophic cardiomyopathy caused by mutations in the cardiac troponin I gene, *Mol Cell Biochem* 263, 99-114.
9. Lassalle, M. (2010) Defective Dynamic Properties of Human Cardiac Troponin Mutations, *Biosci. Biotechnol. Biochem.* 74, 82-91.
10. 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.
11. Murakami, K., Yumoto, F., Ohki, S., Yasunaga, T., Tanokura, M., and Wakabayashi, T. (2005) Structural basis for Ca^{2+} -regulated muscle relaxation at interaction sites of troponin with actin and tropomyosin, *J. Mol. Biol.* 352, 178-201.

12. King, W. A., Stone, D., Timmins, P., Narayanan, T., Von Brasch, A. A., Mendelson, R., and Curmi, P. (2005) Solution structure of the chicken skeletal muscle troponin complex via small-angle neutron and X-ray scattering, *J. Mol. Biol.* 345, 797-815.
13. Vinogradova, M., Stone, D., Malanina, G., Karatzaferi, C., Cooke, R., Mendelson, R., and Fletterick, R. (2005) Ca²⁺-regulated structural changes in troponin, *Proc. Natl. Acad. Sci. U.S.A.* 102, 5038-5043.
14. Cooke, J., Chamoun, J., Howell, M., Curmi, P., Fajer, P., and Brown, L. (2010) Structure and Dynamics of the Mobile Domain of Troponin I by SDSL-EPR, *Biophys. J.* 98, 148A.
15. Tiroli, A., Tasic, L., Oliveira, C., Bloch, C. J., Torriani, I., Farah, C., and Ramos, C. (2005) Mapping contacts between regulatory domains of skeletal muscle TnC and TnI by analyses of single-chain chimeras, *FEBS J.* 272, 779-790.
16. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J. Biomol. NMR* 41, 5-7.
17. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
18. Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., and Karplus, M. (1983) CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comput. Chem.* 4, 187-217.
19. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints, *J. Mol. Biol.* 234, 779-815.
20. Im, W., Beglov, D., and Roux, B. (1998) Continuum Solvation Model: computation of electrostatic forces from numerical solutions to the Poisson-Boltzmann equation, *Comput. Phys. Commun.* 111, 59-75.
21. Mackerell, A., Bashford, D., Bellott, M., Dunbrack, R., Evanseck, J., Field, M., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-Mccarthy, D., Kuchnir, L., Kuczera, K., Lau, F., Mattos, C., Michnick, S., Ngo, T., Nguyen, D., Prodhom, B., Reiher, W., Roux, B., Schlenkrich, M., Smith, J., Stote, R.,

- Straub, J., Watanabe, M., Wiorcikiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* 102, 3586-3616.
22. Mackerell, A., Feig, M., and Brooks, C. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, *J. Comput. Chem.* 25, 1400-1415.
23. Im, W. P., Lee, M., and Brooks, C. (2003) Generalized born model with a simple smoothing function, *J. Comput. Chem.* 24, 1691-1702.
24. Abragam, A. (1961) *The Principles of Nuclear Magnetism*, pp 289-305, Clarendon Press, Oxford.
25. Blumenschein, T., Stone, D., Fletterick, R., Mendelson, R., and Sykes, B. (2005) Calcium-dependent changes in the flexibility of the regulatory domain of troponin C in the troponin complex, *J. Biol. Chem.* 280, 21924-21932.
26. Wishart, D., and Sykes, B. (1994) The ¹³C chemical-shift index: a simple method for the identification of protein secondary structure using ¹³C chemical-shift data, *J. Biomol. NMR* 4, 171-180.
27. Zhukov, I., Bayer, P., Scholermann, B., and Ejchart, A. (2007) ¹⁵N magnetic relaxation study of backbone dynamics of the ribosome-associated cold shock response protein Yfia of Escherichia coli, *Acta Biochim. Pol.* 54, 769-775.
28. Liu, D., Feng, Y., Cheng, Y., and Wang, J. (2004) Human programmed cell death 5 protein has a helical-core and two dissociated structural regions, *Biochem. Biophys. Res. Commun.* 318, 391-396.
29. McKay, R., Tripet, B., Pearlstone, J., Smillie, L., and Sykes, B. (1999) Defining the region of troponin-I that binds to troponin-C, *Biochemistry* 38, 5478-5489.
30. Mercier, P., Ferguson, R. E., Irving, M., Corrie, J. E. T., Trentham, D. R., and Sykes, B. D. (2003) NMR structure of a bifunctional rhodamine labeled N-domain of troponin C complexed with the regulatory "switch" peptide

- from troponin I: implications for in situ fluorescence studies in muscle fibers, *Biochemistry* 42, 4333-4348.
31. 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, 8289-8298.
 32. Dargis, R., Pearlstone, J., Barrette-Ng, I., Edwards, H., and Smillie, L. (2002) Single mutation (A162H) in human cardiac troponin I corrects acid pH sensitivity of Ca²⁺-regulated actomyosin S1 ATPase, *J. Biol. Chem.* 277, 34662-34665.
 33. Day, S., Westfall, M., Fomicheva, E., Hoyer, K., Yasuda, S., La Cross, N. C., D'Alecy, L. G., Ingwall, J., and Metzger, J. (2006) Histidine button engineered into cardiac troponin I protects the ischemic and failing heart, *Nat. Med.* 12, 181-189.
 34. Gagné, S., Tsuda, S., Spyropoulos, L., Kay, L., and Sykes, B. (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, 667-686.

Chapter VII – bPrP^C*Monitoring prion protein stability by NMR*Olivier Julien, Steffen P. Graether, and Brian D. Sykes¹¹ *Department of Biochemistry, University of Alberta, Edmonton, AB, Canada*Publication: Julien et al. 2009. *J. Toxicol. Environ. Health A*. 72:1069-1074.

Contributions: SG designed the clone, AT expressed and purified the protein, OJ prepared and handled the NMR samples, performed the NMR experiments, acquired, processed and analysed the NMR data. OJ and BDS wrote the paper. OJ made all of the figures.

OVERVIEW

Prion diseases, or transmissible spongiform encephalopathies (TSEs), are a group of fatal neurological diseases that affect both humans and animals. At the end of the twentieth century, bovine spongiform encephalopathy (BSE), better known as mad cow disease, was shown to be transmissible to humans. This resulted in considerable concern for public health and a number of questions for scientists. The first question answered was the possible source of the disease, which appears to be the prion protein (PrP). There are two major forms of this protein: the native, non-infectious form (PrP^C), and the misfolded infectious form (PrP^{Sc}). PrP^C is mainly α -helical in structure, whereas PrP^{Sc} aggregates into an assembly of β -sheets, forming amyloid fibrils. Since the first solution structure of the non-infectious form of the mouse prion protein, about 30 structures of the globular portion of PrP^C have been characterized from different organisms. However, only a few minor differences are observed when comparing one PrP^C structure to another. The key to understanding prion formation may then not be in the

structure of PrP^C, but in the mechanism by which the PrP^C is unfolded and then converted into a misfolded fibril state. To identify the possible region(s) of PrP^C responsible for initiating the conversion into the amyloid fibril formation, we are applying nuclear magnetic resonance (NMR) to characterize the stability and structure of PrP^C and intermediate states during the conversion from PrP^C to PrP^{Sc}. We have used urea to induce unfolding, and the analysis of the data revealed region specific structural stabilities that may bring insights into the mechanism of the conversion of the protein to an infectious prion.

INTRODUCTION

Prion diseases, or transmissible spongiform encephalopathies (TSEs), are a group of fatal neurological disorders that affect humans and animals (1). It is characterized by spongiform degeneration. These conditions are commonly known as Creutzfeldt-Jacob disease (CJD) in humans, mad cow disease or bovine spongiform encephalopathies (BSE) in cattle, chronic wasting disease (CWD) in cervids, and scrapie in sheep. These diseases can be sporadic, inherited, or acquired by infectious means. Prion diseases are associated with an accumulation of abnormally folded endogenous protein in the brain, analogous to the formation of plaques by A β peptide aggregation in Alzheimer's disease. This proteinaceous aggregation is thought to be neurotoxic and leads to neurodegeneration. However, the most striking feature is what sets prion and neurodegenerative diseases apart; unlike Alzheimer's, Huntington's or Parkinson's diseases, prion diseases are transmissible.

Bovine spongiform encephalopathy (BSE) appeared in the United Kingdom (UK) in 1987 and rapidly progressed into an epidemic. According to the World Health Organization, 181,376 cases of BSE were confirmed in the UK between November 1986 and November 2002 (2). This epidemic was most likely responsible for a novel human prion disease that appeared in 1995 in the UK, variant CJD (vCJD) (3). vCJD was shown to be caused by a closely similar

protein strain to that seen in BSE (4). In Canada, the first case of BSE from a Canadian-born cow was made public on May 20, 2003. The embargo for Canadian beef jeopardized the beef industry and prompted the Canadian government to take action. Surveillance organizations were formed in order to monitor and contain the spread of the disease.

The generally accepted protein-only hypothesis suggests that the infectious agent causing prion diseases is composed for the most part, if not entirely, of endogenous protein (5,6). There are two major forms of prion protein: the native and non-infectious form (PrP^C), which corresponds mainly to an α -helical structure, and the misfolded infectious form (PrP^{Sc}), which aggregates into an assembly of β -sheets forming an amyloid fibril (7). The solution NMR structure of a domain of the non-infectious form of the mouse prion protein (121-231) was first determined in 1996 (8). Since then, close to 30 structures of the globular portion of PrP^C have been characterized for a variety of organisms, mainly mammals. Surprisingly, only a few minor differences are observed between all the determined structures (Fig. VII-1), which include non-mammalian organisms (e.g. chicken) that share only ~30% sequence identity to the mammalian sequence and are unable to develop TSEs. Thus, the key to understanding prion formation may not be found in the structure of PrP^C, but in the conversion mechanism and the conditions favoring this structural transition.

The prion protein is encoded by the *PRNP* gene located on the short arm of human chromosome 20. PrP^C is an extracellular membrane glycoprotein attached by a glycosyl phosphatidyl inositol (GPI) anchor (9). The native function of PrP^C still remains unclear; however, many functions have been suggested over the years including neuro-protection and metal chelation (see (10) and (11) for reviews). According to the NMR data, the C-terminal region of PrP^C is structured, and contains three α -helices and a two-stranded anti-parallel β -sheet (12). A disulfide bridge also links helix 2 and helix 3. The N-terminal part of the protein (residues 21-121), on the other hand, is unstructured. This section of the protein

contains 4 or 5 octapeptide repeats, which has impeded chemical shift assignment of the N-terminal region of PrP^C. Consequently, only modest information is available for this region. Compared to PrP^C, little is known about PrP^{Sc}. Some of the known characteristics of PrP^{Sc} are partial resistance to PK digestion, insolubility, and formation of aggregates or amyloid-like fibrils (13). The high-resolution structure of PrP^{Sc} is unknown, but several models have been proposed over the years. Based on a molecular dynamics simulation, it was proposed that the prion protein conversion is triggered by low pH and resulted in the growth of the already present anti-parallel β -sheet of PrP^C (14). Using hydrogen/deuterium exchange, site-directed spin labeling coupled with electron paramagnetic resonance spectroscopy and mass spectroscopy analysis, Cobb et al. suggested a parallel in-register β -structure of the amyloid fibrils (15). Finally, based on electron microscopy data from 2D crystals, Govaerts et al. proposed that PrP^{Sc} has a left-handed β -helix trimeric structure (16). The first high-resolution structure of infectious material published for the yeast prion HET-s(218-289) revealed a β -solenoid with a triangular hydrophobic core, a structure similar to the model proposed by Govaerts et al. (17).

The first mutation found in the human *PRNP* gene has been linked to the Gerstmann-Sträussler-Scheinker disease (GSS), a rare autosomal dominant familial neurodegenerative illness. The mutation was identified as a substitution of a proline with a leucine at residue 102 (P102L) (18). Currently, over 20 mutations in *PRNP* have been shown to be associated with familial diseases (see (13) and (1) for reviews). Of particular interest is the mutation D178N identified in 1992 (19). The mutation was first associated with familial fatal insomnia (FFI) disease and shown to be transmissible to animals (19,20). Symptoms include loss of sleep and motor abnormalities, sometimes involving a vegetative state, which can eventually lead to death. The presence of an important polymorphism at codon 129 (M or V, see below) has been shown to regulate the phenotype of the D178N associated-disease mutation: M129/D178N leads to FFI and V129/D178N leads to familial CJD (21). It has been suggested that the D178N mutation could

destabilize the electrostatic side chain interactions with R164 (22), which is located in β -strand 2. This could potentially destabilize the native β -sheet of PrP^C, leading to the higher tendency of the mutant to form β -sheet-rich oligomers *in vitro* (23).

The goal of this research project is to study the structural changes that occur during the molecular conversion of PrP^C into PrP^{Sc}. It was previously shown that the conversion of recombinant prion protein into a β -sheet-rich form with similar physiochemical properties to PrP^{Sc} can be induced by acidic pH, salt and urea (24). Furthermore, it was shown that mouse recombinant PrP expressed in *E. coli*, then polymerized into amyloid fibrils, and inoculated intracerebrally into transgenic mice, caused the mice to develop neurologic disease 380 to 660 days after inoculation (25). In order to identify the possible regions favoring the amyloid fibril formation, we are using ¹H NMR spectroscopic methods to characterize the stability of PrP^C and possibly identify the presence of an intermediate state(s) in the conversion of PrP^C to PrP^{Sc}. We have collected residue-specific structural stabilities of PrP^C to urea denaturation. The identification and characterization of the key elements favoring prion protein transformation is an important step for obtaining a better understanding of protein misfolding diseases.

MATERIALS AND METHODS

The C-terminus of bovine PrP residues 121 to 230 (bPrP₍₁₂₁₋₂₃₀₎) was expressed and purified in *E. coli* as previously described for human PrP (26), with the following modifications. The bPrP(121-230) gene was synthetically synthesized with codon usage optimized for *E. coli* expression. NdeI and BamHI cut sites were included at the 5' and 3' ends of the sequence to facilitate ligation into pRSETa plasmid previously digested with the same restriction enzymes. Ligations were transformed into XL-1 blue cells, and colonies were screened for insertions.

The creation of pRSETa-bPrP(121-230) was confirmed by sequencing in the forward and reverse directions.

All NMR experiments were performed on 600 and 800 MHz Varian INOVA NMR spectrometers, the latter of which was equipped with a cryogenic probe. The NMR sample contained 2 mg of bPrP₍₁₂₁₋₂₃₀₎ dissolved in 500 μ l of deuterated NMR buffer, which contained 0.5 mM of piperazine for pH measurement (27). The pH of the sample was adjusted to 6.3. Deuterated urea (Sigma-Aldrich Co.) was used to unfold the protein, by adding 32.0 mg of urea at every step resulting in a 13 step titration from 0 M to 10 M (after compensation for the increase of the volume of the sample). On the 600 MHz spectrometer, the one-dimensional (1D) ¹H NMR spectra were acquired with 256 transients and a spectral width of 14.0 ppm, while the two-dimensional (2D) ¹H-¹H NOESY spectra were acquired with 32 transients, 192 increments and spectral widths of 14.0 ppm for both proton dimensions. On the 800 MHz spectrometer, the NOESY spectra were acquired with 32 transients, 256 increments and a spectral width of 13.0 ppm for both proton dimensions.

The 1D ¹H NMR spectra presented in Fig. VII-2 were analyzed using the Vnmrj 2.1B software. The spectra were processed using a line broadening of 1.5 Hz and the frequency of the methyl groups of DSS was set to 0 ppm as an internal standard. The vertical scale of each spectrum was adjusted to account for the dilution caused by the addition of urea. The integral of the resonance at each step in the titration was measured using the same spectral width (e.g. 0.10 ppm) centered on the peak. The peak areas were then normalized to the integral measured at 0 M. The stability of the protein to urea denaturation was assumed to be of the form $DG = DG^{H_2O} + m[\text{urea}]$ (Greene and Pace 1974). Accordingly the data was fit using the xcrvfit software to the following sigmoid function:

$$Peak\ area = \frac{a}{1 + e^{-\frac{m}{RT}([\text{urea}] - [\text{urea}]_{1/2})}}$$

where a is a scaling constant, R is the gas constant, T is the temperature, m/RT is the slope, and $[\text{urea}]_{1/2}$ ($= -DG^{H_2O} / m$) is the urea concentration when the peak area is half its initial value.

RESULTS

A molecular weight of $13,093 \pm 5$ Da was measured using mass spectroscopy (compared to 13,097 Da for the predicted molecular weight). The purity of the protein was verified using 1D ^1H NMR spectrum. The spectrum revealed that the protein was folded, given the presence of resonances coming from the small β -sheet of PrP^C around 5.5 ppm and good chemical shift dispersion of the aromatic region (see Fig. VII-2, 0 M urea). We also acquired a 2D ^1H - ^1H NOESY NMR spectrum at 800 MHz in D₂O (Fig. VII-3). The advantage of using D₂O as a solvent is that analysis is simplified since labile protons will exchange with the deuterated solvent and will not be observed in the proton NMR spectra. The comparison of the cross-peak chemical shifts from the NOESY spectrum with the chemical shifts previously deposited in the Biological Magnetic Resonance Data Bank (BMRB entry 4653) confirmed that the structure of the C-terminal domain of bPrP was very similar to the previously determined structure (28).

We performed a series of denaturation experiments to evaluate the stability of the prion protein. Increasing concentrations of deuterated urea were added to the sample containing properly folded PrP^C. For each step, we acquired two NMR spectra: a 1D and a 2D ^1H - ^1H NOESY. The comparison of the aromatic region of the 1D NMR spectra obtained during the urea titration is presented in Fig. VII-2. The use of D₂O and deuterated urea during the titration allow us to monitor the chemical changes of the aromatic and aliphatic protons (in H₂O, the aromatic region is dominated by the presence of the amide NH resonances that resonate at similar frequencies). Most of the resonances found between 6 and 8 ppm correspond to the frequencies of the aromatic protons (H δ and H ϵ) of the eleven tyrosines present in bPrP₍₁₂₁₋₂₃₀₎. Overall, the urea titration shows the high level of stability of the C-terminal region of bPrP compared to most other proteins.

The two resonances observed at 5.59 and 5.34 ppm were first thought to correspond to the two alpha protons (H α) of two residues within the small β -sheet. However, according to the chemical shifts in the BMRB for bPrP, these resonances arise from Y162 (5.548 ppm) and F198 (5.309 ppm). Another well-isolated resonance around 6.15 ppm corresponds to the H δ protons of Y218. These protons are excellent probes to monitor the prion protein local structural stability (i.e. region specific information) since these three residues are located on opposite sides of the structure; Y162 is situated in the β -strand, F178 is located in the loop between helix 1 and helix 2, and Y218 is twelve residues away from the C-terminal end of helix 3. One can see in Fig. VII-2 that the resonance associated with Y162-H α disappears around 7.1 M urea, while the resonances of F198-H α and Y218-H δ are still clearly observable. This suggests that the local stability of the β -sheet is perturbed before helix 3 and the loop between helix 1 and 2. To quantify this observation, the peak area (normalized integral) of each resonance is presented at the bottom of Fig. VII-2 as a function of the urea concentration. We obtained a lower [urea]_{1/2} value of 4.02 M for Y162-H α , compared to 6.50 M and 6.98 M for Y218-H δ and F198-H α , respectively. We obtained a higher *m*-value for F198-H α (1.13 kcal mol⁻¹ M⁻¹) compared to Y162-H α and Y218-H δ (0.56 kcal mol⁻¹ M⁻¹ and 0.54 kcal mol⁻¹ M⁻¹, respectively).

For each step of the urea titration, a 2D ¹H-¹H NOESY NMR spectrum was also acquired to monitor the structural changes that happen during denaturation. The spectrum at 0 M urea is presented in FIG 3. The two arrows labeled “A” and “B” indicates the ¹H chemical shift of the H α of Y162 and F178 on the x-axis. Most of the cross-peaks found for these two protons correspond to intra-residue NOE contacts. However, of special interest is the presence of a long-range NOE between the H α of Y162 and L130 (marked with an asterisk in the spectrum). This NOE is a good probe to monitor the presence of the β -sheet, since these two protons can be close in space only if the β -sheet is properly formed. Our

preliminary analysis suggests that the first secondary structure to be affected during the unfolding process is the β -sheet, reflected by the disappearance of the NOE between Y162 and L130 early in the titration (data not shown, since the NOE disappears with the addition of urea). To confirm these observations, the analysis of the twelve NOESY spectra acquired during the urea titration is underway, which will provide structural information of the PrP conformational changes that happens during the protein unfolding process.

We have also investigated the effect of temperature on the stability of bPrP₍₁₂₁₋₂₃₀₎ using NMR. 2D ¹H-¹H NOESY spectra were acquired at 20, 25, 30, 35 and 40°C to look at the early stages of the unfolding process. The comparison of the spectra only minor changes (data not shown), which was not unexpected since the denaturation of 50% of PrP^C occurs at temperatures higher than 60°C (29).

DISCUSSION

The ¹H NMR data taken during the urea titration have revealed the differential stabilities of three residues, reflecting disparate regions of the PrP^C protein. These have been analyzed according to the accepted linear dependence of the Gibbs' free energy of unfolding as a function of urea concentration. We have identified and analyzed three resonances based on their unambiguous chemical shift assignment and their presence in an isolated part of the spectrum to avoid overlap with other resonances. We have measured the peak area of each resonance as a function of the urea concentration. Our analysis revealed that Y162-H α has a lower [urea]_{1/2} value compared to Y218-H δ and F198-H α . This could be explained by a β -sheet dissociation at an early stage in the urea titration compared to helix 3 and the loop between helix 1 and 2, regions monitored by Y218-H δ and F198-H α , respectively. This would corroborate previous studies suggesting that the β -sheet dissociation of PrP^C is an early and required step in the prion protein structural conversion (Paramithiotis et al. 2003). Interestingly, the shape and [urea]_{1/2} value of the denaturation curve for F198-H α is characteristic of an

intrinsically stable element of structure, but with a strong dependency on the concentration of urea as demonstrated by its higher m -value. For the three residues analyzed herein and others, intermediate or partially unfolded species were not directly detected. This implies that intermediates are in fast exchange on the NMR time scale, possibly due to the pH used to perform the experiments (6.3).

To obtain more structural information about intermediate states forming during the PrP conversion process, we are repeating the urea titration presented using different sample conditions (e.g. pH = 3.5-4.0) in order to induce the formation of an intermediate state at low pH, which was previously shown by other groups (30,31). We are also investigating the outcome of the mutation D178N on the PrP local stability in comparison to wild type protein, and the effect of the polymorphism found at codon 129 in parallel with the mutation D178N to understand the molecular basis of familial CJD and fatal familial insomnia disorders.

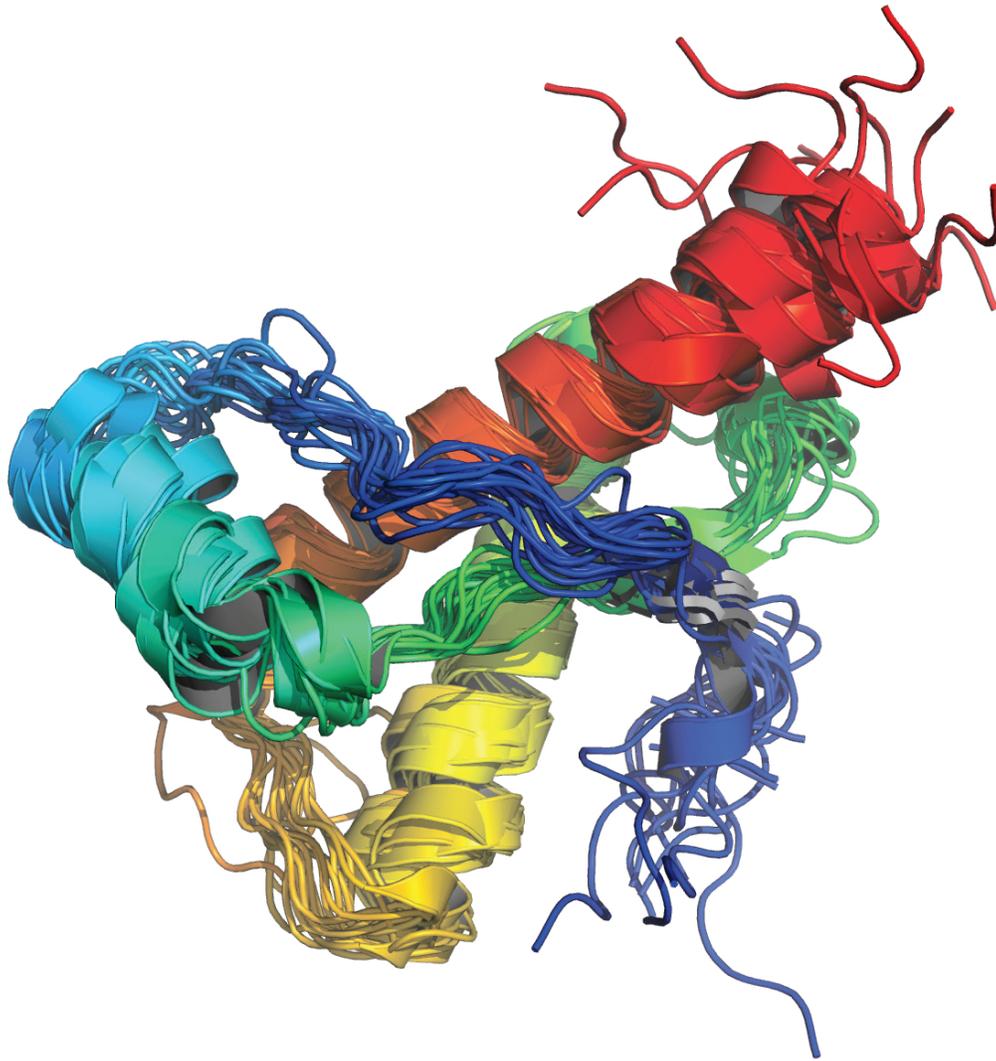
FIGURES

Fig. VII-1. Superimposition of 28 three-dimensional structures of the C-terminus of PrP^C from different animals, at different pH's, containing various mutations and bound to different antibodies (four mouse, elk, two bovine, eight human, six sheep, hamster, chicken, turtle, frog, cat, canine and pig structures). The following Protein Data Bank depositions were used: 1AG2, 1XYX, 1Y15, 1Y16, 1XYW, 1DWZ, 1DX1, 1E1G, 1E1P, 1FO7, 1H0L, 1HJN, 1QLZ, 1QM1, 1QM3, 1TPX, 1TQB, 1TQC, 1Y2S, 1UW3, 1XYU, 1B10, 1U3M, 1U5L, 1XU0, 1XYJ, 1XYK, 1XYQ. This figure was made with PyMol (DeLano Scientific LLC).

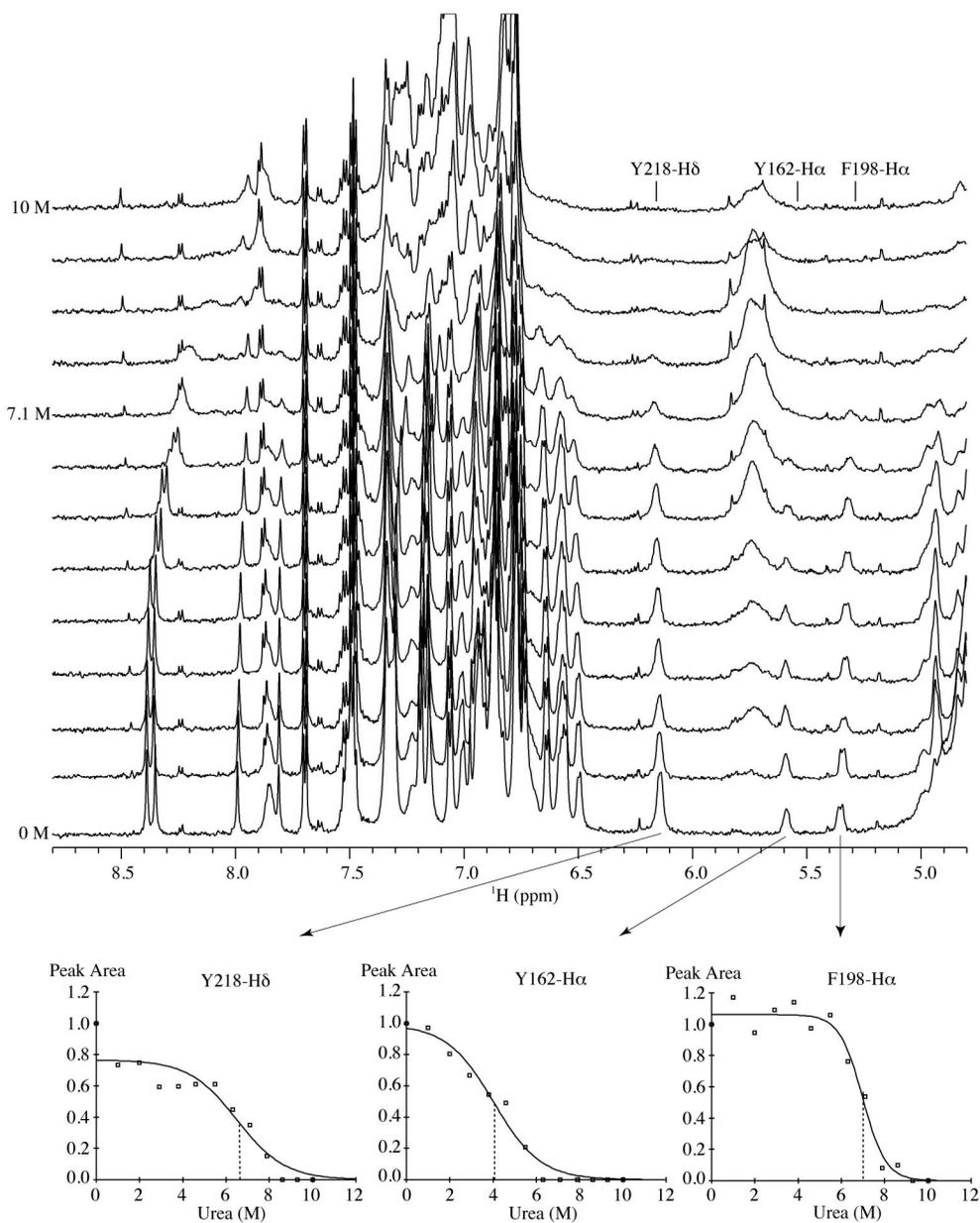


Fig. VII-2. Aromatic region of 1D ¹H NMR spectra of bPrP₍₁₂₁₋₂₃₀₎ in D₂O acquired at 600 MHz as a function of [urea] monitoring the differences from the properly folded (0 M urea) to the unfolded form of bPrP₍₁₂₁₋₂₃₀₎ (10 M urea). The resonance at 7.75 ppm comes from the urea which is partially suppressed. The peak area of three distinct resonances (Y218-Hδ, Y162-Hα and F198-Hα) is plotted as a function of the urea concentration at the bottom of the figure.

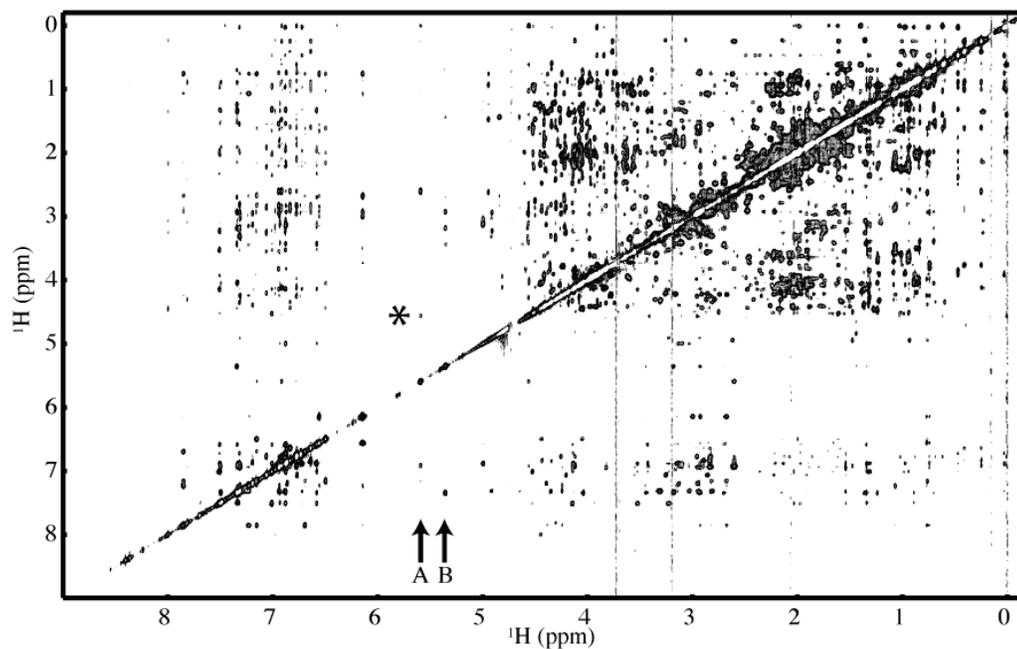


Fig. VII-3. 2D ^1H - ^1H NOESY NMR spectrum of bPrP₍₁₂₁₋₂₃₁₎ acquired at 800 MHz. The arrows labeled with an “A” and a “B” corresponds to the chemical shifts (x-axis) of the H α of residues Y162 (5.59 ppm) and F198 (5.30 ppm), respectively. The cross-peak identified with an asterisk originates from an NOE between Y162-H α and L130-H α , both located in the β -sheet of PrP^C.

REFERENCES

1. Collinge, J. (2001) Prion diseases of humans and animals: their causes and molecular basis, *Annu. Rev. Neurosci.* 24, 519-550.
2. World Health Organization,
<http://www.who.int/mediacentre/factsheets/fs113/en/>
3. Will, R. G., Ironside, J. W., Zeidler, M., Cousens, S. N., Estibeiro, K., Alperovitch, A., Poser, S., Pocchiari, M., Hofman, A., and Smith, P. G. (1996) A new variant of Creutzfeldt-Jakob disease in the UK, *Lancet* 347, 921-925.
4. Collinge, J., Sidle, K. C., Meads, J., Ironside, J., and Hill, A. F. (1996) Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD, *Nature* 383, 685-690.
5. Griffith, J. S. (1967) Self-replication and scrapie, *Nature* 215, 1043-1044.
6. Prusiner, S. B. (1982) Novel proteinaceous infectious particles cause scrapie, *Science* 216, 136-144.
7. Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., and Prusiner, S. B. (1993) Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, *Proc. Natl. Acad. Sci. USA* 90, 10962-10966.
8. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., and Wuthrich, K. (1996) NMR structure of the mouse prion protein domain PrP(121-321), *Nature* 382, 180-182.
9. Stahl, N., Borchelt, D. R., Hsiao, K., and Prusiner, S. B. (1987) Scrapie prion protein contains a phosphatidylinositol glycolipid, *Cell* 51, 229-240.
10. Martins, V. R., Linden, R., Prado, M. A., Walz, R., Sakamoto, A. C., Izquierdo, I., and Brentani, R. R. (2002) Cellular prion protein: on the road for functions, *FEBS Lett.* 512, 25-28.
11. Roucou, X., and Leblanc, A. C. (2005) Cellular prion protein neuroprotective function: implications in prion diseases, *J. Mol. Med.* 83, 3-11.

12. Zahn, R., Liu, A., Luhrs, T., Riek, R., Von Schroetter, C., Lopez Garcia, F., Billeter, M., Calzolari, L., Wider, G., and Wuthrich, K. (2000) NMR solution structure of the human prion protein, *Proc. Natl. Acad. Sci. USA* 97, 145-150.
13. Prusiner, S. B. (1998) Prions, *Proc. Natl. Acad. Sci. USA* 95, 13363-13383.
14. Demarco, M. L., and Daggett, V. (2007) Molecular mechanism for low pH triggered misfolding of the human prion protein, *Biochemistry* 46, 3045-3054.
15. Cobb, N. J., Sonnichsen, F. D., Mchaourab, H., and Surewicz, W. K. (2007) Molecular architecture of human prion protein amyloid: a parallel, in-register beta-structure, *Proc. Natl. Acad. Sci. USA* 104, 18946-18951.
16. Govaerts, C., Wille, H., Prusiner, S. B., and Cohen, F. E. (2004) Evidence for assembly of prions with left-handed beta-helices into trimers, *Proc. Natl. Acad. Sci. USA* 101, 8342-8347.
17. Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core, *Science* 319, 1523-1526.
18. Hsiao, K., Baker, H. F., Crow, T. J., Poulter, M., Owen, F., Terwilliger, J. D., Westaway, D., Ott, J., and Prusiner, S. B. (1989) Linkage of a prion protein missense variant to Gerstmann-Straussler syndrome, *Nature* 338, 342-345.
19. Medori, R., Tritschler, H. J., Leblanc, A., Villare, F., Manetto, V., Chen, H. Y., Xue, R., Leal, S., Montagna, P., Cortelli, P., Tinuper, P., Avoni, P., Mochi, M., Baruzzi, A., Hauw, J., Lugaresi, E., Autilio-Gambetti, L., and Gambetti, P. (1992) Fatal familial insomnia, a prion disease with a mutation at codon 178 of the prion protein gene, *New Engl. J. Med.* 326, 444-449.
20. Tateishi, J., Brown, P., Kitamoto, T., Hoque, Z. M., Roos, R., Wollman, R., Cervenakova, L., and Gajdusek, D. C. (1995) First experimental transmission of fatal familial insomnia, *Nature* 376, 434-435.
21. Goldfarb, L. G., Haltia, M., Brown, P., Nieto, A., Kovanen, J., McCombie, W. R., Trapp, S., and Gajdusek, D. C. (1991) New mutation in scrapie

- amyloid precursor gene (at codon 178) in Finnish Creutzfeldt-Jakob kindred, *Lancet* 337, 425.
22. Zuegg, J., and Gready, J. E. (1999) Molecular dynamics simulations of human prion protein: importance of correct treatment of electrostatic interactions, *Biochemistry* 38, 13862-13876.
 23. Apetri, A. C., Vanik, D. L., and Surewicz, W. K. (2005) Polymorphism at residue 129 modulates the conformational conversion of the D178N variant of human prion protein 90-231, *Biochemistry* 44, 15880-15888.
 24. Swietnicki, W., Morillas, M., Chen, S. G., Gambetti, P., and Surewicz, W. K. (2000) Aggregation and fibrillization of the recombinant human prion protein huPrP90-231, *Biochemistry* 39, 424-431.
 25. Legname, G., Baskakov, I. V., Nguyen, H. O., Riesner, D., Cohen, F. E., Dearmond, S. J., and Prusiner, S. B. (2004) Synthetic mammalian prions, *Science* 305, 673-676.
 26. Zahn, R., Von Schroetter, C., and Wuthrich, K. (1997) Human prion proteins expressed in *Escherichia coli* and purified by high-affinity column refolding, *FEBS Lett.* 417, 400-404.
 27. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J Biomol NMR* 41, 5-7.
 28. Lopez Garcia, F., Zahn, R., Riek, R., and Wuthrich, K. (2000) NMR structure of the bovine prion protein, *Proc. Natl. Acad. Sci. USA* 97, 8334-8339.
 29. Swietnicki, W., Petersen, R. B., Gambetti, P., and Surewicz, W. K. (1998) Familial mutations and the thermodynamic stability of the recombinant human prion protein, *J. Biol. Chem.* 273, 31048-31052.
 30. Apetri, A. C., and Surewicz, W. K. (2002) Kinetic intermediate in the folding of human prion protein, *J. Biol. Chem.* 277, 44589-44592.
 31. Gerber, R., Tahiri-Alaoui, A., Hore, P. J., and James, W. (2008) Conformational pH dependence of intermediate states during oligomerization of the human prion protein, *Protein Sci.* 17, 537-544.

CHAPTER VIII – bPrP^C*Differential stability of the bovine prion protein upon urea unfolding*

Olivier Julien,¹ Subhrangsu Chatterjee,¹ Angela Thiessen,¹ Steffen P. Graether,¹
and Brian D. Sykes¹

¹ *Department of Biochemistry, University of Alberta, Edmonton, AB, Canada*

² *Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON, Canada*

Publication: Julien et al. 2009. *Protein Science*, 18:2172-82.

Contributions: SG designed the clone, AT expressed the protein, OJ prepared and handled the NMR samples, performed the experiments, acquired and processed the NMR data, analysed the NMR data (with SC) and BDS performed the NOE buildup curves. OJ and BDS wrote the paper. OJ made all of the figures.

OVERVIEW

Prion diseases, or transmissible spongiform encephalopathies, are a group of infectious neurological diseases associated with the structural conversion of an endogenous protein (PrP) in the central nervous system. There are two major forms of this protein: the native and non-infectious cellular form, PrP^C; and the misfolded, infectious and proteinase K-resistant form, PrP^{Sc}. The C-terminal domain of PrP^C is mainly α -helical in structure, whereas PrP^{Sc} is known to aggregate into an assembly of β -sheets, forming amyloid fibrils. To identify the regions of PrP^C potentially involved in the initial steps of the conversion to the infectious conformation, we have used high resolution NMR spectroscopy to characterize the stability and structure of bovine recombinant PrP^C (residues 121 to 230) during unfolding with the denaturant urea. Analysis of the 800 MHz ¹H NMR spectra reveals region specific information about the structural changes occurring upon unfolding. Our data suggest that the dissociation of the native β -

sheet of PrP^C is a primary step in the urea-induced unfolding process, whilst strong hydrophobic interactions between helices $\alpha 1$ and $\alpha 3$, and between $\alpha 2$ and $\alpha 3$, stabilize these regions even at very high concentrations of urea.

INTRODUCTION

Prion diseases, or transmissible spongiform encephalopathies, are a group of fatal neurological disorders affecting both humans and animals that share in common spongiform degeneration (1-6). Prion diseases can be sporadic, inherited, or acquired by infectious means. These disorders are commonly known as kuru, fatal familial insomnia and Creutzfeldt-Jacob disease in humans, mad cow disease or bovine spongiform encephalopathy in cattle, chronic wasting disease in cervids, and scrapie in sheep. Independent of the species, these conditions are associated with an accumulation in the brain of an abnormally folded endogenous protein, and these proteinaceous aggregates are thought to lead to neurodegeneration.

Studies of the molecular agent responsible for scrapie revealed that the entity was extremely resistant to UV and ionizing radiation, treatments that normally inactivated nucleic acids (7,8). These observations lead to the generally accepted “protein-only” hypothesis suggesting that the infectious agent causing prion disease is composed for the most part, if not entirely, of endogenous protein (1,9). There are two major forms of prion protein: the native and non-infectious form (PrP^C), which has mainly an α -helical structure, and the misfolded infectious form (PrP^{Sc}), which aggregates into an assembly of β -sheets forming amyloid fibrils (10-12). The first NMR structure of the mouse PrP^C revealed an unstructured N-terminal region (23-121) followed by a globular domain (121-230) containing three α -helices, a small β -sheet and a disulfide bridge between residues C179 and C214 (13). Subsequently, more than 30 structures of the globular portion of PrP^C have been characterized using NMR and X-ray crystallography from a variety of organisms, including mammals and non-mammals. Surprisingly, only a few minor

differences are observed amongst all of these structures (14), that include protein structures from organisms that seem to be immune against prion disease (e.g. rabbit). These results suggest that an understanding of prion disease and the presence of species barriers may not be fully revealed by the inspection of the structures PrP^C and/or PrP^{Sc}, but through an understanding of the mechanism by which the PrP^C protein is converted into the misfolded and infectious form. Surewicz and co-workers have recently shown that recombinant PrP^C can be converted using native-like conditions (i.e. physiologically relevant) into amyloid fibrils, consisting of proteins with parallel and in-register β -structures (15,16). It was shown that the conversion of recombinant prion protein into a β -sheet-rich form with similar physicochemical properties to PrP^{Sc} could be induced by acidic pH, salt and urea (17).

Over the years, many models were proposed in order to highlight the different characteristics of PrP^{Sc}. These include insolubility, partial proteinase K digestion and formation of amyloid-like fibrils (2). Using molecular dynamics simulation, Demarco et al. proposed a model for the formation of PrP^{Sc} molecules involving mainly a growth of the native β -sheet of PrP^C using the flexible N-terminal tail of the protein (18). On the other hand, based on experimental data, Cobb et al. suggested a parallel in-register β -structure for the amyloid fibrils using hydrogen/deuterium exchange, site-directed spin labeling coupled with electron paramagnetic resonance spectroscopy and mass spectroscopy analysis (15). One of the most popular models was presented by Govaerts et al., which is based on electron microscopy data of 2D crystals that shows a left-handed β -helix trimeric structure (19). The feasibility of this model is supported by the high-resolution solid-state NMR structure of infectious material published recently for the yeast prion HET-s(218-289) (20), which revealed a β -solenoid with a triangular hydrophobic core that is similar to the model proposed by Govaerts et al.

In order to obtain a residue specific structural description of the potential initial steps in the conversion process, we are using 800 MHz ¹H 1D and ¹H-¹H 2D NOESY NMR methods to characterize the differential stability of regions of the PrP^C protein in presence of a denaturant. The use of denaturants to characterize protein unfolding is a widely used technique that has been extensively exploited to evaluate and compare protein stabilities (21). According to the linear free energy model, the changes in free energy that accompany protein unfolding are linearly dependent on the concentration of denaturant (22). It has been shown that an uncharged denaturant like urea is a good choice to monitor protein unfolding compared to other denaturants (23), and this is convenient since denaturants like guanidine hydrochloride are not ideal for NMR studies because of their ionic properties. The characterization of protein stability by NMR using urea as denaturant has been used previously to evaluate the stability of other proteins, such as the 434-repressor by Wüthrich and co-workers (24). Nicholson et al. have also compared the global stability of Syrian hamster PrP (shPrP) and the mouse doppel protein using optical spectroscopy, and monitored the effect of urea on the amide hydrogen exchange rates using NMR (25). Several other studies on the stability of the C-terminal domain of the prion protein can be found in the literature; some of them using molecular dynamics simulations (26,27), and biophysical techniques like Trp fluorescence, NMR and CD spectroscopies (12,25,28-31). The identification and characterization of the key elements favoring the prion protein transformation is an important step to obtain a more complete understanding of protein misfolding diseases.

MATERIALS AND METHODS

Protein purification

The bPrP₁₂₁₋₂₃₀ gene was synthetically synthesized with codon usage optimized for *E. coli* expression. *Nde*I and *Bam*HI cut sites were included at the 5' and 3' ends of the sequence to facilitate ligation into pRSETa plasmid previously digested with the same restriction enzymes. Ligations were transformed into XL-1

blue cells, and colonies were screened for insertions. The creation of pRSETa-bPrP₁₂₁₋₂₃₀ was confirmed by sequencing in the forward and reverse directions. The C-terminus region of bovine PrP containing residues 121 to 230 (bPrP₁₂₁₋₂₃₀) was expressed and purified in *E. coli* based on previously described methodology for human PrP (32). A molecular weight of $13,093 \pm 5$ Da was measured using mass spectroscopy compared to 13,097 Da for the predicted molecular weight. The purity of the protein was also verified using 1D ¹H NMR spectrum.

NMR samples

The NMR sample was prepared with ~2 mg of bPrP₁₂₁₋₂₃₀ in 500 µl of deuterated NMR buffer containing 10 mM of sodium acetate, 0.3 mM of deuterated 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS-*d*₆) and 0.012 % (w/v) NaN₃. The pH of the sample was 6.3 according to the chemical shift of the acetate methyl group based on previously described methodology (33). The urea-induced unfolding of bPrP₁₂₁₋₂₃₀ was performed by adding solid deuterated urea (urea-*d*₄, Sigma-Aldrich Co.) directly into the NMR tube. The urea concentration was increased by 1 M at each titration step until 10 M, since the solubility of urea in water is 10.49 M at 25 °C. In order to account for the volume increase caused by the addition of solid denaturant, the length of the NMR sample was measured at each step, and an increased amount of urea-*d*₄ was added at each step to reach the correct molarity.

NMR spectroscopy

The NMR experiments were all performed on an 800 MHz Varian INOVA NMR spectrometer equipped with a cryogenic probe. The one-dimensional (1D) proton NMR spectra were acquired with 256 transients, a spectral width of 15.0 ppm and a time delay of 2 s, while the two-dimensional (2D) ¹H-¹H NOESY spectra were acquired with 32 transients, 256 increments, spectral widths of 15.0 ppm for both proton dimensions, a time delay of 1 s and a mixing time 100 ms. A total of 50 hours were needed to acquire the NMR data between the time at which urea-*d*₄

was added for the first point until the acquisition of the last spectrum of the titration.

Data processing

The processing and visualization of the 1D spectra was done with the VnmrJ software v2.1B (Varian inc.). Each spectrum was processed with a line broadening of 1.5 Hz and the vertical scale (and integral scale) was adjusted according to the volume increase of the sample caused by the addition of solid urea. The peak areas were then calculated by isolating a given resonance in a window of approximately 0.1 ppm, applying a drift correction on the baseline, and measuring the area under the curve by using the *dli* command in the Vnmrj software (display list of integral). Intensities, integral and deconvolution techniques were used to analyze the resonances, but only the results of the peak area are reported here (see Discussion). The peak area of a given resonance was normalized to the first titration point (0 M), then plotted as a function of the urea concentration and fit using the xcrvfit software (34).

The processing of the 2D spectra was performed using nmrPipe (35). Linear prediction was used in the *t2* dimension to add resolution for a maximum of half the number of acquired points. Sine functions of 90° and 75° were applied in the *t1* and *t2* dimensions, respectively, and zero-filling was used to double the number of points in both dimensions before the Fourier Transform. NMRViewJ (One Moon Scientific inc.) was used to analyze the 2D NOESY spectra and obtain the cross-peak volumes from the peak analysis tools. Prior to the volume measurement, a box was manually adjusted around each identified resonances to define the limits of the cross-peaks, allowing a more accurate measurement of the peak volumes.

Data analysis

Assuming a two-state model of unfolding and a linear dependency of the Gibbs energy of unfolding (ΔG_{N-U}) as a function of the concentration of urea (21,22):

$$\Delta G_{N-U} = \Delta G_{N-U}^{\circ} - m[D] \quad [1]$$

where ΔG_{N-U}° is the free energy of unfolding extrapolated to zero denaturant concentration, m is the slope at midpoint of the unfolding transition and $[D]$ is the concentration of denaturant. Since ΔG_{N-U} is equal to zero at midpoint of the denaturation, ΔG_{N-U}° is by definition equal to $m \cdot [D]_{1/2}$, where $[D]_{1/2}$ is the concentration of denaturant at midpoint of unfolding. This leads to the following formula to fit the data and directly extract the thermodynamic properties of the protein upon denaturant-induced unfolding:

$$\frac{I}{I_0} = \frac{a}{1 + \exp\{m([D] - [D]_{1/2})/RT\}} \quad [2]$$

Where I is the peak area or peak volume at a given denaturant concentration, I_0 is the peak area or peak volume at zero concentration denaturant, the m -value reflects the dependence of the free energy of unfolding on the denaturant concentration, $[D]$ is the concentration of denaturant, $[D]_{1/2}$ is the concentration of denaturant at midpoint of unfolding, a is a scaling constant, R is the gas constant and T is the temperature.

RESULTS

We used NMR to compare our purified bPrP^C₁₂₁₋₂₃₀ with those of a previous study. Two regions of the 1D ¹H NMR spectrum of bPrP^C₁₂₁₋₂₃₀ are shown in the lower traces of Fig. VIII-1A and VIII-1B. D₂O was used as a solvent to ensure a better detection of the H_α and aromatic protons, which would otherwise be buried under a strong water signal and overlapping amide protons, respectively. This spectrum was compared to those obtained by Wüthrich and co-workers for recombinant bPrP^C₂₃₋₂₃₀ and bPrP^C₂₃₋₂₃₀ isolated from healthy calf brains (36). The very close similarity between the spectra (e.g. methyl region presented in Fig. VIII-1A) suggests that the proteins have a similar three-dimensional globular fold

and that the D₂O did not affect the protein structure. We also acquired a 2D ¹H-¹H NOESY NMR spectrum at 800 MHz. The comparison of the cross-peaks with the chemical shifts previously deposited in the Biological Magnetic Resonance Data Bank (BMRB entry 4653) confirmed that the structure of the C-terminal domain of bPrP was properly folded and essentially identical to the previously determined structure (37). We then performed a urea denaturation titration to evaluate the stability of the prion protein. Increasing concentrations of deuterated urea (from 0 to 10 M) were added to the sample containing bPrP^C. By using deuterated urea we would avoid re-association of labile protons with the protein. For each step, two NMR spectra were acquired: a 1D ¹H and a 2D ¹H-¹H NOESY. The results and analysis of the two sets of experiments are presented below.

The 1D ¹H NMR spectra monitor the effect of urea on the chemical environment of various residues, which affects their position or chemical shift in the NMR spectrum, which in turn reflects the protein structure changes caused by the presence of the denaturant. Stacked plots of the 1D ¹H spectra acquired during the urea-induced denaturation titration for the aliphatic and aromatic regions are presented in Fig. VIII-1A and VIII-1B, respectively. For all spectra presented, the peak intensities were carefully corrected for the dilution during the urea titration from the partial specific volume of the urea added (see Materials and Methods for details). In order to quantify the changes in the spectra, one can utilize the chemical shifts or the peak intensity/integral of the resonances. In the presence of fast exchange (i.e. when the exchange rate is on the μs time scale), one may use chemical shifts to monitor the structural changes since the peak intensities are likely to be constant. However, in the presence of slow conformational exchange (i.e. when the exchange rate is on the s time scale), the chemical shifts are relatively unperturbed while the peak intensities or peak areas can be used to monitor the unfolding. We have simulated the line shapes and chemical shift perturbations that would occur in the presence of fast or slow exchange (see Appendix C, Fig. C-1). Using a full NMR lineshape analysis (38), we have

determined that the chemical exchange between the native and unfolded forms is in the NMR slow exchange limit with a lifetime on the order of approximately 1 second, as was previously shown by Wüthrich and co-workers for a protein of similar stability (39). This justifies the use of peak integral as a measure of the fraction of the native conformation for the analysis. It also precludes a simple interpretation of the observed chemical shifts. The use of the peak integrals is a better choice than the measurement of the peak intensities for three reasons: 1) the peak integral will take into account any effect of line-broadening of the resonance, 2) it will include all frequencies of coupled protons for a given resonance, and 3) it will be less affected by the data processing method, like drift correction, apodization and line broadening.

We have identified the eight resonances with which to follow structural changes (I139.H δ ₁, I182.H δ ₁, I182.H γ ₂, F198.H α , Y162.H α , Y218.H δ , Y163.H δ and Y150.H δ). These resonances were chosen since they show no or minimal overlap with other peaks during the urea titration for a quantitative analysis. Conveniently, the eight most distinguishable resonances identified in the 1D ¹H spectrum are from residues well dispersed in different regions of PrP. By plotting the peak area of these resonances as a function of the urea concentration (Fig. VIII-2), one can compare the relative stability of the protein as reflected by these residues. The data were fit using a sigmoidal function (see Eq. [2]) from which the two thermodynamic parameters that reflect the sensitivity of a given residue to urea can be extracted: $[D]_{1/2}$, the urea concentration required to diminish the peak area by half, and the m -value, the sensitivity of the free energy of unfolding on the urea concentration. Assuming a two-state mechanism of unfolding and a linear dependency of ΔG°_{N-U} as a function of the concentration of the denaturant, one can extrapolate to 0 M and obtain the ΔG°_{N-U} in absence of urea. The thermodynamic data for these eight resonances are presented in Table VIII-1. The resonances were ranked according to their increasing $[D]_{1/2}$ value. The two residues most sensitive to urea are Y162.H α and Y163.H δ , showing the two lowest $[D]_{1/2}$ values (5.4 ± 0.1 M and 6.2 ± 0.2 M), while the other six resonances

showed similar $[D]_{1/2}$ ranging from 7.1 to 7.7 M. The resonances Y162.H α and Y163.H δ also showed the two highest m -values ($4.9 \pm 0.3 \text{ kJ mol}^{-1} \text{ M}^{-1}$ and $3.9 \pm 0.3 \text{ kJ mol}^{-1} \text{ M}^{-1}$). Interestingly, these two residues are both located in the in β -sheet of bPrP^C. This result indicates that the β -sheet region is perturbed at relatively low urea concentration compared to other residues that are located away from the β -sheet. Using these thermodynamic parameters, we have simulated the native structure populations of each residue as a function of the concentration of urea (Fig. VIII-3A), and shown the location of each residue shown on the bPrP^C structure (Fig. VIII-3B). Finally, in order to obtain a global energy of unfolding for the C-terminal domain of bPrP^C, we have used a global fit on the individual denaturation curves to obtain the single best m -value and $[D]_{1/2}$ that fit all datasets. We obtained a m -value of $2.4 \text{ kJ mol}^{-1} \text{ M}^{-1}$, a $[D]_{1/2}$ of 7.2 M, which lead to a global ΔG°_{N-U} of 17.3 kJ mol^{-1} for bPrP^C. This energy of unfolding is comparable to the ΔG°_{N-U} of $4.6 \text{ kcal mol}^{-1}$ ($\sim 19.2 \text{ kJ mol}^{-1}$) obtained by Nicholson et al. for shPrP₉₀₋₂₃₀, also using urea as denaturant but CD spectroscopy to quantify the structural changes.

For several steps of the urea titration (0, 2, 5 and 8 M) a 2D ¹H-¹H NOESY NMR spectrum was also acquired to monitor the presence and disappearance of ¹H-¹H cross-peaks as a function of the denaturant concentration (Fig. VIII-4). The chemical shift assignment was based on deposited chemical shifts from the Wuthrich's group (BMRB 4563) (37). The presence of a cross-peak between two protons in a protein is dependent upon a number of factors, but can generally be taken as an indication that these protons are close in space (typically less than 6 Å) and show limited motions in the protein. This is made more complex if there is conformational exchange between two states such as native and unfolded with different distances and motional properties. We have used a full relaxation matrix analysis to simulate the effect of urea denaturation on the known structure of the bovine PrP^C, and shown that the magnitude of the NOE parallels the fraction of the native conformation (see Appendix C, Fig. C-2). We have used peak volumes to obtain an accurate measurement of the NOEs, since the peak intensities would

not take into account changes in cross-peak lineshape or proton coupling. For each spectrum, we have measured the peak volumes as a function of the urea concentration. Only unambiguous inter-residue NOEs showing no overlap with other resonances were kept for analysis. In order to visualize the location of these ¹H-¹H distances, we map the different NOEs on the three-dimensional structure of bPrP^C. Only four titration points (at 0, 2, 5 and 8 M) are available for each inter-residue NOE, limiting the accuracy of individual fit to obtain thermodynamic data. In order to obtain the *m* and $[D]_{1/2}$ values that describes best the data, a global fit using Eq. 2 was performed for sets of NOEs representing each region. The *a* value was held constant, while the *m*-value and $[D]_{1/2}$ were iterated by 0.1 between 1 to 5 kJ mol⁻¹ M⁻¹ and 1.0 and 7.0 M, respectively.

We have divided the identified NOEs in three distinct regions of bPrP^C: the β1-strand region (residues 128-133), the α1-helix region (residues 138-157) and the β2-strand region (residues 160-164). The first region corresponds to inter-residual NOEs contacts between the residues of the β1-strand, and the residues found in α2-helix and β2-strand (Fig. VIII-5). These NOEs are indicative of the proper formation of the small β-sheet of bPrP^C. The global fit of the NOEs for β1-strand region revealed a *m*-value of 1.5 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 3.5 M (the contour plot of error of the fit is presented in the inset of the graph shown in Fig. VIII-5A). Of special interest is the presence of multiple long-range NOEs between Y162 and L130; out of the 14 inter-residues NOEs identified for β1-strand, 5 of them are between Y162 and L130, with only one of them detectable at 5 M urea. These NOEs are a good probe for detecting the presence of the β-sheet, since these two protons are only close in space when the β-sheet is properly formed. The disappearance of these NOEs suggests that the β1 and β2-strands dissociate at a relatively low urea concentration. However, whilst β1-strand dissociates from its β2-strand partner, the β2-strand remains strongly in contact with helices α2 and α3 as monitored by the second region (Fig. VIII-6). The global fit obtained for the NOEs between β2-strand and helices α2 and α3 defined this region with an *m*-value of 2.6 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 5.7 M (Fig. VIII-6A). The presence of

these strong NOEs, many still present at 8 M urea, suggests that this region is highly stable even in the presence of high concentrations of denaturant. The third region that was investigated corresponds to the residues surrounding α 1-helix (residues 138-157), which also showed the characteristics of a well stable region (Fig. VIII-7). The global fitting of the NOEs of this region revealed an m -value of 3.6 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 5.7 M, similar to the β 2-strand region mentioned above. Only a few NOEs vanished early in the titration for this region, mostly long-range NOEs that were already weak in the absence of urea. Since a few NOEs were surprisingly still identifiable at 8 M urea, we have mapped these contacts on the structure of bPrP^C (Fig. VIII-8). The global fitting of the eight identified NOEs revealed a m -value of 2.2 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 6.3 M. Three of these NOEs were found in the α 1-helix region and another three NOEs were found between β 2-strand and α 2-helix. Interestingly, the last two NOEs found to be still present at 8 M urea are between the side chains of residues V166 and Y218, linking the loop₁₆₅₋₁₇₂ with the end of helix α 3. No NOE contacts were found between strands β 1 and β 2 at 8 M urea.

A different analysis of the data is presented in Appendix C, dividing the 44 inter-residue NOEs identified into two distinct categories: the stable (Fig. C-3) and less stable NOEs (Fig. C-4). The stable NOEs corresponds to the NOEs having a peak volume at 5 M urea larger than 50% of the original peak volume at 0 M urea (typically $[D]_{1/2} > 5$ M). The second group contains the less stable NOEs (Fig. C-4), corresponding to the NOEs having a peak volume at 5 M that was lower than 50% of the peak volume at 0 M urea (typically $[D]_{1/2} < 5$ M). Again, in order to obtain an m -value and $[D]_{1/2}$ that describes the data best, a global fit was performed for each set of NOEs (i.e. stable and less stable) using Eq. 2. We obtained a global minimum with an m -value of 3.0 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 5.8 M for the stable NOEs, compared to an m -value of 4.4 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 2.2 M for the less stable NOEs. Most of these NOEs are located in the α 1-helix region, which makes strong contacts with the N-terminal portion of α 3-helix (Fig. C-3). This complex network of NOEs is also accompanied by stable NOEs

between F198 situated in the loop before the α 3-helix, and M206 and Y157. On the other side of the protein, many stable NOEs can be found between β 2-strand and the adjacent helices α 2 and α 3, and only a few in the β 1-strand. With the exception of three weak NOEs found in the α 1-helix region, all of the less stable NOEs are between residues located in the β 1 and β 2-strands.

We have also performed a protein-refolding experiment. Taking the sample obtained after the urea titration from 0 to 10 M urea, we buffer exchanged the NMR sample and removed most of the urea, as indicated by the presence of a low intensity peak of urea in the 1D ¹H spectrum. We compared the acquired 1D ¹H NMR spectrum of this sample to the original 0 M urea spectrum to see if the bPrP^C was refolded. Both spectra were very similar without significant changes in regards to the chemical shifts. The peak intensities were reduced, most likely due to a loss of protein during the urea extraction, but it could also suggest an irreversible state of the aggregates observed by electron microscopy (data not shown). The similarities between the two spectra indicate that the refolded bPrP^C is identical to the original starting structure, confirming that the protein primary sequence remained unmodified during the denaturation process.

DISCUSSION

We have used ¹H NMR spectroscopy to investigate the stability of bPrP^C₁₂₁₋₂₃₀ using urea as a denaturant. The analysis of the 1D and 2D ¹H NMR spectra showed that the titration curves for all residues observed were monotonic and adequately fit using a two-state unfolding model to extract residue specific thermodynamic parameters. Using this method, we obtained individual $[D]_{1/2}$ and m -values for various residues of bPrP^C. No intermediates were stable enough, or at high enough concentrations, to be observed in any of the denaturation curves. However, the analysis revealed different stabilities for the different residues monitored, which implies that the denaturation process must involve many microscopic intermediates. Similar observations have been reported for human

PrP^C (29), and Syrian hamster PrP^C (40). Kinetics studies of the global folding of murine PrP^C showed that PrP^C rapidly folds into its native conformation without the presence of structural intermediates (41), while kinetic studies of human PrP^C (42,43) observed a folding intermediate. It has been shown in a previous study that the denaturation curve of Syrian hamster PrP^C using urea and monitored by circular dichroism exhibits the classic sigmoidal shape associated with two-state unfolding (25). However, the authors did not exclude the possibility of an undetected intermediate state.

From the 1D ¹H NMR spectra analysis, we have identified residues Y162 and Y163 as being perturbed at lower concentration of urea relative to other residues located in other secondary structures of the protein. While these two tyrosines are the most sensitive residues to urea among those examined, they also possess the two highest ΔG°_{N-U} . This implies that in the absence of denaturant, the β -sheet region is a stable region of PrP^C, but in the presence of urea is the most easily perturbed region of the protein. The NOE data also suggest that the β -sheet region is a sensitive region of bPrP^C upon addition of urea. Alternative explanations would be that some of these titrations involve close, undetected multiple steps which would lower the observed *m*-value, or that the assumption that the ΔG_{N-U} is only linearly dependent upon urea concentration is inadequate. These observations are reflected in the discrepancies found in the literature about the stability of the β -sheet. For example, molecular dynamics simulations by DeMarco et al. suggest that the extension/growth of the anti-parallel β -sheet of PrP^C is the principal conformational change leading to PrP^{Sc} (18). Kachel et al. also reported that the β 1 and β 2-strand structures are insensitive to the application of high pressure (29). However, several other studies, experimental as well as theoretical, suggest that the β -sheet of PrP^C is one of the less stable parts of the C-terminal region of PrP. Using a collective dynamic simulation approach, Blinov et al. suggested that the locations of relatively low stability may be associated with the β -sheet formed by β 1 and β 2-strands and the adjacent loops, whereas α 3-helix

appears to be a relatively stable part of the protein (26). Using coarse-grained protein molecular dynamics simulations, Chebaro et al. found that α 1-helix is remarkably stable and concluded that the stability of the native β -sheet is the principal region affected by the different pathogenic mutations (44). Using molecular dynamics simulations, Hirschberger et al. suggested that the hydrophobic interaction between helix α 1 to helix α 3 is required for PrP^C to correctly fold into its stable native structure (45). Also using computer simulations, Barducci et al. found that the antiparallel β -sheet is significantly weaker in the pathogenic D178N mutant than in the wild-type PrP^C, because of hydrogen bonding network involving residues N178 with R164 and Y128 side chains (46,47). These simulations are experimentally supported by antibody binding essays that suggest exposure of a pathological epitope YYR during the conversion of PrP^C into PrP^{Sc} (48). The evidence points to the YYR motif found in the β 2-strand, which would also implies a β -sheet dissociation. Finally, a recent NMR study of the dynamics of PrP^C reported slow exchange motions in the short anti-parallel β -sheet (31).

The analysis of our experimental results using NMR spectroscopy suggests that the β -sheet of PrP^C shows an important conformational change upon denaturation as monitored by the disappearance of key NOEs between β 1 and β 2-strands. In general, the NOE is a complex phenomenon that can be difficult to interpret, and the disappearance of an NOE can reflect a number of factors (see Supplementary Material). For example, if two residues are involved in an inter-residue NOE, but only one moves, the NOE might disappear. The loss of a NOE can reflect such change in distances, but can also result from a change in correlation time towards faster motions than the overall tumbling of the protein, which is often responsible for the loss of an intra-residue NOE, for example. Obviously the strongest support for a conformational change in a given structural element is the disappearance of several NOEs reflecting several ^1H - ^1H distances in that region. According to our 2D NOESY analysis, the α 1-helix region is highly resistant to urea denaturation, and the β 2-strand region makes strong interactions with α 2 and α 3 helices. Even

at 8 M urea, NOEs are still observable for these two stable regions, in addition to two NOEs contacts remaining present between residue V166, located in the loop₁₆₅₋₁₇₂, and Y218 located in the α 3-helix. These last two NOEs are part of an important network of intramolecular interactions found in PrP^C that defines the stability of the loop₁₆₅₋₁₇₂ and the α 3-helix together. This has been the focus of a recent study by the Wuthrich group (49), suggesting that these well-defined long-range interactions could be important for the functional specificity of PrP^C, and that this loop is a potential recognition area for the hypothetical "protein X".

Another particular residue is F198, located in the loop₁₉₄₋₂₀₀ between helices α 2 and α 3 (see Fig. VIII-7B). Among others, the aromatic protons of F198 make stable NOE contacts with the methyl of M206 and with the aromatic ring of Y157. These NOEs are still intense at 5 M urea. The mutation of this residue (F198S) is associated with the familial Gerstmann-Straussler-Scheinker disease. It has been shown that the F198S variant of human PrP^C has a dramatically increased propensity to self-associate into β -sheet-rich oligomers (50). The role of this aromatic ring seems to be to anchor the loop₁₉₄₋₂₀₀ into the hydrophobic core PrP^C. It is reasonable to think that the substitution of this bulky hydrophobic residue for a smaller and polar residue like serine would abolish the hydrophobic interactions of the aromatic ring of F198 with the core of the protein.

One could wonder about the effect of not having the flexible N-terminal tail of bPrP^C on the stability of the protein. According to a previous study using high pressure NMR spectroscopy (29), huPrP₂₃₋₂₃₀ and huPrP₁₂₁₋₂₃₀ have a similar ΔG°_{N-U} . The authors observed virtually no difference between both constructs, suggesting that the N-terminal tail does not affect the stability of the globular C-terminal domain of PrP^C.

We have quantified the stability of the C-terminal domain of the bPrP^C on a per residue basis using NMR spectroscopy. Upon urea denaturation, we have

observed an early perturbation of the β -sheet region compared to the other secondary structures, suggesting a dissociation of β 1-strand from β 2-strand at low urea concentrations. While β 1-strand detaches from β 2-strand, the β 2-strand remains strongly in contact with α 2 and α 3 as evidenced by the presence of strong NOE contacts between the two regions even at a high concentration of urea. We have also noticed strong hydrophobic interactions between the α 1-helix region and the N-terminal section of the α 3-helix, as well as between the loop₁₆₅₋₁₇₂ with the C-terminal end of α 3-helix.

FIGURES AND TABLES

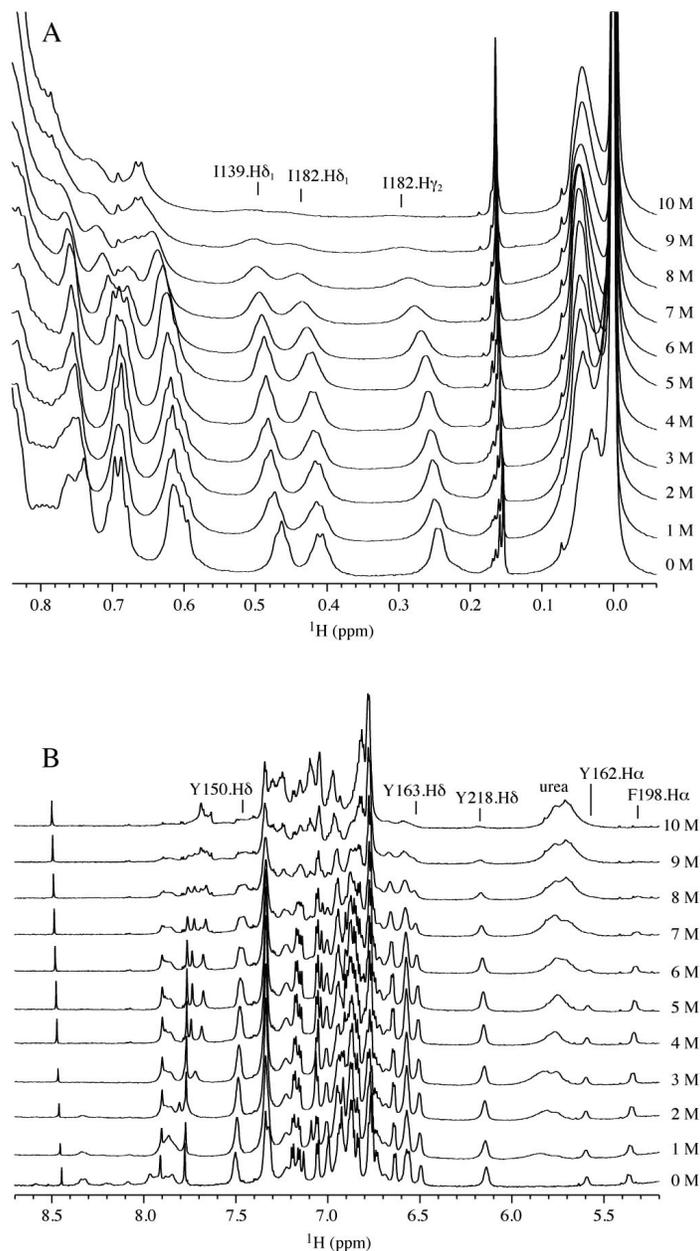


Fig. VIII-1. Titration from 0 to 10 M urea of recombinant bPrP₁₂₁₋₂₃₀ followed by 1D ¹H NMR recorded at 800 MHz. A) Stacked plot focusing on the methyl region from 0 to 0.8 ppm. I139.Hδ₁, I182.Hδ₁ and I182.Hγ₂ resonances are identified. B) Stacked plot focusing on the aromatic region from 5.2 to 8.7 ppm. F198.Hα, Y162.Hα, Y218.Hδ, Y163.Hδ and Y150.Hδ resonances are identified. The spectra were all referenced on the DSS signal at 0 ppm.

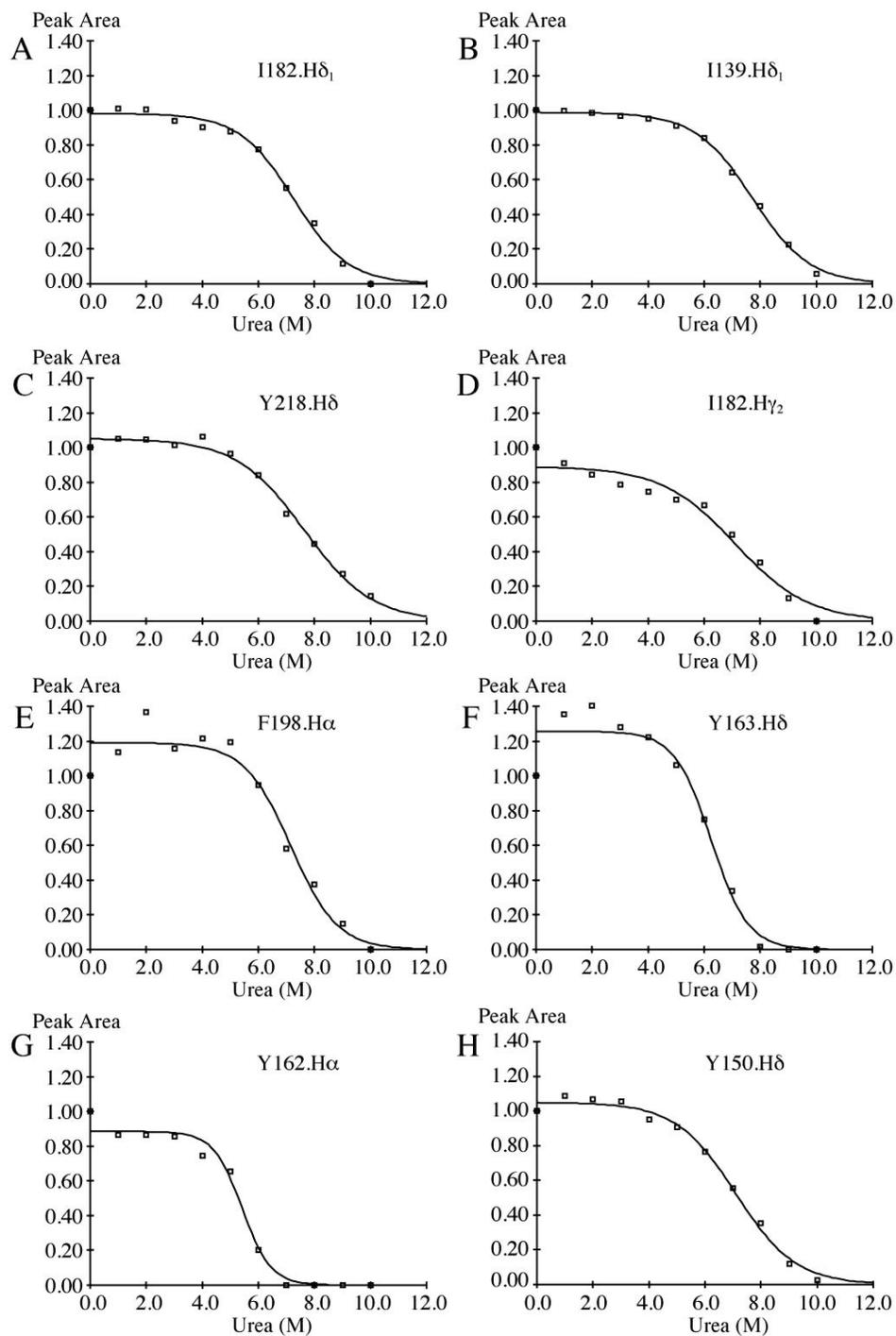


Fig. VIII-2. Peak area of the eight resonances identified in Fig. VIII-1, plotted as a function of the urea concentration: A) I182.H δ_1 , B) I139.H δ_1 , C) Y218.H δ , D) I182.H γ_2 , E) F198.H α , F) Y163.H δ , G) Y162.H α , and H) Y150.H δ . The $[D]_{1/2}$ and m -value are reported for each residue.

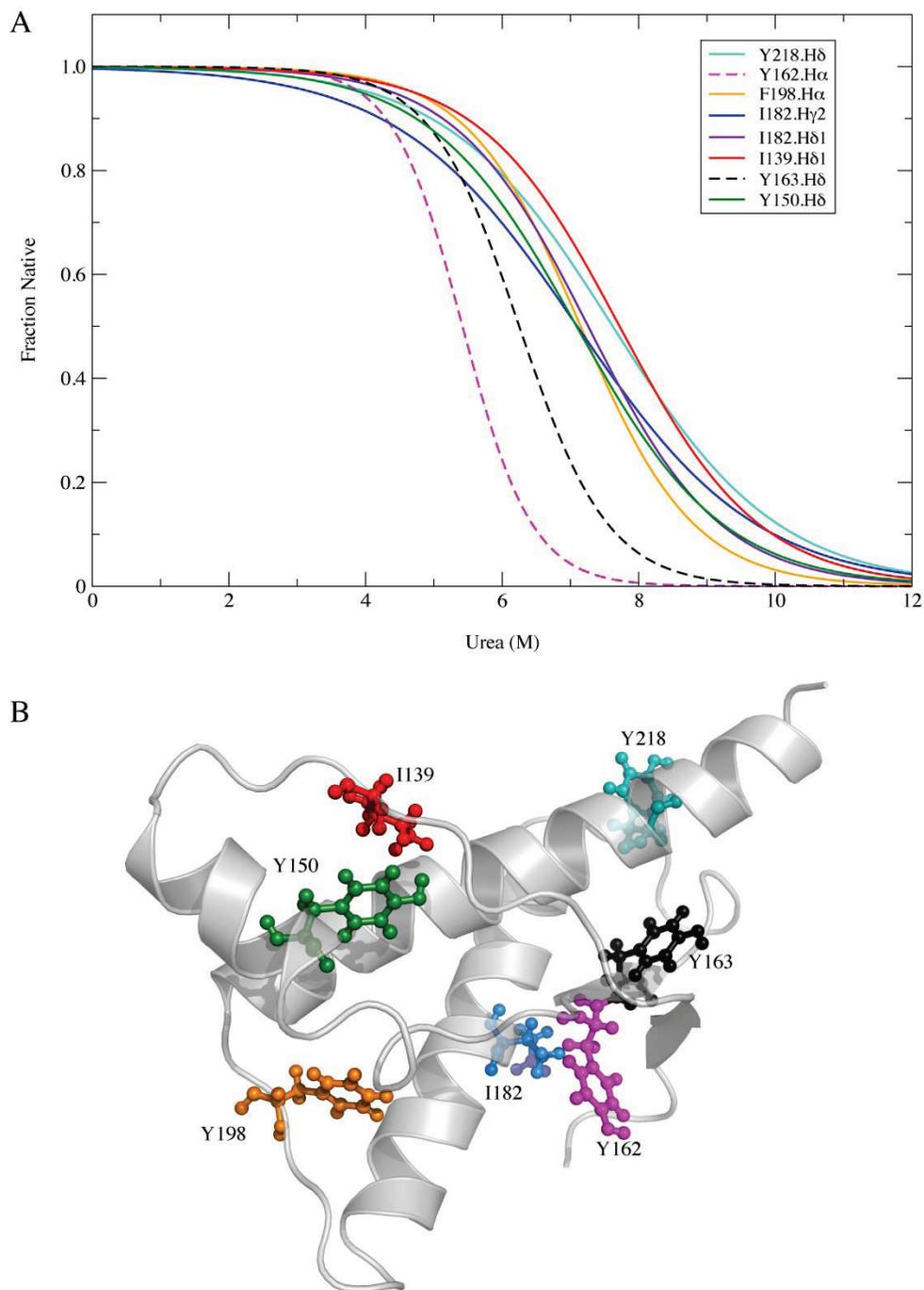


Fig. VIII-3. A) Fraction of the native conformation of different residues of bPrP₁₂₁₋₂₃₀ as a function of the urea concentration. The dash lines correspond to the denaturation curves of resonances Y162.Hα and Y163.Hδ. B) Cartoon representation of bPrP^C (PDB 1dyz). The different residues are displayed in ball-and-stick in the structure, and colored as in A).

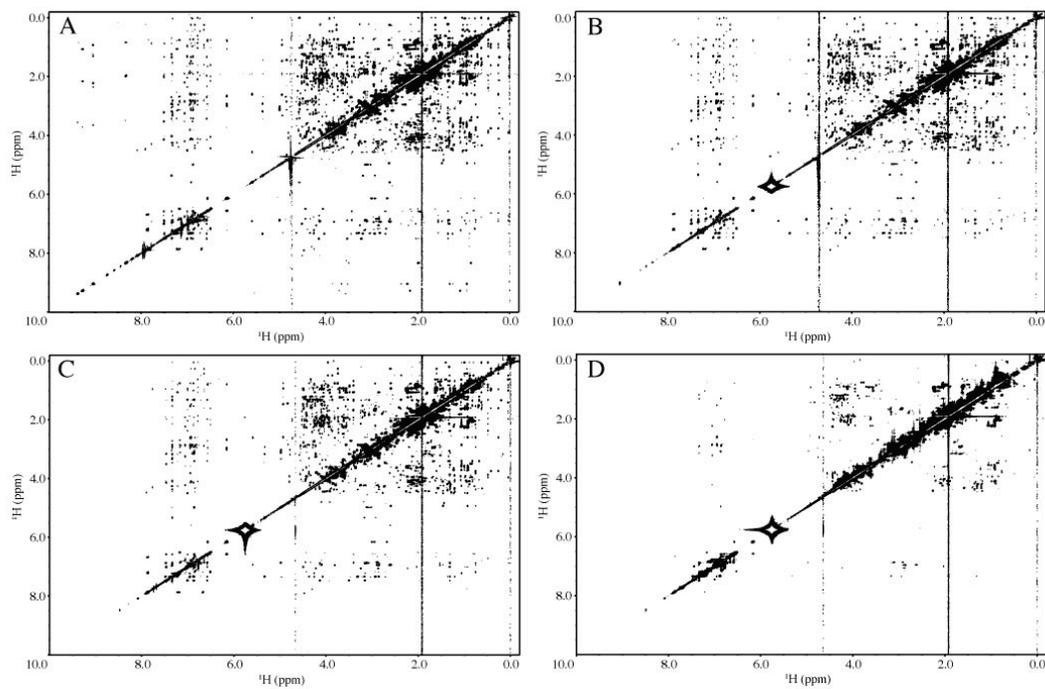


Fig. VIII-4. Four 2D ^1H - ^1H NOESY spectra of bPrP^C acquired at 800 MHz. The spectra show the decreased number of NOE cross-peaks, as the concentration of urea is increased: A) 0 M urea, B) 2 M urea, C) 5 M urea, D) 8 M urea.

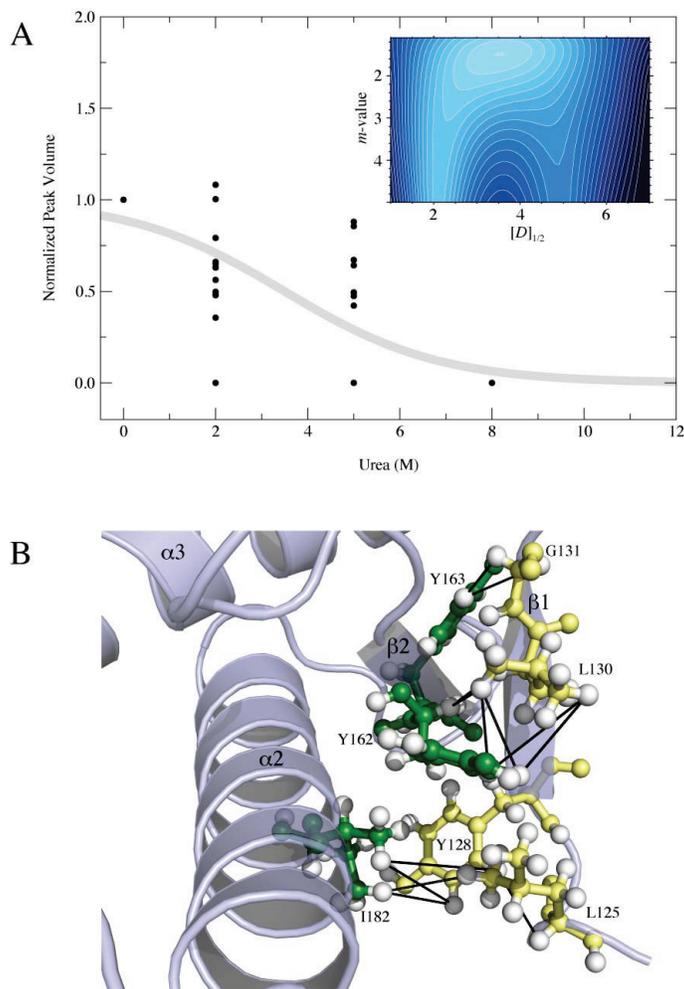


Fig. VIII-5. Inter-residue NOEs found in the $\beta 1$ -strand region of bPrP^C. A) Normalized NOE cross-peak volumes as a function of the concentration of urea for residues located in the $\beta 1$ -strand. The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of $1.5 \text{ kJ mol}^{-1} \text{ M}^{-1}$ and a $[D]_{1/2}$ of 3.5 M. The grey line on the graph corresponds to a denaturation curve with these thermodynamic parameters. B) Cartoon representation of the β -sheet region of bPrP^C. The following inter-residue ^1H - ^1H NOEs contacts are shown with black lines: 128.H δ -182.H δ_1 , 162.H δ -130.H δ_2 , 162.H α -130.H δ_1 , 162.H ϵ -130.H δ_1 , 162.H δ -130.H δ_1 , 128.H ϵ -182.H γ_2 , 128.H ϵ -182.H δ_1 , 128.H δ -182.H γ_2 , 162.H ϵ -130.H δ_2 , 163.H ϵ -131.H α_1 , 128.H δ -125.H α , 162.H α -130.H α , 163.H ϵ -131.H α_2 . The heavy atoms of the residues found in the $\beta 1$ -strand are colored in yellow, while the remaining atoms from other regions are colored in green. The hydrogen atoms are all colored in white.

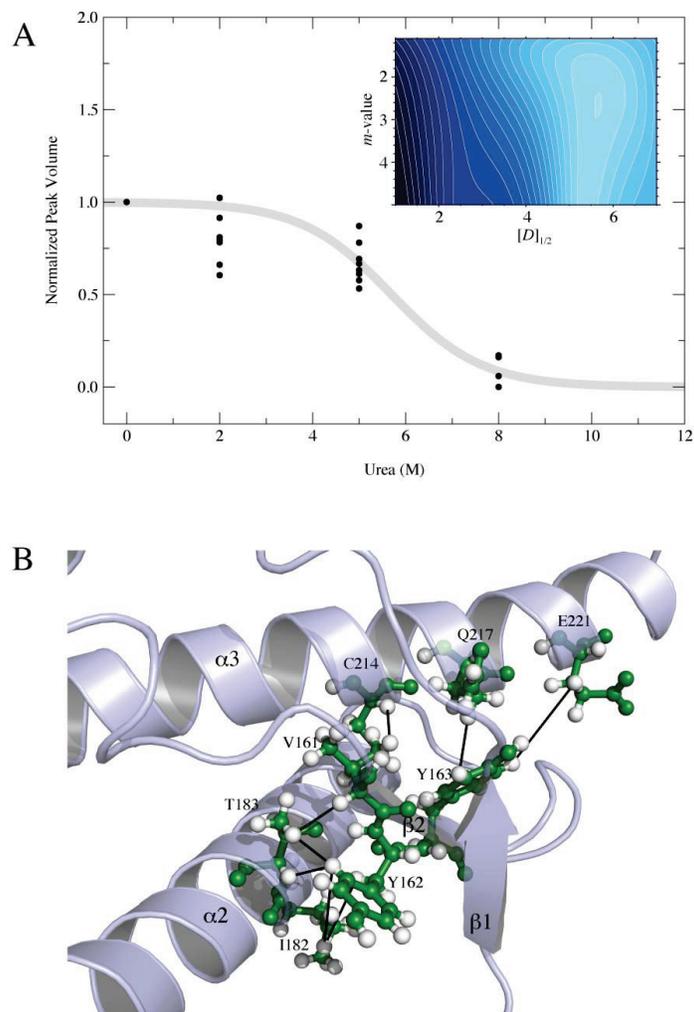


Fig. VIII-6. Inter-residue NOEs found in the β 2-strand region of bPrP^C. A) Normalized NOE cross-peak volumes as a function of the concentration of urea, for residues located in the β 2-strand region but not making contacts with residues located in the β 1-strand. The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of 2.6 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 5.7 M. The grey line on the graph corresponds to a denaturation curve with these thermodynamic parameters. B) Cartoon representation of the β -sheet region of bPrP^C. The following inter-residue ¹H-¹H NOEs contacts are shown with black lines: 162.H β ₁-182.H γ ₂, 214.H α -161.H γ ₂, 161.H α -183.H γ ₂, 162.H δ -182.H γ ₂, 162.H δ -183.H γ ₂, 163.H ϵ -217.H γ , 163.H ϵ -221.H γ , 162.H δ -183.H α . The heavy atoms are colored in green and the hydrogen atoms are colored in white.

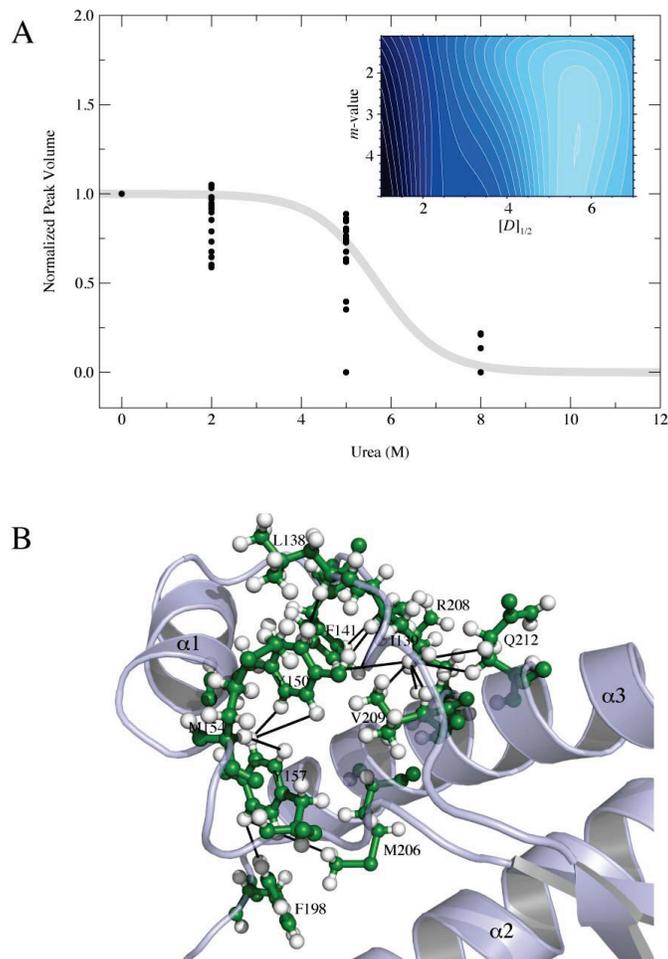


Fig. VIII-7. Inter-residue NOEs found in the $\alpha 1$ -helix region of bPrP^C. A) Normalized NOE cross-peak volumes as a function of the concentration of urea, for residues located in the $\alpha 1$ -helix region (residues 138-157). The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of $3.6 \text{ kJ mol}^{-1} \text{ M}^{-1}$ and a $[D]_{1/2}$ of 5.7 M . The grey line on the graph corresponds to a denaturation curve with these thermodynamic parameters. B) Cartoon representation of the $\alpha 1$ region of bPrP^C. The following inter-residue ¹H-¹H NOE contacts are shown with black lines: 139.H δ_1 -212.H γ , 139.H δ_1 -212.H β , 209.H α -139.H δ_1 , 139.H δ_1 -209.H γ_2 , 139.H δ_1 -208.H β , 139.H δ_1 -212.H γ , 206.H ϵ -157.H ϵ , 141.H ζ -139.H γ_2 , 141.H ϵ -139.H γ_2 , 141.H δ -139.H γ_2 , 141.H ζ -139.H δ_1 , 141.H ϵ -139.H γ_1 , 157.H ϵ -154.H α , 157.H δ -154.H α , 150.H ϵ -138.H α , 150.H ϵ -157.H ϵ , 198.H δ -157.H ϵ , 150.H δ -157.H ϵ . The heavy atoms are colored in green and the hydrogen atoms are colored in white.

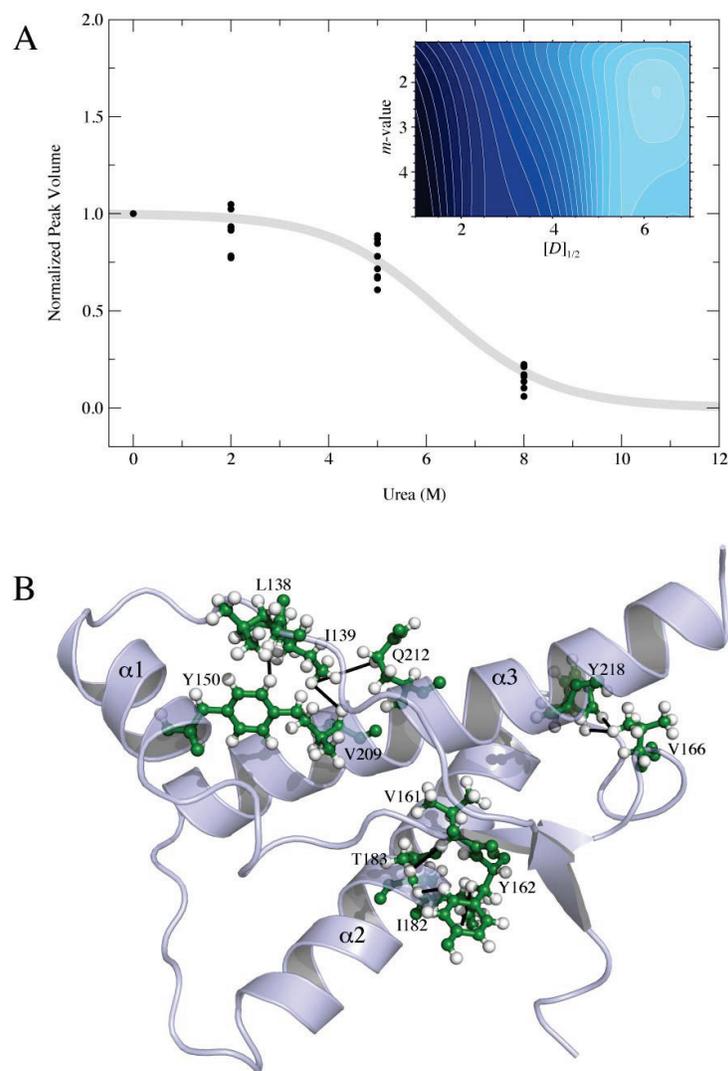


Fig. VIII-8. Inter-residue NOEs of bPrP^C identified at 8 M of urea. A) Normalized NOE cross-peak volumes as a function of the concentration of urea, for NOEs still identifiable at 8 M urea. The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of 2.2 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 6.3 M. The grey line on the graph corresponds to a denaturation curve with these thermodynamic parameters. B) Cartoon representation of bPrP^C. The following inter-residue ¹H-¹H NOEs contacts are shown with black lines: 162.H β -182.H γ_2 , 209.H α -139.H δ_1 , 161.H α -183.H γ_2 , 139.H δ_1 -212.H γ , 218.H δ -166.H γ_2 , 218.H ϵ -166.H γ_2 , 162.H δ -183.H α , 150.H ϵ -138.H α . The heavy atoms are colored in green and the hydrogen atoms are colored in white.

Table VIII-1. Thermodynamic properties of different residues of recombinant bPrP₁₂₁₋₂₃₀ as measured by 1D ¹H NMR

Resonance	$[D]_{1/2}$	m -value	ΔG°_{N-U}
	M	$\text{kJ mol}^{-1} \text{M}^{-1}$	kJ mol^{-1}
Y162.H α	5.4 ± 0.1	4.9 ± 0.3	26.7 ± 2.0
Y163.H δ	6.2 ± 0.2	3.9 ± 0.4	24.1 ± 3.5
Y150.H δ	7.1 ± 0.2	2.4 ± 0.2	16.8 ± 2.0
I182.H γ_2	7.1 ± 0.1	1.9 ± 0.7	13.7 ± 5.2
F198.H α	7.2 ± 0.1	3.0 ± 0.5	21.8 ± 4.3
I182.H δ_1	7.3 ± 0.1	2.6 ± 0.2	18.8 ± 1.9
Y218.H δ	7.6 ± 0.1	2.1 ± 0.3	15.9 ± 2.1
I139.H δ_1	7.7 ± 0.3	2.5 ± 0.1	19.1 ± 1.1

REFERENCES

1. Prusiner, S. B. (1982) Novel proteinaceous infectious particles cause scrapie, *Science* 216, 136-144.
2. Prusiner, S. B. (1998) Prions, *Proc. Natl. Acad. Sci. USA* 95, 13363-13383.
3. Collinge, J. (2001) Prion diseases of humans and animals: their causes and molecular basis, *Annu. Rev. Neurosci.* 24, 519-550.
4. Weissmann, C. (2004) The state of the prion, *Nat. Rev. Microbiol.* 2, 861-871.
5. Caughey, B., and Baron, G. (2006) Prions and their partners in crime, *Nature* 443, 803-810.
6. Aguzzi, A., Sigurdson, C., and Heikenwaelder, M. (2008) Molecular mechanisms of prion pathogenesis, *Annu. Rev. Pathol.* 3, 11-40.
7. Hadlow, W. (1959) Scrapie and kuru, *Lancet* 254, 289-290.
8. Klatzo, I., Gajdusek, D., and Zigas, V. (1959) Pathology of Kuru, *Lab. Invest.* 8, 799-847.
9. Griffith, J. S. (1967) Self-replication and scrapie, *Nature* 215, 1043-1044.
10. Caughey, B., Dong, A., Bhat, K., Ernst, D., Hayes, S., and Caughey, W. (1991) Secondary structure analysis of the scrapie-associated protein PrP 27-30 in water by infrared spectroscopy, *Biochemistry* 30, 7672-7680.
11. Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., and Prusiner, S. B. (1993) Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, *Proc. Natl. Acad. Sci. U.S.A.* 90, 10962-10966.
12. Safar, J., Roller, P., Gajdusek, D., and Gibbs, C. J. (1993) Thermal stability and conformational transitions of scrapie amyloid (prion) protein correlate with infectivity, *Protein Sci.* 2, 2206-2216.
13. Riek, R., Hornemann, S., Wider, G., Billeter, M., Glockshuber, R., and Wuthrich, K. (1996) NMR structure of the mouse prion protein domain PrP(121-231), *Nature* 382, 180-182.

14. Julien, O., Graether, S., and Sykes, B. (2009) Monitoring prion protein stability by NMR, *J. Toxicol. Environ. Health A*. 72, 1069-1074.
15. Cobb, N. J., Sonnichsen, F. D., Mchaourab, H., and Surewicz, W. K. (2007) Molecular architecture of human prion protein amyloid: a parallel, in-register beta-structure, *Proc. Natl. Acad. Sci. U.S.A.* 104, 18946-18951.
16. Cobb, N., Apetri, A., and Surewicz, W. (2008) Prion protein amyloid formation under native-like conditions involves refolding of the C-terminal alpha-helical domain, *J. Biol. Chem.* 283, 34704-34711.
17. Swietnicki, W., Morillas, M., Chen, S. G., Gambetti, P., and Surewicz, W. K. (2000) Aggregation and fibrillization of the recombinant human prion protein huPrP90-231, *Biochemistry* 39, 424-431.
18. Demarco, M., and Daggett, V. (2004) From conversion to aggregation: protofibril formation of the prion protein, *Proc. Natl. Acad. Sci. U.S.A.* 101, 2293-2298.
19. Govaerts, C., Wille, H., Prusiner, S. B., and Cohen, F. E. (2004) Evidence for assembly of prions with left-handed beta-helices into trimers, *Proc. Natl. Acad. Sci. U.S.A.* 101, 8342-8347.
20. Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A. B., Riek, R., and Meier, B. H. (2008) Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core, *Science* 319, 1523-1526.
21. Pace, C. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol.* 131, 266-280.
22. Greene, R. J., and Pace, C. (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin, *J. Biol. Chem.* 249, 5388-5393.
23. Monera, O., Kay, C., and Hodges, R. (1994) Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions, *Protein Sci.* 3, 1984-1991.

24. Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor, *Science* 257, 1559-1563.
25. Nicholson, E., Mo, H., Prusiner, S., Cohen, F., and Marqusee, S. (2002) Differences between the prion protein and its homolog Doppel: a partially structured state with implications for scrapie formation, *J. Mol. Biol.* 316, 807-815.
26. Blinov, N., Berjanskii, M., Wishart, D., and Stepanova, M. (2009) Structural domains and main-chain flexibility in prion proteins, *Biochemistry* 48, 1488-1497.
27. Zuegg, J., and Gready, J. E. (1999) Molecular dynamics simulations of human prion protein: importance of correct treatment of electrostatic interactions, *Biochemistry* 38, 13862-13876.
28. Apetri, A., and Surewicz, W. (2003) Atypical effect of salts on the thermodynamic stability of human prion protein, *J. Biol. Chem.* 278, 22187-22192.
29. Kachel, N., Kremer, W., Zahn, R., and Kalbitzer, H. (2006) Observation of intermediate states of the human prion protein by high pressure NMR spectroscopy, *BMC Struct. Biol.* 6, 16.
30. Viles, J., Donne, D., Kroon, G., Prusiner, S., Cohen, F., Dyson, H., and Wright, P. (2001) Local structural plasticity of the prion protein. Analysis of NMR relaxation dynamics, *Biochemistry* 40, 2743-2753.
31. O'Sullivan, D. B., Jones, C., Abdelraheim, S., Brazier, M., Toms, H., Brown, D., and Viles, J. (2009) Dynamics of a truncated prion protein, PrP(113-231), from (15)N NMR relaxation: order parameters calculated and slow conformational fluctuations localized to a distinct region, *Protein Sci.* 18, 410-423.
32. Zahn, R., Von Schroetter, C., and Wuthrich, K. (1997) Human prion proteins expressed in *Escherichia coli* and purified by high-affinity column refolding, *FEBS Lett.* 417, 400-404.

33. Baryshnikova, O., Williams, T., and Sykes, B. (2008) Internal pH indicators for biomolecular NMR, *J. Biomol. NMR* 41, 5-7.
34. <http://www.bionmr.ualberta.ca/bds/software/xcrvfit>.
35. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
36. Hornemann, S., Schorn, C., and Wuthrich, K. (2004) NMR structure of the bovine prion protein isolated from healthy calf brains, *EMBO Reports* 5, 1159-1164.
37. Lopez Garcia, F., Zahn, R., Riek, R., and Wuthrich, K. (2000) NMR structure of the bovine prion protein, *Proc. Natl. Acad. Sci. U.S.A.* 97, 8334-8339.
38. Mckay, R. T., Tripet, B., Pearlstone, J., Smillie, L., and Sykes, B. (1999) Defining the region of troponin-I that binds to troponin-C, *Biochemistry* 38, 5478-5489.
39. Neri, D., Wider, G., and Wuthrich, K. (1992) ¹H, ¹⁵N and ¹³C NMR assignments of the 434 repressor fragments 1-63 and 44-63 unfolded in 7 M urea, *FEBS Lett.* 303, 129-135.
40. Kuwata, K., Li, H., Yamada, H., Legname, G., Prusiner, S., Akasaka, K., and James, T. (2002) Locally disordered conformer of the hamster prion protein: a crucial intermediate to PrPSc?, *Biochemistry* 41, 12277-12283.
41. Wildegger, G., Liemann, S., and Glockshuber, R. (1999) Extremely rapid folding of the C-terminal domain of the prion protein without kinetic intermediates, *Nat. Struct. Biol.* 6, 550-553.
42. Apetri, A. C., and Surewicz, W. K. (2002) Kinetic intermediate in the folding of human prion protein, *J. Biol. Chem.* 277, 44589-44592.
43. Apetri, A. C., Surewicz, K., and Surewicz, W. (2004) The effect of disease-associated mutations on the folding pathway of human prion protein, *J. Biol. Chem.* 279, 18008-18014.

44. Chebaro, Y., and Derreumaux, P. (2009) The Conversion of Helix H2 to beta-Sheet Is Accelerated in the Monomer and Dimer of the Prion Protein upon T183A Mutation, *J. Phys. Chem. B* 113, 6942-6948.
45. Hirschberger, T., Stork, M., Schropp, B., Winklhofer, K., Tatzelt, J., and Tavan, P. (2006) Structural instability of the prion protein upon M205S/R mutations revealed by molecular dynamics simulations, *Biophys. J.* 90, 3908-3918.
46. Barducci, A., Chelli, R., Procacci, P., and Schettino, V. (2005) Misfolding pathways of the prion protein probed by molecular dynamics simulations, *Biophys. J.* 88, 1334-1343.
47. Barducci, A., Chelli, R., Procacci, P., Schettino, V., Gervasio, F., and Parrinello, M. (2006) Metadynamics simulation of prion protein: beta-structure stability and the early stages of misfolding, *J. Am. Chem. Soc.* 128, 2705-2710.
48. Paramithiotis, E., Pinard, M., Lawton, T., Laboissiere, S., Leathers, V., Zou, W., Estey, L., Lamontagne, J., Lehto, M., Kondejewski, L., Francoeur, G., Papadopoulos, M., Haghighat, A., Spatz, S., Head, M., Will, R., Ironside, J., O'Rourke, K., Tonelli, Q., Ledebur, H., Chakrabarty, A., and Cashman, N. (2003) A prion protein epitope selective for the pathologically misfolded conformation, *Nat. Med.* 9, 893-899.
49. Christen, B., Hornemann, S., Damberger, F., and Wuthrich, K. (2009) Prion Protein NMR Structure from Tammar Wallaby (*Macropus eugenii*) Shows that the beta2-alpha2 Loop is Modulated by Long-Range Sequence Effects, *J. Mol. Biol.* 389, 833-845.
50. Vanik, D., and Surewicz, W. (2002) Disease-associated F198S mutation increases the propensity of the recombinant prion protein for conformational conversion to scrapie-like form, *J. Biol. Chem.* 277, 49065-49070.

CHAPTER IX – Cidofovir DNA duplex*Solution Structure of a DNA Duplex Containing the Potent
Anti-Poxvirus Agent Cidofovir*

Olivier Julien,¹ James R. Beadle,² Wendy C. Magee,³ Subhrangsu Chatterjee,¹
Karl Y. Hostetler,² David H. Evans,³ and Brian D. Sykes¹

¹ *Department of Biochemistry, University of Alberta, Edmonton, AB, Canada.*

² *Veterans Medical Research Foundation and the Department of Medicine, University of California, San Diego, La Jolla, California, USA.*

³ *Department of Medical Microbiology and Immunology, Li Ka Shing Institute of Virology, University of Alberta, Edmonton, AB, Canada.*

Publication: Julien et al. 2011. *J. Am. Chem. Soc.*, 79:1240-1250.

Contributions: OJ and JRB contributed equally to this work. JRB and KYH synthesized the DNA duplexes. OJ, SC and BDS acquired NMR experiments and assigned chemical shifts. OJ determined the structures, made all of the figures (except scheme IX-1 and figure IX-2) and wrote the manuscript with WCM, DHE and BDS.

OVERVIEW

Cidofovir (1-(*S*)-[3-hydroxy-2-(phosphonomethoxy)propyl]cytosine, CDV) is a potent inhibitor of orthopoxvirus DNA replication. Prior studies have shown that when CDV is incorporated into a growing primer strand, it can inhibit both the 3'-to-5' exonuclease and the 5'-to-3' chain extension activities of vaccinia virus DNA polymerase. This drug can also be incorporated into DNA, creating a significant impediment to *trans*-lesion DNA synthesis in a manner resembling DNA damage. CDV and deoxycytidine share a common nucleobase but CDV lacks the deoxyribose sugar. The acyclic phosphonate bears a hydroxyl moiety that is

equivalent to the 3'-hydroxyl of dCMP and permits CDV incorporation into duplex DNA. To study the structural consequences of inserting CDV into DNA, we have used ^1H NMR to solve the solution structures of a dodecamer DNA duplex containing a CDV molecule at position 7 and of a control DNA duplex. The overall structures of both DNA duplexes were found to be very similar. We observed a decrease of intensity (>50%) for the imino protons neighboring the CDV (G6, T8) and the cognate base G18, and a large chemical shift change for G18. This indicates higher proton exchange rates for this region, which was confirmed using NMR monitored melting experiments. DNA duplex melting experiments monitored by circular dichroism revealed a lower T_m for the CDV DNA duplex (46°C) compared to the control (58°C) in 0.2 M salt. Our results suggest that the CDV drug is well accommodated and stable within the dodecamer DNA duplex, but the stability of the complex is less than the control suggesting increased dynamics around the CDV.

INTRODUCTION

The antiviral agent cidofovir (1-(*S*)-[3-hydroxy-2-(phosphonomethoxy)propyl]cytosine, CDV) is the prototype of the acyclic nucleoside phosphonate class of drugs. This compound is a deoxycytidine monophosphate (dCMP) analog in which the phosphate has been replaced by an isosteric phosphonate moiety (Fig. IX-1A). CDV is an effective inhibitor of a broad spectrum of double stranded DNA viruses, including poxviruses, herpesviruses and adenoviruses, and is currently approved for the treatment of cytomegalovirus retinitis in AIDS patients (1). The main drawbacks encountered during clinical use of CDV are its poor oral bioavailability requiring that it be administered by intravenous infusion, and its tendency to concentrate in the kidney proximal tubule, resulting in nephrotoxicity (2,3). However, a lipid conjugate of CDV, hexadecyloxypropyl CDV (CMX001), appears to overcome these limitations (4). Interestingly, CMX001 possesses increased antiviral activity over the parent compound against vaccinia virus, cowpox virus, ectromelia virus,

cytomegalovirus, herpes simplex virus and adenovirus (5-9). CMX001 is currently being developed as an orally active drug that could be used in case of an outbreak of smallpox (10) or for various double stranded DNA virus infections.

We have been studying how CDV and related acyclic nucleoside phosphonates inhibit orthopoxvirus DNA replication (11-14). CDV is taken up into cells by fluid-phase endocytosis (15) and then converted by cellular kinases into CDV diphosphate (CDVpp), an analog of dCTP (16). The enhanced antiviral efficacy of CMX001 is due to increased uptake as well as by better metabolic conversion into CDVpp (17). We have used vaccinia virus DNA polymerase to show that CDVpp is a substrate for orthopoxvirus polymerases *in vitro*, although it is used less efficiently than dCTP (11) and, unlike many drugs based upon nucleotide analogs, adding CDV to a growing DNA strand does not cause immediate chain termination (11). Instead, incorporating this drug into DNA has three different effects on reactions catalyzed by the DNA polymerase (11,12). First, after CDV and one more dNMP (CDV+1) are incorporated into a growing primer strand, addition of the next nucleotide is greatly slowed. Second, this CDV+1 structure blocks the 3'-to-5' proofreading exonuclease activity, preventing drug excision. Finally, when CDV is incorporated into what is destined to become the next template strand, the polymerase cannot extend new primers beyond the site of CDV incorporation, effectively blocking further rounds of replication. More recent studies have shown that virus DNA forms aberrant structures that are packaged poorly when vaccinia virus replicates in the presence of CDV.(18) Interestingly, the related compound 9-(*S*)-[3-hydroxy-(2-phosphonomethoxy)propyl]adenine [(*S*)-HPMPA] is incorporated much more efficiently into DNA by vaccinia DNA polymerase than is CDV. However, it also creates a profound block to replication when encountered by the polymerase in the template strand (12). This difference in substrate properties may be reflected in the relatively greater antiviral activity of (*S*)-HPMPA over CDV (19).

Vaccinia virus has also been used to study the development of resistance to CDV and other related nucleoside phosphonate drugs. We, and others, have mapped resistance mutations to both the proofreading exonuclease and DNA polymerase domains of vaccinia DNA polymerase (13,20,21), reflecting the drug's complex mechanism of action. Although viruses encoding these mutations are 3- to 14-fold more resistant to CDV than wild type virus *in vitro*, they are attenuated *in vivo* (13,20,21), and these mutations do not preclude still using CDV to treat CDV-resistant virus infections in mice (13,20).

Structural studies of a molecule containing an embedded CDV could provide better insights into why these drugs create such a profound impediment to orthopoxvirus DNA synthesis. Therefore, we synthesized a single-stranded dodecamer containing CDV using a combination of “reverse” phosphoramidite DNA synthesis and methods previously described by Birkus *et al.* (22). The duplex form d(5'-CGCATG-CDV-TACGC-3')•d(5'-GCGTACGATGCG-3') was prepared in sufficient quantity for detailed studies by two dimensional nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy. In this report we compare the three dimensional structure and dynamics of the CDV-containing DNA duplex to an isosequential control DNA.

EXPERIMENTAL SECTION

General Synthetic Chemistry

¹H NMR spectra were recorded on a Varian HG spectrophotometer operating at 400 MHz and are reported in units of ppm relative to internal tetramethylsilane at 0.00 ppm. Electrospray ionization mass spectra were recorded on a Finnigan LCQDECA spectrometer at the small molecule facility, Department of Chemistry, University of California, San Diego. Chromatographic purification was done using the flash method and silica gel 60 (EMD Chemicals, Inc., 230–400 mesh). Purity (> 98%) of the target compounds was assessed by thin layer chromatography (TLC) using Analtech silica gel-GF (250 μm) plates. The

products were visualized with UV light, phospray (Supelco, Bellefonte, PA, USA) and charring.

Synthesis of the CDV Monomer

Dimethoxytrityl chloride (8.5 g, 25 mmol) was added to a solution of diethyl (*S*)-4-*N*-benzoyl-1-[3-hydroxy-2-(phosphonomethoxy)propyl]cytosine (**1**, 8.9 g, 20.25 mmol, prepared according to Brodfuehrer *et al.*(23)) and 4-dimethylaminopyridine (200 mg, 1.6 mmol) in anhydrous pyridine (100 ml). The mixture was stirred at r. t. for 18 h, then quenched with H₂O (2 ml) and concentrated under reduced pressure. The residue was dissolved in CH₂Cl₂ (150 ml) and washed with saturated aq. NaHCO₃. The organic layer was concentrated and the crude product was purified by column chromatography on silica gel 60. Elution with 1:1 hexanes/ethyl acetate yielded 13.3 g diethyl (*S*)-4-*N*-benzoyl-1-[3-(dimethoxytrityloxy)-2-(phosphonomethoxy)propyl]cytosine (**2**) as a glassy solid (82% yield). ¹H NMR (DMSO-*d*₆) δ 8.05 (d, 1H), 8.03 (d, 1H), 7.66 (t, 1H), 7.55 (t, 2H), 7.46 (d, 2H), 7.36 (t, 2H), 7.34 – 7.21 (m, 7H), 6.94 (d, 4H), 4.14 – 3.88 (m, 9H), 3.77 (s, 6H), 3.29 (d, 1H), 3.00 (d, 1H), 1.23 (t, 3H), 1.21 (t, 3H).

Bromotrimethylsilane (830 mg, 5.4 mmol) was added to a solution of diethyl ester **2** (1.0 g, 1.35 mmol) and 2,6-lutidine (1.15 g, 10.8 mmol) in anhydrous acetonitrile (25 ml). The mixture was stirred at r. t. overnight then concentrated in vacuo. Water (5 ml) was added to the residue and the mixture was frozen and lyophilized. The crude phosphonic acid [(*S*)-4-*N*-benzoyl-1-[3-(dimethoxytrityloxy)-2-(phosphonomethoxy)propyl]cytosine (**3**)] was used for the next step without further purification.

N,N-dicyclohexylcarbodiimide (DCC, 618 mg, 3 mmol) was added to a mixture of phosphonic acid **3** (1.06 g, 1.35 mmol) and 4-methoxy-1-oxido-2-pyridylmethanol (314 mg, 2 mmol, prepared according to Rejman *et al.* (24)) in dry pyridine (10 ml), and the reaction mixture was stirred overnight at r. t. then quenched with H₂O (0.5 ml) and concentrated under reduced pressure. The product [4-Methoxy-1-oxido-2-picolyl (*S*)-4-*N*-benzoyl-1-[3-(dimethoxytrityloxy)-2-

(phosphonomethoxy)propyl]cytosine (4) was purified by column chromatography on silica gel 60 using an elution gradient from 100% CH₂Cl₂ to 20% EtOH/CH₂Cl₂. ¹H NMR (DMSO-*d*₆) δ 8.13 (d, 1H), 8.11 (d, 1H), 8.03 (d, 2H), 7.65 (t, 1H), 7.55 (t, 2H), 7.43 (d, 2H), 7.33 (t, 1H), 7.29 (d, 4H), 7.23 (t, 1H), 7.17 (d, 1H), 7.13 (d, 1H), 6.90 (d, 4H), 4.90 (d, 2H), 3.77 (s, 3H), 3.75 (s, 6H), 3.50-3.64 (m, 4H), 3.40 (dd, 2H), 3.0-2.8 (m, 3H). Mass spectrum (electrospray ionization): m/z 821.38 [M-H]⁻.

Oligonucleotide Synthesis

The single strand CDV-containing oligonucleotide was prepared by TriLink BioTechnologies, Inc. (San Diego, CA). The oligonucleotide was synthesized on a 10 μmol scale in the 5'-to-3' direction using 5'-phosphoramidite monomers. CDV monomer 4 was incorporated into the oligonucleotide using the phosphotriester coupling method (22). The synthetic oligonucleotide was purified by reverse-phase HPLC and analyzed by PAGE and mass spectroscopy. An isosequential control and complementary oligonucleotides were synthesized by the University of Calgary Core DNA Services (Calgary, AB) on a 15 μmole scale and purified by standard desalting. Additional oligonucleotides used for primer extension analyses were purchased from IDT (Coralville, IA).

CD Spectroscopy

Samples of the CDV and control DNA duplexes were prepared at a final concentration of 3.25 μM in 10 mM sodium phosphate (pH 7.3), 185 mM sodium chloride. These samples were scanned from 205 to 360 nm with a 1-nm increment using a 10 mm path length and at 20 °C to 92 °C in 3-degree increments. A sample containing only buffer was used as a control. All scans were performed on an Olis DSM 17 Circular Dichroism Spectrophotometer. The CD spectra were corrected to subtract the solvent blank (acquired at 20 °C) and plotted using GraphPad Prism software.

NMR Spectroscopy of the DNA Duplexes

The NMR samples were prepared in 95% H₂O/5% D₂O or 99.99% D₂O, pH between 7.0-7.2, 50 mM NaCl, 20 mM Na₂HPO₄, 1 mM EDTA and 0.25 mM DSS-*d*₆. According to UV absorbance measurements, the final concentration of control and CDV DNA dodecamers in the sample was approximately 2 mM for both samples. All NMR spectra used in this study were collected at 25 °C (unless otherwise specified) using Varian Inova 500 and 600 MHz NMR spectrometers equipped with triple resonance probes with Z-pulsed field gradients and a computer-controlled variable temperature (VT) module to regulate the temperature. The NOESY spectra were acquired at 600 MHz with a spectral width of 20 ppm (H₂O samples) or 14 ppm (D₂O samples) in both t_2 and t_1 dimensions, with 2048 (ω_2) and 512 (ω_1) complex points, a saturation delay of 1.5 s, and a mixing time of 80 ms. DQF-COSY spectra were acquired at 600 MHz with a spectral width of 20 ppm in both t_2 and t_1 dimensions, with 4096 (ω_2) and 512 (ω_1) complex points and a saturation delay of 1.5 s. Natural abundance ¹H, ¹³C HSQC spectra were acquired at 500 MHz with spectral widths of 10 ppm and 30 ppm in t_2 and t_1 dimensions (carrier at 56 ppm in t_1), 610 (ω_2) and 256 (ω_1) complex points, and a saturation delay of 1.4 s.

Temperature series (from 5 °C to 40 °C) were performed for both the control and CDV DNA duplexes and monitored by NMR spectroscopy at 600 MHz. The spectra were acquired with the water pulse sequence of Biopack (Varian Inc.) with a spectral width of 30 ppm to insure a flat baseline, 128 transients, a relaxation delay of 2 s and an acquisition delay of 2 s. The spectra were processed and plotted with VnmrJ 2.1B, using a line broadening of 1.5 Hz. Backward linear prediction was used to correct the first two points of the free induction decay.

All two-dimensional NMR spectra were processed with nmrPipe v4.9 (25) and analyzed with NMRviewJ v8.0.b30 (One Moon Scientific Inc.). For the NOESY and HSQC spectra, a sine-bell function shifted by 90° or 75° was applied to the free induction decays in each dimensions, followed by zero-filling to a maximum of twice the number of complex points and Fourier transform. For the DQF-

COSY spectra, an unshifted sine-bell function was applied in both ^1H dimensions before continuing with the processing mentioned above.

Structure Calculations

A model of the control and CDV B-DNA 12-mers were built using Nucleic Acid Builder (26), LEap and Antechamber (27) included in AmberTools 1.3. In order to build the CDV DNA duplex, the x-ray structure of CDV (28) was superimposed on C7 of the control DNA duplex. The coordinates of C7 were then removed. The aliphatic backbone dihedral angles of the new X7 nucleotide were adjusted to obtain appropriate P5' and O3' atom positions with PyMOL (Delano Scientific Inc.). Finally, a short minimization was performed on both DNA dodecamers with Amber 10 (29) to obtain a starting point for the structure calculations using experimental restraints. Antechamber was used to obtain the topology information (bond lengths, angles and torsion angles force constants) of the modified nucleotide by itself and when incorporated into DNA, based on the X-ray coordinates of CDV (Crystallography Information File ab0159). The non-standard terms that were missing were estimated from parameters found in other molecules.

The calibration of the NOE cross-peak intensities obtained with NMRViewJ was performed with MARDIGRAS (30) using a full-relaxation matrix approach. All NOE cross-peak intensities involving the cytosine H6 protons were doublets because of the $^3J_{\text{H}_5\text{H}_6}$ coupling observed in the spectra, while the intensities of the T and G imino contacts with the A and C amino protons were divided by 3 and 2, respectively, to compensate for chemical exchange. An initial run of MARDIGRAS was performed using the B-DNA model of the control and CDV DNA duplex described above. The RMS error between the model and the calibrated NOE distances was 0.57 Å for both duplexes. The upper bounds obtained from MARDIGRAS were multiplied by a factor 1.25 before being exported to Amber 10 for a 25 ps simulated annealing protocol (25,000 steps). The simulated annealing protocol using the pairwise generalized Born model (31) was as follows: the temperature of the system was kept constant at 600 K during

the first 5 ps, cooled down slowly to 100 K between 5-18 ps, and cooled down to 0 K for the last 7 ps. The protocol was repeated 50 times to obtain an NMR ensemble with the 10 lowest RMSD structures. The structure with the lowest RMSD of the ensemble was put back into MARDIGRAS for a second and third cycle of NOE calibration and structure calculations.

For the control DNA duplex, 181 NOEs distance restraints were used in addition to 19 pseudorotation phase angle restraints of $162 \pm 20^\circ$ based on the measurement of $^3J_{\text{H1H2}'}$ and $^3J_{\text{H1H2}''}$ coupling constants. A total of 52 standard Watson-Crick inter-strand distance restraints were added to keep the two strands together during the simulated annealing. A total of 158 α , β , γ , δ , ϵ , and ζ backbone angle restraints ($\pm 30^\circ$) based on the Dickerson-Drew B-DNA dodecamer X-ray structure (32) were used to keep the molecule in a loose B-DNA duplex conformation. For the CDV DNA duplex, 195 NOEs distance restraints, and 17 pseudorotation phase angle restraints based on $^3J_{\text{H1H2}'}$ and $^3J_{\text{H1H2}''}$ coupling constant measurements were used. A total of 150 α , β , γ , δ , ϵ , and ζ backbone angle restraints ($\pm 30^\circ$, none given for the X7), and 52 standard Watson-Crick inter-strand distance restraints were used during the calculations. The helical parameters of the DNA structures were calculated using the program X3DNA (33).

RESULTS

Synthesis of a CDV-Containing Oligonucleotide

Oligonucleotides containing isosteric phosphonate residues have been described in several previous papers (34,35). In particular, Birkus *et al.* (22) described the preparation of oligonucleotides containing 9-(*S*)-[3-hydroxy-2-(phosphonomethoxy)propyl]adenine residues. Employing a similar strategy, we synthesized protected CDV monomer **4** and used it to incorporate CDV at the center of a dodecanucleotide. As shown in Scheme 1, 3'-*O*-dimethoxytritylation of the 4-*N*-benzoyl derivative of cidofovir (**1**) yielded fully protected CDV

derivative 2. Removal of the ethyl ester groups with bromotrimethylsilane followed by esterification of phosphonic acid 3 with 4-methoxy-1-oxido-2-pyridylmethanol afforded the CDV monomer (4). The modified oligonucleotide (dCGCATG-X₇-TACGC), where X₇=CDV, was synthesized from the 5'-to-3'-end using reversed 5' phosphoramidites; CDV monomer 4 was introduced into the oligonucleotide chain using the phosphotriester method. The complementary and dCMP-containing control strands (Fig. IX-1B) were all synthesized using standard chemistry.

We had previously used enzymatic methods to prepare a CDV-containing DNA template and showed that the drug blocks chain extension by vaccinia DNA polymerase (*12*). To show that the chemically synthesized substrate behaved the same way, we compared the properties of the CDV- and dCMP-containing template strands in primer extension assays (see Supporting Information). Each 12-mer was first ligated to a template-extending oligonucleotide in the presence of an IRDye700® labelled primer, to create a longer and more stable primer-template substrate (Fig. D-1A). The assay showed that template-primer pairs composed of the dCMP-containing control dodecamer permitted efficient primer extension to the ends of the template (Fig. D-1B). In contrast, negligible primer extension was detected when the template contained a CDV molecule (Fig. D-1B).

CD Spectroscopy of the CDV DNA Duplex

We next used CD spectroscopy to see whether the substitution of CDV for dCMP would radically perturb the structure of a DNA duplex. Fig. D-2 shows the CD spectra for the two DNA duplexes at 20 °C in ~0.2 M salt, and spanning wavelengths from 205 to 360 nm. The two spectra are quite similar, and exhibit features characteristic of B-DNA. There are, however, some differences in the peak intensities between these two spectra at 255 and 282 nm, suggesting a minor alteration in the structure of the CDV-containing duplex. We also obtained melting profiles of the two duplexes at different wavelengths and spanning

temperatures ranging from 20 °C to 92 °C. Fig. IX-2 shows the change in molar CD at 246 nm. Curves fitted to these and other data (collected at different wavelengths) indicated T_m values of 46 ± 1 °C and 58 ± 1 °C for the CDV- and dCMP-containing DNA duplexes, respectively. The lower melting point of the CDV-containing molecule reflects differences in both the van't Hoff melting enthalpies ($\Delta\Delta H_{\text{CDV-dCMP}} = 3.6$ Kcal/mol) and entropies ($\Delta\Delta S_{\text{CDV-dCMP}} = 13$ cal/mol·deg) under these salt and DNA concentration conditions.

NMR Spectra

The ^1H NMR spectra of the control and CDV DNA duplexes were remarkably similar in all respects – chemical shift dispersion, line widths, intensities, and resolution – implying homologous structures for both duplexes.

Exchangeable Imino Protons

The hydrogen bonded imino protons from the G and T nucleotides observed in the very low field region of the ^1H NMR spectrum are highly sensitive indicators of the Watson-Crick base pairing and stability of DNA duplexes in solution. The imino regions of the 1D ^1H NMR spectra of the control and CDV DNA duplexes are compared in Fig. IX-1C. The imino protons of the six G and four T of the duplexes are clearly observable, with the exception of weak G13 and G24 resonances located at the extremities of the dodecamers. The presence of all imino protons and the strong similarities between the control and CDV duplex spectra indicate normal base stacking for both duplexes. Most resonances have very similar intensities and chemical shifts between both molecules, with the exception of G6, T8 and G18 for the CDV duplex that show approximately half the intensities found in the control. These three resonances are in contact with X7 in the CDV duplex. Moreover, G18, the complementary base to X7, shows by far the largest chemical shift change (0.3 ppm observed between both duplexes at 25 °C). To make sure that the CDV base makes canonical hydrogen bonds to its complimentary base, we acquired a 2D ^1H , ^1H NOESY spectrum of the CDV DNA duplex at 10 °C to identify the inter-strand NOE contacts (Fig. IX-3). Six

imino-imino, ten imino-amino and the four imino-methyl NOEs contacts were identified. Of special interest, two strong NOEs were found between the imino protons of G18 with the amino protons of X7, confirming that X7 is properly hydrogen bonded with G18 in a Watson-Crick manner.

Temperature Series of the CDV and Control DNA Duplexes

1D ^1H NMR spectra of both DNA duplexes were acquired at 5, 10, 15, 20, 25, 30, 35, 40 and 45 °C. The imino regions between 12 and 15.5 ppm are shown in Fig. IX-4A,B. The peak intensities of four resonances of interest (G6, T8, G18 and G15) are plotted as a function of temperature in Fig. IX-4C-F. The peak intensity of a given imino resonance is dependent on multiple complex phenomena including molecular rotational correlation time which influences NMR relaxation, hydrogen bonding strength which influences the rate of chemical exchange of the proton, and conformational changes. For a DNA duplex, it is typical to observe an increase in the NMR signal with temperature as molecular motion increases, and then a decrease as the temperature approaches the melting point of the base pair in question, resulting in a parabolic profile. From Fig. IX-4C-E it is clear that the imino proton resonances of G6, T8 and G18 are less stable in the CDV DNA duplex compared to the control by 10-20 °C, but those of base pairs further away from the inserted drug such as G15 and G22 have a very similar stability (Fig. IX-4F).

Chemical Shift Assignments of the CDV and Incorporation Into DNA

The chemical shift assignment of the proton and carbon resonances of the CDV drug was accomplished using natural abundance ^1H , ^{13}C HSQC and DQF-COSY spectra (Fig. IX-5A,B). The ^1H and ^{13}C chemical shifts of the drug are in good agreement with previously published chemical shifts for CDV-phosphocholine extracted from rat kidney (36). Of particular interest are the resonances of C4', H4' and H4''. Since ^{31}P decoupling was not used during the acquisition of the HSQC spectra, the $^{13}\text{C}4'$ nucleus showed a splitting of 153 Hz characteristic of a ^{13}C - ^{31}P coupling constant ($^1J_{13\text{C}-31\text{P}}$). A chemical shift difference of 9.5 Hz

between the two equivalent protons H4' and H4'' was also observed in the HSQC spectrum. This results from the coupling constant between the H4 protons with the ^{31}P nucleus ($^2J_{\text{1H-31P}}$), as confirmed in the 1D ^1H NMR spectrum. This E.COSY like pattern (37) depends upon the relative signs of the coupling constants involved and leads to an apparent offset in the chemical shift of the coupled protons. The recognition of this particular J-coupling pattern was essential to the assignment of the CDV resonances in the DNA duplex (Fig. IX-5C,D). Unlike the situation for the CDV drug, the protons H4' and H4'' are non-equivalent in the duplex. This was confirmed by the NOE contacts found in the 2D NOESY spectrum of the duplex. These experiments were necessary in order to carefully assign the X7 resonances, particularly because protons H1'', H4'' and H3' all possess very similar chemical shifts, making the assignment of the NOE contacts more prone to error. The natural abundance $^1\text{H},^{13}\text{C}$ HSQC and DQF-COSY spectra of the control DNA duplex are presented in Fig. IX-5E,F for comparison.

Sequential Assignment of the Non-exchangeable Protons in the Control and CDV DNA Duplexes.

Sequential assignment of the non-exchangeable base protons (purine H8 and pyrimidine H6) was performed using through-space connectivities with the H1'-ribose protons in the 2D ^1H - ^1H NOESY spectra. The sequential assignment for the control and CDV DNA duplexes are shown in Fig. IX-6. The connectivities of the first strand (nucleotides 1-12) and the second strand (nucleotides 13-24) are identified with solid and dashed lines, respectively. In Fig. IX-6, the strong intra-residue NOEs between cytosines H5 and H6 protons are labeled with a superscript #, and the base H8 protons with the base H5 protons of the previous residue are grouped together in the grey areas. All expected H6/H8-base and H1'-ribose connectivities were found in the NOESY spectra. Strong NOEs were found for the H6 protons of X7 and T8, making connections with the H1' and H1'' protons of X7 (see below for a more detailed description the NOEs found for X7). The H2' and H3' sugar protons were subsequently assigned based on these assignments.

Solution Structures of the Control and CDV DNA Duplexes

In order to measure the effect of the incorporation of the CDV drug on the structure of the DNA duplex, we performed restrained molecular dynamic simulations on both the control and CDV DNA duplexes (see Experimental Section for details). Distance restraints were obtained by calibration of the NOEs using MARDIGRAS (30). A detailed visualization of the NOE contacts involving the CDV incorporated into DNA is shown in Fig. D-3. The position of X7 in the DNA duplex is defined by 18 experimental NOEs.

The sugar pucker pseudo-rotation angles were determined by measuring the $^3J_{\text{H1H2}'}$ and $^3J_{\text{H1H2}''}$ coupling constants to estimate the percentage of C2'- and C3'-endo populations (see Experimental Section) (38). Backbone dihedral restraints based on the Dickerson-Drew B-DNA structure (32) and standard Watson-Crick inter-strand distance restraints were added to keep the two strands together during the simulated annealing. The molecular dynamics simulations were repeated 50 times for each duplex, and the 10 structures with the lowest root mean square deviation (RMSD) were kept for analysis. The NMR ensembles of the control and CDV DNA duplexes containing the 10 lowest RMSD structures are presented in Fig. IX-7A and 7B, respectively. The superimposition of the representative structure of the control (green) and CDV (blue) DNA duplexes is shown in Fig. IX-7C. The RMSD between both structures is 1.5 Å over all common atoms, confirming a similar overall B-DNA structure. The comparison of the helical parameters of the control and CDV DNA duplexes is presented in Fig. IX-8.

In order to validate the structure determination protocol, a back-calculation of the NOESY spectrum based on the NMR structure was performed using a complete relaxation matrix analysis method (CORMA (39)) using a 80 ms mixing time and a theoretical isotropic correlation time of 2 ns. The high-level of similarity between the back-calculated and the experimental NOESY spectra of the control DNA duplex presented in Fig. D-4 confirm that the structures determined here

well represent the experimental data. Moreover, structure calculations were performed with and without the presence of the dihedral restraints, and with and without the presence of the NOEs to evaluate the effect of the respective restraints on the ensembles. A typical B-DNA duplex ensemble (but with higher RMSD) was obtained without the use of the dihedral restraints, while a fully distorted structure was obtained without the NOE restraints. No dihedral restraints were used in any calculation for X7. No NOE violations over 0.3 Å were observed in any structure, and only one NOE over 0.2 Å was violated in more than one structure. The average total distance and torsion penalties obtained for the ensembles were 7.92 and 0.92 for the control DNA duplex, compared to 3.35 and 1.33 for the CDV DNA duplex.

DISCUSSION

The structure of an asymmetric DNA duplex containing a single molecule of the antiviral agent CDV has been compared to a DNA duplex containing dCMP at the equivalent site. The CDV-containing oligonucleotide was prepared using a reverse synthesis method (22) and permitted the efficient synthesis of milligram quantities of DNA without constraining our choice of flanking sequence. We subsequently tested what effects a CDV molecule, located in the template strand, would have on reactions catalyzed by vaccinia virus DNA polymerase and observed that it blocked DNA synthesis across the drug lesion (Fig. D-1B). Thus these fully synthetic substrates behave just like the enzymatically synthesized oligonucleotides that we have used previously (12). We then used CD and two-dimensional NMR spectroscopy to examine why CDV-containing DNA can so negatively affect *trans*-lesion synthesis catalyzed by vaccinia DNA polymerase.

The CDV and control DNA duplexes exhibited CD spectra characteristic of B-DNA under near physiological salt and temperature conditions (Fig. D-2) and differed only slightly in the intensities of the absorbance peaks at 255 and 282 nm. This suggested some minor differences in base stacking between the two

duplexes. CD methods have previously been used to examine the structures of DNAs containing the nucleoside analogs ganciclovir (40) and cyclohexenyladenine (41) and with similar results. For example, Marshalko *et al.* (40) studied a self-complementary 10-mer containing two molecules of the antiviral drug ganciclovir. Ganciclovir is an acyclic nucleoside analog possessing an unmodified guanine base and these authors showed that the ganciclovir and control DNA duplexes produced similar CD spectra characteristic of B-DNA (40). We concluded that CDV causes only minor perturbations of the overall double-stranded DNA structure.

CD analysis was also used to determine the melting properties of the two oligonucleotides. The data indicated that the incorporation of a single molecule of CDV into this duplex lowers the overall T_m by ~ 12 °C in 0.2 M salt (46 ± 1 °C and 58 ± 1 °C for CDV and dCMP duplexes, respectively). The 1D ^1H NMR temperature melting series showed that the instability is localized to the region around the CDV molecule. The nucleotides surrounding CDV (G6, T8 and G18) are less stable by 10 to 14 °C compared to those surrounding a dCMP control (Fig. IX-4C-E), whereas nucleotides located further away show melting profiles similar to that seen in the control duplex (Fig. IX-4F). Similar effects have been reported for other forms of modified DNAs. For example the two molecules of ganciclovir noted above lowered the T_m by 13 °C relative to a control 10-mer duplex (40) and a glycerol-based acyclic nucleoside linked to a thymine base reduced the T_m of a 9-mer DNA by 15 °C (42). Interestingly the 12-mer DNA containing a normal furanose sugar replaced by a cyclohexenyl ring exhibited a T_m only 0.4 °C different from the dAMP-containing control.(41) These data show that modified nucleotides can be readily incorporated into B-DNA structures, but that as Marshalko *et al.* have suggested (40), a cyclic sugar moiety may provide an additional important stabilizing element.

The solution structures of the CDV and control DNA duplexes confirmed the B-DNA configuration suggested by the CD spectra, and the similarities between the

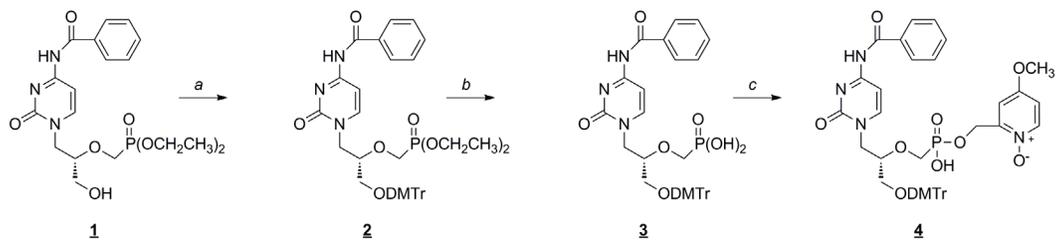
two structures are illustrated by an RMSD of only 1.5 Å over all common atoms between both duplexes. The ready accommodation of CDV into B-DNA contrasts with the structure of a DNA containing ganciclovir (43). Although the guanine base of ganciclovir can form a normal Watson-Crick pair with a cytosine on the opposite strand, the ganciclovir-containing DNA duplex still exhibits a kink in the sugar-phosphate backbone on the 3' side of the drug. Closer inspection of the region surrounding the CDV molecule in our structure showed only a small conformational change in the position of the cytosine base of CDV (X7) relative to that of dCMP (C7) (Fig. IX-9). This is associated with a significant alteration in the base pair stagger, rise, and tilt at C7/X7 although the deviation does not markedly exceed the natural variation in these parameters that can be observed elsewhere in the DNA (Fig. IX-8). The next nucleotide, T8, superimposes well in both the CDV and control DNA duplex structures although again one does see some changes in the helical parameters (Fig. IX-8). One also still sees some alteration in the positioning of the 5'-phosphonate and 3'-phosphates flanking the CDV molecule. The movement of the phosphorus in the 5'-phosphonate linkage, relative to its phosphate homolog, may cause some perturbation of the G6 deoxyribose in the CDV-containing structure and is reflected in a change in rise (Fig. IX-8). One also sees displacement of the phosphorus linking X7 and T8, although this is accommodated within the helix without further perturbing the T8 deoxyribose. We acquired ³¹P NMR spectra (data not shown), but the individual phosphates were not sufficiently resolved for either DNA to be assigned with the exception of the phosphonate from the CDV. That said, slightly larger ³¹P chemical shift dispersion was observed for the CDV DNA duplex.

Vaccinia virus DNA polymerase has been shown to be the target of the anti-poxvirus activity of CDV based on analyses of mutants resistant to the drug (13,20,21); our lab has also shown that the incorporation of CDV into DNA affects both the 3'-to-5' exonuclease and 5'-to-3' polymerase activities of the enzyme (11,12). The alterations seen in the CDV-containing structure provide some insights into how this nucleotide analog affects the activities of the DNA

polymerase although with the *caveat* that we are extrapolating from a duplex structure to enzyme-bound and partially single-stranded DNA substrates. For example, the 3'-to-5' exonuclease activity cannot attack the phosphodiester bond linking CDV to another 3' nucleotide (the CDV+1 product) (11). Based upon the conservation of sequence, this reaction will presumably depend upon a metal ion coordinated nucleophilic attack (44) on the phosphorus linking the CDV to the 3'-terminal dNMP (i.e. between X7 and T8). The displacement of this atom, perhaps further exacerbated by an altered configuration around the 5'-phosphonate linkage in a displaced single-strand of DNA (Fig. IX-9), could severely interfere with the reaction geometry within the exonuclease active site. The same arguments could be marshaled to explain how these structures partially inhibit DNA polymerization from a CDV+1 terminated primer (11) or completely block synthesis across a CDV in the template strand (12). Although the position of the T8 3'-oxygen can be nearly superimposed in the two stacked structures, the several small distortions we see might be exacerbated in the absence of a fully duplex structure and disturb the positioning of the 3'-hydroxyl residue. This would disfavor 3'-chain extension from a CDV+1 terminated primer. The block to synthesis across a CDV residue is nearly absolute (Fig. D-1) and suggests that a CDV molecule in the template strand creates much greater problems for virus polymerases. Two features of the DNA could account for this effect. For example, any enhancement of the perturbations we see in the rise (Fig. IX-8) would interfere with H-bonding of an incoming nucleotide to the G6 residue. Furthermore, the altered positions of the phosphorus atoms flanking the CDV molecule (Fig. IX-9) might interfere with movement within the polymerase's DNA binding site, much as a nut cannot easily move on a damaged bolt. Both effects would inhibit polymerization across a CDV residue.

Although one can select viruses encoding polymerase mutations that independently promote enhanced drug excision (45) and probably inhibit CDVpp binding (13), even double-mutant viruses exhibit only a 15-fold increase in CDV resistance. High-level resistance has proven impossible to obtain without severe

effects on virulence. This may reflect the fact that no simple evolutionary path exists for producing DNA polymerases that can accommodate these insidious forms of drug-damaged DNA templates, and illustrates the promising clinical value of these non-chain-terminating antivirals.

FIGURES

Scheme IX-1. Synthesis of the monomer used to prepare a CDV-containing oligonucleotide. Reagents: a) dimethoxytrityl chloride, 4-dimethylaminopyridine, pyridine; b) 1) bromotrimethylsilane, acetonitrile; 2) H₂O; c) 4-methoxy-1-oxido-2-pyridylmethanol, N,N-dicyclohexylcarbodiimide, pyridine.

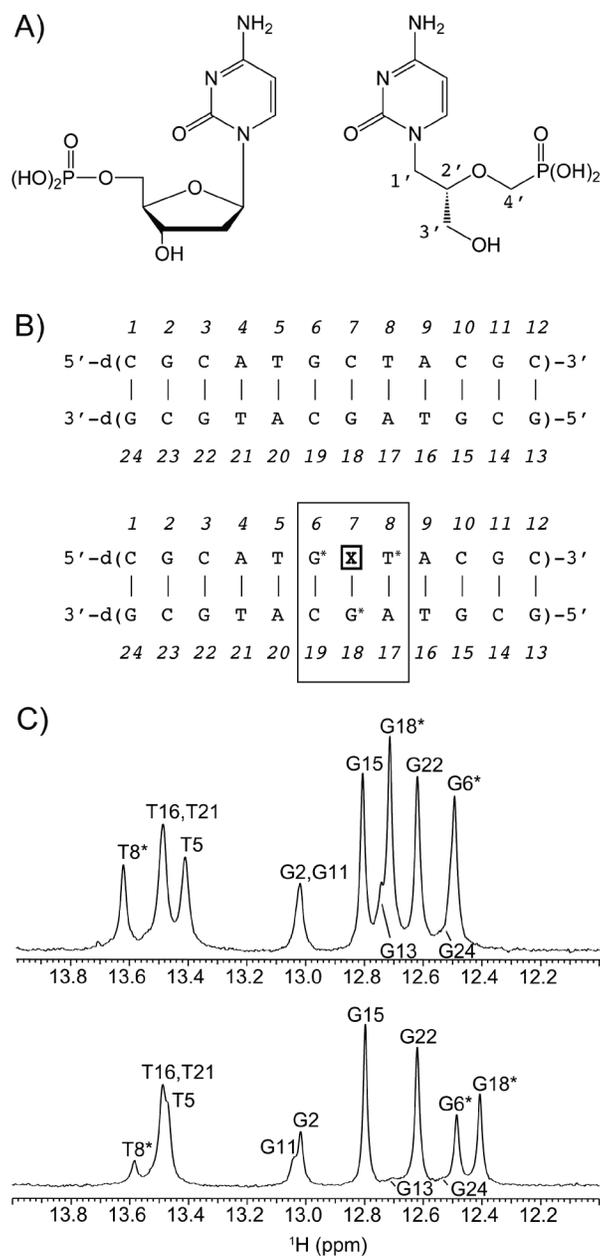


Fig. IX-1. A) Structures of deoxycytidine monophosphate (left) and CDV (right). B) Sequences of the control DNA duplex (top) and the CDV DNA duplex (bottom), showing insertion at position 7 (X7). C) 1D ^1H NMR spectra of the imino proton regions acquired at 25°C on a 600 MHz NMR spectrometer for the control DNA duplex (top) and CDV DNA duplex (bottom). The most perturbed imino protons G6, T8 and G18 are marked with an asterisk. The largest chemical shift change is for G18, the complementary base to CDV in the dodecamer.

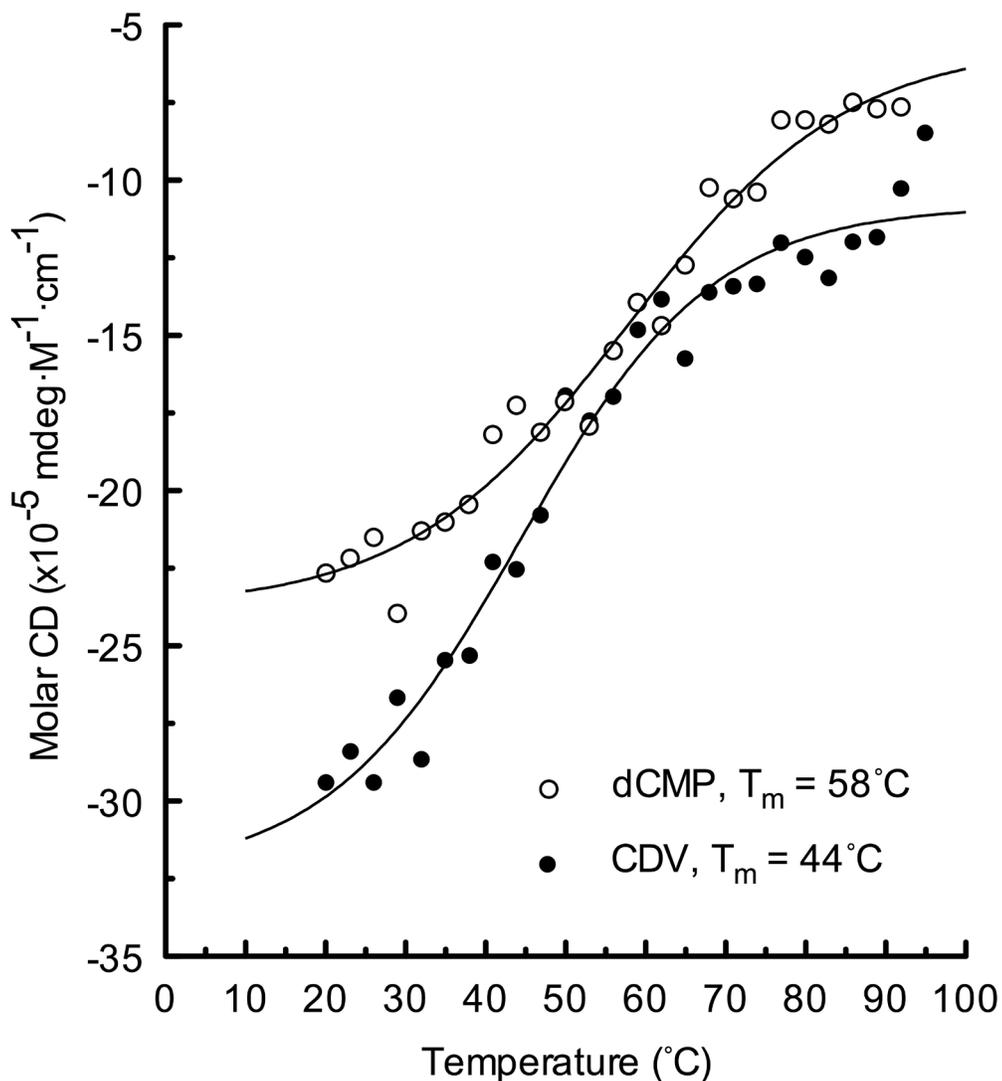


Fig. IX-2. Melting profiles for CDV- and dCMP-containing duplexes. Each sample was dissolved in 10 mM sodium phosphate (pH 7.3), 185 mM sodium chloride and CD spectra were acquired at each of the indicated temperatures. The Fig. IX-shows the CD signal recorded at 246 nm for each duplex and the curve fits suggested T_m values of 58° C and 44° C for dCMP and CDV-containing duplexes, respectively. Nearly identical melting points were determined over a spectral range spanning 245-255 nm ($58 \pm 1^{\circ}$ C and $46 \pm 1^{\circ}$ C [s.e.m.] for dCMP and CDV-containing duplexes, respectively).

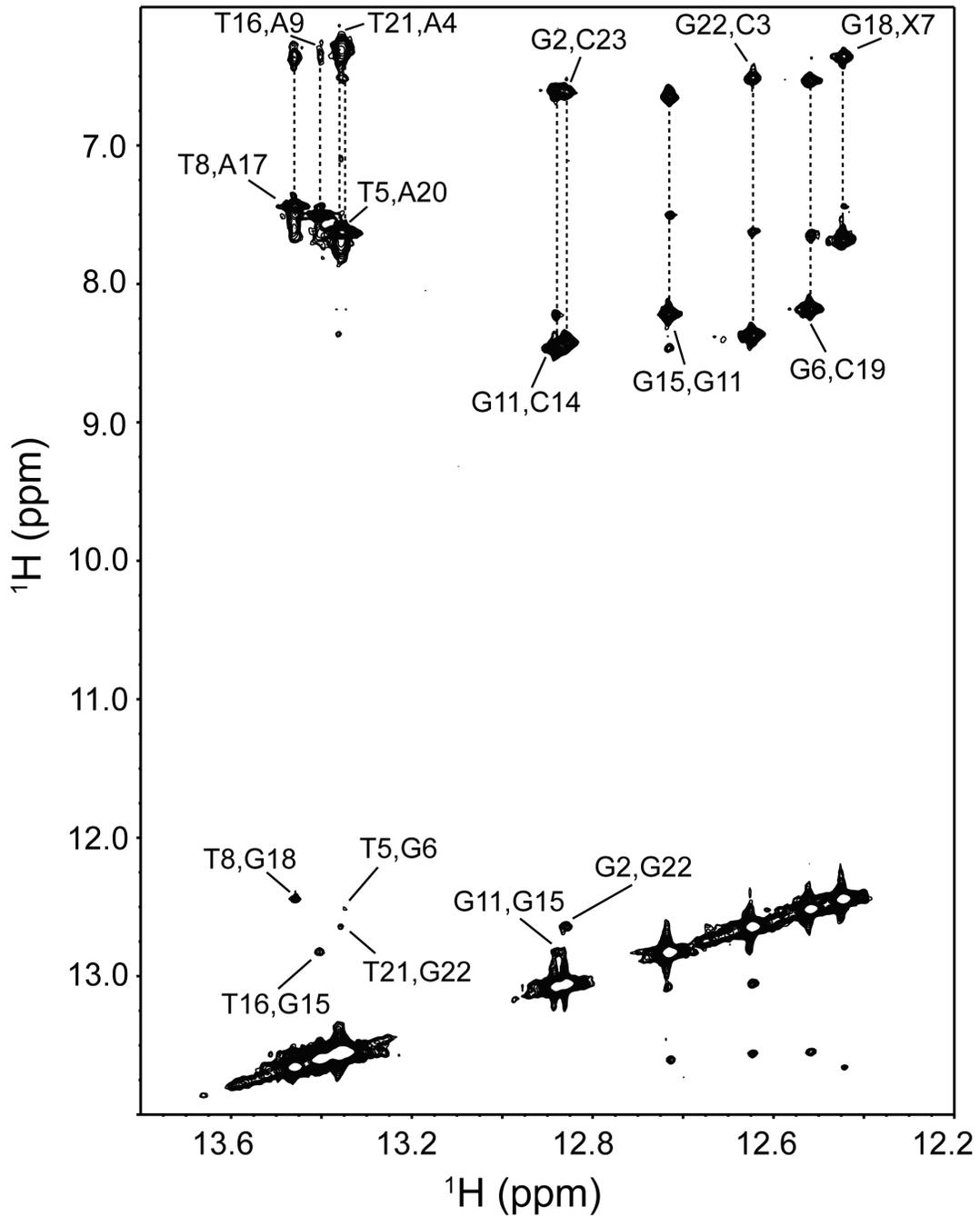


Fig. IX-3. 2D ^1H - ^1H NOESY spectrum of the imino proton regions of the CDV DNA duplex at 10°C. Assigned imino-imino and imino-amino contacts are indicated, including G18,X7 contacts.

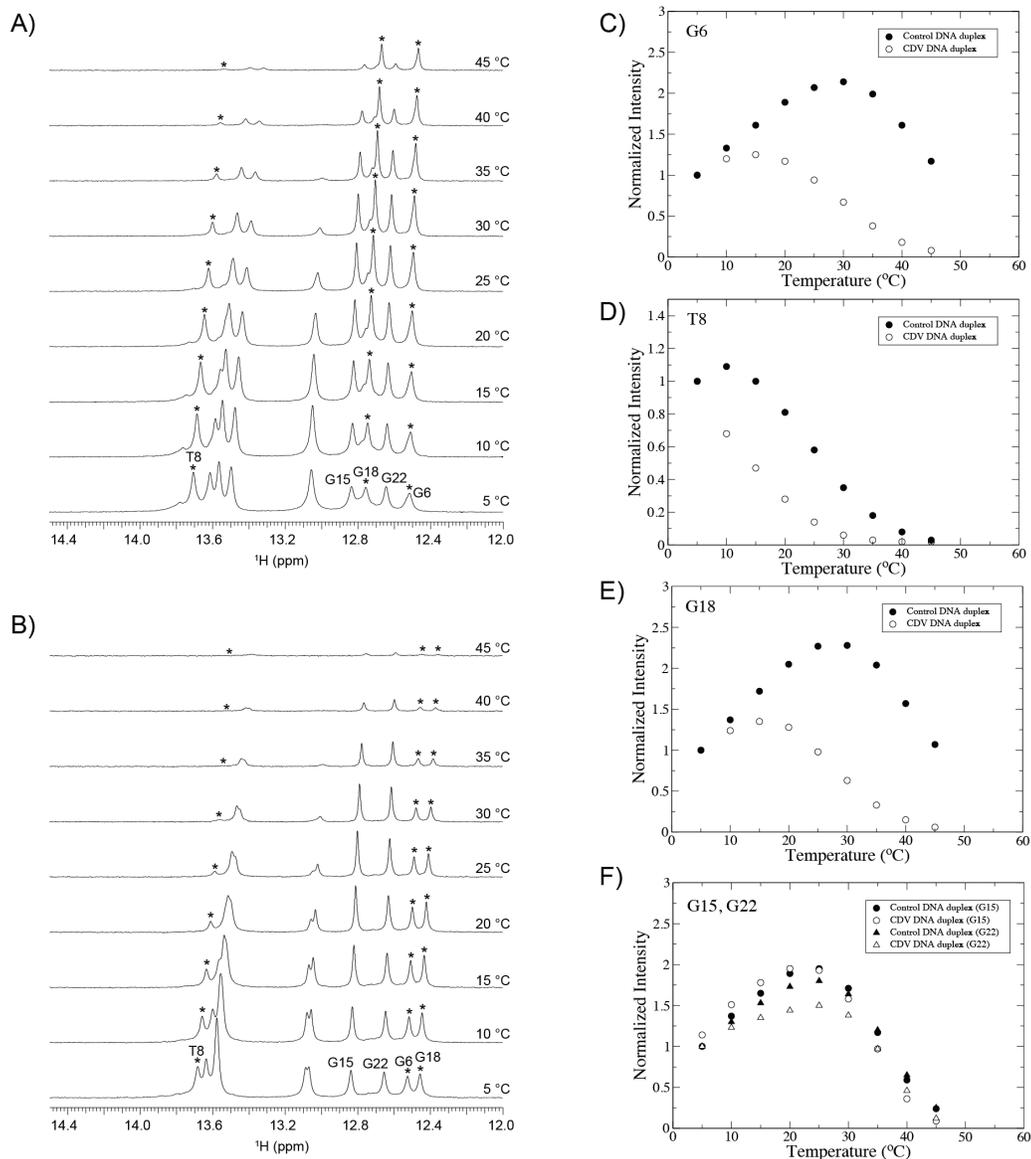


Fig. IX-4. Temperature series monitored by 1D ^1H NMR spectra. A) Control DNA duplex. B) CDV DNA duplex. As reflected by the plots of peak intensities of the imino protons of G6, T8 and G18 as a function of temperature (C-E), the CDV DNA duplex shows a decrease in the local stability of the nucleotides surrounding the modified base. The imino protons from bases farther away, like G15 and G22 (F), show very similar melting profiles for both duplexes.

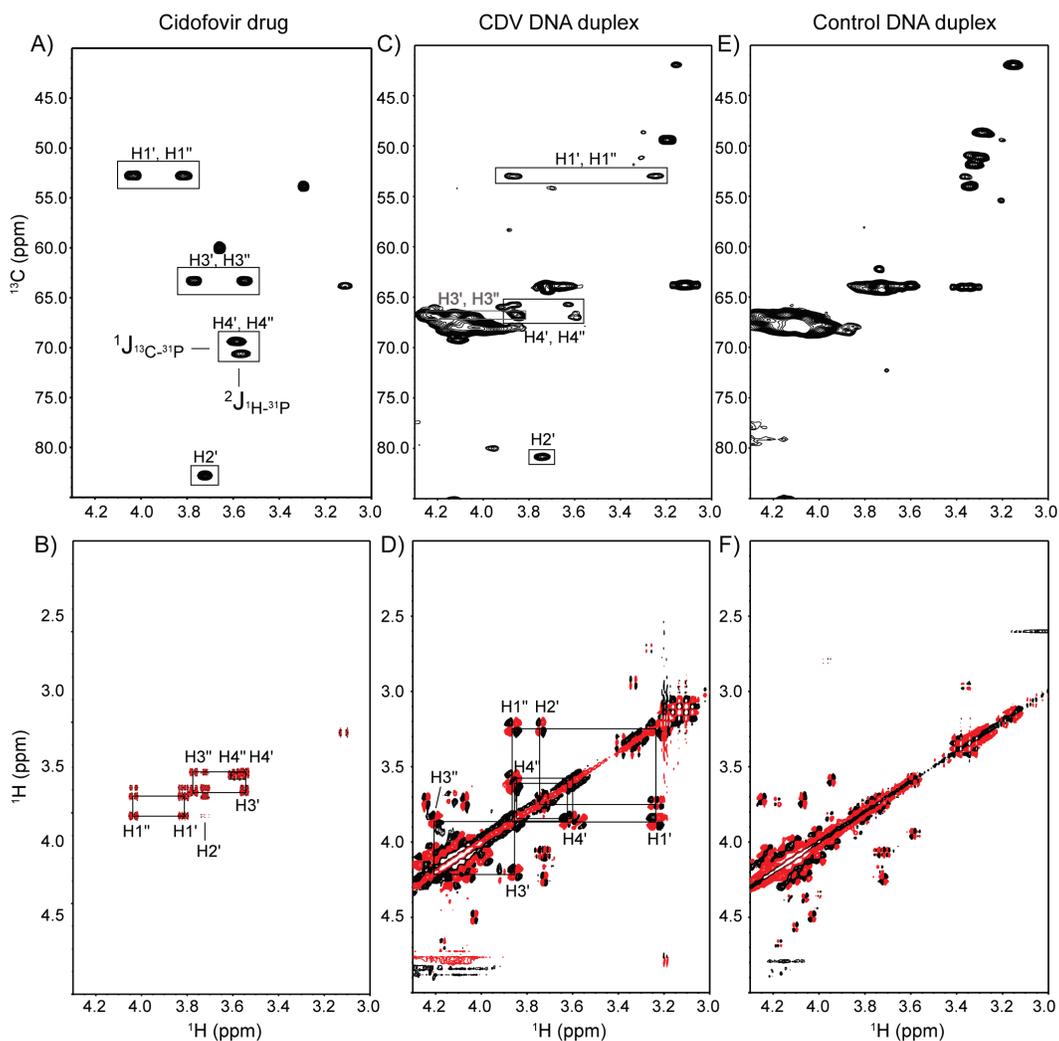


Fig. IX-5. Natural abundance ^1H - ^{13}C HSQC (top) and DQF-COSY (bottom) showing the chemical shift assignments of Cidofovir. A-B) CDV drug. C-D) CDV DNA duplex. E-F) Control DNA duplex. The $\text{H4}'$, $\text{H4}''$ protons of the phosphonate show an E.COSY like pattern reflecting coupling to the ^{31}P . The ^1H - ^{13}C HSQC and DQF-COSY were acquired at 500 and 600 MHz, respectively.

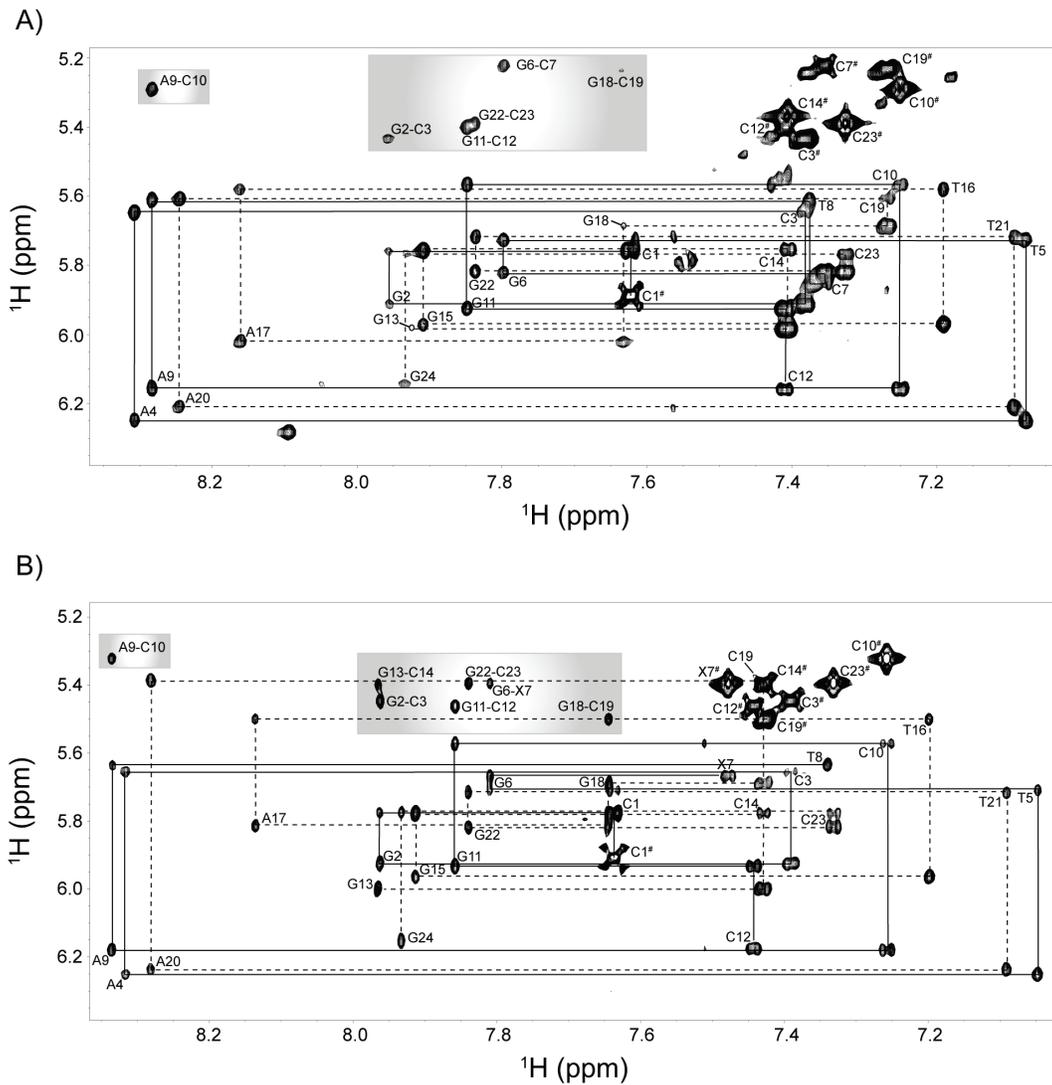


Fig. IX-6. 2D ^1H - ^1H NOESY NMR spectra showing the sequential assignment of the base H6/H8 protons with the H1'-ribose protons. A) Control DNA duplex. B) CDV DNA duplex. Solid and dashed lines indicate the assignments of both strands. The H8,H5 NOEs are located in the grey areas, while the H6,H5 NOEs are marked with #.

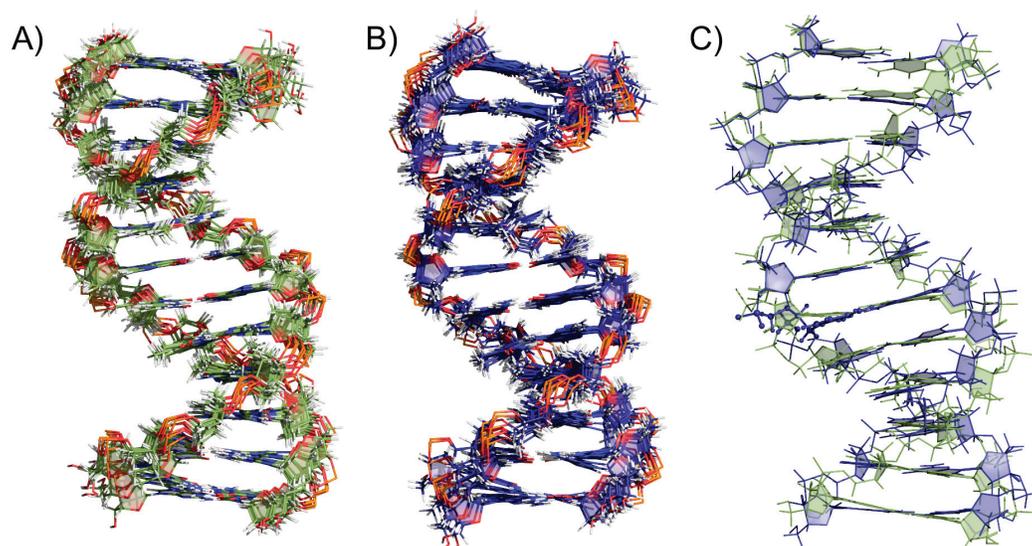


Fig. IX-7. NMR structures of DNA duplexes A) Control DNA duplex: ensemble of 10 lowest RMSD structures. B) CDV DNA duplex: ensemble of 10 lowest RMSD structures. C) Superimposition of the control (green) and CDV (blue) DNA duplexes. The RMSD between both structures is 1.5 Å over all common atoms (PDB accession codes 2L8P and 2L8Q).

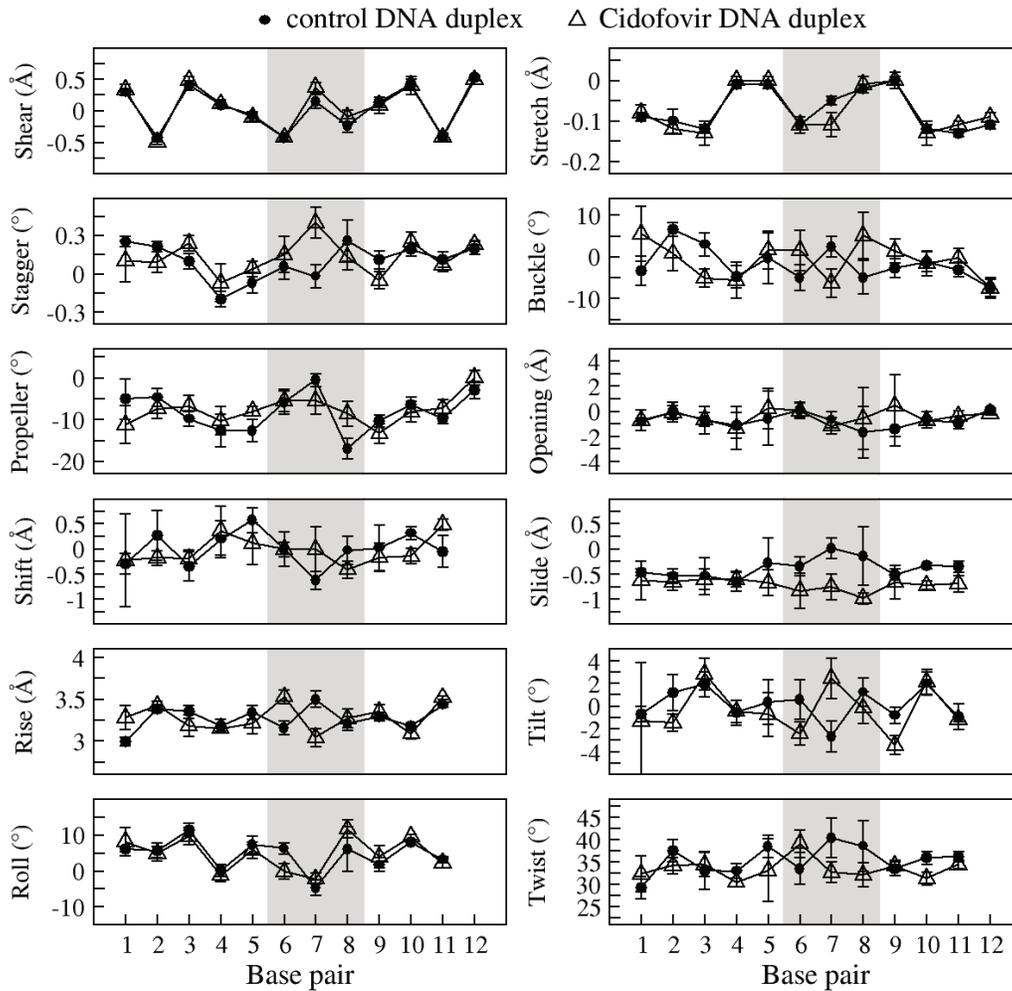


Fig. IX-8. Comparison of the helical parameters for the ensemble of 10 structures obtained for the control DNA duplex (●) and the CDV DNA duplex (Δ). The error bars correspond to the standard deviations of the ensembles. The parameters were calculated using the program X3DNA.³³

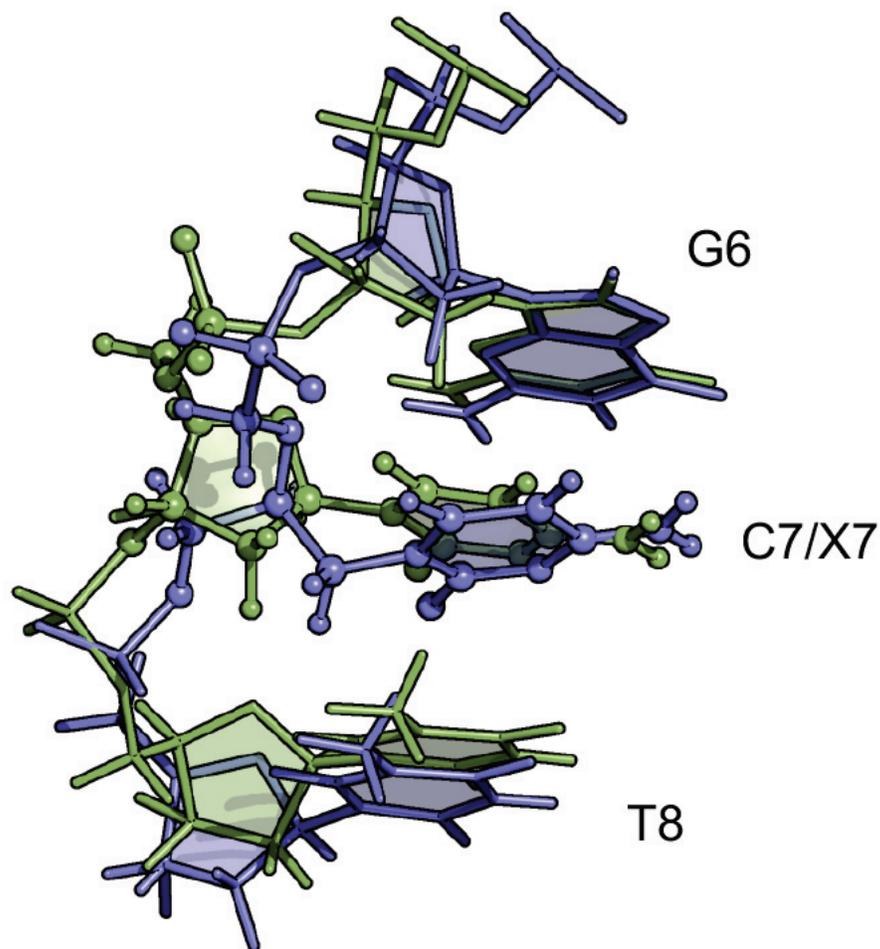


Fig. IX-9. Structural comparison of C7 (green) and X7 (blue) in the middle of the DNA duplexes. The nucleobases C7 and X7 superimposed relatively well on each other with a RMSD of 0.88 Å for the heavy atoms of the ring. The absence of the ribose ring for X7 leads to a small conformational change in the positioning of the base of X7. The ring of X7 is positioned with a tilt between G6 and T8. Base pairs G6-C19, X7/C7-G18 and T8-A17 were superimposed to obtain this alignment.

REFERENCES

1. De Clercq, E., and Holý, A. (2005) Acyclic nucleoside phosphonates: a key class of antiviral drugs, *Nat. Rev. Drug. Discov.* 4, 928-940.
2. Wachsman, M., Petty, B., Cundy, K., Jaffe, H., Fisher, P., Pastelak, A., and Lietman, P. (1996) Pharmacokinetics, safety and bioavailability of HPMPC (cidofovir) in human immunodeficiency virus-infected subjects, *Antiviral Res.* 29, 153-161.
3. Lalezari, J. P., Drew, W. L., Glutzer, E., James, C., Miner, D., Flaherty, J., Fisher, P. E., Cundy, K., Hannigan, J., Martin, J. C., and Jaffe, H. S. (1995) (S)-1-[3-hydroxy-2-(phosphonylmethoxy)propyl]cytosine (Cidofovir): results of a phase I/II study of a novel antiviral nucleotide analogue, *J. Infect. Dis.* 171, 788-796.
4. Ciesla, S. L., Trahan, J., Wan, W., Beadle, J., Aldern, K., Painter, G., and Hostetler, K. (2003) Esterification of cidofovir with alkoxyalkanols increases oral bioavailability and diminishes drug accumulation in kidney, *Antiviral Res.* 59, 163-171.
5. Kern, E., Hartline, C., Harden, E., Keith, K., Rodriguez, N., Beadle, J., and Hostetler, K. (2002) Enhanced inhibition of orthopoxvirus replication in vitro by alkoxyalkyl esters of cidofovir and cyclic cidofovir, *Antimicrob. Agents Chemother.* 46, 991-995.
6. Quenelle, D., Collins, D., Wan, W., Beadle, J., Hostetler, K., and Kern, E. (2004) Oral treatment of cowpox and vaccinia virus infections in mice with ether lipid esters of cidofovir, *Antimicrob. Agents Chemother.* 48, 404-412.
7. Buller, R. M., Owens, G., Schriewer, J., Melman, L., Beadle, J., and Hostetler, K. (2004) Efficacy of oral active ether lipid analogs of cidofovir in a lethal mousepox model, *Virology* 318, 474-481.
8. Beadle, J., Hartline, C., Aldern, K., Rodriguez, N., Harden, E., Kern, E., and Hostetler, K. (2002) Alkoxyalkyl esters of cidofovir and cyclic cidofovir exhibit multiple-log enhancement of antiviral activity against

- cytomegalovirus and herpesvirus replication in vitro, *Antimicrob. Agents Chemother.* 46, 2381-2386.
9. Toth, K., Spencer, J., Dhar, D., Sagartz, J., Buller, R., Painter, G., and Wold, W. (2008) Hexadecyloxypropyl-cidofovir, CMX001, prevents adenovirus-induced mortality in a permissive, immunosuppressed animal model, *Proc. Natl. Acad. Sci. U.S.A.* 105, 7293-7297.
 10. Painter, G., and Hostetler, K. (2004) Design and development of oral drugs for the prophylaxis and treatment of smallpox infection, *Trends Biotechnol.* 22, 423-427.
 11. Magee, W., Hostetler, K., and Evans, D. (2005) Mechanism of inhibition of vaccinia virus DNA polymerase by cidofovir diphosphate, *Antimicrob. Agents Chemother.* 49, 3153-3162.
 12. Magee, W., Aldern, K., Hostetler, K., and Evans, D. (2008) Cidofovir and (S)-9-[3-hydroxy-(2-phosphonomethoxy)propyl]adenine are highly effective inhibitors of vaccinia virus DNA polymerase when incorporated into the template strand, *Antimicrob. Agents Chemother.* 52, 586-597.
 13. Andrei, G., Gammon, D., Fiten, P., De Clercq, E., Opdenakker, G., Snoeck, R., and Evans, D. (2006) Cidofovir resistance in vaccinia virus is linked to diminished virulence in mice, *J. Virol.* 80, 9391-9401.
 14. Gammon, D., Snoeck, R., Fiten, P., Krecmerova, M., Holy, A., De Clercq, E., Opdenakker, G., Evans, D., and Andrei, G. (2008) Mechanism of antiviral drug resistance of vaccinia virus: identification of residues in the viral DNA polymerase conferring differential resistance to antipoxvirus drugs, *J. Virol.* 82, 12520-12534.
 15. Connelly, M., Robbins, B., and Fridland, A. (1993) Mechanism of uptake of the phosphonate analog (S)-1-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine (HPMPC) in Vero cells, *Biochem. Pharmacol.* 46, 1053-1057.
 16. Cihlar, T., and Chen, M. (1996) Identification of enzymes catalyzing two-step phosphorylation of cidofovir and the effect of cytomegalovirus infection on their activities in host cells, *Mol. Pharmacol.* 50, 1502-1510.

17. Aldern, K. A., Ciesla, S., Winegarden, K., and Hostetler, K. (2003) Increased antiviral activity of 1-O-hexadecyloxypropyl-[2-(14)C]cidofovir in MRC-5 human lung fibroblasts is explained by unique cellular uptake and metabolism, *Mol. Pharmacol.* 63, 678-681.
18. Jesus, D., Costa, L., Goncalves, D., Achete, C., Attias, M., Moussatche, N., and Damaso, C. (2009) Cidofovir inhibits genome encapsidation and affects morphogenesis during the replication of vaccinia virus, *J. Virol.* 83, 11477-11490.
19. Beadle, J. R., Wan, W., Ciesla, S., Keith, K., Hartline, C., Kern, E., and Hostetler, K. (2006) Synthesis and antiviral evaluation of alkoxyalkyl derivatives of 9-(S)-(3-hydroxy-2-phosphonomethoxypropyl)adenine against cytomegalovirus and orthopoxviruses, *J. Med. Chem.* 49, 2010-2015.
20. Kornbluth, R., Smee, D., Sidwell, R., Snarsky, V., Evans, D., and Hostetler, K. (2006) Mutations in the E9L polymerase gene of cidofovir-resistant vaccinia virus strain WR are associated with the drug resistance phenotype, *Antimicrob. Agents Chemother.* 50, 4038-4043.
21. Becker, M., Obraztsova, M., Kern, E., Quenelle, D., Keith, K., Prichard, M., Luo, M., and Moyer, R. (2008) Isolation and characterization of cidofovir resistant vaccinia viruses, *Virol. J.* 5, 58.
22. Birkus, G., Rejman, D., Otmar, M., Votruba, I., Rosenberg, I., and Holy, A. (2004) The substrate activity of (S)-9-[3-hydroxy-(2-phosphonomethoxy)propyl]adenine diphosphate toward DNA polymerases alpha, delta and epsilon, *Antivir. Chem. Chemother.* 15, 23-33.
23. Brodfuehrer, P., Howell, H. G., Sapino, C., and Vemishetti, P. (1994) A practical synthesis of (S)-HPMPC, *Tetrahedron Lett.* 35, 3243-3246.
24. Rejman, D., Erbs, J., and Rosenberg, I. (2000) Large-scale synthesis of a key catalytic reagent for phosphorus protection in building blocks for isopolar phosphonate oligonucleotide preparation, *Org. Proc. Res. Dev.* 4, 473-476.

25. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR* 6, 277-293.
26. Macke, T., and Case, D. A. (1998) Modeling unusual nucleic acid structures. In *Molecular Modeling of Nucleic Acids*, pp 379-393, N.B. Leontes and J. SantaLucia, Jr., eds., Washington, DC.
27. Wang, J., Wolf, R., Caldwell, J., Kollman, P., and Case, D. (2004) Development and testing of a general amber force field, *J. Comput. Chem.* 25, 1157-1174.
28. Symerský, J., and Holý, A. (1991) Structure of 1-(S)-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine; an antiviral agent, *Acta Crystallogr C* 47, 2104-2107.
29. Case, D., Cheatham, T. R., Darden, T., Gohlke, H., Luo, R., Merz, K. J., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. (2005) The Amber biomolecular simulation programs, *J. Comput. Chem.* 26, 1668-1688.
30. Borgias, B., and James, T. (1989) Two-dimensional nuclear Overhauser effect: complete relaxation matrix analysis, *Methods Enzymol.* 176, 169-183.
31. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, *J. Phys. Chem.* 100, 19824-19839.
32. Drew, H. R., Wing, R., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R. (1981) Structure of a B-DNA dodecamer: conformation and dynamics, *Proc. Natl. Acad. Sci. U.S.A.* 78, 2179-2183.
33. Lu, X. J., and Olson, W. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures, *Nucleic Acids Res.* 31, 5108-5121.
34. Točík, Z., Barvik, I., Buděšínský, M., and Rosenberg, I. (2006) Novel isosteric, isopolar phosphonate analogs of oligonucleotides: preparation and properties, *Biopolymers* 83, 400-413.

35. Pav, O., Protivinska, E., Pressova, M., Collinsova, M., Jiracek, J., Snasel, J., Masojidkova, M., Budesinsky, M., and Rosenberg, I. (2006) Activation of murine RNase L by isopolar 2'-phosphonate analogues of 2',5' oligoadenylates, *J. Med. Chem.* 49, 3955-3962.
36. Eisenberg, E., Lynch, G., Bidgood, A., Krishnamurty, K., and Cundy, K. (1998) Isolation and identification of a metabolite of cidofovir from rat kidney, *J. Pharm. Biomed. Anal.* 16, 1349-1356.
37. Griesinger, C., Sørensen, O. W., and Ernst, R. R. (1985) Two-dimensional correlation of connected NMR transitions, *J. Am. Chem. Soc.* 107, 6394-6396.
38. Rinkel, L. J., and Altona, C. (1987) Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants: a graphical method, *J. Biomol. Struct. Dyn.* 4, 621-649.
39. Keepers, J. W., and James, T. L. (1984) A theoretical-study of distance determinations from NMR - Two-dimensional nuclear overhauser effect spectra, *J. Magn. Reson.* 57, 404-426.
40. Marshalko, S., Schweitzer, B., and Beardsley, G. (1995) Chiral chemical synthesis of DNA containing (S)-9-(1,3-dihydroxy-2-propoxymethyl)guanine (DHPG) and effects on thermal stability, duplex structure, and thermodynamics of duplex formation, *Biochemistry* 34, 9235-9248.
41. Wang, J., Verbeure, B., Luyten, I., Lescrinier, E., Froeyen, M., Hendrix, C., Rosemeyer, H., Seela, F., Van Aerschot, A., and Herdewijn, P. (2000) Cyclohexene nucleic acids (CeNA): Serum stable oligonucleotides that activate RNase H and increase duplex stability with complementary RNA, *J. Am. Chem. Soc.* 122, 8595-8602.
42. Schneider, K. C., and Benner, S. A. (1990) Oligonucleotides containing flexible nucleoside analogues, *J. Am. Chem. Soc.* 112, 453-455.
43. Foti, M., Marshalko, S., Schurter, E., Kumar, S., Beardsley, G., and Schweitzer, B. (1997) Solution structure of a DNA decamer containing the antiviral drug ganciclovir: combined use of NMR, restrained molecular

-
- dynamics, and full relaxation matrix refinement, *Biochemistry* 36, 5336-5345.
44. Brautigam, C., and Steitz, T. (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes, *Curr. Opin. Struct. Biol.* 8, 54-63.
 45. Gammon, D., and Evans, D. (2009) The 3'-to-5' exonuclease activity of vaccinia virus DNA polymerase is essential and plays a role in promoting virus genetic recombination, *J. Virol.* 83, 4236-4250.

CHAPTER X - Conclusions

After Six Years of Structural Biology, Where Am I?

GENERAL DISCUSSION AND CONCLUSIONS

During my Ph.D., my supervisor must have realised early on that in order for him to keep me away from Avenue Pizza, he had to keep me busy with more than one project at a given time. This could explain in part why my thesis contains a wide variety of research topics. The question I address herein is in the end, what did the work presented in this thesis bring to the different areas of research that I studied? The importance of each independent chapter is relative, depending on the criteria considered; my contribution to the work presented? The impact factor of the journal where the work has been published? The number of times the article is or will be cited? The kind of response received by scientific peers? The quality of the research? The elegance of the written document? I would argue that it is probably a mixture of all these criteria, although I strongly disagree with arguments supporting the value of the impact factor.

The first paper presented in Chapter II (*I*), is the result of a long-time collaboration between the Sykes laboratory at the University of Alberta, and the laboratories of Drs. Trentham, Irving and Corrie at the NIMR or at King's College London, United Kingdom. Our collaborators have been successful over the years in measuring the orientation of several protein domains in muscle fibers using fluorescent bifunctional rhodamine (BR) probes synthesized by Dr. Corrie. Subsequently, a new bifunctional sulforhodamine (BSR) probe which is supposedly more rigid than previous probes became commercially available. This raised questions as to whether the new probe would give results similar to the BR probes, would be more or less reactive than the BR probes, or would perturb the structure of TnC. The results showed that BSR labelling does not significantly affect the secondary structure of the protein or its dynamics in solution, and that

both BR and BSR probes accurately report the orientation of the C-helix of TnC to which they are attached. The BSR dipole had smaller orientational dispersion, consistent with less flexible linkers between the rhodamine dipole and cysteine-reactive groups. The paper on its whole supports the use of bifunctional rhodamine probes to accurately report the mean *in situ* orientation of the vector joining the two cysteines to which they are attached.

While Chapter II showed that the structure and dynamics of TnC is not affected by the presence of the BR probe, no investigation of the dynamics of the probe itself was ever made using NMR. The importance of the work presented in Chapter III has broad implications, relating to the use of fluorescence probes in general to measure distances (i.e. FRET) or protein domain orientations. The measurement of similar order parameters for the backbone protein ($S^2 \sim 0.85$) and for the methyl groups of the covalently attached BR probe ($S^2 \sim 0.79$) suggests that the amplitude of independent reorientation of the BR probe is small in magnitude, consistent with results from fluorescence polarization measurements in reconstituted muscle fibers. What makes this paper (2) one of the highlights of my Ph.D for me is not only the conclusion – showing that BR probes are more rigid than one would have thought and supporting the use of BR probe to measure protein domain orientations – but the technical difficulties of the NMR experiments described in Chapter III that we had to overcome in order to obtain reliable results.

While working on the BR project, I was also working on the determination of the structure of a Phe-to-Trp mutation (F77W) of cNTnC. I was introduced to NMR data acquisition, calcium and peptide titrations, measurement of dissociation constants, and NMR relaxation analysis. By comparing the ^{15}N - T_2 of the mutant with wild-type cNTnC, we recognized the presence of dimerization in solution because of the lower average ^{15}N - T_2 for the mutant. A TFE titration was performed, and the addition of TFE resulted in better chemical shift dispersion in the ^1H - ^{15}N HSQC NMR spectrum, and a slight improvement in the ^{15}N - T_2 . A full

set of NMR data required for structure calculation was then acquired in 19% TFE. During the chemical shift assignment, I was having a hard time assigning the chemical shifts of the alanines (which have a distinct $\Delta\delta$ C β of ~ 17 ppm). I could identify 6 alanines (one too many) according to my CBCACONH and HNCACB spectra. After several weeks of looking carefully at the spectra and hearing everyone telling me “*make sure you did it right*”, the presence of an extra mutation (V82A) was confirmed when we re-sequenced the plasmid. Displeased with the discovery of this extra mutation, we verified the original sequencing (raw fluorescence data) obtained by Dr. Lindhout (who cloned the mutant), and the mutation V82A was nowhere to be seen. To this day, the only hypothesis as to where the V82A mutation came from is that the *E. coli* screwed up. Since all of the NMR data was already acquired for the double mutant F77W-V82A-cNTnC, it took only a few more months to obtain the NMR structure. We found that the insertion of the Trp in the hydrophobic core of cNTnC increased the calcium affinity of the protein slightly, but decreased the cSp binding affinity by a factor of ten. Of particular interest, the orientation of the Trp indole in the double mutant was found to be in opposite orientation compared to homologous calcium binding proteins (3).

The last result led to Chapter V, where the structure of F77W-cNTnC (no V82A) bound to its TnI binding partner (residues 144-163) was solved. The project took an interesting route when we found out that the single mutant showed a Trp indole orientation in opposite direction to the double mutant F77W-V82A. In our opinion, there were three potential causes: 1) the presence of TnI in the new structure, 2) the presence of TFE in the previous structure, 3) the presence of the V82A mutation in the previous structure. I remember being convinced that TnI was responsible for this difference, while Brian thought the reason was the presence/absence of the mutation V82A. As is sometimes the case in science, it turned out that neither of our hypotheses was correct. The presence of TFE was shown to invert the Trp rotamer population from a favourable $t\text{-}105^\circ$ conformation to a $t90^\circ$ conformation with same χ_1 but different χ_2 (4). Once this

work was published, it came to the attention of the Tryptophan fluorescence lifetime ‘crowd’ who hypothesize that the observation of multiple lifetimes is associated with different Trp rotamer populations, so that this case might provide a test of their theories. We are currently in the process of writing a small report describing the effect of TFE on the Trp lifetime populations of the two mutants monitored by fluorescence lifetime spectroscopy (collaborator Dr. Engelborghs, University of Leuven, Belgium, data not shown).

Changing focus from TnC to TnI, the main goal of Chapter VI is the structural characterization of the C-terminal region of TnI, which has been proposed to be an intrinsically disordered region (IDR) and this in turn important in the regulation of the troponin complex (5). Because various structures have been proposed for this region, we wanted to determine if any nascent structure exists in this IDR. I initially looked at the whole troponin complex, and these results are included in an appendix. To avoid the experimental constraints created by the size of the troponin complex (52 kDa, see Appendix E), we used a skeletal TnC-TnI chimera (20 kDa) that has many advantages over the troponin complex. Overall, our results indicate that the use of the TnC-TnI chimera is a good strategy to probe the dynamics of TnI. Our relaxation data showed that the N-domain of sTnC is well folded and interacts with its sTnI binding partner in the chimera as expected. The NMR relaxation and chemical shift data of the C-terminal region of sTnI show the presence of a flexible and disordered region. On balance, our experimental and MD data argues against any stable secondary structure significantly populated in the IDR of sTnI, supporting the role of a ‘fly-casting’ mechanism for regulation of muscle contraction. Of particular interest, the fact that the plateau observed in the relaxation data of the chimera can also be observed for other proteins suggests that this is an intrinsic property of extended mobile IDRs. The next step of this project is the construction of a cardiac chimera and the study of the TnC-TnI interaction in the cardiac system.

I then proposed that I switch topics from troponin and muscle regulation to prion diseases. My interest in prions started in 2005 with my (obligatory) enrolment in the 671 course, where I decided to present three back-to-back papers about prions that were published in the journal *Nature* (6-8). At the time, I didn't know much about these proteins, other than that they caused mad cow disease and were the responsible agents for the transmissibility of prion diseases. In an attempt to understand the initial structural changes accompanying the conversion of the cellular form of the prion protein (PrP^{C}) to its fibril infectious counterpart (PrP^{Sc}), we decided to study recombinant bovine PrP^{C} and monitor its structural conversion using what we do best, NMR spectroscopy. Chapters VII and VIII describe the use of ^1H NMR spectroscopy to monitor on a per-residue basis the stability of the bovine prion protein upon urea-induced unfolding (9,10). Our data suggests that the dissociation of the native β -sheet of PrP^{C} is a primary step in the structural conversion, whilst strong hydrophobic interactions between helices $\alpha 1$ and $\alpha 3$, and between $\alpha 2$ and $\alpha 3$, stabilize these regions even at very high concentrations of urea. These observations are key for the identification and characterization of the elements favoring the prion protein transformation, and also bring insights about mutations found in the prion protein that are associated with familial genetic diseases. The study of this system allowed me to get more familiar with the protein folding/unfolding field, and to better understand these fascinating disorders that are prion diseases.

The common theme of the different parts of the work presented so far is the investigation of protein structure and dynamics using NMR spectroscopy. We were approached to determine the structure of a DNA construct. This work is presented in Chapter IX, where the structure and dynamics of DNA duplexes are investigated. This section reports the chemical synthesis of a dodecamer DNA duplex containing the Cidofovir (CDV) drug inserted at position 7 (11). The solution structure of the CDV DNA duplex was determined by NMR spectroscopy, and the effect of the CDV insertion was compared to a control DNA duplex containing a deoxycytidine. Our results show that the drug molecule

is well accommodated within the dodecamer DNA duplex, although destabilization of the nucleotides surrounding the CDV was observed during DNA melting experiments. Thus, even though CDV causes only subtle changes to the DNA structure, it still suffices to create a profound impediment to *Orthopoxvirus* DNA replication. This discovery has important implications for researchers pursuing rational antiviral drug design strategies.

These results form the last piece of the mosaic of this thesis, the investigation of biomolecular structure and dynamics by NMR spectroscopy.

REFERENCES

1. Julien, O., Sun, Y., Knowles, A., Brandmeier, B., Dale, R., Trentham, D., Corrie, J., Sykes, B., and Irving, M. (2007) Toward protein structure in situ: comparison of two bifunctional rhodamine adducts of troponin C, *Biophys. J.* 93, 1008-1020.
2. Julien, O., Mercier, P., Spyropoulos, L., Corrie, J., and Sykes, B. (2008) NMR studies of the dynamics of a bifunctional rhodamine probe attached to troponin C, *J. Am. Chem. Soc.* 130, 2602-2609.
3. Julien, O., Sun, Y., Wang, X., Lindhout, D., Thiessen, A., Irving, M., and Sykes, B. (2008) Tryptophan mutants of cardiac troponin C: 3D structure, troponin I affinity, and in situ activity, *Biochemistry* 47, 597-606.
4. Julien, O., Mercier, P., Crane, M., and Sykes, B. (2009) The effect of the cosolvent trifluoroethanol on a tryptophan side chain orientation in the hydrophobic core of troponin C, *Protein Sci.* 18, 1165-1174.
5. Julien, O., Mercier, P., Allen, C., Fiset, O., Ramos, C., Lagüe, P., Blumenschein, T., and Sykes, B. (2011) Is There Nascent Structure in the Intrinsically Disordered Region of Troponin I?, *Proteins: Structure, Function, and Bioinformatics*, (in press).
6. Ritter, C., Maddelein, M., Siemer, A., Luhrs, T., Ernst, M., Meier, B., Saupe, S., and Riek, R. (2005) Correlation of structural elements and infectivity of the HET-s prion, *Nature* 435, 844-848.
7. Nelson, R., Sawaya, M., Balbirnie, M., Madsen, A., Riek, C., Grothe, R., and Eisenberg, D. (2005) Structure of the cross-beta spine of amyloid-like fibrils, *Nature* 435, 773-778.
8. Krishnan, R., and Lindquist, S. (2005) Structural insights into a yeast prion illuminate nucleation and strain diversity, *Nature* 435, 765-772.
9. Julien, O., Graether, S., and Sykes, B. (2009) Monitoring prion protein stability by NMR, *J. Toxicol. Environ. Health A* 72, 1069-1074.

-
10. Julien, O., Chatterjee, S., Thiessen, A., Graether, S., and Sykes, B. D. (2009) Differential stability of the bovine prion protein upon urea unfolding, *Protein Sci.* 18, 2172-2182.
 11. Julien, O., Beadle, J., Magee, W., Chatterjee, S., Hostetler, K., Evans, D., and Sykes, B. (2011) Solution Structure of a DNA Duplex Containing the Potent Anti-Poxvirus Agent Cidofovir, *J. Am. Chem. Soc.*, (in press).

Appendix A – Supporting Information for Chapter II

MATERIAL AND METHODS

A calcium titration was performed to measure the calcium binding affinity of sNTnC•BSR₅₆₋₆₃. The NMR sample containing sNTnC•2Ca²⁺•TnI₁₁₅₋₁₃₁•BSR₅₆₋₆₃ was recycled into an apo sNTnC•BSR₅₆₋₆₃ sample. 2 mM of EDTA were added to remove the calcium from the two binding sites of sNTnC. The sample was passed through a G-25 column to separate the sNTnC•BSR₅₆₋₆₃ from EDTA•2Ca²⁺ and TnI₁₁₅₋₁₃₁. The protein solution obtained was then lyophilized overnight. To minimize the loss of sNTnC•BSR₅₆₋₆₃, 475 μ l of NMR buffer (100 mM KCl and 10 mM Imidazol), 6 μ l of protease inhibitor and 25 μ l of a solution containing 5 mM DSS-D6 and 0.2 % NaN₃ in D₂O were added directly into the lyophilized tube. The solution obtained was then filtered by centrifugation with 0.22 μ m centrifuge tube filter (Corning, NY) and finally transfer into a 5mm NMR tube.

During the calcium titration, eight 2D $\{^1\text{H}-^{15}\text{N}\}$ -HSQC spectra were acquired on a Varian INOVA 600-MHz for different calcium concentrations (0 to 1.88 mM). The spectra were acquired with 1024 (¹H) \times 512 (¹⁵N) complex points, 32 transients and spectral widths of 11,990 Hz and 3242 Hz for the first and second dimension respectively.

RESULTS

The HSQC spectra acquired during the calcium titration showed broad line widths, low signal-to-noise ratio and did not show good peak dispersion. We observed the same phenomena during the preparation of our first NMR sample, which is caused by both the low protein concentration used and the absence of the TnI₁₁₅₋₁₃₁ peptide. For those reasons, only the peaks of Gly35 and Gly43 were used to calculate the calcium binding affinity of sNTnC•BSR₅₆₋₆₃. They were about the only peaks having unambiguous chemical shifts for both ¹H and ¹⁵N

dimensions during the calcium titration. The calcium binding constants for the two binding sites were calculated by plotting the normalized chemical change in function of the ratio of calcium and protein concentration for each step of the titration (figure S1). The calcium concentrations were adjusted to compensate for residual EDTA in the sample. The curves in red represent the best fit for Gly35 and Gly43. Both curve used exactly the same Ca^{2+} dissociation constant, which were determined to be $K_{d1} = 13 \mu\text{M}$ and $K_{d2} = 145 \mu\text{M}$ for the first and second calcium binding respectively (site II and site I respectively). In comparison, the binding affinity of sNTnC has been previously determined to be $0.8 \mu\text{M} \leq K_{d1} \leq 3 \mu\text{M}$ and $5 \mu\text{M} \leq K_{d2} \leq 23 \mu\text{M}^1$. As we can see, the calcium binding affinity of sNTnC•BSR₅₆₋₆₃ is evidently diminished. Because of the reduced quality of the spectra available used to calculate the binding constants, the conclusion we can get from those data is qualitative only. However, those results show clearly that the calcium affinity of sNTnC is decreased by the presence of the bifunctional sulforhodamine on the C helix, and are consistent with the fluorescence data.

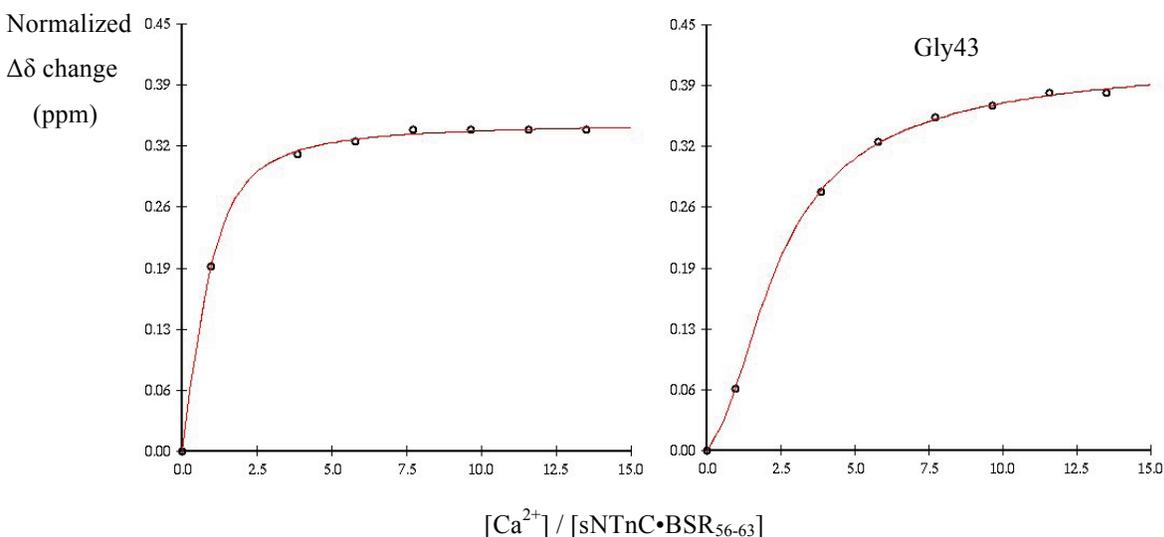


Fig. A-1. Calcium titration plots of sNTnC•BSR₅₆₋₆₃ for Gly35 and Gly43. The chemical shift changes were obtained from eight 2D $\{^1\text{H}-^{15}\text{N}\}$ -HSQC spectra and are plotted in function of the $[\text{Ca}^{2+}] / [\text{sNTnC}\cdot\text{BSR}_{56-63}]$. The curves (in red) are the best fit calculated for Gly35 and Gly43. Both residues show a K_{D1} of $\sim 13 \mu\text{M}$ and a K_{D2} of $\sim 145 \mu\text{M}$.

REFERENCES

1. Li, M. X., S. M. Gagné, S. Tsuda, C. M. Kay, L. B. Smilie, and B. D. Sykes. 1995. Calcium binding to the regulatory N-domain of skeletal muscle troponin C occurs in a stepwise manner. *Biochemistry*. 34:8330-40

Appendix B – Supporting Information for Chapter III

Abbreviations:

R_2^{obs} = ^{15}N -transverse relaxation rate (s^{-1}).

R_2^{M} and R_2^{D} = theoretical ^{15}N - R_2 values for cNTnC in the monomer and dimer states (s^{-1}).

f_{M} and f_{D} = fraction of monomer and dimer in solution.

τ_{ex} = chemical exchange term.

τ_{M} and τ_{D} = lifetime of the monomer and dimer states respectively.

MATERIAL AND METHODS

^{15}N -Transverse Relaxation vs Concentration of F77W-V82A-cNTnC

18 mg of [^{15}N]F77W-V82A-cNTnC were dissolved in 550 μl of NMR buffer for a final protein concentration of ~ 3.37 mM (from amino acid analysis). To the sample were added 0 mM CaCl_2 to saturate the calcium binding site and 20 mM DTT to prevent cysteine oxidation; 0.03 % sodium azide and protease inhibitors were also added to prevent sample degradation. Deuterated-DSS (0.25 mM) was added to reference the NMR spectra. The pH was adjusted to 6.6 according to the imidazole signal in a 1D ^1H NMR spectrum. 500 μl were put in the NMR tube. A set of six 2D ^1H , ^{15}N HSQC spectra using the BioPack pulse sequence (Varian Inc.) with different T_2 relaxation time delays ($\tau = 10, 30, 50, 70, 90, 110$ ms) was acquired for each of the six dilution steps. A dilution step consists of replacing 250 μl of NMR sample by 250 μl of NMR buffer. The ^{15}N -transverse relaxation rates of the protein amide residues were extracted for the different protein concentrations used: 3.37, 1.69, 0.84, 0.42, 0.21 and 0.11 mM of F77W-V82A-cNTnC. An amino acid analysis was performed for the first and last steps of the experiment, and intermediate protein concentrations were determined by interpolation. The spectra were acquired on a Varian INOVA 600 MHz NMR spectrometer.

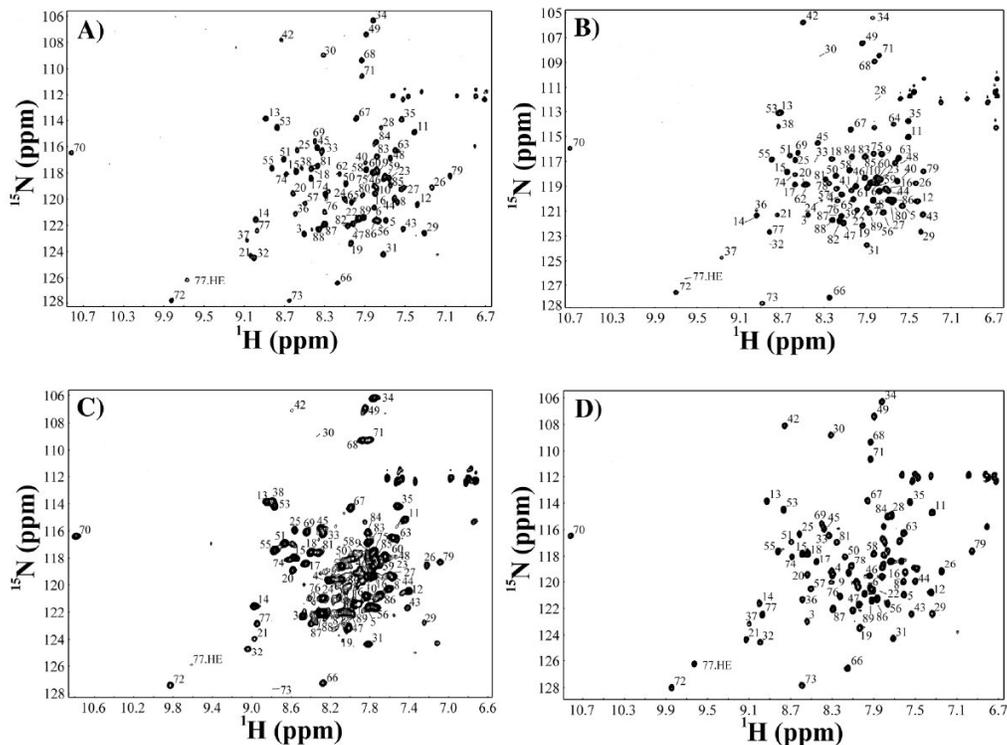


Fig. B-1. $^1\text{H},^{15}\text{N}$ HSQC NMR spectra of F77W and F77W-V82A-cNTnC•Ca $^{2+}$. (A) 0.11 mM of F77W-V82A-cNTnC, (B) 1.5 mM of F77W-V82A-cNTnC in 19% TFE, (C) 3.37 mM of F77W-V82A-cNTnC (D) 0.4 mM of F77W-cNTnC.

RESULTS

Dimerization of F77W-V82A-cNTnC•Ca $^{2+}$

The cross peaks in the 2D $^1\text{H},^{15}\text{N}$ HSQC NMR spectra of F77W-V82A-cNTnC•Ca $^{2+}$ are broader at higher protein concentration (Fig. B-1, panel A versus panel C), and are narrower at high concentration in the presence of TFE (Fig. B-1, panel B), which has previously been used in this system as a denaturant of quaternary structure (1). The calcium saturated form of sNTnC has been shown to weakly dimerize, and cNTnC less so (2); and F29W-sTnC has been shown to have a concentration dependent calcium affinity consistent with dimerization (3). To evaluate the monomer:dimer equilibrium of F77W-V82A-cNTnC•Ca $^{2+}$, the backbone amide ^{15}N -relaxation rates ($^{15}\text{N-R}_2$) were measured as a function of protein concentration.

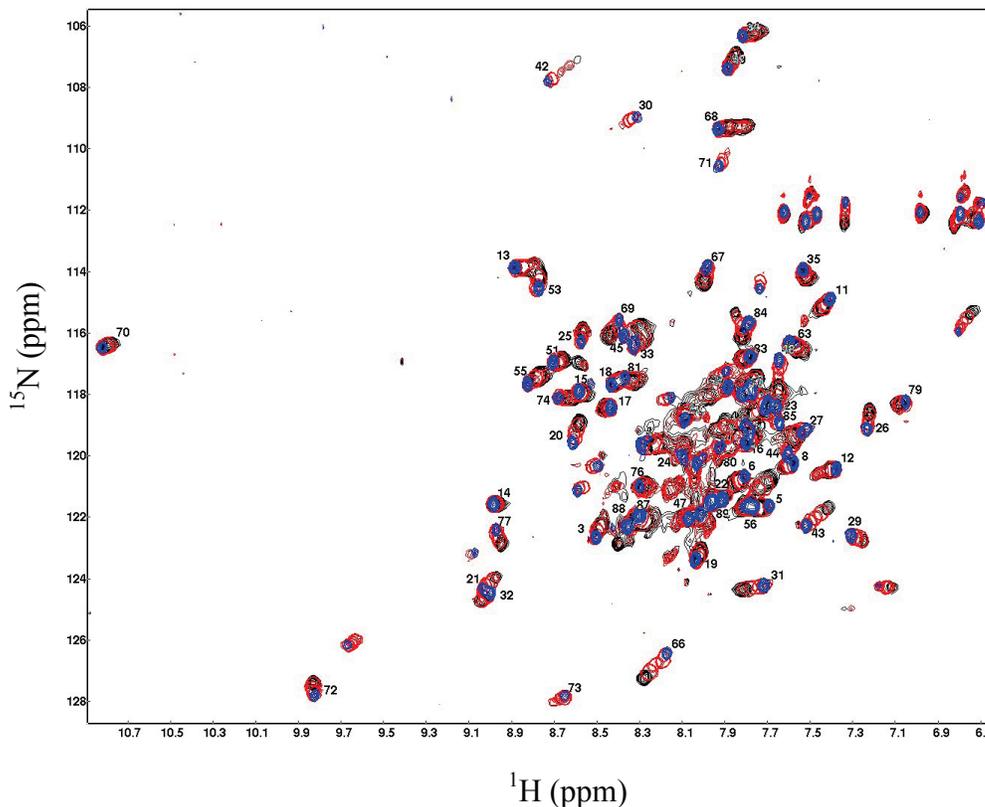


Fig. B-2. Six ^1H , ^{15}N HSQC NMR spectra of F77W-V82A-cNTnC:Ca $^{2+}$ at different protein concentration. The first step of the dilution series is shown in black (3.37 mM), step 2 to step 5 are shown in red, and the spectrum at lowest protein concentration is shown in blue (0.11 mM). 67 residues were unambiguously identified.

The ^{15}N - R_2 of a molecule is proportional to its rotational correlation time in solution, which can be directly correlated to its apparent molecular weight. In the presence of a monomer-dimer equilibrium, the observed ^{15}N - R_2 is given by:

$$R_2^{\text{obs}} = f_M R_2^{\text{M}} + f_D R_2^{\text{D}} \quad [1]$$

where f_M and f_D are the fraction of monomer and dimer, respectively, and R_2^{M} and R_2^{D} are the relaxation rates of the monomer and dimer, respectively. Exchange

broadening can also increase the observed relaxation rate:

$$R_2^{\text{obs}} = f_M R_2^M + f_D R_2^D + f_M f_D \tau_{\text{ex}} \Delta\omega^2 \quad [2]$$

where τ_D is the lifetime of the dimer and $\Delta\omega$ is the chemical shift difference for the nucleus studied. The term $f_M f_D \tau_{\text{ex}} \Delta\omega^2$ can also be written as $f_M^2 f_D \tau_D \Delta\omega^2$ because $\tau_{\text{ex}} = \tau_M \tau_D / (\tau_M + \tau_D)$ and $f_M = \tau_M / (\tau_M + \tau_D)$. Using the linear correlation for different molecules of the troponin complex (4), the expected ^{15}N - R_2 values of cNTnC are 6.0 s^{-1} for a monomer and 12 s^{-1} for a dimer. The dimerization constant (K_{dimer}) for calcium saturated cNTnC has been shown to be 7.3 mM (2). In comparison, the K_{dimer} for the N-domain of calcium saturated skeletal TnC (sNTnC) is 1.3 mM , indicating a higher tendency for dimerization. To look at the effect of the mutation F77W-V82A on the dimerization of cNTnC, the average backbone ^{15}N - R_2 was determined as a function of the protein concentration (Fig. B-3). For each step of the dilution, the average ^{15}N - R_2 was calculated for at least 54 resonances out of 80 possibilities (residues 5 to 86). The C- and N-terminal ends were excluded because of the mobility of residues in these regions. As expected, the fastest ^{15}N - R_2 (11.25 s^{-1}) was obtained at high concentration ($>1.7 \text{ mM}$) and the longest ^{15}N - R_2 (7.4 s^{-1}) at the lowest concentration allowing sufficient signal-to-noise for accurate relaxation measurements (0.11 mM). As one can see from this data presented in Fig. B-3, F77W-V82A-cNTnC ($K_{\text{dimer}} = 5.0 \text{ mM}$) is more similar to sNTnC and has a higher tendency to dimerize than wild type cNTnC. Several peaks were impossible to identify at high concentration in the ^1H , ^{15}N HSQC, but appeared as unambiguous resonances at low concentration. In fact, almost all of the amide resonances are observed at 0.11 mM , and more than 20 are missing at 3.37 mM . This is the result of extensive exchange broadening of the NMR cross peaks. Similar results were obtained for single mutant F77W-cNTnC (data not shown), indicating that the mutation V82A is not responsible alone for the increase of dimerization. The ^{15}N - R_2 values for several individual residues plotted as a function of protein concentration show the considerable differences in exchange

broadening of different cross peaks, with the residues most affected by dimerization generally broadening the most.

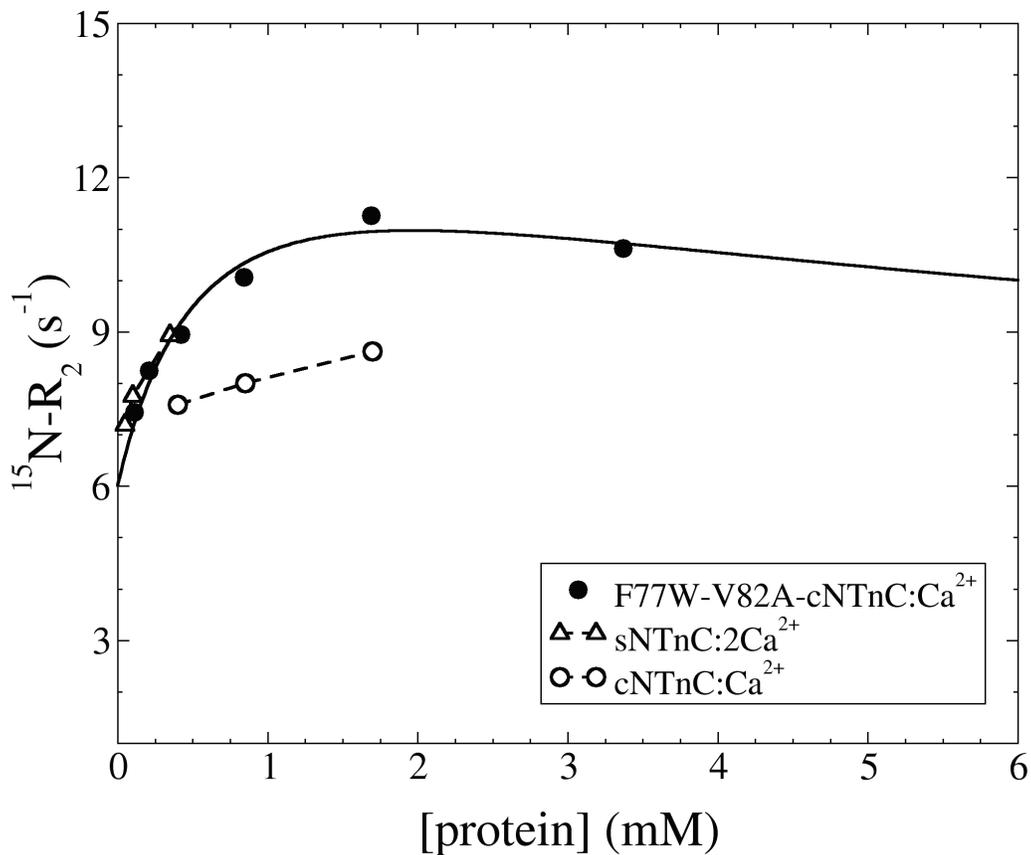


Fig. B-3. Average $^{15}\text{N-R}_2$ relaxation rates for F77W-V82A-cNTnC•Ca²⁺ (filled circles) as a function of the protein concentration compared to cNTnC•Ca²⁺ (triangles) and sNTnC•2Ca²⁺ (empty circles). The data for wild type proteins were taken from (2) and the dashed lines are only to guide one's eye. A K_{dimer} of 5.0 mM and a value of 46 for $\tau_{\text{ex}}\Delta\omega^2$ were used to fit the data of F77W-V82A-cNTnC (continuous line).

The relaxation rates of residues L29, A31 and W77 were measured as a function of the protein concentration of mutant F77W-V82A-cNTnC:Ca²⁺ (Fig. B-4) and the fits were performed using equation [2] presented above. The three fits were plotted using the theoretical values for R_2^{M} and R_2^{D} of 6 and 12 s⁻¹ respectively, a

K_{dimer} of 9.0 s^{-1} , and $\tau_D \Delta\omega^2$ values of 200, 100 and 15 for residues L29, W77 and A31, respectively. Those three residues are unambiguous resonances in isolated area of the spectra, meaning that no overlap could affect the relaxation rates, neither the precision of the chemical shift measurements. The $\tau_D \Delta\omega^2$ values were determined using xcrvfit (in-house curve fitting program) by keeping all other parameters constant. Having in hand the $\tau_D \Delta\omega^2$ values for each residues, one can determined the lifetime of the dimer (τ_D) by measuring the ^{15}N chemical shift changes ($\Delta\omega$) as following (results shown in Fig. B-4):

$$\tau_D = [\tau_D \Delta\omega^2]_i / [\Delta\omega^{\text{obs}}]^2 \quad [3]$$

where $\Delta\omega^{\text{obs}} = (2\pi \times |\Delta\delta_{\text{initial}} - \Delta\delta_{\text{final}}|)^2$. The $\Delta\delta$ values (Hz) were obtained by measuring the ^{15}N chemical shift difference between the spectra at 3.37 and 0.11 mM of protein. In theory, the lifetime of the dimer should be equal for every residue if only one phenomenon is responsible for the chemical exchange.

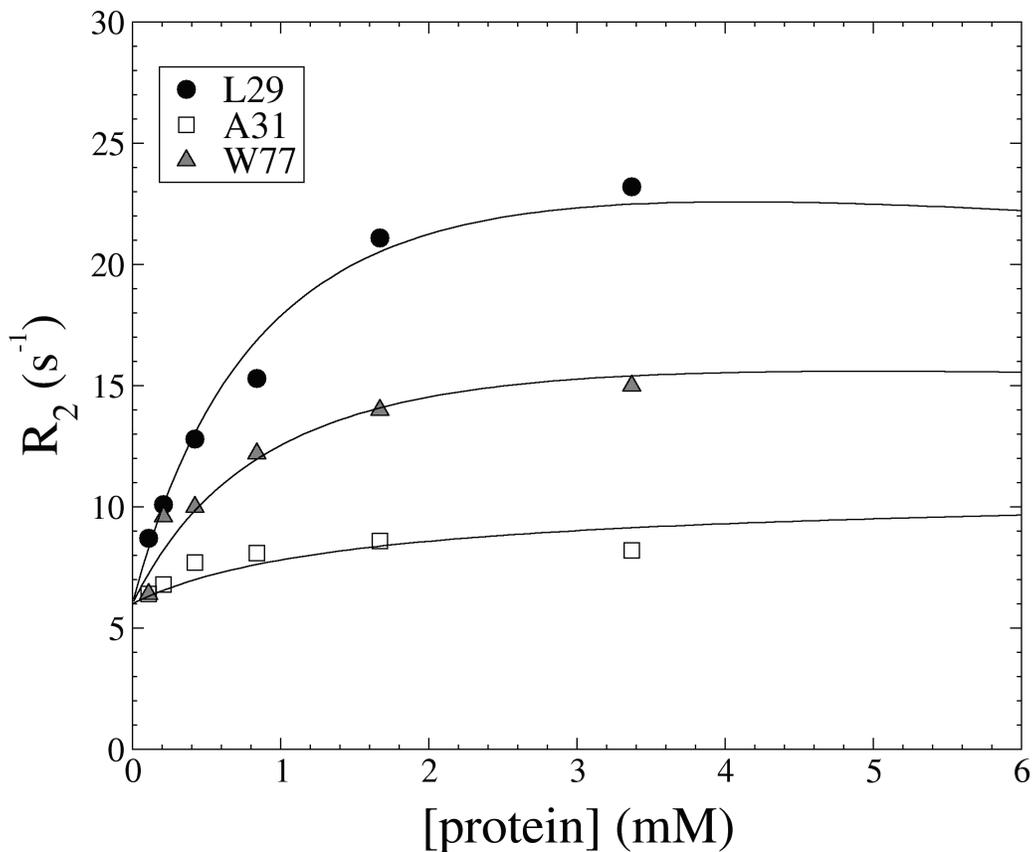


Fig. B-4. Relaxation rates of residues L29, A31 and W77 as a function of the protein concentration for mutant F77W-V82A-cNTnC:Ca²⁺. A K_{dimer} of 9.0 s^{-1} was used for the three fits, and $\tau_{\text{ex}}\Delta\omega^2$ values of 200, 100 and 15 were used for residues L29, W77 and A31 respectively.

The data found in Table I are in agreement with Fig. B-1, where one can see that L29 shows a large R_2^{obs} variation as a function of the protein concentration, while A31 shows only a little and W77 is in between. We also confirmed with the calculation of τ_D . However, the lifetime are not consistent to one and other, as seen in Fig. B-5. This suggests that more than one phenomena is responsible for the exchange broadening observed in the ¹H, ¹⁵N HSQC spectra.

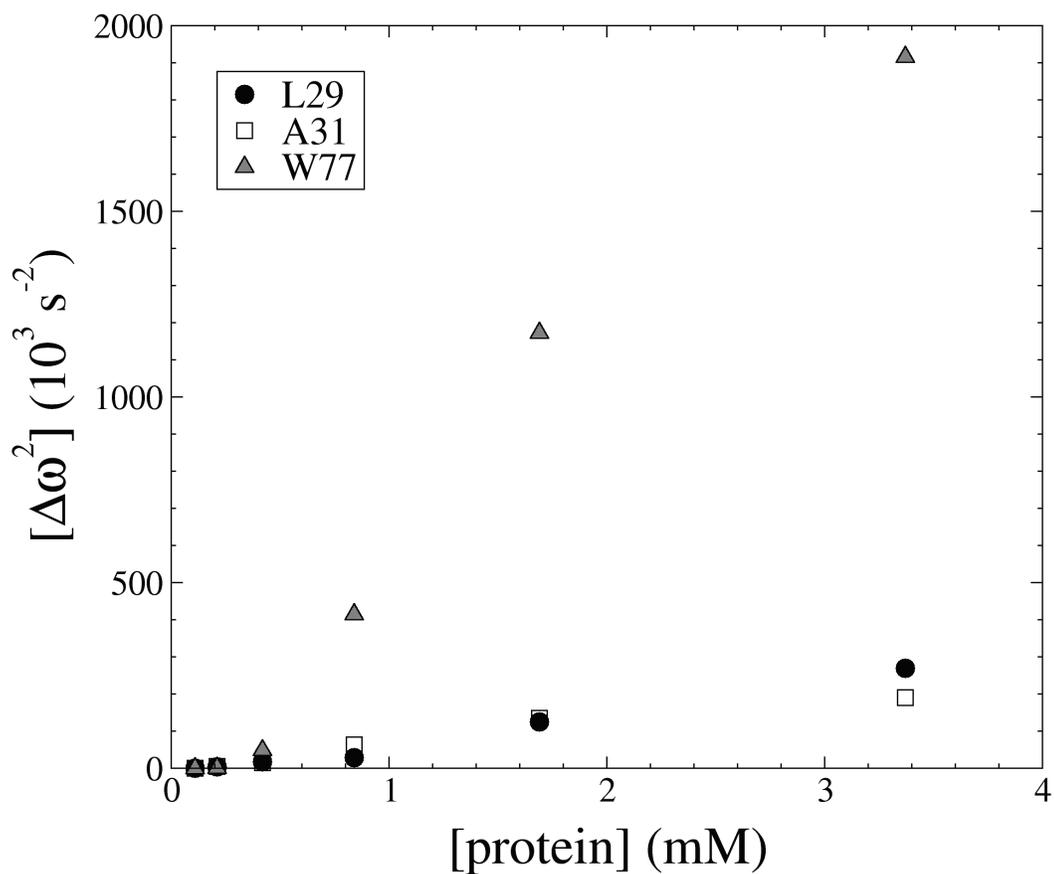


Fig. B-5. Chemical shift variation ($\Delta\omega^2$) of the ^{15}N backbone amide of residues L29, A31 and W77 as a function of the protein concentration.

Table I. Calculation of the τ_{ex} values for residues L29, A31 and W77

Residues	$\tau_{\text{D}} \Delta\omega^2$	$ \Delta\delta_{\text{initial}} - \Delta\delta_{\text{final}} $ (ppm)	$\Delta\omega^2$ (10^3 s^{-2})	τ_{D} (10^{-3} s)
L29	195.6	0.16497	268.60	0.728
A31	15.4	0.13870	189.87	0.081
W77	99.9	0.44055	1915.54	0.052

DISCUSSION

It is known that the N-domain of TnC dimerizes in solution, and that the tendency to dimerize is greater for skeletal ($K_{\text{dimer}} = 1.3 \text{ mM}$) than cardiac ($K_{\text{dimer}} = 7.3 \text{ mM}$) NTnC (2). Dimerization affects the apparent molecular weight of the protein in solution, broadening the NMR resonances and diminishing the resolution in the spectrum. Any study of the dynamics of the N-domain of TnC should therefore be done at a relatively low protein concentration to reduce dimerization. The first extensive NMR study of the dynamics of cNTnC was done at 0.15 mM, in contrast to the $\sim 1 \text{ mM}$ required to solve the NMR structure (5). Another group also solved the NMR structure of cNTnC at the same time (6), but (to this day) never deposited their structure in the Protein Data Bank. They also recognized the presence of dimerization in their NMR sample (i.e. slightly high $^{15}\text{N-R}_2$), explaining the use TFE for the structure determination and dynamics characterisation. However, they did not realize that the protein concentration was the caused of their problems.

Comparison of two ^1H - ^{15}N HSQC for F77W-V82A-cNTnC•Ca $^{2+}$ at low and high concentrations shows the effect of dimerization on the 2D spectra (lower spectral dispersion and broader line widths for Fig. B-2 panel C compared to panel A), and Fig. B-3 shows the effect of dimerization on the average $^{15}\text{N-R}_2$ relaxation rates of the backbone amides of different N-domains of TnC. The F77W substitution in the core of the N-domain of cTnC modifies the monomer-dimer equilibrium of the cardiac protein, and makes the protein concentration dependence of the $^{15}\text{N-R}_2$ relaxation rates more similar to that of skeletal NTnC. The structures of F77W-V82A-cNTnC•Ca $^{2+}$ and sNTnC•Ca $^{2+}$ are also more open when compared to cNTnC•Ca $^{2+}$ (see Chapter III). The opening of the N-domain expands the hydrophobic surface area exposed to solvent, potentially favouring the formation of dimer. The analysis of NMR $^{15}\text{N-R}_2$ data of F77W-V82A-cNTnC•Ca $^{2+}$ on a per residue basis as a function of protein concentration showed different chemical exchange (R_{ex}) contribution. However, a clear correlation was not observed

between the ^{15}N - R_{ex} contributions and the differences in the ^{15}N chemical shifts between the monomer and dimer forms of the protein. Using the equations presented above, we obtained different τ_{D} values among the residues, suggesting that more than one mechanism is responsible for the exchange broadening observed in the spectra. We conclude that a complex mixture of dimerization, open-closed conformational exchange and calcium exchange is responsible for the chemical exchange broadening observed in the NMR spectrum, in agreement with the detailed studies of Eichmüller et al. (7).

REFERENCES

1. Slupsky, C. M., Kay, C. M., Reinach, F. C., Smillie, L. B., and Sykes, B. D. (1995) *Biochemistry* 34:7365-7375
2. Spyrapoulos, L. (2001) *Know your protein: effects of weak dimerization*, Edited by Oleg Jardetzky and Michael D. Finucane Ed., IOS Press
3. Li, M. X., Chandra, M., Pearlstone, J. R., Racher, K. I., Trigo-Gonzalez, G., Borgford, T., Kay, C. M., and Smillie, L. B. (1994) *Biochemistry* 33:917-925
4. Lindhout, D. (2005) An NMR investigation of troponin I within the cardiac troponin complex and the effects on the calcium induced cardiac muscle contraction. In. *Department of Biochemistry*, University of Alberta
5. Spyrapoulos, L., Gagné, S. M., Li, M. X., and Sykes, B. D. (1998) *Biochemistry* 37:18032-18044
6. Paakkonen, K., Annala, A., Sorsa, T., Pollesello, P., Tilgmann, C., Kilpelainen, I., Karisola, P., Ulmanen, I., and Drakenberg, T. (1998) Solution structure and main chain dynamics of the regulatory domain (Residues 1-91) of human cardiac troponin C, *J. Biol. Chem.* 273, 15633-15638.
7. Eichmüller, C., and Skrynnikov, N. R. (2005) *J. Biomol. NMR* 32:281-293

Appendix C – Supporting Information for Chapter VIII

Lineshape Analysis of the 1D ^1H NMR spectra

In order to determine the rate of the conformational exchange observed in the 1D ^1H NMR spectra during the urea-induced denaturation of bPrP^C, we have performed a full-lineshape analysis as previously described (1). We have simulated the lineshape of Y218.H δ in the case of fast ($k_{N-U} = 1000 \text{ s}^{-1}$) and slow exchange ($k_{N-U} = 1 \text{ s}^{-1}$). Linewidths of 20 Hz and 10 Hz were used for the native and unfolded conformations, respectively, and a chemical shift difference of approximately 1 ppm at 800 MHz was measured for the two states. Clearly, the observed lineshape of the resonance of the H δ protons of Y218 matches the simulated lineshape as a function of the concentration of urea better with a k_{N-U} in the order of 1 s^{-1} (see Fig. C-1). This confirms that the unfolding equilibria is in the slow exchange limit on the NMR timescale.

Definitions:

Reaction is of the type	Folded	<---->	Unfolded
labels:	A		B
populations	pA		pB
lifetimes:	tauA = 1/kfu		tauB = pB*tauA/pA
chemical shifts:	vA		vB
linewidths	wA		wB

(* **deltaG = deltaG0 - m [urea]** *)

R = 0.008314472;

T = 303.15;

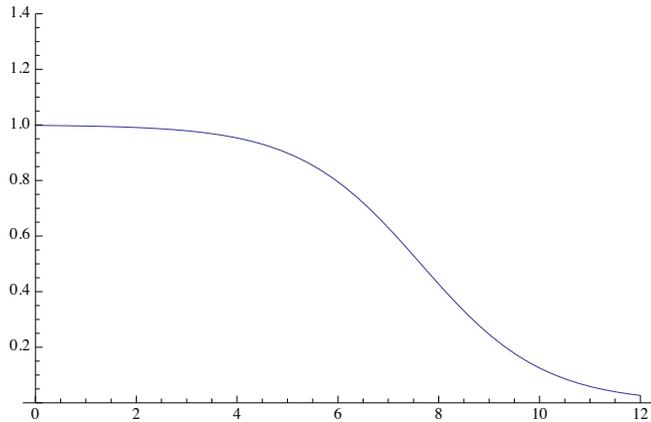
deltaG = 15.9/(R * T);

m = 2.08/(R * T);

pA := 1/(1 + Exp[(-deltaG + m*urea)]);

pB := (1 - pA);

Plot[{pA}, {urea, 0, 12}, PlotRange -> {0, 1.4}]



2 | UreaDenaturePrPc_Y218-slow.nb

```

tauA := 1/kfu;
(* kfu in sec -
  1 is rate of foled to unfolded which is constant Prot. Sci 10 1685-88 (2001) *)
tauB := (pB) tauA / (pA);
tau := pB tauA;

vA = -200; (* shifts in Hz *)
vB = -1000;
dv := vA - vB;
Dv := 0.5(vA + vB) - v;

wA = 20; (* linewidths in Hz *)
wB = 10;
T2A := 1 / (Pi wA);
T2B := 1 / (Pi wB);

P := tau((1/ T2A T2B) - 4 Pi Pi Dv Dv + Pi Pi dv dv) + pB/T2B + pA/T2A;

Q := tau(2 Pi Dv - Pi dv(pA - pB));

R :=
  2 Pi Dv(1 + tau(1/T2A + 1/T2B)) + Pi dv tau(1/T2B - 1/T2A) + Pi dv(pA - pB);

Int := (P(1 + tau(pB/T2A + pA/T2B)) + QR) / (PP + RR);

kfu = 1;
offset = 0.01;

urea = 0.0;
plot1 = Plot[{Int}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
  Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

urea = 1.0;
plot2 =
  Plot[{Int+offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
  Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

urea = 2.0;
plot3 =
  Plot[{Int+2*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
  Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

urea = 3.0;
plot4 =
  Plot[{Int+3*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
  Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

urea = 4.0;
plot5 =

```

```

    Plot[{Int+4*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
      Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 5.0;
    plot6 =
      Plot[{Int+5*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 6.0;
    plot7 =
      Plot[{Int+6*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 7.0;
    plot8 =
      Plot[{Int+7*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 8.0;
    plot9 =
      Plot[{Int+8*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 9.0;
    plot10 =
      Plot[{Int+9*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 10.0;
    plot11 =
      Plot[{Int+10*offset}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
        Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    urea = 11.0;
    plot12 = Plot[{0.13}, {v, -1200, 0}, PlotRange -> All, PlotPoints -> 100,
      Frame -> True, AxesLabel -> {"frequency ( Hz) ", "Intensity"}];

    Print[" Off rate constant (sec -1) =", " ", kfu]

    Show[{plot1, plot2, plot3, plot4, plot5,
      plot6, plot7, plot8, plot9, plot10, plot11, plot12}]

Off rate constant (sec -1) = 1

```

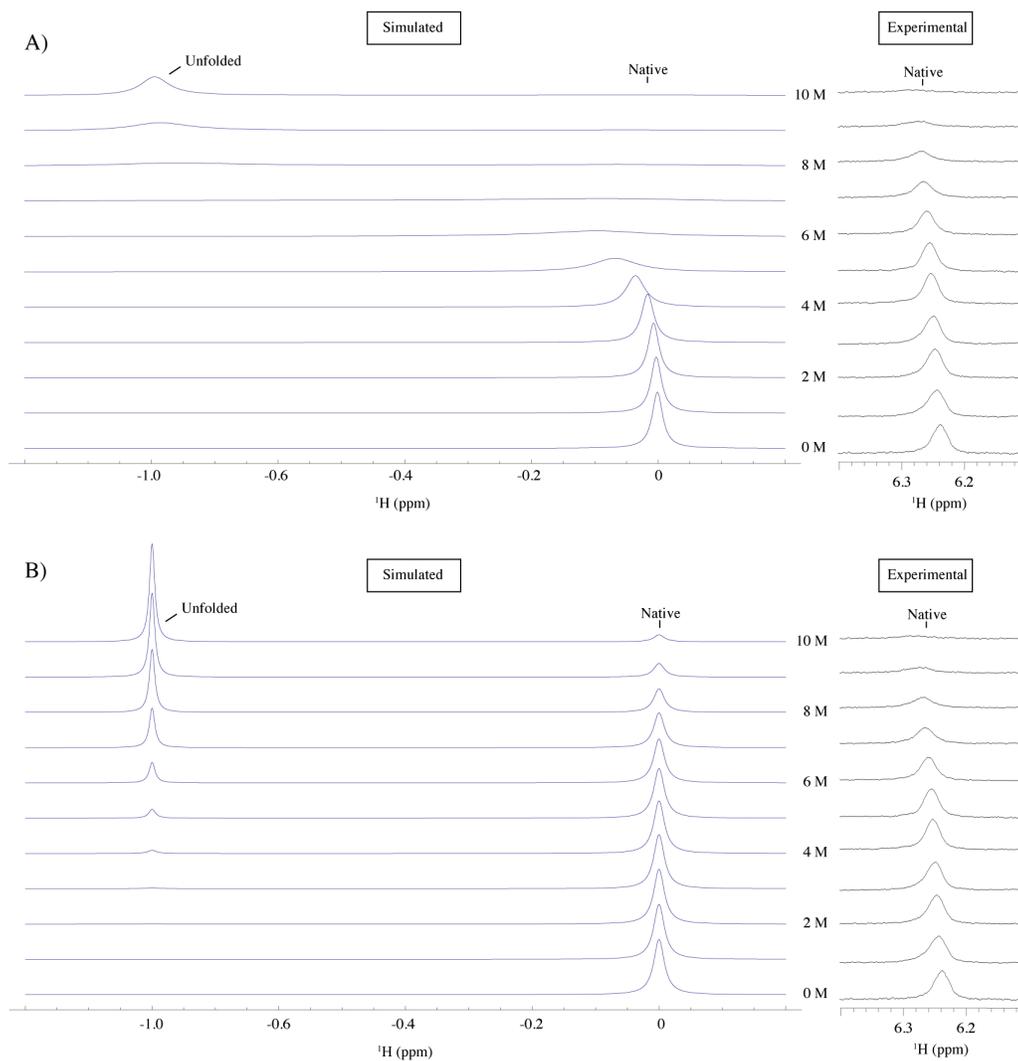


Fig. C-1 Lineshape analysis of Y218.H δ during the urea-induced denaturation of bPrP^C. A) The theoretical lineshape (blue) corresponding to fast exchange ($k_{N-U} = 1000 \text{ s}^{-1}$) between the native and unfolded states is plotted as a function of the urea concentration. In comparison, the acquired 1D ¹H NMR spectra are shown on the right (black). B) The theoretical lineshape (blue) corresponding to slow exchange ($k_{N-U} = 1 \text{ s}^{-1}$) is plotted as a function of the urea concentration, and the 1D ¹H NMR spectra shown on the right (black) confirmed the presence of slow exchange.

NOESY Building Curve Analysis

In general, the size of a NOE observed between two protons in a protein is dependant upon a number of factors; these include the rotational correlation time of the protein, internal motions of either proton, the distance between the two protons, whether or not the distance is modulated by the motions of the two partners, the efficacy of other relaxation pathways for either proton, and the mixing time of the NMR experiment. This is made even more complex if there is conformational exchange between two states, such as native and unfolded conformations, so that the observed NOE is now also a function of the fractional populations of the states involved and the rate of exchange between states.

To evaluate whether the observed NOEs of bPrP^C reflect the population of the native and unfolded forms during the urea denaturation experiment, we used a relaxation-matrix analysis to calculate the magnitude of the transferred NOE for exchange between two conformations as previously described (2,3). The geometry used for the conformations of the native form corresponds to the β -sheet region of bPrP^C (PDB ID 1dwy) including protons from L130 (H α , H β_1 , H β_2), G131 (HN), Y162 (H α , H β_1 , H β_2) and Y163 (HN). In order to create an unfolded conformation, the distance between the two β -strands was doubled to bring them apart from each other. Both native and unfolded forms were considered rigid. The input parameters for the relaxation matrix calculation were the inter-atomic distance matrix and the rotational correlation times for each conformation, the unfolding rate constant, the external relaxation contributions for each conformation (assumed equal for all protons), the spectrometer frequency, and the mixing time. The fraction of the native protein is given by:

$$\frac{1}{1 + \exp\{(-\Delta G^\circ_{N-U} + m[D])/RT\}} \quad \text{Eq. [1]}$$

Using the arbitrary parameters $\Delta G^\circ_{N-U} = 15$ kJ, $m = 3$ kJ mol⁻¹ M⁻¹, $k_{N-U} = 1$ s⁻¹, $\tau_N = 6$ ns, $\tau_U = 3$ ns, a spectrometer frequency of 800 MHz, and a mixing time of 0.1 s, the NOE L130.H α -162H α was calculated as a function of urea concentration

(Figure C-2). The calculated NOEs as a function of [urea] reflect the fractional populations from the simulated urea-induced unfolding curve, with a slight leftward shift due to the reduction of the NOE of the folded species due to chemical exchange. This supports our use of NOESY experiments to monitor the conversion from the native and to the unfolded state.

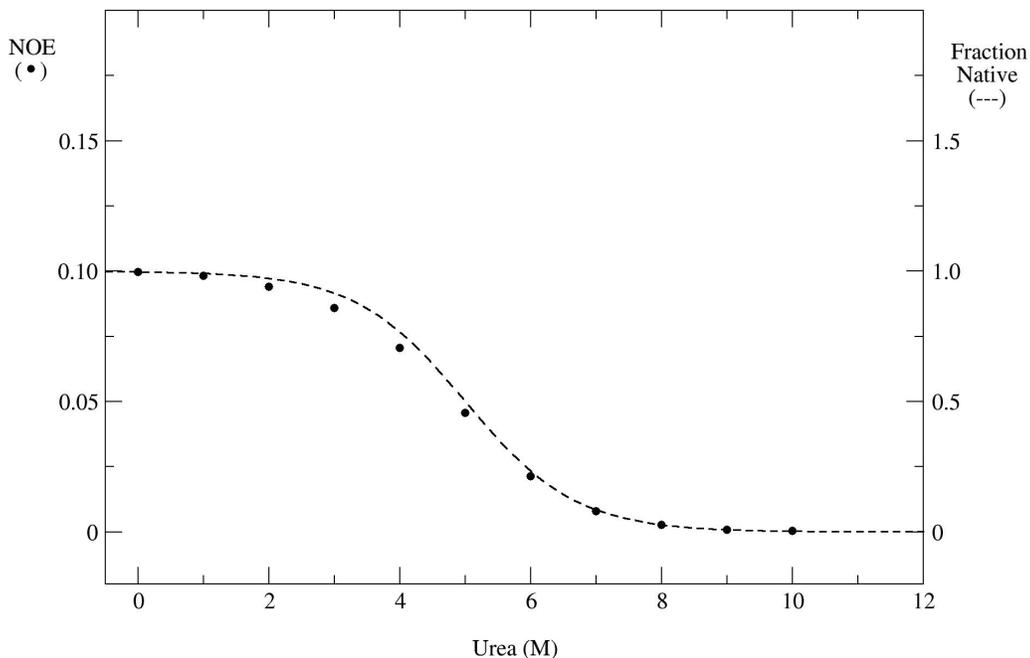
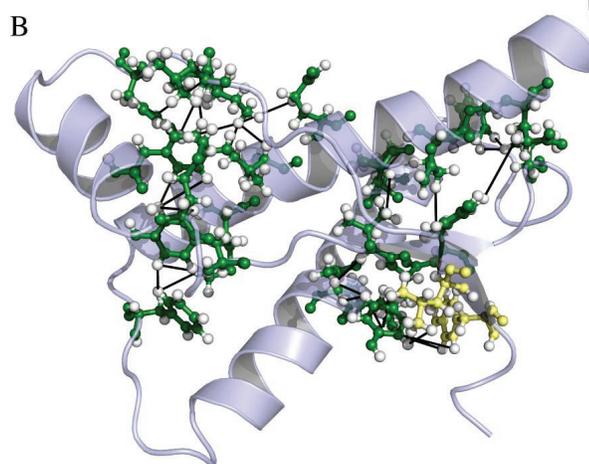
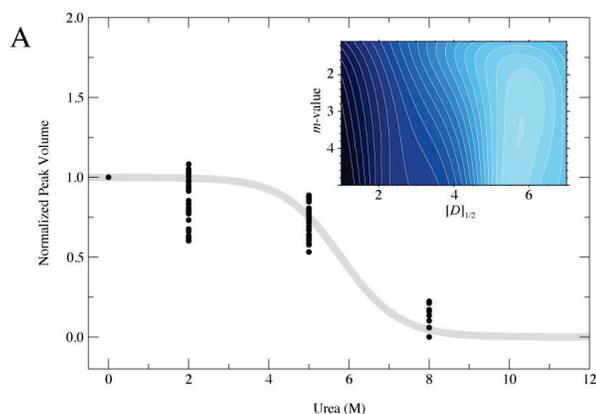


Fig. C-2 Simulation of the magnitude of the NOE (L130.H α -Y162.H α) as a function of the urea concentration. The fraction of the native conformation calculated from Eq. [1] is shown with a dash line for comparison.

Fig. C-3 Stable inter-residues NOEs of bPrP^C. A) Normalized NOEs cross-peak volumes as a function of the concentration of urea, for NOEs having a peak volume at 5 M urea larger than 50% of the original peak volume at 0 M urea (typically $[D]_{1/2} > 5\text{M}$). The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of $3.0 \text{ kJ mol}^{-1} \text{ M}^{-1}$ and a $[D]_{1/2}$ of 5.8 M. The grey line on the graph corresponds to a denaturation curve with these thermodynamics parameters.

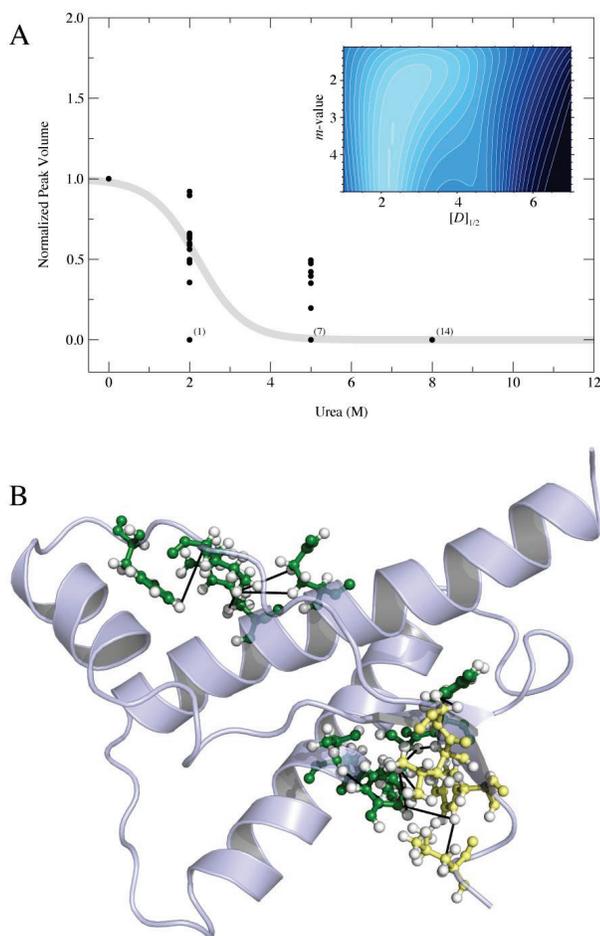


B) Cartoon representation of bPrP^C showing the stable inter-residue ^1H - ^1H NOEs contacts with black lines: Y150.H δ -Y157.H ϵ , F198.H δ -Y157.H ϵ , Y150.H ϵ -Y157.H ϵ , Y150.H ϵ -L138.H α , Y162.H δ -T183.H α , Y157.H δ -M154.H α , Y157.H ϵ -M154.H α , Y163.H ϵ -E221.H γ , Y163.H ϵ -Q217.H γ , Y162.H δ -T183.H γ_2 , F141.H ϵ -I139.H γ_1 , Y218.H ϵ -V166.H γ_2 , Y218.H δ -V166.H γ_2 , Y162.H ϵ -L130.H δ_2 , F141.H ζ -I139.H δ_1 , Y162.H δ -I182.H γ_2 , Y128.H δ -I182.H γ_2 , Y128.H ϵ -I182.H δ_1 , Y128.H ϵ -I182.H γ_2 , F141.H δ -I139.H γ_2 , F141.H ϵ -I139.H γ_2 , M206.H ϵ -F198.H δ , M206.H ϵ -Y157.H ϵ , I139.H δ_1 -Q212.H γ , V161.H α -T183.H γ_2 , I139.H δ_1 -V209.H γ_2 , C214.H α -V161.H γ_2 , C214.H β -V176.H γ_2 , V209.H α -I139.H δ_1 , Y162.H β -I182.H γ_2 .

Fig. C-4 Less stable inter-residues NOEs of bPrP^C. A) Normalized NOEs cross-peak volumes as a function of the concentration of urea, for NOEs having a peak volume at 5 M urea smaller than 50% of the original peak volume at 0 M urea (typically $[D]_{1/2} < 5\text{M}$).

The contour plot (inset) shows the sum of the square of the error for a global fit of the data yielding a m -value of 4.4 kJ mol⁻¹ M⁻¹ and a $[D]_{1/2}$ of 2.2 M. The grey line on the graph corresponds to a denaturation curve with these thermodynamics parameters,

and the numbers in parenthesis indicate the number of superimposed points. B) Cartoon representation of bPrP^C showing the stable inter-residue ¹H-¹H NOEs contacts with black lines: Y163.Hε-G131.Hα₂, Y162.Hα-L130.Hα, Y128.Hδ-L125.Hα, Y163.Hε-G131.Hα₁, F141.Hζ-I139.Hγ₂, Y162.Hδ-L130.Hδ₁, Y162.Hε-L130.Hδ₁, Y162.Hα-L130.Hδ₁, I139.Hδ₁-R208.Hβ, T183.Hα-I182.Hγ₂, Y162.Hδ-L130.Hδ₂, Y128.Hδ-I182.Hδ₁, I139.Hδ₁-Q212.Hβ, I139.Hδ₁-Q212.Hγ. The residues involved are represented in ball-and-sticks, with the heavy atoms are colored in green and the hydrogen atoms are colored in white.



REFERENCES

1. McKay RT, Tripet BP, Pearlstone JR, Smillie LB, Sykes BD (1999) Defining the region of troponin-I that binds to troponin-C. *Biochemistry* 38:5478-5489.
2. London RE, Perlman ME, Davis DG (1992) Relaxation-matrix Analysis of the transferred nuclear Overhauser effect for finite exchange-rates. *J. Magn. Reson.* 97:79-98.
3. Landy, SB and Rao, BDN (1989) Dynamical NOE in Multiple-Spin S systems Undergoing Chemical Exchange. *J. Magn. Reson.* 81:371-377.

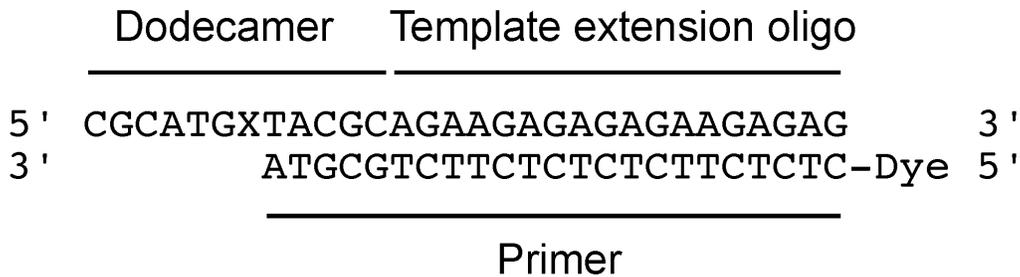
Appendix D – Supporting Information for Chapter IX

Primer Extension Analysis

The CDV-containing and control dodecamers were first separately ligated to an 18-mer template extension oligonucleotide [5'-phosphate-d(AGAAGAGAGAGAAGAGAG)-3'] in the presence of an IRDye700®-labelled primer [5'-IRDye700®-d(CTCTCTTCTCTCTTCTGCGTA)-3'] to create primer-template substrates (Figure S1A). Briefly, 21 nmol of template extension oligonucleotide, 21 nmol of CDV-containing or control oligonucleotide and 7 nmol of labelled primer were mixed, incubated at 55°C for 5 min and cooled to room temperature. One unit of T4 DNA ligase (Fermentas) was added and the reactions incubated overnight at 16°C. Negative control reactions without ligase were also prepared in parallel. Each reaction mixture was purified by phenol extraction, the aqueous layers desalted using G-50 MicroSpin columns (GE Healthcare), and the DNA collected by ethanol precipitation. After resuspension in 49 µl ddH₂O, primer extension reaction mixtures were prepared by the addition of polymerase buffer [30 mM Tris-HCl (pH 7.9), 5 mM MgCl₂, 70 mM NaCl, 1.8 mM dithiothreitol, 80 µg/ml bovine serum albumin], 50 µM of each dNTP and 25 ng/ml purified vaccinia virus DNA polymerase.¹ Reaction mixtures were incubated at 37°C and at 0, 1, 2, 5, 10 and 15 min, 10 µl aliquots were removed and added to 5 µl gel loading buffer [80% formamide, 10 mM EDTA (pH 8.0), 1 mg/ml bromophenol blue]. Reaction products were separated on a 15% denaturing polyacrylamide gel run at 50 W for 1 hr in Tris-borate-EDTA. The gel was transferred to cellophane and scanned using an Odyssey scanner (Li-Cor) at 169 µm resolution, medium quality and an intensity of 5.0.

FIGURES

A.



B.

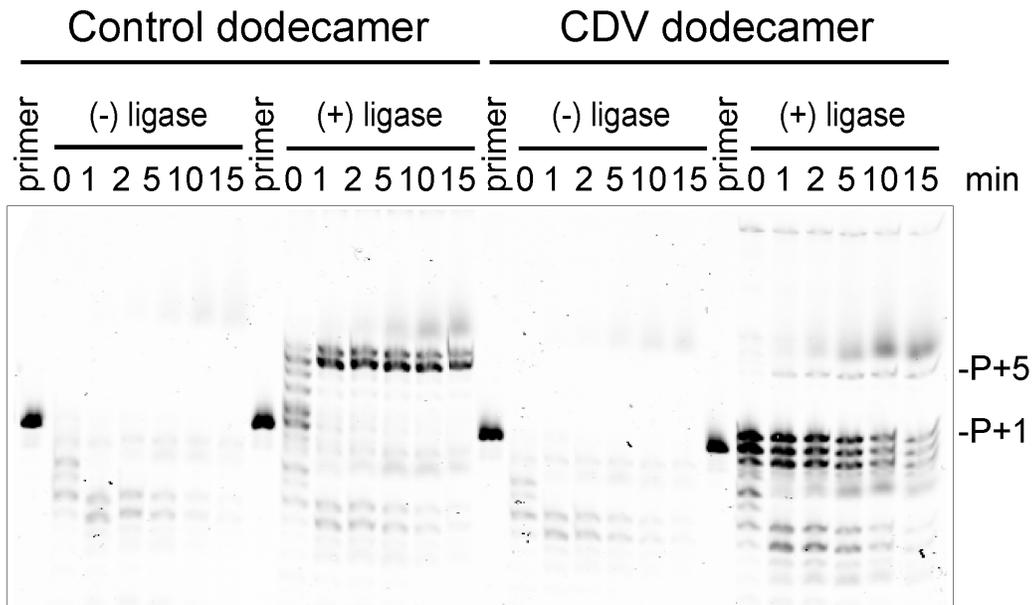


Fig. D-1. The presence of CDV in the template strand inhibits DNA synthesis across the drug lesion. (A) Primer-template substrates. “X” in the dodecamer indicates CDV; this residue is replaced by dCMP in the control DNA dodecamer. The dodecamers and template extension oligonucleotide are not ligated in the negative control substrates. (B) Primer extension assay. The primer-template substrates shown in (A) were incubated with vaccinia virus DNA polymerase and 200 μ M dNTPs at 37°C. Samples were taken at 0, 1, 2, 5, 10 and 15 min, stopped with gel loading buffer and the reaction products resolved by denaturing PAGE.

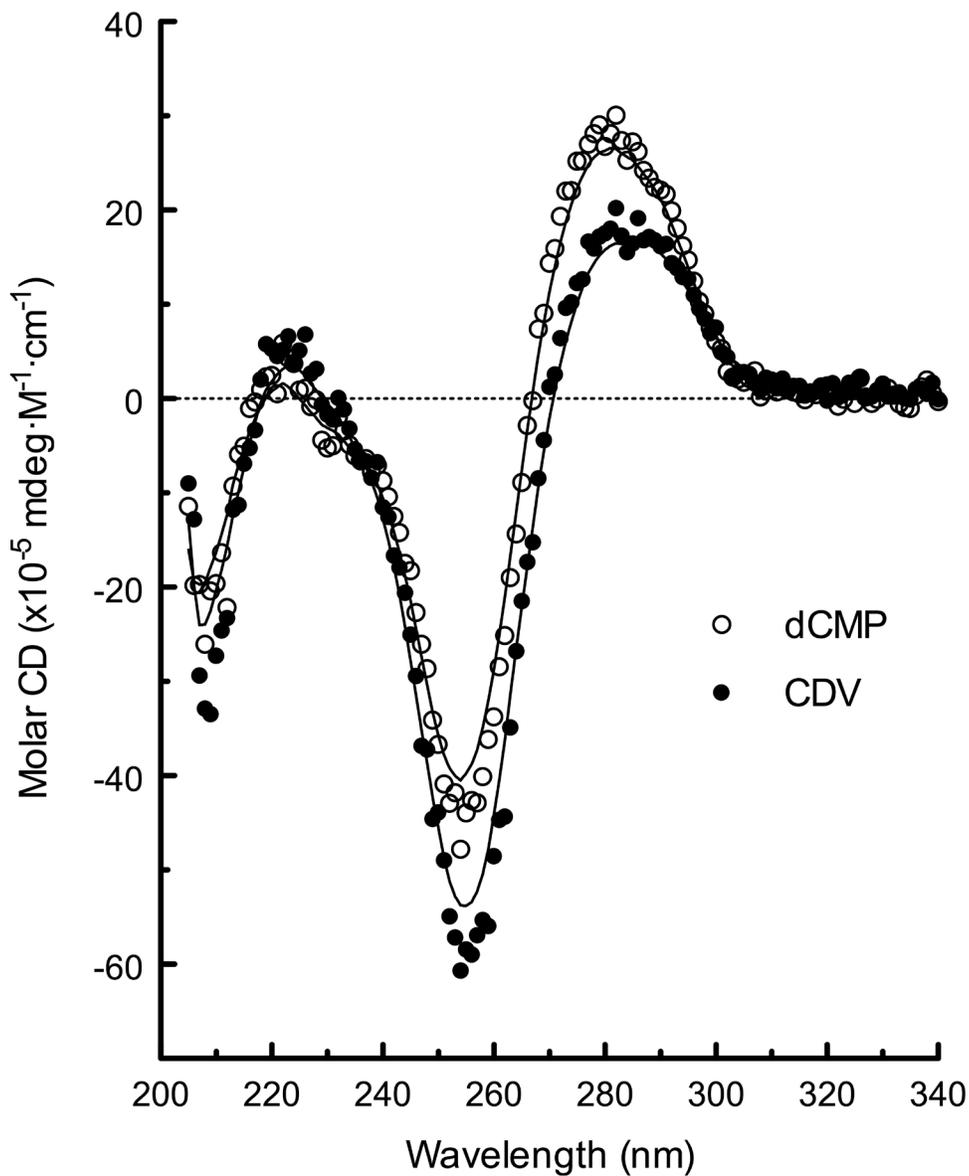


Fig. D-2. CD spectra of CDV- and dCMP-containing DNA duplexes at 20°C. Each sample was dissolved in 10 mM sodium phosphate (pH 7.3), 185 mM sodium chloride and scanned from 205 to 360 nm in 1-nm increments. A buffer baseline spectrum was also recorded and subtracted from the data presented here. The CDV-containing DNA exhibits spectral properties characteristic of a well-ordered B-DNA duplex, although some perturbation of the structure is illustrated by alterations in the peak intensities at 255 and 282 nm.

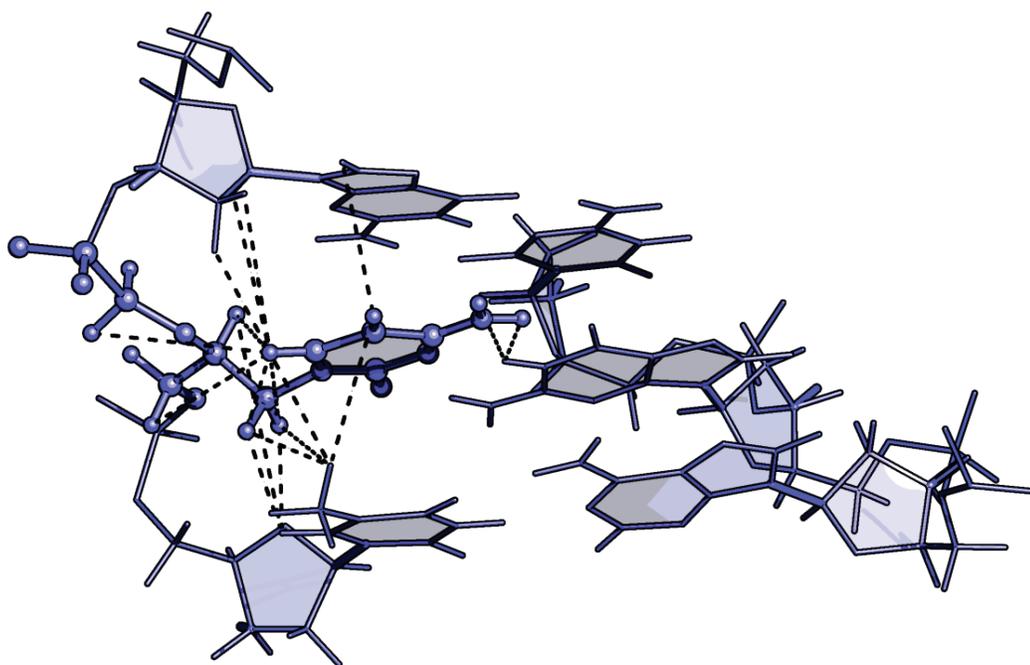


Fig. D-3. The position of X7 in the DNA duplex is defined by 18 NOE distances represented by black dashed lines: (7.H5, 6.H8), (7.H6, 6.H1'), (7.H6, 6.H2'), (7.H6, 6.H2''), (7.H6, 7.H4'), (7.H6, 7.H2'), (7.H6, 7.H3''), (7.H6, 7.H1'), (7.H6, 7.H1''), (8.H6, 7.H2'), (8.H6, 7.H1'), (8.H6, 7.H1''), (8.M7, 7.H1'), (8.M7, 7.H1''), (8.M7, 7.H5), (8.M7, 7.H6), (18.H1, 7.H42), (18.H1, 7.H41).

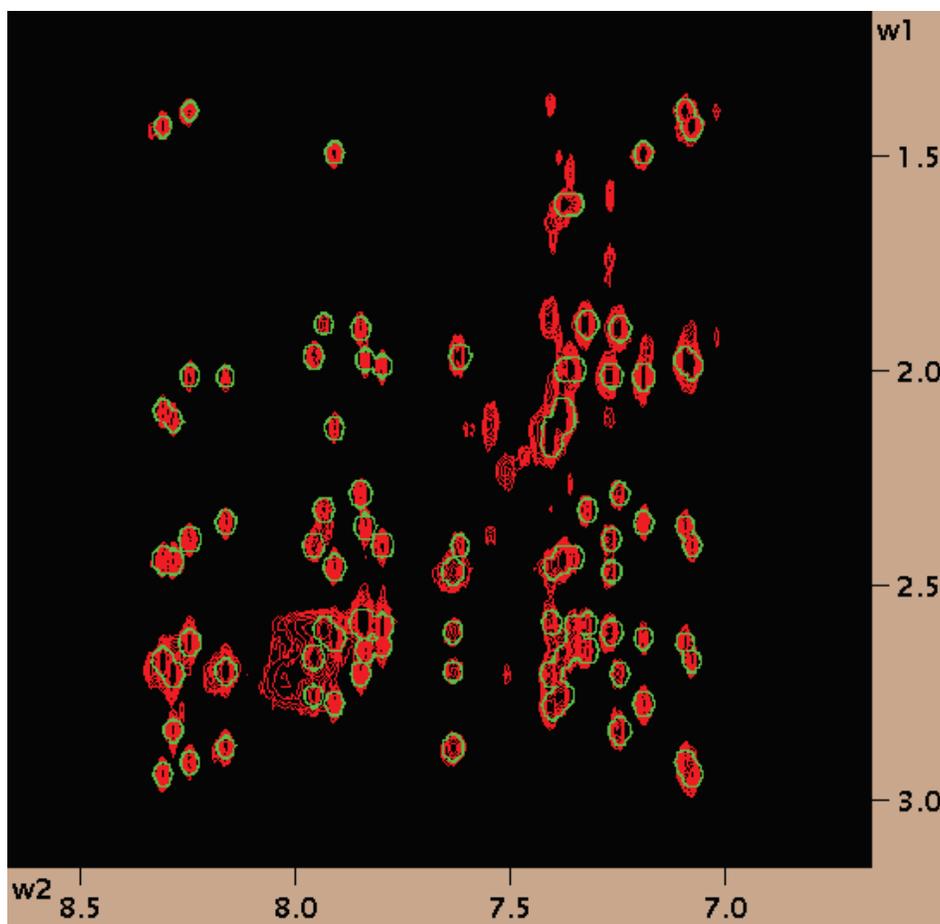


Fig. D-4. Portion of the back-calculated NOESY spectrum for the control DNA duplex. The observed spectrum is colored in red and the reconstructed one in green. Extra peaks in the observed spectrum correspond to unassigned resonances. The back-calculation was performed using a complete relaxation matrix analysis method (CORMA²) using a 80 ms mixing time and a theoretical isotropic correlation time of 2 ns. The low values of the sixth root squared R factor ($0.128 \pm 0.009 \text{ \AA}$ for the control DNA duplex, and $0.117 \pm 0.007 \text{ \AA}$ for the CDV DNA duplex) reflects the high level of similarity between the experimental and back-calculated spectra.³ The spectra overlay was generated using SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco).

REFERENCES

- (1) McDonald, W. F.; Traktman, P. *Protein Expr. Purif.* 1994, 5, 409-421.
- (2) Keepers, J. W.; James, T. L. *J. Magn. Reson.* 1984, 57, 404-426.
- (3) Thomas, P. D.; Basus, V. J.; James, T. L. *Proc. Natl. Acad. Sci. U.S.A.* 1991, 88, 1237-1241.

Appendix E– Troponin Complex

This appendix present the NMR spectra obtained for the skeletal troponin complex:

$[U\text{-}^2\text{H}, ^{13}\text{C}, ^{15}\text{N}] \text{sTnI}_{1-137}$

$[U\text{-}^2\text{H}] \text{sTnC}_{1-162}$

$[U\text{-}^2\text{H}] \text{sTnT-T2 Domain}$

MATERIAL AND METHODS

NMR sample preparation: sample TnI-1

$[U\text{-}^2\text{H}, ^{13}\text{C}, ^{15}\text{N}] \text{sTnI}$

100 mM KCl, 10 mM Imidazole, 0.3 mM DSS, 0.015% NaN_3 , 6.3% D_2O , pH = 5.

NMR sample preparation: sample TnI-2

10 mg of $[U\text{-}^2\text{H}, ^{13}\text{C}, ^{15}\text{N}] \text{sTnI} \cdot [U\text{-}^2\text{H}] \text{sTnC} \cdot [U\text{-}^2\text{H}] \text{sTnT}$

250 mM KCl, 6.25 mM CaCl_2 , 9.4 mM DTT, 0.3 mM DSS, 0.015% NaN_3 , protease inhibitor, 6.3% D_2O , pH = 6.7.

NMR sample preparation: sample TnI-3

+10 mg of $[U\text{-}^2\text{H}, ^{13}\text{C}, ^{15}\text{N}] \text{sTnI} \cdot [U\text{-}^2\text{H}] \text{sTnC} \cdot [U\text{-}^2\text{H}] \text{sTnT} = 20 \text{ mg of protein}$

250 mM KCl, 6.25 mM CaCl_2 , 9.4 mM DTT, 0.3 mM DSS, 0.015% NaN_3 , protease inhibitor, 6.3% D_2O , pH = 6.7. January 2008. For many reasons, the project was put on hold. The sample was kept at frozen at -20°C until July 2009.

NMR sample preparation: sample TnI-7

The NMR tube TnI-3 was broken in the freezer during the storage time. The protein was recycle, but the HSQC was never as good as it was before. Consequently, the sample was denatured in 6 M urea, and subsequently dialysed against: **1)** 20 mM Imidazole (pH 7.5, 4°C), 1 M NaCl, 1 mM CaCl_2 , 2 mM DTT, 0.1 mM NaN_3 , 0.1 mM PMSF, **2)** 50 mM Imidazole (pH 7.5, 4°C), 1 mM

CaCl₂, 2 mM DTT, 0.1 mM NaN₃, 0.1 mM PMSF, and finally **3)** 100 mM NH₄HCO₃, 0.1 mM DTT, protease inhibitors. The protein sample was then lyophilised and re-suspended in new filtered NMR buffer:

~10 mg of [*U*-²H, ¹³C, ¹⁵N]sTnI•[*U*-²H]sTnC•[*U*-²H]sTnT

250 mM KCl, 10 mM Imidazole, 6 mM CaCl₂, 10 mM DTT, 0.3 mM DSS, 0.015% NaN₃, protease inhibitor, 6% D₂O, pH ~ 6.

RESULTS

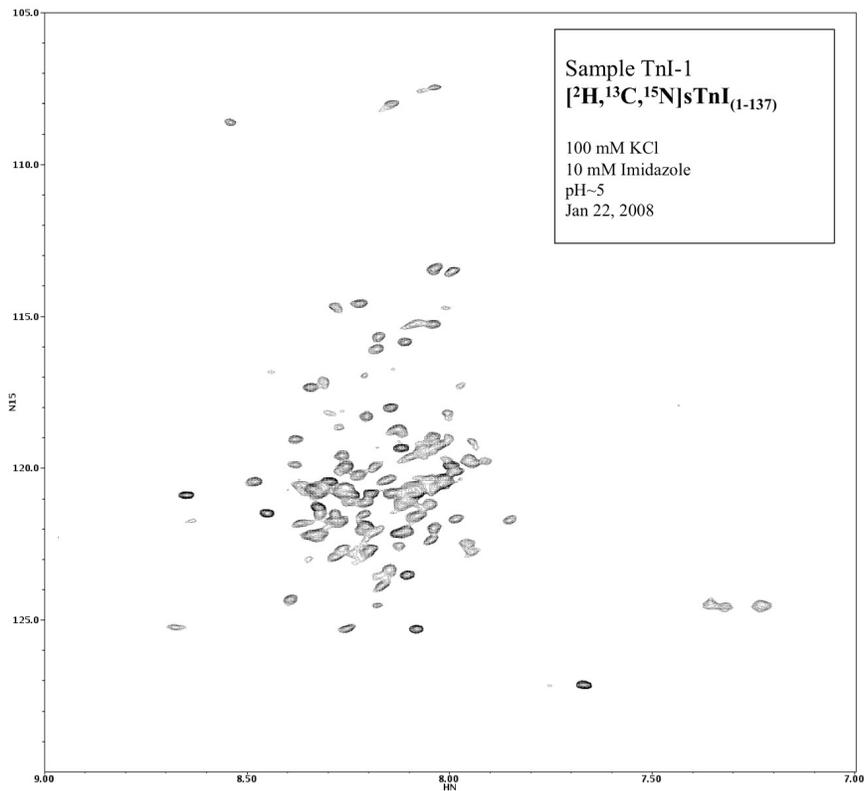


Fig. E-1 2D ¹H, ¹⁵N HSQC of [¹⁵N]sTnI₁₋₁₃₇. Sample TnI-1. The spectra of sTnI show the typical characteristics of an unfolded protein in the absence of its binding partners sTnC and sTnT.

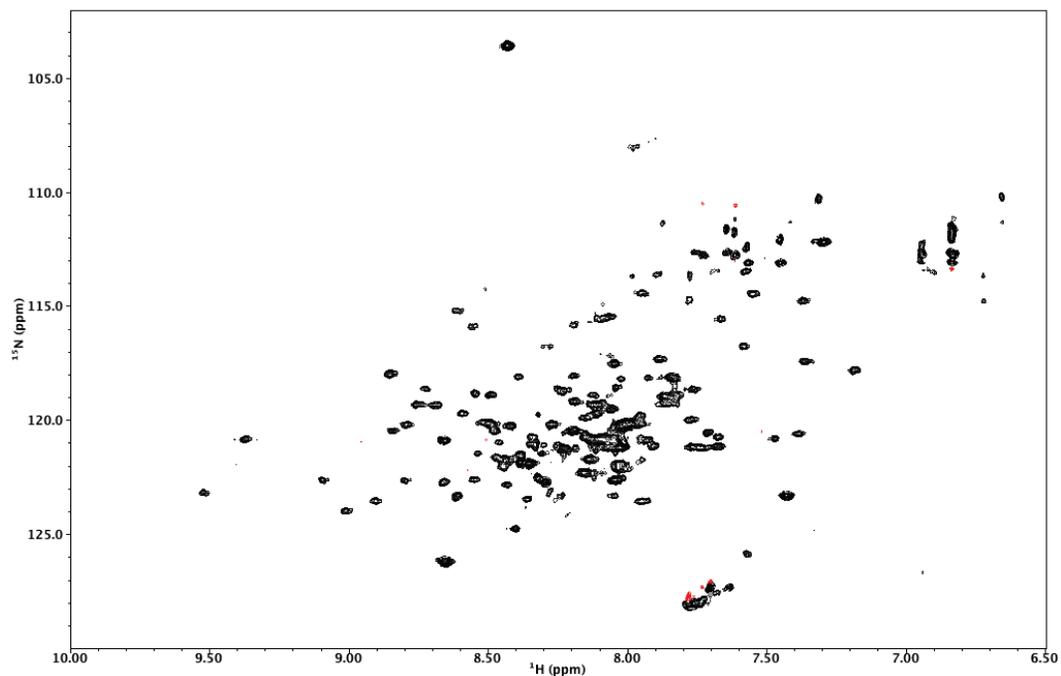


Fig. E-1 2D ^1H , ^{15}N TROSY HSQC acquired at 800 MHz. Sample TnI-2. Parameters: $nt = 16$, $ni = 256$, $dI = 1.5$ s, time = 3h46. The signal-to-noise ratio was too low to acquire good 3D experiments.

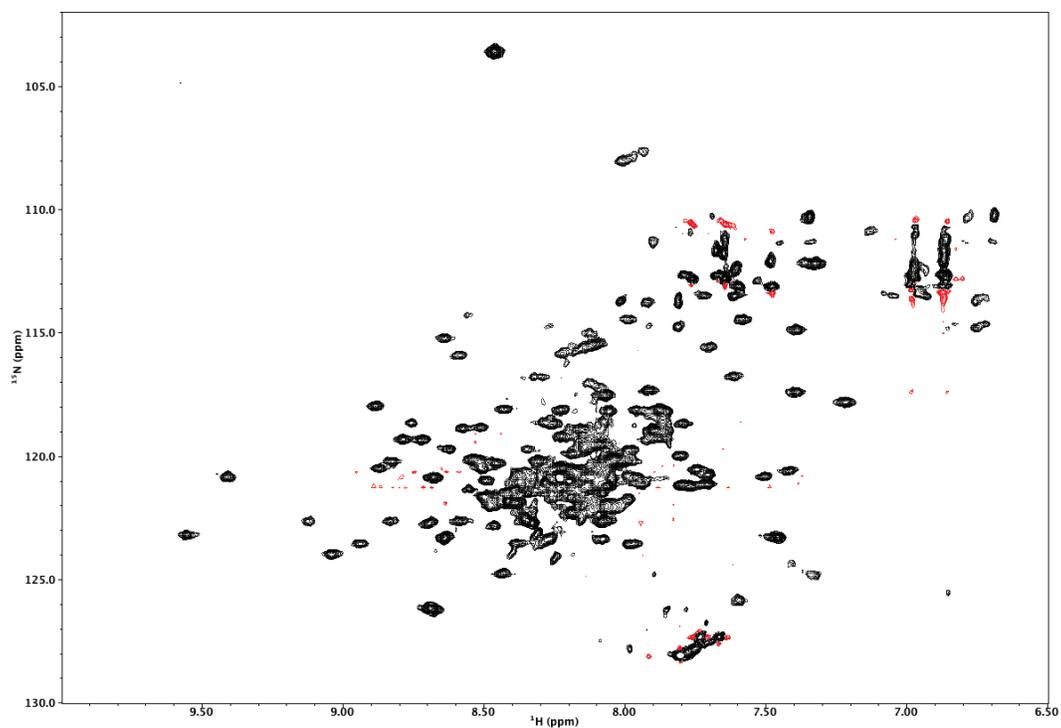


Fig. E-2 2D ^1H , ^{15}N TROSY HSQC acquired at 800 MHz B) Sample TnI-3. Parameters: $nt = 96$, $ni = 196$, $dI = 1.0$ s, time = 11h41.

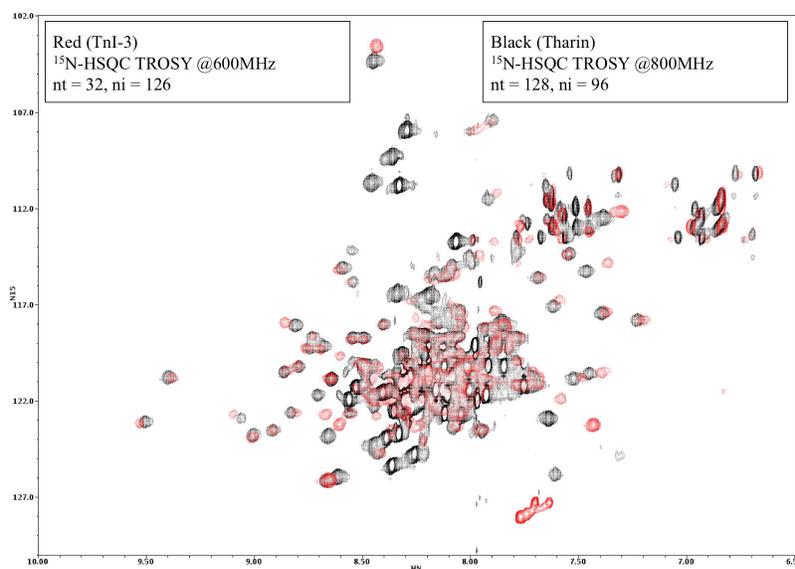


Fig. E-3 Superimposition of 2D ^1H , ^{15}N TROSY HSQC for sTnI₁₋₁₃₇ (red) and sTnI₁₈₂ (black) when part of the Tn complex. The figure shows the advantage of not having the flexible region 138-182 which dominate the HSQC (black).

Sample TnI-7

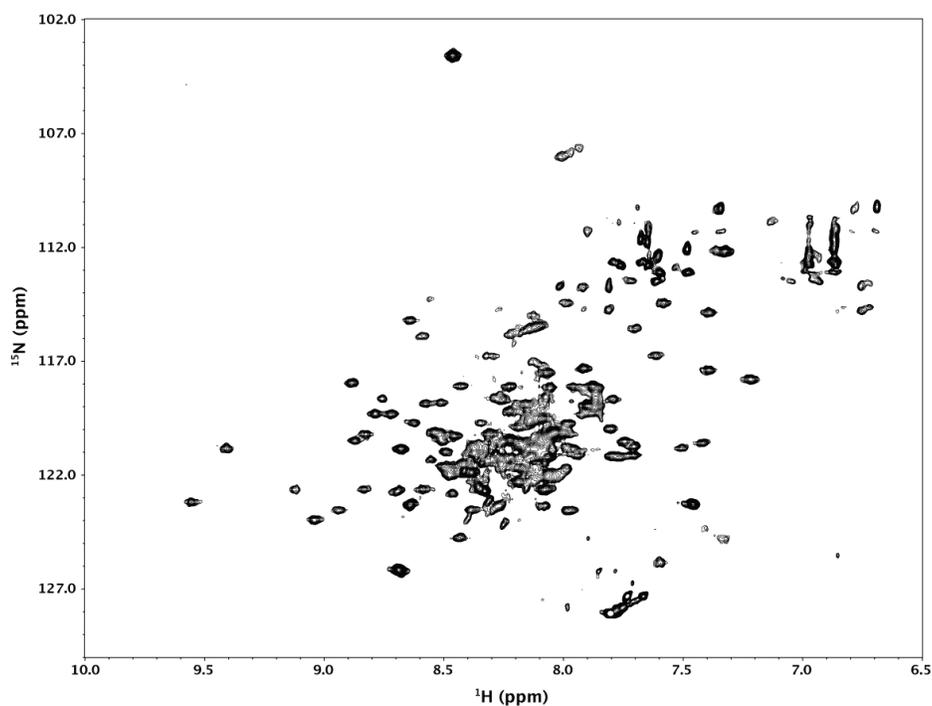


Fig. E-4 2D ^1H , ^{15}N TROSY HSQC acquired at 800 MHz B) Sample TnI-7. Parameters: $nt = 96$, $ni = 196$, $dI = 1.0$ s, time = 11h41.

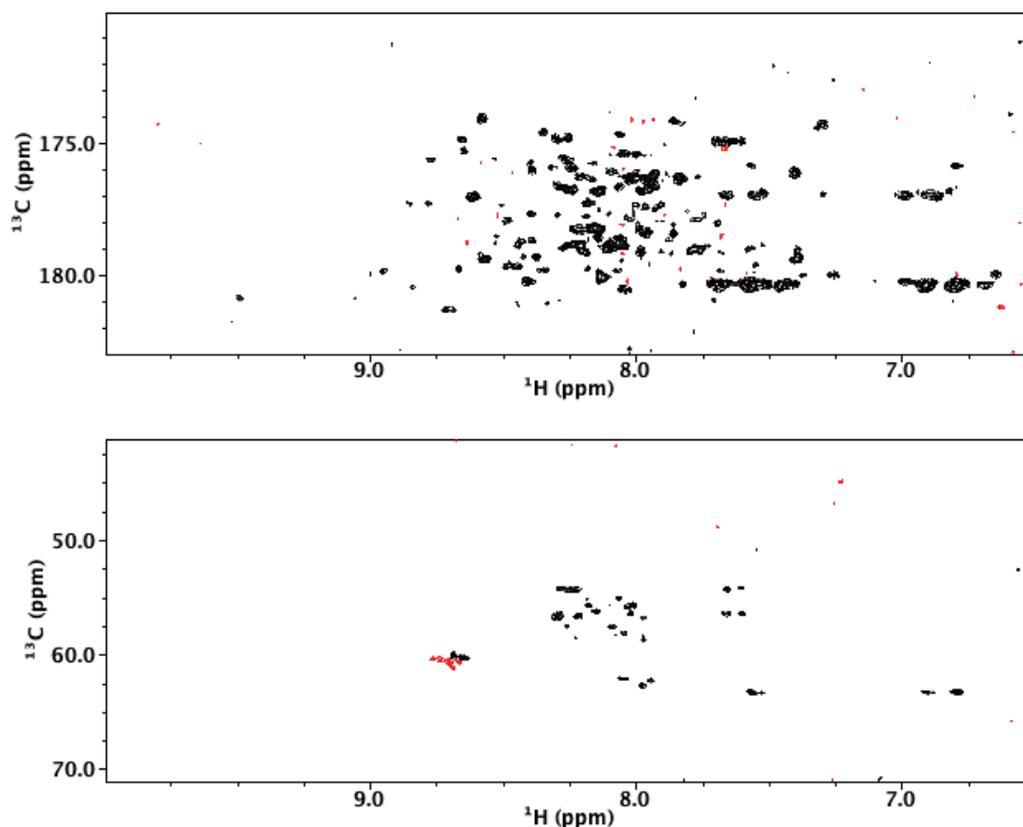


Fig. E-4 2D TROSY HNCO (top) and 2D TROSY HNCA of TnI-7. The HNCO spectra could have been used for chemical shift assignment, but the lack of resonances in the HNCA (and HNCOCA) made the backbone assignment impossible.

Eventually, the sample was given to Lewis Kay and his group at the University of Toronto in the hope that he would be able to acquire better 3D spectra on his 800 MHz with no cryo-probe (better for salty samples) and with his in-house pulse sequences instead of the ones we used (Biopack). They were never able to reproduce a good HSQC, suggesting that the sample was most likely degraded. The project came to an end.

CONCLUSION

The lessons one can take from this project: **1)** It is hard to acquire good NMR data on such a big complex (52 kDa). This supports the use of TnC-TnI chimera to study TnI structure and dynamics (Chapter VI). **2)** The complex was fully deuterated to avoid ^1H cross-relaxation, which allowed the acquisition of TROSY NMR spectra. **3)** The presence of high salt concentration (required to avoid aggregation) certainly didn't help the acquisition of NMR data (long pulse widths). **4)** The $^1\text{H}, ^{15}\text{N}$ TROSY HSQC were promising, but the acquisition of good 3D NMR spectra required for chemical shift assignment was never achieved.