

**NMR STRUCTURAL STUDIES ON THE
PERIPLASMIC DOMAIN OF
CitA and DcuS**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Vinesh Vijayan
aus VAIKOM, INDIA

Göttingen 2007

D7

Referent: Prof. Dr. Christian Griesinger

Korreferent: Prof. Dr. Martin Suhm

Tag der mündlichen Prüfung:

Publication lists

This thesis is based on the following papers:

Chapter 3

- {1} V. Vijayan and M. Zweckstetter, Simultaneous measurement of protein one-bond residual dipolar couplings without increased resonance overlap, *Journal of Magnetic Resonance*, **174**(2): 245–253, 2005.

Chapter 4

- {2} L. Pappalardo, and I.G. Janausch, V. Vijayan, E. Zientz, J. Junker, W. Peti, M. Zweckstetter, G. Unden, and C. Griesinger, The NMR structure of the sensory domain of the membranous Two-Component fumarate sensor (histidine protein kinase) Dcus of *Escherichia Coli*, *Journal of biological chemistry*, **278**(40): 39185–38188, 2003.
- {3} H. Kneuper, I.J. Janausch, V. Vijayan, M. Zweckstetter, V. Bock, C. Griesinger, and G. Unden, The nature of the stimulus and of the fumarate binding site of the fumarate sensor Dcus of *Escherichia Coli*, *Journal of biological chemistry*, **280**(21): 20596–20603, 2005.
- {4} M. Sevvana, V. Vijayan, M. Zweckstetter, S. Reinelt, D.R. Madden, R. Herbst-Immer, G.M. Sheldrick, M. Bott, C. Griesinger and S. Becker, Ligand-induced switch mechanism regulates signal transduction in sensor histidine kinase CitA, *submitted*, 2007.

Related papers :

- {5} S. Rumpel, A. Razeto, C. M. Pillar, V. Vijayan, A. Taylor, K. Giller, M.S. Gilmore, S. Becker, and M. Zweckstetter, Structure and DNA-binding properties of the cytolysin regulator Cylr2 from *Enterococcus Faecalis*, *EMBO journal*, **23**(18): 3632–3642, 2004.

- {6} M. Bayrhuber, V. Vijayan, M. Ferber, R. Graf, J. Korukottu, J. Imperial, J.E. Garrett, B.M. Olivera, H. Terlau, M. Zweckstetter, and S. Becker, Conkunitzin-S1 as the first member of a new Kunitz-type neurotoxin family. Structural and functional characterization, *Journal of biological chemistry*, **280**(25) : 23766–23770, 2005.
- {7} S. Rumpel, H.Y. Kim, V. Vijayan, S. Becker, and M. Zweckstetter, Backbone resonance assignment of the homodimeric, 35 Kda chaperone Cest from enteropathogenic *Escherichia Coli*, *Journal of Biomolecular NMR*, **31**(4) : 377–378, 2005.
- {8} P. Montaville, H.Y. Kim, V. Vijayan, S. Becker, and M. Zweckstetter, $^1\text{H}^N$, ^{15}N , and ^{13}C Resonance assignment of the C2a Domain of Rabphilin3a, *Journal of Biomolecular NMR*, **36**(5) : 20, 2006.
- {9} J. Korukottu, M. Bayrhuber, P. Montaville, V. Vijayan, Y.S. Jung, S. Becker, M. Zweckstetter, Fast high-resolution protein structure determination by using unassigned NMR data, *Angewandte Chemie International Edition English*, **46**(7) : 1176–1179, 2007.
- {10} J. Korukottu, A. Lange, V. Vijayan, R. Schneider, O. Pongs, S. Becker, M. Baldus, and M. Zweckstetter, Conformational plasticity in ion channel recognition of a peptide toxin. *to be communicated*, 2007.

Acknowledgments

Now that I have reached the end of a long journey in crafting this thesis, its time to express my gratitude to all those who have been a part of this travel. I have constantly wondered as to how I would feel at this exact moment ever since I landed in Germany in the fall of 2002 (I was then naive, full of hope, and optimistic). A rather quick four and half years later, as I write this, I am filled with mixed emotions. In retrospect, the last four and a half years in my life has been credible in that it has been intellectually satisfying. Thanks to all the help I received. On the other hand, I would have been furthermore at ease if I had been little more systematic. It should have definitely helped me overcome the sleepless nights I have had owing to the incredible stress.

I express my profound gratitude to Prof. Christian Griesinger, for his excellent guidance and support throughout the course of my thesis. The discussions we had, were immensely stimulating and thought provoking. I am indebted to him for the affection he showered on me and it was a great pleasure to work with him.

I would like to thank Dr. Markus Zweckstetter, my group leader, for his help with many of my projects, for his immense faith and confidence in me, in spite of all my slip-ups.

I thank my collaborators, Prof. Gotfried Uden, Prof. Micheal Bott and their group, in DcuS and CitA projects. I owe my thanks to Dr. Stefan Becker of our molecular biology lab for his outstanding guidance in preparing excellent protein samples for my NMR measurements. A thanks doesn't seem sufficient to Dr. Stefan Becker and Karin Giller for the stable CitAP protein they produced. But it is said with lots of appreciation.

The road to my graduation has been long and winding. There are quite a list of people with whom I have had the privilege to work during this 4.5 years. I am grateful to all the

present and former members of our group, who have contributed in every minute way to make my stay in the group a very memorable one. Mere words are insufficient to express my thanks to Pierre Montaville, who is my trusted friend, my counselor and a brotherly figure, all rolled into one. It was a pleasure working with him throughout the entire tenure of my Ph.D. It was through his great assistance that I managed to overcome my inhibition for biological systems. A special thanks to Pierre for having taken a close look at both the english grammar and scientific content, correcting both and offering suggestions for improvement. Many thanks to Jung Sang Jung. He made my first three years of stay in Göttingen a pleasant and memorable one. The valuable time he spent with me for long chocolate and coffee breaks especially during the weekends will always be treasured.

During my entire stay, I was fortunate enough to share my work space with wonderful people: Dirk Lennartz, my first room mate. If not for his extended support and help, getting myself adapted to German environment would have been more difficult. Sigrun Rumpel, for her timely coffee breaks that drove away my boredom. Monika Bayrhuber, along with Sigrun for the innumerable translations they patiently did. Daniela Fischer and Julien Orts for their pleasant company. Many thanks to all of them. I had the great pleasure of working together with them in few of their projects and I appreciate their feed backs in my projects as well.

And then there are all the people who have made Göttingen a very special place over this four years. Nils, Marco, Fernando, Kerstin, Minkyu, Hai-Young, Pinar, Dirk Bockelmann, Christophe, Nicolas, Jochen, Devan, Raghav, Jegannath, Peter, Valerie, Lukasz, Karel, Jörg, Edith, Melanie and the rest of my friends in our group. I acknowledge their friendly affection, especially Nils, Marco, Fernando, Kerstin, Minkyu and Hai-Young, for lending a patient ear to all my problems and for their warm support all along the way. Thanks to Dr. Donghan Lee for his interesting discussions on NMR and other topics. I also like to specifically thank, Dirk Bockelmann, Christophe, Nils, Minkyu and Lukasz for their kind help in spectrometer maintenance.

My sincere thanks to Mrs. Silberer, secretary to Prof. Griesinger for her ever helping attitude in dealing with bureaucratic and other paper works.

I am indebted to Prof. Dr. Martin Suhm for being the Korreferent of this thesis work. In addition I would like to thank Prof Grubmüller for kindly accepting the examiner role in my Rigorosum examination. I would also like to thank Prof. Diederichsen, Prof. Buback and Prof. Sheldrick for their kind consent to be a member of my Prüfungskommission.

I sincerely thank Dr. Thomas Jovin and Dr. Donna Arndt-Jovin for the engaging discussions, we have had, although few in numbers. I gratefully acknowledge Dr. Reinhard Klement at the Department of Molecular biology, for his well-timed help in fixing my computer, when it failed to cooperate just a week before the D-day. Many heartfelt thanks to Prof. Elizabeth Jares-Erijman. It was through her that my wife and I got introduced to the delicious world of “Dulce de Leche”. Eversince, I have been an addict to it, which I don’t regret.

My heartfelt thanks to all my buddies outside NMR department Rebecca, Harshad, Guy and especially to my Argentinean ”amigos dulces” Fito, Valeria, Carlos, Marianela, Claudio, Sole C, Sole G for the copious supplies of Dulce de Leche they provided.

I would like to thank my cousin Nisanth for his encouragement during my PhD, Shankarayathri for her timely help, particularly during the drafting of my thesis. A special word of thanks to my junior C.Subbu, for the unconditional love he showers on me, and whom I consider as my younger brother.

I dedicate this thesis to my father and my mother whose unlimited love, patience has seen me through all the frustrations of being alone during the first half of my PhD tenure. I thank my sister Veena and her husband Shyam for their understanding, support and affection. Last but not least, my wife Shyamala, love of my life. Without her, I would not be where I am today.

Zusammenfassung

Zwei-Komponenten regulatorische Systeme sind die häufigsten Systeme für transmembrane Signaltransduktion in Bakterien und spielen eine Hauptrolle bei der zellulären Adaptation an die Bedingungen der Umwelt und Stress. Sie bestehen aus zwei verschiedenen Proteinen, einer sensorischen Histidinkinase, die normalerweise in der Membran lokalisiert ist, und einem verwandten Antwortregulator im Cytoplasma. Für diese Systeme gibt es eine Fülle von molekularbiologischen Studien. Trotzdem sind keine Strukturinformationen über die transmembrane Signaltransduktion vorhanden. Das Ziel dieser Untersuchung war es, Informationen über die Struktur und Dynamik der Signalerkennung und -transduktion von der periplasmatischen sensorischen Domäne über den membranständigen zwei-Komponenten Sensor in die cytoplasmatische Domäne zu erhalten. In dieser Arbeit werden Strukturuntersuchungen mittels NMR an der periplasmatischen Domäne von zwei Histidinkinasen präsentiert.

DcuS und CitA sind bakterielle Sensorhistidinkinasen, die eine transmembrane Domäne besitzen. Sie sind Teil eines zwei-Komponenten Signaltransduktionssystems, das den Transport und Metabolismus von Di- und Tri-Carboxylaten in Abhängigkeit ihrer Konzentration in der Umgebung regulieren. Ihre periplasmatischen Domänen (DcuS-PD und CitAP) sind homolog, haben eine PAS-Domäne und eine Bindungsstelle für die Carboxylate. CitA fungiert als ein hochspezifischer Citratrezeptor während DcuS von einer Reihe von C4-Dicarboxylaten wie Fumarat und Succinat stimuliert wird. Als ein erster Schritt in Hinblick auf die Aufklärung des Signaltransduktionsprozesses wurde die Lösungs-NMR-Struktur der periplasmatischen Domäne von DcuS gelöst. Die Struktur wurde mit residualen dipolaren Kopplungen (RDCs), die über eine neuartige Strategie zur simultanen Messung von RDCs mit minimalem Resonanzüberlap gemessen wurden, verfeinert. Die Bindungstasche von DcuS-PD für einige C4 Di-Carboxylate wurde mittels ^{15}N - ^1H HSQC basierter Titrations definiert. Der Einfluss der Ligandenbindung an DcuS-PD war schwach. Weder Veränderungen

der chemischen Verschiebungen noch Anstieg der Signalintensitäten für Reste außerhalb der Bindungstasche wurden beobachtet. Deshalb blieb der Mechanismus der Signaltransduktion ungewiss.

Lösungs-NMR-Strukturen von CitAP konnten aufgrund starker Linienverbreiterung, die in den NMR-Spektren beobachtet wurde, nicht gelöst werden. Konformationeller Austausch war der Hauptgrund der Linienverbreiterung. Die Kristallstrukturen der citrat-freien und -gebundenen Form von CitAP konnten aufgeklärt werden. Hauptunterschiede wurden in der Citratbindungsregion und in der C-terminalen Region des Proteins beobachtet. Zusätzlich veränderten sich die chemischen Verschiebungen und die HetNOE-Werte in diesen Teilen des Proteins stark. In der citrat-gebundenen Struktur wurde ein Na^+ -Ion zwischen die N-terminale Helix und die β -Faltblätter gesetzt. Das wurde auch durch NMR-Titrations bestätigt. Damit könnte CitAP in Lösung sowohl an der Erkennung von Citrat als auch von Na^+ beteiligt sein. Überraschenderweise passen die für citratfreies CitAP gemessenen RDCs besser zu der citratgebundenen Struktur von CitAP. Das deutet darauf hin, dass in Lösung eine vorgeformte Bindungstasche von CitAP vorliegt. Nichtsdestotrotz ermöglichten die spezifischen strukturellen Unterschiede zwischen der citratfreien und den -gebundenen Strukturen den Vorschlag eines Modells für den Mechanismus der Signaltransduktion. Dieses Modell passt zu den verfügbaren NMR-Daten und ist auch ähnlich zu dem für Aspartatsensoren beschriebenen Mechanismus der Signaltransduktion .

Abstract

Two-component regulatory system represent the most frequent system for transmembrane signaling in bacteria and play a major role in the cellular adaptation to environmental conditions and stress. They consist of two separate proteins, a sensory histidine protein kinase which is located typically in the membrane, and a cognate response regulator in the cytoplasm. Despite the wealth of molecular biological studies in these systems, no structural informations are available on the signal transduction by these systems across the membrane. The aim of the study was to gain structural and dynamic information on signal perception and signal transduction from the periplasmic sensor domain of the two component membraneous sensor into the cytoplasmic domain. In this thesis NMR structural studies on the periplasmic domain of two histidine kinase are presented.

DcuS and CitA are bacterial membraneous sensory histidine kinases. They are part of a two component signal transduction systems that regulate the transport and metabolism of di- and tri-carboxylates in response to their environmental concentration. Their periplasmic domains (DcuS-PD and CitAP), are homologous, share a PAS fold, and contain the binding site for the carboxylates. CitA works as a highly specific citrate receptor whereas DcuS uses a wider range of C4 dicarboxylates like fumarate, succinate etc as stimulus. As a first step to understand the signal transfer process, the NMR solution structure of periplasmic domain of DcuS was determined. The structure was refined with residual dipolar couplings (RDCs), measured using a novel strategy for simultaneous measurement of RDCs with minimum resonance overlap. The binding pocket of DcuS-PD for C4 di-carboxylates was defined using ^{15}N - ^1H HSQC based titrations. The effect of the ligand binding to DcuS-PD was weak. No chemical shift changes or intensity increase for residues were observed outside the binding pocket and hence the signal transduction mechanism remained undetermined. Therefore the

sensory domain of CitAP which binds citrate more specifically was studied to obtain a better understanding of the conformational changes that lead to signal transduction.

The NMR solution structures of CitAP could not be determined because of the large number of missing peaks due to severe line broadening observed in the NMR spectra. Conformational exchange was the major cause of line broadening. However the X-ray structures of citrate free and bound form of CitAP could be determined. The major conformational changes were observed in the citrate binding region and in the C-terminal region of the protein. Large chemical shift changes and Het-NOE values were also observed in these parts of the protein. In the citrate bound structure, a Na⁺ ion was tentatively localized between N terminal helix and the β sheets. This was also confirmed by NMR titrations. Hence CitAP may be involved in sensing both citrate and Na⁺ ion in solution. Surprisingly the RDCs measured for citrate free CitAP fit better with citrate bound structure of CitAP. This indicates a pre-formed binding pocket of CitAP in solution. Nevertheless, the specific structural differences between the citrate free and bound structures allowed to formulate a model for the mechanism of signal transduction. This model is consistent with available NMR data and also very similar to the signal transduction mechanism described for aspartate sensors.

Table of Contents

	Page
Publication lists	i
Zusammenfassung	vii
Abstract	ix
List of Abbreviations	xxi
CHAPTER	
1 Introduction	1
1.1 General Introduction	1
1.2 Backbone chemical shift assignment	2
1.2.1 HNCO	5
1.2.2 HN(CA)CO	6
1.2.3 HNCACB	7
1.2.4 CBCA(CO)NH	8
1.2.5 Assignment strategy	8
1.3 Residual Dipolar Couplings (RDC)	9
1.3.1 Alignment tensor determination for a rigid molecule	11
1.3.2 Measurement of the RDCs	14
2 Materials and Methods	15
2.1 Materials	15
2.1.1 Ubiquitin	15
2.1.2 Periplasmic domain of DcuS (DcuS-PD)	15
2.1.3 Periplasmic domain of CitA (CitAP)	16

2.1.4	Equipment	17
2.1.5	Software	17
2.2	NMR spectroscopy	17
2.2.1	NMR experiments	18
2.2.2	Chemical shift assignment	18
2.2.3	Backbone chemical shift assignment	18
2.2.4	Secondary structure determination	21
2.2.5	Residual dipolar couplings (RDCs)	21
2.2.6	NMR relaxation data	23
2.2.7	^{15}N - ^1H steady state NOE	24
2.2.8	Measurement of rotational correlation time (τ_c)	24
2.2.9	Chemical shift mapping of binding surfaces	27
3	Simultaneous measurement of protein one-bond residual dipolar couplings without increased resonance overlap .	29
3.1	Abstract	29
3.2	Introduction	30
3.3	Description of the pulse sequences	31
3.3.1	TROSY-HNCO experiment	32
3.3.2	CBCA(CO)NH experiment	33
3.4	Data collection	36
3.5	Results and discussion	37
3.5.1	Sensitivity consideration for CBCA(CO)NH experiment	37
3.5.2	Sensitivity consideration for TROSY-HNCO experiment	39
3.5.3	Error estimation	41
3.5.4	Correlation of measured RDCs with structure	43
3.6	Conclusions	43
4	NMR structural studies on the periplasmic domain of DcuS and CitA	45
4.1	Introduction	45

4.1.1	Histidine Kinase (HK)	46
4.2	Periplasmic domain of the sensory domain of the two component fumarate sensor DcuS	51
4.2.1	Material and Methods	52
4.2.2	Results and discussions	54
4.2.3	Structure of DcuS-PD	54
4.2.4	Binding studies on DcuS-PD	57
4.2.5	Conclusion	62
4.3	Periplasmic domain of the sensory domain of the two component citrate sensor CitA	64
4.3.1	Introduction	64
4.3.2	Materials and Methods	66
4.3.3	Results and Discussion	69
4.3.4	X-ray structure of citrate bound form of CitAP without molybdate	72
4.3.5	X-ray structure of citrate free form of CitA	76
4.3.6	NMR studies on the citrate free form of CitAP	77
4.3.7	NMR studies on the citrate bound form of CitAP	80
4.3.8	Sodium binding to CitAP	82
4.3.9	Residual dipolar coupling analysis	83
4.3.10	Comparison of citrate free CitAP and citrate bound CitAP using X-ray and NMR	86
4.3.11	Mechanism of Signal Transduction	91
4.3.12	Comparison of structures of DcuS-PD and CitAP	93
4.3.13	Conclusion	96
5	Summary and outlook	99
	References	103
	APPENDIX	
A	Time evolution of product operators	121

B	Chemical shift assignments and residual dipolar couplings.	125
B.1	Chemical shift Assignments	125
B.2	Residual dipolar coupling	130
C	Bruker Pulseprogramme and MATHEMATICA scripts	151
C.1	MATHEMATICA scripts	159

List of Figures

1.1	Examples of coherence transfer pathways	3
1.2	Spin system of a peptide backbone	4
1.3	Pulse sequence for HNCO experiment	5
1.4	Magnetization transfer pathway for HNCO experiment	6
1.5	Magnetization transfer pathway for HN(CA)CO experiment	7
1.6	Magnetization transfer pathway for HNCACB experiment	8
1.7	Magnetization transfer pathway for CBCA(CO)NH experiment	9
1.8	Coordinate system for internuclear vector orientation with external magnetic field	12
2.1	Pulse sequence for measuring cross correlated relaxation rate of ^{15}N	25
3.1	Pulse scheme of the modified 3D TROSY-HNCO experiment for simultaneous measurement of $^1J_{C'N}$ and $^1J_{NH}$ couplings	34
3.2	Pulse scheme of the 3D CBCA(CO)NH quantitative J_{CH} and $J_{C\alpha C'}$ experiment	35
3.3	Selected region of TROSY-HNCO spectrum	37
3.4	Selected region of CBCA(CO)NH spectrum	38
3.5	Correlation between observed dipolar couplings and values back-calculated from the crystal structure of ubiquitin	42
3.6	Correlation of dipolar and J couplings measured for analysing systematic errors	42
4.1	Two component system	46
4.2	Histidine kinase	47
4.3	The kinase core of Histidine Kinases	47
4.4	Structure of PYP PAS domain	49
4.5	DcuS-DcuR system	52
4.6	Secondary structure of DcuS-PD	54

4.7	Solution structure of DcuS-PD	55
4.8	Comparison of the secondary structure of DcuS-PD and PYP	56
4.9	Plot of Het-NOE values measured for DcuS-PD	57
4.10	Ratio of peak intensities in fumarate titration with DcuS-PD	58
4.11	Chemical shift changes in tartrate titrations with DcuS	58
4.12	Structure of DcuS-PD with residues most affected by fumarate binding and tartrate binding.	59
4.13	Electrostatic surface potential of DcuS-PD	59
4.14	Comparison of the amino acid sequences of the periplasmic sensor domains of C4-dicarboxylate or citrate sensory histidine kinases	61
4.15	HSQCs of DcuS-PD mutants	63
4.16	The cit regulon of <i>Klebsiella pneumoniae</i>	64
4.17	X-ray structure of GJ dimer of CitAP in complex with citrate and molybdate	65
4.18	^{15}N - ^1H HSQC spectra of different construct of CitAP used in the study and the ^{15}N - ^1H HSQC spectrum of the CitAP shorter construct at a salt concentration of 300 mM	67
4.19	Overlay of ^{15}N - ^1H HSQC spectra of molybdate titration	71
4.20	^{15}N - ^1H HSQC spectrum of citrate free CitAP with assignments	71
4.21	^{15}N - ^1H HSQC spectrum of citrate bound CitAP with assignments	72
4.22	CitAP bound structure	73
4.23	X-ray structure of citrate bound CitAP	73
4.24	Superposition of GJ type dimer of 1P0Z and the dimer of new citrate bound CitAP structure	75
4.25	X-ray structure of Citrate-free CitAP	76
4.26	^{13}C - ^1H HSQC spectrum of the methyl region and the $^{13}\text{C}_\alpha$ - $^1\text{H}_\alpha$ projection of HACACO experiment	78
4.27	Plot of measured correlation time of individual residues of citrate free CitAP	79
4.28	Plot of Heteronuclear ^{15}N - ^1H NOE values of citrate free CitAP	80
4.29	Secondary structure of citrate bound-CitAP	81
4.30	Citrate bound CitAP dimer with the unassigned residues	81

4.31	Effect of Na ⁺ on citrate bound-CitAP	83
4.32	Correlation between measured H-N RDCs of citrate bound CitAP with back calculated RDCs from different crystal structures of CitAP	84
4.33	Correlation of measured RDCs of citrate free-CitAP with back calculated RDCs from different Xray structures of CitAP and their combinations	87
4.34	Conformational differences between citrate-bound and citrate-free CitAP	88
4.35	Difference between citrate free CitAP and citrate bound CitAP	89
4.36	Structural and dynamic changes in CitAP upon binding to citrate as moni- tored by multidimensional NMR spectroscopy	90
4.37	Signal transduction mechanism in CitA	92
4.38	Superposition of DcuS-PD structure with citrate-free and bound structures of CitAP	94
4.39	Conformational exchange line broadening of DcuS-PD and CitAP	95

List of Tables

2.1	Table of software used in the study	17
2.2	Table of experiments measured for the citrate free CitAP sample that was produced in Jülich	19
2.3	Table of experiments measured for the citrate bound CitAP sample that was produced in Jülich	19
2.4	Table of experiments measured for the citrate free CitAP sample that was produced in our laboratory	20
2.5	Table of experiments measured for the citrate bound CitAP sample that was produced in our laboratory	20
2.6	Compilation of physical constants	26
A.1	Basis operators for $q = 0, 1$ or 2 for an IS system of spin $1/2$	122
A.2	Time evolution of common product operators for a two spin systems under different spin hamiltonians	122
B.1	Backbone chemical shifts of citrate free CitAP at 310K.	125
B.2	Backbone chemical shifts of citrate bound CitAP at 310K.	127
B.3	Residual dipolar couplings measured for DcuS-PD in phages	130
B.4	Residual dipolar couplings measured for Ubiquitine in Otting phase	139
B.5	Residual dipolar couplings measured for citrate bound form of CitAP in phages	145
B.6	Residual dipolar couplings measured for citrate free form of CitAP in phages	147

List of Abbreviations

ADP	Adenosine di-phosphate
ADP	Adenosine tri-phosphate
CBCA(CO)NH	NMR experiment observing peptide ^{15}N , $^1\text{H}^N$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$
CCR	Cross correlated relaxation
COS-CT	Coherence order transfer
CPMG	Carr-Purcell-Meiboom and Gill experiment for relaxation studies
CSA	Chemical shift anisotropy
FID	Free induction decay
HACACO	NMR experiment observing correlations between peptide $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$ and CO
HCCH-TOCSY	TOCSY experiment using H-C-C-H magnetization transfer pathway
HK	Histidine Kinase
HNCA	NMR experiment observing peptide ^{15}N , $^1\text{H}^N$ and $^{13}\text{C}_\alpha$
HNCACB	NMR experiment observing peptide ^{15}N , $^1\text{H}^N$, $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$
HN(CA)CO	NMR experiment observing peptide ^{15}N , $^1\text{H}^N$ and $^{13}\text{C}'$
HNCO	NMR experiment observing peptide ^{15}N , $^1\text{H}^N$ and previous CO
HSQC	Heteronuclear single-quantum correlation
INEPT	Insensitive nuclei enhanced by polarization transfer
IPAP	Inphase-Antiphase experiment to measure J couplings
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	NOE Spectroscopy

PAS	Principal axis system and also Per-Arnt Sim domain
PD	Periplasmic domain
pdb	Protein data bank
PYP	Photoactive yellow protein
RDC	Residual dipolar coupling
rf	Radio frequency
ROESY	Rotational frame NOE Spectroscopy
SVD	Singular Value Decomposition
TROSY	Transverse relaxation optimized spectroscopy

Chapter 1

Introduction

1.1 General Introduction

Nuclear Magnetic Resonance (NMR) that originated approximately 60 years ago primarily as a potentially accurate method for measuring nuclear magnetogyric ratios (Felix Bloch and E.M Purcell, Nobel prize in physics, 1952) [1, 2]. But it turned out to be a drawback, when it transpired that the rf magnetic susceptibility measured could be a quite a complicated function, exhibiting many sharp, close lying resonances.

However, when it was realized that this complexity rather subtly reflected exceedingly fine characteristics of the electronic environment in which the nuclei are embedded [3], NMR began being developed as a high resolution spectroscopic technique for the elucidation of molecular structure, dynamics and spatial distribution (*i.e* NMR imaging).

Development of NMR for biological application, mainly protein structural studies were hampered by the poor sensitivity and crowding of signals in the NMR spectra due to the multitude of resonances arising from the protein. The first, three dimensional structure of a small protein was solved in the lab of Kurt Wuthrich (Nobel prize in chemistry, 2002) [4] in 1985. This was made possible because of the invention of two dimensional fourier transform spectroscopy (Richard R Ernst, Noble prize in chemistry 1991)[5, 6] . Since then in last two decades major improvements in NMR hardware (magnetic field strength, cryo-probes) and NMR methodology, combined with the availability of molecular biology and biochemical methods for the preparation, and isotope labeling of recombinant proteins have dramatically increased the application of NMR spectroscopy for the characterization of structure and

dynamics of biological molecules in solution. These improvements were designed to overcome the main problems with NMR of biomolecules, namely the signal to noise ratio and spectral overlap. Isotope labeling (incorporating magnetically active spin like ^{13}C and ^{15}N into biomolecules) [7, 8, 9, 10, 11, 12, 13, 14, 15, 16], and multi dimensional NMR experiments [17, 18, 19] have become a common procedure for protein NMR analysis. A suite of multidimensional NMR experiments have become routine experiments for the spectroscopic characterization of molecules [20]. Experiments have been developed for the determination of distance restraints derived from quantification of NOESY, and ROESY spectra as well as torsional angle restraints from measurement of scalar 3J coupling constants [21]. These traditional NMR parameters provide, short to medium range structural restraints for structural calculation of proteins. During the last decade these traditional NMR parameters were augmented by measurement and interpretation of anisotropic parameters such as residual dipolar couplings (RDCs) and chemical shift anisotropy for structural studies and structural refinement [22].

In this chapter an introduction to three dimensional experiments measured for protein backbone assignment is given. A brief description of RDCs is also presented with emphasis on the determination of the alignment tensor for proteins with already a known structure.

1.2 Backbone chemical shift assignment

The first step in NMR investigation of biomolecules is to assign the frequencies of all atoms in a molecule. Backbone chemical shift assignments of protein serves as a starting point for studies of its structure, dynamics and binding properties. Powerful multidimensional experiments have been developed by which the resonance position of all the NMR active, spin half atoms, namely ^1H , ^{15}N and ^{13}C for a protein can be assigned. The underlying principles for these experiments are correlation experiments that employ the transfer of coherences.

Coherence transfer is realized using modules in pulse sequence. The most common coherence transfer module used in multidimensional NMR experiment for backbone assignment is

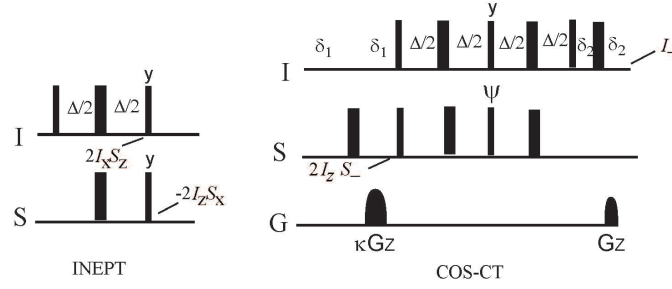


Figure 1.1: INEPT(left) and COS-CT(right) pulse sequence modules that realize coherence transfers, that are incorporated into many multidimensional NMR experiments. In the COS-CT module, concomitant with the change of the pulse with phase ψ (y to $-y$), sign of either one of the gradient is changed to give two specific coherence transfer scheme (see Equation 1.2 and 1.3). The value of κ depends on the ratio of gyromagnetic ratios of I and S spins. $\Delta = 1/2J_{IS}$

the INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) [23] coherence transfer scheme (refer Figure 1.1). This can be written in product operator term [24] as

$$2I_x S_z \xrightarrow{90_y^{(I,S)}} -2I_z S_x \quad (1.1)$$

where I and S are the heteronuclear (having different frequencies; different gyromagnetic ratios) spin half particles. $90_y^{(I,S)}$ is the radio frequency pulse(rf) applied along the y axis on both the spins simultaneously with a power corresponding to the rotation of initial magnetization by 90° .

Another type of coherence transfer used extensively for signal enhancement in multidimensional NMR, makes use of a module which changes the coherence order (denoted by p) of the initial and final operator in a specific (designed) way.

$$\text{Antiecho} : I_z S_- e^{i\Omega_S t_1} \xrightarrow{\text{COS}, t_2} I_- e^{i\Omega_S t_1} e^{i\Omega_I t_2} \quad (1.2)$$

$$\text{Echo} : I_z S_+ e^{-i\Omega_S t_1} \xrightarrow{\text{COS}, t_2} I_- e^{-i\Omega_S t_1} e^{i\Omega_I t_2} \quad (1.3)$$

where $S_\pm = S_x \pm iS_y$ and $I_\pm = I_x \pm iI_y$ and $\Omega_{I,S}$ is the chemical shift (frequency) of corresponding nucleus its associated with. In equation 1.2 coherence order transfer is from

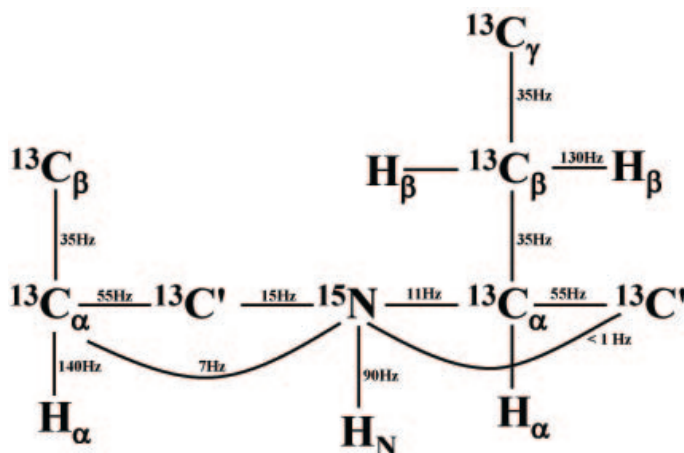


Figure 1.2: Spin system of a peptide backbone and the size of 1J and 2J coupling constants that are used for magnetization transfer in ^{13}C -, ^{15}N -labeled proteins

$p = -1$ to $p = -1$ and for equation 1.3 from $p = +1$ to $p = -1$. This type of transfer is commonly used in three dimensional experiments to increase signal sensitivity. Such modules are called PEP (Preservation of Equivalent Pathways) or COS (Coherence Order Selective transfer)[25, 26, 27, 28]. Usually these modules are combined with gradient selection for maximum sensitivity and water suppression (Figure 1.1). For multi-dimensional experiments with only one COS transfer, a gain of $\sqrt{2}$ in S/N is obtained with gradient selection [26].

Assignment strategies for non isotopically labeled proteins realizes on the transfer of magnetization in the homonuclear spins (spins of same γ 's, here ^1H)[29]. The sequential connectivity relies on the occurrence of resonance frequency of $\text{H}_\alpha(i)$ proton of amino acid (i) in the $^1\text{H}^N(i)$, $\text{H}_\alpha(i)$ cross peak in the COSY spectrum and in $^1\text{H}^N(i + 1)$, $\text{H}_\alpha(i)$ cross peak in NOESY spectrum. Such an assignment strategy is limited to small proteins because of the poor resolution of H_α spins (2 ppm), and the conformational dependence of the NOESY spectra. These difficulties are overcome in the assignment strategies for ^{13}C -, ^{15}N -labeled proteins, since coherence is transferred via the one bond coupling (1J) which is considerably larger and less prone to conformational dependency than the 3J coupling used in the

homonuclear approach. Signals can also be dispersed in multi dimensions, hence providing lesser crowding in the spectra. Larger 1J coupling also shortens the coherence transfer delays, making it favorable for the measurement of larger system in which the magnetization relaxes relatively faster. Figure 1.2 shows the spin system of peptide backbone and indicates the size of coupling constants used for magnetization transfer in a doubly ^{13}C -, ^{15}N -labeled proteins. For small to medium size proteins (4 kDa–20 kDa), a set of four experiments are to be measured for the backbone assignments. These are briefly described in the following subsections.

1.2.1 HNC0

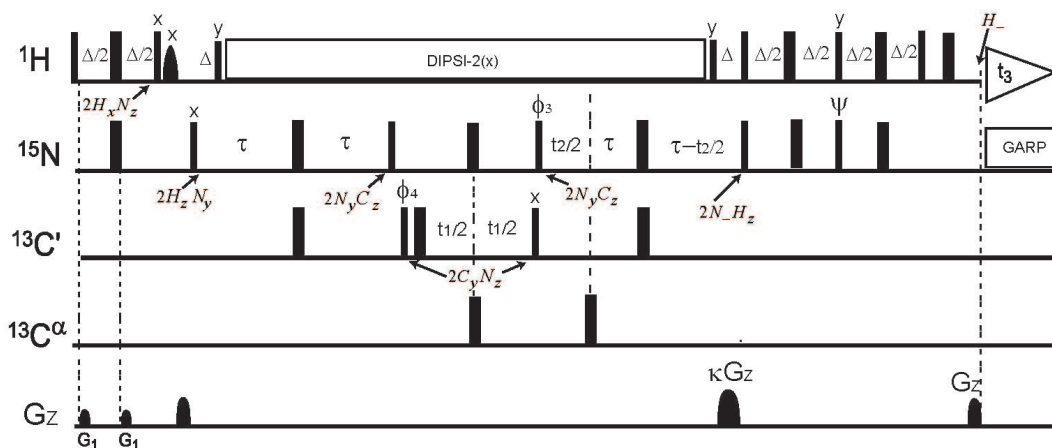


Figure 1.3: Pulse scheme for a sensitivity enhanced three dimensional HNC0 experiment with relevant product operators given in different time point of the sequence. 180° and 90° pulses are represented by broad and narrow bars. $\Delta = 1/2^1J_{NH}$ and $2\tau = 1/2^1J_{C'N}$. The phases of all pulses are x unless specified. $\phi_3 = 2(x), 2(-x)$, $\phi_4 = x, -x$, $\psi = y$ and receiver = $x, 2(-x), x$. Quadrature detection in $^{13}\text{C}'$ dimension is obtained by STATES-TPPI method by incrementing the phase of ϕ_4 . Echo-anti echo method is used in the ^{15}N dimension by shifting the phase of ψ from y to $-y$ with the inversion of the sign of one of the gradient G_z . G_z is the gradient strength in G/cm with $\kappa = \gamma_H/\gamma_N = \pm 10$

The HNC0 experiment is the prototype of all triple resonance experiments. It correlates the $^{13}\text{C}'$ resonances of an amino acid residue with the $^1\text{H}^N$ and ^{15}N resonances of the following residue. Schematic representation of the pulse sequence for the HNC0 experiment is given in the Figure 1.3. Magnetization transfer pathway is shown in Figure 1.4. Starting at an

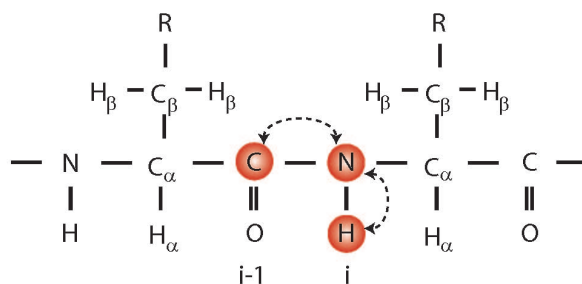


Figure 1.4: HNCOC experiment: The magnetization is transferred from the ${}^1\text{H}^N(i) \rightarrow {}^{15}\text{N}(i) \rightarrow {}^{13}\text{C}'(i-1)$ and then comes back to ${}^1\text{H}^N(i)$ along the same path. The frequencies of ${}^1\text{H}^N(i)$, ${}^{15}\text{N}(i)$ and ${}^{13}\text{C}'(i-1)$ (red) are observed.

${}^1\text{H}^N$, the magnetization is transferred to the directly attached ${}^{15}\text{N}$ through an INEPT step. The initial operator for the INEPT transfer is obtained by the evolution of the proton magnetization under the heteronuclear weak scalar coupling hamiltonian ($2^1J_{NH}H_zN_z$) (refer appendix A). After the INEPT transfer, ${}^{15}\text{N}$ magnetization evolves under the ${}^1J_{C'N}$ coupling to the $N_xC'_z$ term, which is converted to $N_zC'_y$ with the next INEPT transfer. The C' magnetization ($N_zC'_y$) is frequency labeled and transferred back to ${}^{15}\text{N}$ using a reverse INEPT step. During the subsequent constant time delay the ${}^1J_{C'N}$ coupling is refocused together with the frequency labeling of ${}^{15}\text{N}$. During the following delays and pulses the ${}^{15}\text{N}$ anti-phase magnetization (H_zN_-) is transferred to its directly attached ${}^1\text{H}^N$ single quantum coherence (H_-) using the COSY or PEP transfer. This is an out and back experiment [30] with the magnetization starting on ${}^1\text{H}^N$ and being detected also in the same ${}^1\text{H}^N$. This scheme uses echo-antiecho [31] with sensitivity enhancement for quadrature detection in the ${}^{15}\text{N}$ dimension.

1.2.2 HN(CA)CO

The HN(CA)CO experiment provides sequential correlations between the ${}^1\text{H}^N$ and ${}^{15}\text{N}$ chemical shifts of one amino acid residue and the ${}^{13}\text{C}'$ chemical shift of the same as well as preceding residue by transferring coherence via the intervening ${}^{13}\text{C}_\alpha$ spin (Figure 1.5). In this

experiment, $C'(i)$, $C'(i-1)$, $N(i)$, and $H^N(i)$ resonances are observed. Just like the HNCO experiment, the $^1\text{H}^N$ magnetization is transferred to its directly attached ^{15}N using an INEPT step. The ^{15}N magnetization evolves with $^1J_{C_\alpha N}$ and $^2J_{C_\alpha N}$ coupling constants to its directly attached $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\alpha(i-1)$ of preceding amino acids respectively, and from there to their directly attached $^{13}\text{C}'$ spins. After frequency labeling the $^{13}\text{C}'$ resonances, the magnetization is transferred back to $^{13}\text{C}_\alpha$ resonances and then back to ^{15}N using reverse INEPT transfer steps using the $^1J_{C_\alpha N}$ and $^2J_{C_\alpha N}$ couplings. Like in HNCO experiment the ^{15}N anti-phase magnetization is transferred to its directly attached $^1\text{H}^N$ spins using the COS-CT transfer step.

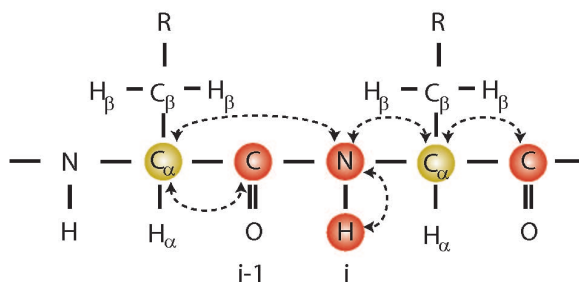


Figure 1.5: HN(CA)CO experiment: The magnetization is transferred from the $^1\text{H}^N(i) \rightarrow ^{15}\text{N}(i) \rightarrow ^{13}\text{C}_\alpha(i) \rightarrow ^{13}\text{C}'(i-1)/^{13}\text{C}'(i)$ and then comes back to $^1\text{H}^N(i)$ along the same pathway. The $^{13}\text{C}_\alpha$ (yellow) acts only as relay nucleus, its frequency is not detected. The frequencies of $^1\text{H}^N$, ^{15}N and $^{13}\text{C}'$ (red) are observed.

1.2.3 HNCACB

The HNCACB experiment correlates the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ resonances with the $^1\text{H}^N$ and ^{15}N resonances of the same residue and the $^1\text{H}^N$ and ^{15}N resonances of the neighboring residue (Figure 1.6). The magnetization transfer is similar to the HNCACO experiment till the $^{13}\text{C}_\alpha$ resonance, after which the $C_\alpha(i)$ and $C_\alpha(i-1)$ resonances are partially transferred to their attached $^{13}\text{C}_\beta$ resonances via the $^1J_{C_\alpha C_\beta}$ coupling. $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ resonances are frequency labeled and then transferred to the attached ^{15}N spins and from there to $^1\text{H}^N$ spins using similar magnetization scheme to HNCO. In this experiment, $C^\beta(i)$, $C^\beta(i-1)$, $C^\alpha(i)$, $C^\alpha(i-1)$, $N(i)$, and $H^N(i)$ resonances are observed. For a medium-sized protein (\sim

15 kDa), this experiment alone can provide virtually complete sequential assignment of the $^1\text{H}^N$, ^{15}N , $^{13}\text{C}_\alpha$, and $^{13}\text{C}_\beta$ resonances, because in addition to the sequential connectivities, the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts provide information on the amino acid type.

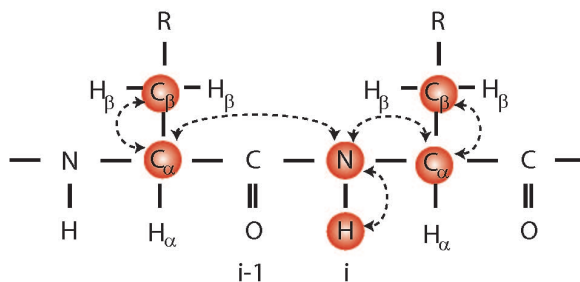


Figure 1.6: HNCACB experiment: The magnetization is transferred from the $^1\text{H}^N(i) \rightarrow ^{15}\text{N}(i) \rightarrow ^{13}\text{C}_\alpha(i)/^{13}\text{C}_\alpha(i-1)$ and then from there to $^{13}\text{C}_\beta(i)/^{13}\text{C}_\beta(i-1)$. After which it comes back to $^1\text{H}^N(i)$ along the same path. The frequencies of $^1\text{H}^N(i)$, $^{15}\text{N}(i)$, $^{13}\text{C}_\alpha(i)$, $^{13}\text{C}_\beta(i)$, $^{13}\text{C}_\alpha(i-1)$ and $^{13}\text{C}_\beta(i-1)$ (red) are observed.

1.2.4 CBCA(CO)NH

The CBCA(CO)NH experiment correlates both the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ resonances of an amino acid residue with the $^1\text{H}^N$ and ^{15}N resonances of the preceding residue (Figure 1.7). In this experiment instead of starting on the $^1\text{H}^N$, magnetization transfer starts with $^1\text{H}_\alpha$ and $^1\text{H}_\beta$ spins. Magnetization is transferred to its directly attached $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ with an initial INEPT step followed by chemical shift evolution and then using another INEPT transfer to directly attached $^{13}\text{C}'$ spins. The $^{13}\text{C}'$ spins evolves with the $^1\text{J}_{\text{NC}'}$ to its directly attached ^{15}N which is then frequency labeled and then transferred to the $^1\text{H}^N$ magnetization using the COS-CT step.

1.2.5 Assignment strategy

From the combination of CBCA(CO)NH and HNCACB experiments backbone resonance assignments and the sequential connectivities can be obtained. These experiments are sensitive enough for medium size proteins (~ 15 kDa, 130 amino acids) and provide the $^{13}\text{C}_\alpha$ and

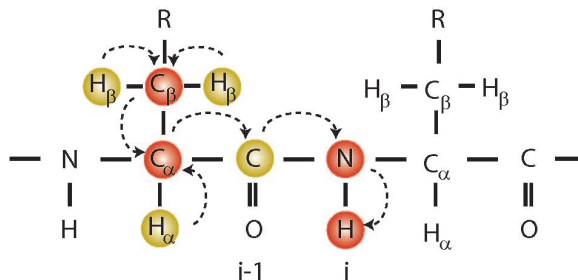


Figure 1.7: CBCA(CO)NH experiment: The magnetization is transferred from the $^1\text{H}_\alpha(i-1)/^1\text{H}_\beta(i-1) \rightarrow ^{13}\text{C}_\alpha(i-1)/^{13}\text{C}_\beta(i-1) \rightarrow ^{13}\text{C}'(i-1) \rightarrow ^{15}\text{N}(i) \rightarrow ^1\text{H}^N(i)$. The $^1\text{H}_\alpha$, $^1\text{H}_\beta$ and $^{13}\text{C}'$ (yellow) act only as relay nucleus, their frequency are not detected. The frequencies of $^1\text{H}^N(i)$, $^{15}\text{N}(i)$, $^{13}\text{C}_\alpha(i-1)$ and $^{13}\text{C}_\beta(i-1)$ (red) are observed.

$^{13}\text{C}_\beta$ chemical shifts to establish the sequential link between neighboring residues. Furthermore, when both the $^{13}\text{C}_\alpha$ and $^{13}\text{C}_\beta$ chemical shifts are provided at the same time, it gives important information about the amino acid type and secondary structure (*e.g.* α -helix and β -strand). $^{13}\text{C}'$ resonances from HNCO and HN(CA)CO experiments are used to resolve the assignment ambiguities for residues for which $^{13}\text{C}_\alpha$ $^{13}\text{C}_\beta$ degeneracies exist.

This type of assignment strategy is used extensively in the proteins studied in this thesis.

1.3 Residual Dipolar Couplings (RDC)

One of the major breakthrough in the last decade has been the use of anisotropic interactions like RDC's, CSAs and pseudo contact shifts in solution state NMR for structural studies of biological molecules [22, 32, 33, 34]. This was made possible with the discovery of tunable alignment of macromolecules as solutes in a dilute aqueous liquid crystal [35, 36, 37, 38, 39, 40]. In these media the anisotropic interaction do not average to zero and the residual effect could be measured easily through NMR.

Anisotropic interactions are magnetic properties that depend on the orientation of the molecule with respect to external magnetic field. Hence valuable information can be obtained

about the shape and orientation of the molecule if one could exploit these anisotropic interaction. RDCs are one such parameter which is comparatively very easy to measure once the molecules are partially aligned with the external magnetic field. A brief description on RDC is given below. Equations can be formulated for the calculation of alignment tensor from a known molecule structure and measured residual dipolar couplings. Equations can also be obtained for predicting dipolar coupling if the three principle component of the alignment tensor are known. The order matrix approach described below is commonly used in the structural validation using RDCs. A good correlation of RDCs measured for a protein in solution and those calculated from a known structure (X-ray or NMR) would indicate a similar structure for molecule in solution also.

Hamiltonian for dipolar coupling between two spins (I_1, I_2) is given by

$$H_{I_1 I_2}^{DD} = D_{\max} \left\{ 3 \frac{1}{r_{I_1 I_2}^3} (\mathbf{I}_1 \cdot \mathbf{r}_{I_1 I_2})(\mathbf{I}_2 \cdot \mathbf{r}_{I_1 I_2}) - \mathbf{I}_1 \cdot \mathbf{I}_2 \right\} \quad (1.4)$$

where $D_{\max} = -\mu_0 \frac{\gamma_I \gamma_S \hbar}{4\pi r_{I_1 I_2}^3}$ (Refer Table 2.6 in materials and methods chapter 2) for the values of all the physical constants). Expressing \mathbf{r} in equation 1.4 in spherical coordinates and taking equation 1.4 to a rotating frame, and using the secular approximation, the dipolar hamiltonian for two homonuclear spins will be

$$H_{I_1 I_2}^{\text{hom}} = D_{\max} \left\{ \left[2I_{1z} I_{2z} - \frac{1}{2}(I_1^+ I_2^- + I_1^- I_2^+) \right] \frac{(3 \cos^2 \theta - 1)}{2} \right\} \quad (1.5)$$

where θ is the angle between the magnetic field B_0 and the internuclear vector connecting spins I_1 and I_2 . For a weak coupling limit, only the first term in equation 1.4 need to be considered. For two heteronuclear spins ($I_1=I, I_2=S$) the dipolar hamiltonian in the rotating frame would be

$$H_{IS}^{\text{het}} = D_{\max} \{ (2I_z S_z) P_2(\cos \theta) \} \quad (1.6)$$

where $\frac{(3 \cos^2 \theta - 1)}{2} = P_2(\cos \theta)$ is the second order Legendre Polynomial.

In solution, $P_2(\cos \theta)$ is time averaged due to fast tumbling of the molecule ($\langle P_2(\cos \theta) \rangle$ -angular brackets denote time averaging). In an isotropic solution all possible orientation of the dipolar couplings are possible (probability of all orientations in a sphere are same) and hence the orientational dependence drops to zero ($\int_0^{2\pi} d\phi \int_0^\pi P_2(\cos \theta) d(\cos \theta) = 0$). In an

alignment media, the orientational dependence does not average to zero and hence could be measured with NMR.

Including the time average for orientational dependence in equation 1.6, the RDC for two heteronuclear spins is,

$$D_{IS}^{rdc} = D_{\max} [\langle P_2(\cos \theta) \rangle] \quad (1.7)$$

1.3.1 Alignment tensor determination for a rigid molecule

For a rigid molecule for which the atom positions are already known, the alignment tensor for the molecule can be readily derived [41]. The direction of the internuclear vector \mathbf{r}_{IS} can be given with respect to a cartesian molecular coordinate system by the angles ϕ_x , ϕ_y and ϕ_z (Figure 1.8). In the case of rigid object, the projections $c_x = \cos\phi_x$, $c_y = \cos\phi_y$ and $c_z = \cos\phi_z$ to axes x, y and z do not change with time and are identical for every molecule with identical structure. This coordinate system is fixed in the molecule and this reorients as molecule tumble in solution. Consequently, the axes x, y and z makes instantaneous angles $\beta_x(t)$, $\beta_y(t)$ and $\beta_z(t)$ with the magnetic field. The time dependence is associated with corresponding projections $\langle C_x \rangle = \cos\beta_x$, $\langle C_y \rangle = \cos\beta_y$ and $\langle C_z \rangle = \cos\beta_z$ along \mathbf{B} .

Remembering that the scalar product between two unit vectors is identical to the cosine of the angle θ between the two vectors we can write $\cos\theta$ in equation 1.7 as:

$$\langle \cos \theta \rangle = \mathbf{B}^T \cdot \mathbf{r} = (\langle C_x \rangle \quad \langle C_y \rangle \quad \langle C_z \rangle) \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix}$$

Hence the orientational dependence of the dipolar coupling takes the form

$$\langle P_2(\cos \theta) \rangle = \frac{(3\langle \cos^2 \theta \rangle - 1)}{2} = \frac{1}{2} \left[3(\langle C_x \rangle^2 c_x^2 + \langle C_y \rangle^2 c_y^2 + \langle C_z \rangle^2 c_z^2 + 2\langle C_x C_y \rangle c_x c_y) \right. \\ \left. + 2\langle C_y C_z \rangle c_y c_z + 2\langle C_z C_x \rangle c_z c_x) - 1 \right] \quad (1.8)$$

Separating the time dependent and time independent variables, a concise notation is reached by defining a 3×3 matrix known as Saupe order matrix[42] S such as

$$\langle P_2(\cos \theta) \rangle = \sum_{ij=\{x,y,z\}} \frac{1}{2} [3\langle C_i C_j \rangle - \delta_{ij}] c_i c_j = \sum_{ij=\{x,y,z\}} S_{ij} c_i c_j \quad (1.9)$$

where δ_{ij} is the kronecker delta function. Since $\langle C_i C_j \rangle = \langle C_j C_i \rangle$ and $\langle C_x \rangle^2 + \langle C_y \rangle^2 + \langle C_z \rangle^2 = 1$ the order matrix is symmetric and traceless. Consequently, there are only five independent matrix elements which means that \mathbf{S} can be determined when at least five non-redundant RDCs are available. Substituting $P_2(\cos \theta)$ of equation 1.9 in equation 1.7, the measured RDC will be

$$D_{IS}^{rdc} = D_{\max} \sum_{ij=\{x,y,z\}} S_{ij} c_i c_j = D_{\max} \sum_{ij=\{x,y,z\}} S_{ij} \cos \phi_i \cos \phi_j \quad (1.10)$$

Order matrix analysis is basically setting up and solving a system of linear equation of the

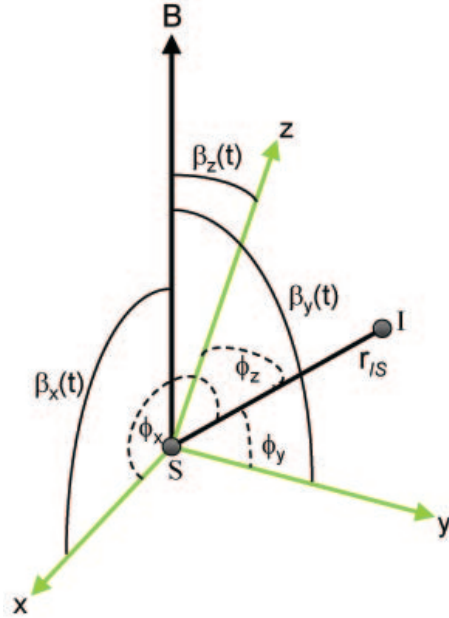


Figure 1.8: Internuclear vector \mathbf{r}_{IS} makes fixed angle ϕ_x , ϕ_y and ϕ_z with the molecular coordinate system (green). As the molecule tumble in solution, the angles of the axes x , y , and z denotes by β_x , β_y and β_z with respect to field \mathbf{B} change with time.

type $\mathbf{AX} = \mathbf{b}$ which can be explicitly written in matrix form.

$$\begin{pmatrix} c_{1,z}^2 - c_{1,x}^2 & c_{1,y}^2 - c_{1,x}^2 & 2c_{1,x}c_{1,y} & 2c_{1,y}c_{1,z} & 2c_{1,x}c_{1,z} \\ c_{2,z}^2 - c_{2,x}^2 & c_{2,y}^2 - c_{2,x}^2 & 2c_{2,x}c_{2,y} & 2c_{2,y}c_{2,z} & 2c_{2,x}c_{2,z} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_{n-1,z}^2 - c_{n-1,x}^2 & c_{n-1,y}^2 - c_{n-1,x}^2 & 2c_{n-1,x}c_{n-1,y} & 2c_{n-1,y}c_{n-1,z} & 2c_{n-1,x}c_{n-1,z} \\ c_{n,z}^2 - c_{n,x}^2 & c_{n,y}^2 - c_{n,x}^2 & 2c_{n,x}c_{n,y} & 2c_{n,y}c_{n,z} & 2c_{n,x}c_{n,z} \end{pmatrix} \begin{pmatrix} S_{yy} \\ S_{zz} \\ S_{xy} \\ S_{xz} \\ S_{yz} \end{pmatrix} = \begin{pmatrix} D_{red}^1 \\ D_{red}^2 \\ \cdot \\ \cdot \\ D_{red}^{n-1} \\ D_{red}^n \end{pmatrix} \quad (1.11)$$

where D_{red} is obtained by the division of the measured dipolar coupling with D_{max} .

Hence using equation 1.11 elements of the Saupe matrix can be obtained with the knowledge of dipolar coupling and the directional cosines associated with it. Elements of the overdetermined \mathbf{S} are best determined by singular value decomposition (SVD) since the order matrix is symmetric and real [43]. Alternatively it can also be obtained by a gradient search or by a fitting program that minimizes the difference between the observed and predicted coupling (details are given in chapter 2).

By diagonalizing the order matrix, it is possible to reduce order parameter description to a principal order parameter, S_{zz} and an asymmetric parameter $\eta = S_{yy} - S_{xx}/S_{zz}$ where $|S_{zz}| > |S_{yy}| > |S_{xx}|$. The coordinate system in which the Saupe matrix is diagonal is called principal axis system or PAS. In PAS the RDCs are given by

$$D(\phi_x, \phi_y, \phi_z) = D_{IS}^{\max} \sum_{i=\{x,y,z\}} S_{ii} c_i^2$$

By converting the vector direction (c_i 's) to polar coordinate; $c_x^2 = \sin^2 \vartheta \cos^2 \varphi$, $c_y^2 = \sin^2 \vartheta \sin^2 \varphi$ and $c_z^2 = \cos^2 \vartheta$, RDCs can be written as a function of ϑ and φ

$$D(\vartheta, \varphi) = D_{IS}^{\max} [S_{zz} \cos^2 \vartheta + S_{xx} \sin^2 \vartheta \cos^2 \varphi + S_{yy} \sin^2 \vartheta \sin^2 \varphi] \quad (1.12)$$

Making use of the definition of η and using trigonometric identities like $2 \sin^2 \varphi = 1 - \cos 2\varphi$ and $2 \cos^2 \varphi = 1 + \cos 2\varphi$ equation 1.12 can be recast into

$$D(\vartheta, \varphi) = D_{IS}^{\max} S_{zz} \left[\frac{1}{2} (3 \cos^2 \vartheta - 1) + \frac{1}{2} \eta \sin^2 \vartheta \cos 2\varphi \right] \quad (1.13)$$

Equation 1.13 can be expressed in terms of alignment matrix \mathbf{A} which is simply Saupe matrix \mathbf{S} scaled by 2/3. Writing $A_a = S_{zz}$ as the axial component and $\eta = S_{xx} - S_{yy}/S_{zz} = 3/2R$ equation 1.13 can be written as

$$D(\vartheta, \varphi) = D_{IS}^{\max} \frac{A_a}{2} \left[(3 \cos^2 \vartheta - 1) + \frac{3}{2} R \sin^2 \vartheta \cos 2\varphi \right] \quad (1.14)$$

where R is the measure of non axuality or rhombicity of the alignment tensor. Many proteins are asymmetric in their shape and, thus, their alignment is more or less rhombic, where as stretch of double stranded DNA is nearly axial.

From equation 1.14 it is clear that one cannot obtain specific directions of an internuclear vector but only possible combinations of ϑ and φ values along the surface of a cone is obtained. Furthermore since the form of D is an even function of ϑ and φ it is impossible to distinguish the symmetry alternative $(\pm \vartheta, \pm \varphi)$. Multivalued $D(\vartheta, \varphi)$ can be overcome by measuring RDCs in two different alignment media. Writing $D_{IS}^{\max} \frac{A_a}{2} = D_a$ as the magnitude of the residual dipolar coupling, the measured RDCs can be expressed in the most popular form.

$$D(\vartheta, \varphi) = D_a \left[(3 \cos^2 \vartheta - 1) + \frac{3}{2} R \sin^2 \vartheta \cos 2\varphi \right] \quad (1.15)$$

1.3.2 Measurement of the RDCs

RDCs are measured in the same way as scalar coupling of comparable strength. Because of their larger size, heteronuclear one bond dipolar couplings are usually measured in biomolecules. For heteronuclear spins the hamiltonian for evolution of dipolar couplings (see equation 1.6) is very similar to hamiltonian of weak coupling $2\pi J I_z S_z$. Hence for biological macromolecules, the total evolution of $J + D$ interaction under the operator $2I_z S_z$ is conveniently obtained from frequency resolved or intensity modulated heteronuclear multi-dimensional correlation spectra. A simple strategy for simultaneous measurement of RDC is given in chapter 3. These pulse sequences have been used for proteins studied in this thesis and also for a number of other proteins.

Chapter 2

Materials and Methods

In this chapter, a general description for the materials and methods employed in this work is given.

2.1 Materials

Protein samples used in the study were mostly prepared by our collaborators or prepared in the molecular biology lab supervised by Dr. Stefan Becker in our department.

2.1.1 Ubiquitin

^{15}N - ^{13}C ubiquitin sample was used in the measurement of residual dipolar couplings for the evaluation of the pulse sequence described in chapter 3. ^{15}N - ^{13}C ubiquitin was initially bought as a lyophilized powder from VLI research Philadelphia, USA (<http://www.vli-research.com>). Around 0.8 mM of sample solution was obtained by mixing 1 mg of the lyophilized ubiquitin in 250 μl of buffer solution. 50 mM of sodium phosphate buffer containing 95% H_2O and 5% D_2O at pH 6.5 was used for this study. To obtain anisotropic condition for the measurement of RDCs, lyophilized ubiquitin was directly dissolved into the above mentioned buffer solution, additionally containing 5% C12E5/n-hexanol (r=0.96) mixture [38].

2.1.2 Periplasmic domain of DcuS (DcuS-PD)

^{15}N - ^{13}C labeled samples of periplasmic domain of DcuS (DcuS-PD) was produced in the lab of Prof. Gottfried Unden in University of Mainz, Germany. It was expressed in *E-coli* strain carrying the DcuS_{45–180} plasmid [44]. ^{15}N labeled mutants namely H110A, F120A and R147A were also produced in the same lab. The buffer condition for the wild type and mutants were

50 mM Sodium Phosphate at pH 7.0 containing 200 mM NaCl, 0.8 mM CHAPS(a detergent), 50 mM Glycine, 50 pM PefablocSC, 0.01% NaN₃ and H₂O/D₂O in 90/10%. Properly buffered di-sodium salt of Fumaric Acid (Fluka) and D-Tartaric acid (Aldrich) were used for ligand titration. The ligands were dissolved in the same buffer condition as the protein. The pH was adjusted by drop wise addition of NaOH(0.1 mM) to the buffer containing the ligand. Anisotropic condition for measurement of RDCs for wild type DcuS-PD was obtained by the addition of around 10 mg/ml of Pfl1 phage(Asla) to the protein solution. The quadrupolar splitting of Deuterium was 9.2 Hz.

2.1.3 Periplasmic domain of CitA (CitAP)

¹⁵N and ¹⁵N-¹³C labeled protein was initially produced in the lab of Prof. Michael Bott in the Institut für Biotechnologie, Forschungszentrum Jülich. ¹⁵N labeled samples were used to optimize the NMR experimental condition for the protein. Two constructs were made: Cit_{39–185} and Cit_{45–176}. Both the construct had a 6 membered Histidine tag in the C terminal. Shorter construct, Cit_{45–176}, was used for the assignment experiments as it was more stable, less prone to aggregation and the peaks in the HSQC spectrum of that construct had more favorable line widths (see Chapter 4 Figure 4.18). The NMR buffer was: 50 mM Sodium Phosphate at pH 6.00 containing H₂O/D₂O in 90/10%. Properly buffered tri-sodium citrate di-hydrate(Fluka) and sodium molybdate (Na₂MoO₄)(Aldrich) were used for ligand titration. These ligands were dissolved in the same NMR buffer. The histidine tag on the protein was cleaved for the double labeled ¹⁵N-¹³C CitAP protein. All three dimensional experiments used for the backbone assignments were measured at 298 K. Measurements had to be repeated for the citrate bound form of CitAP.

Later ¹⁵N-¹³C labeled CitAP protein was prepared in our molecular biology lab. These samples were more stable than the ones prepared in Jülich. The new double labeled samples were used to re-measure the three dimensional experiments for protein backbone assignment with and without the ligand. Owing to the higher stability, the temperature for the measurements was set to 310 K. The new double labeled samples were also used to measure the RDCs in anisotropic media. Citrate free and citrate bound CitAP samples were aligned separately in

filamentous bacterial phages (Pf1). Concentration of Pf1 phage in citrate free CitAP sample was 10 mg/ml and in citrate bound CitAP was 14 mg/ml.

2.1.4 Equipment

All NMR spectra were measured in Bruker spectrometers (Bruker Karlsruhe). Following instruments were used to measure the spectra: AVANCE 400, AVANCE 600, DRX 600, AVANCE 700, DRX 800 and AVANCE 900. AVANCE 600, DRX 800 and AVANCE 900 was equipped with a cryo probe for increased signal sensitivity in proton dimension.

2.1.5 Software

The software used for processing and analysing NMR spectra, for calculating and analysing structures and for preparing figures is summarized in Table 2.1

PROGRAM	REFERENCE/ORIGIN
DALI	[45]
MARS	[46]
MOLMOL	[47]
MATHEMATICA	Wolfram Research, http://www.wolfram.com
NMRPipe/NMRDraw	[48]
PALES	[49]
PyMOL	http://www.pymol.org
Sparky 3	T. D. Goddard and D. G. Kneller, University of California, San Francisco
X-WINNMR 3.5	Bruker, Karlsruhe, Germany

Table 2.1: Table of software used in the study

2.2 NMR spectroscopy

Here a brief description of all the methods and the NMR experiments measured are presented.

2.2.1 NMR experiments

All NMR experiments were acquired using Bruker AVANCE 600, 700, and 900 or DRX 600 and 800 spectrometers running under X-WINNMR 3.5. NMR spectra were processed using X-WINNMR 3.5 or NMRPipe/NMRDraw [48] and analysed using Sparky 3 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

2.2.2 Chemical shift assignment

All NMR experiments performed for chemical shift assignments of the different protein samples are summarized in the Tables 2.2, 2.3, 2.4, 2.5.

2.2.3 Backbone chemical shift assignment

Backbone C_{α} , C' , N and $^1H^N$ assignments were obtained using standard triple resonance experiments used for proteins smaller than 30 kDa [50]. The assignments were, derived from 3D HNCA, CBCA(CO)NH, HNCACB, HNCO and HN(CA)CO experiments. Relaxation optimized experiments like TROSY [51] was also measured for citrate free CitAP protein. The manual assignment of backbone resonances was assisted by using the automatic assignment program MARS [46]. Experiment for the assignments of side chain resonances like 3D HCCH-TOCSY, C(CO)NH-TOCSY and HC(CO)NH-TOCSY spectra were also used for the backbone assignment of CitAP. Refocusing and mixing of ^{13}C magnetization in the HCCH-TOCSY experiment was performed using adiabatic pulses [52, 53]. Additional data provided by 3D [$^{15}N, ^1H$]-NOESY-HSQC and [$^{13}C, ^1H$]-NOESY-HSQC experiments were used for further assignment as well as confirmation of the through-bond data. For citrate bound form of CitAP, simultaneous ^{13}C , ^{15}N edited NOESY experiment were measured [54, 55]. Following specific experiments were also measured with the citrate free CitAP.

HSQC-CPMG

It is a ^{15}N - 1H HSQC experiment designed to reduce chemical exchange broadening of amide proton signals. It was recognized that dephasing of spin by chemical exchange processes can be reduced by CPMG(Carr-Purcell-Meiboom and Gill)-like pulse trains if the rate of the

180° spin-echo is on the order of or exceeds the rate of the exchange process. In this HSQC the INEPT and reverse-INEPT periods for coherence transfer through scalar coupling are modified by the inclusion of space modulated spin echo pulse train [56]. A Bruker pulse sequence (hsqctf3gpxy.2) was used for the measurement of this experiment.

HACACO

The three dimensional HACACO experiment was an “out and back” experiment with H_α detection in the direct dimension. This experiment connects the $^{13}C_\alpha$ and $^{13}C'$ of the i^{th} residue to its H_α . H_α magnetization is first transferred to its one bond coupled $^{13}C_\alpha$ using the $^1J_{C_\alpha H_\alpha}$ coupling. The $^{13}C_\alpha$ magnetization is transferred to its directly attached $^{13}C'$ using the $^1J_{C_\alpha C'}$ coupling. After frequency labeling in $^{13}C'$, the magnetization is transferred back to $^{13}C_\alpha$. It is frequency labeled in $^{13}C_\alpha$ in constant time manner and magnetization is transferred back to directly attached H^α . H^α is detected in the direct dimension. Citrate free CitAP protein was used for the measurement of this pulse sequence. The pulse sequence used was hacacogp3d.vi.

Experiment	Spectrometer	Pulse program
2D ^{15}N HSQC	AVANCE 900 cryo probe	hsqc.vi
3D HNCA	AVANCE 900 cryo probe	hnca-ge.mz.txt
3D HNCACB	AVANCE 900 cryo probe	hncacb-cpd.mz.txt
3D CBCA(CO)NH	AVANCE 900 cryo probe	cbcaconh-J.mz

Table 2.2: Table of experiments measured for the citrate free CitAP sample that was produced in Jülich

Experiment	Spectrometer	Pulse program
2D ^{15}N HSQC	AVANCE 900 cryo probe	hsqc.vi
3D HNCO	AVANCE 900 cryo probe	hnco-cpd.fb.vi
3D HNCACO	AVANCE 900 cryo probe	hncacogp3d
3D HNCACB	AVANCE 900 cryo probe	hncacb-cpd.mz.txt
3D CBCA(CO)NH	AVANCE 900 cryo probe	cbcaconh-J.mz

Table 2.3: Table of experiments measured for the citrate bound CitAP sample that was produced in Jülich

Experiment	Spectrometer	Pulse program
2D ^{15}N HSQC	AVANCE 900 cryo probe	hsqc.vi
2D ^{15}N HSQC-CPMG	AVANCE 900 cryo probe	hsqcetf3gpxy.2
2D ^{13}C HSQC-METHYL	AVANCE 600 cryo probe	ct-hsqc-methyl.vi
3D TROSY-HNCO	AVANCE 600 cryo probe	trhncogp3d
3D TROSY-HNCA	AVANCE 900 cryo probe	trhncagp3d
3D HNCACO	DRX 800 cryo probe	hncacogp3d
3D HNCACB	DRX 800 cryo probe	hncacbgp3d
3D CBCA(CO)NH	DRX 800 cryo probe	cbcaconhgp3d.vi
3D HC(CO)NH-TOCSY	AVANCE 700	hccconhgp3d2
3D C(CO)NH-TOCSY	AVANCE 700	hccconhgp3d3
3D HCCH-TOCSY	DRX 800 cryo probe	hcchatgp3d.vi
3D ^{15}N NOESY-HSQC	AVANCE 900 cryo probe	noesyhsqcetf3gp3d
3D ^{13}C NOESY-HSQC	DRX 800 cryo probe	noesyetgp3d13c
2D ^1H - ^{15}N HetNOE	AVANCE 600 cryo probe	invinoef3gpsi
3D HCACO	AVANCE 900 cryo probe	hcacogp3d.vi
2D ^1H - ^{15}N CCR experiment	DRX 800 cryo probe	hsqccc3gp3d.vi

Table 2.4: Table of experiments measured for the citrate free CitAP sample that was produced in our laboratory

Experiment	Spectrometer	Pulse program
2D ^{15}N HSQC	AVANCE 900 cryo probe	hsqc.vi
3D HNCO	AVANCE 600 cryo probe	hncogp3d
3D HNCA	AVANCE 900 cryo probe	hncagp3d
3D HNCACO	AVANCE 600 cryo probe	hncacogp3d.vi
3D HNCACB	AVANCE 900 cryo probe	hncacbgp3d
3D CBCA(CO)NH	DRX 800 cryo probe	cbcaconhgp3d.vi
3D HC(CO)NH-TOCSY	AVANCE 600 cryo probe	hccconhgp3d
3D C(CO)NH-TOCSY	AVANCE 600 cryo probe	hccconhgp3d.vi
3D HCCH-TOCSY	DRX 800 cryo probe	hcchatgp3d.vi
3D ^{15}N , ^{13}C Simultaneous NOESY-HSQC	AVANCE 600 cryo probe	noesyhsqcetf3gp3d
2D ^1H - ^{15}N Het-NOE	AVANCE 600 cryo probe	invinoef3gpsi

Table 2.5: Table of experiments measured for the citrate bound CitAP sample that was produced in our laboratory

2.2.4 Secondary structure determination

Secondary structure elements in a protein can be identified by calculating the difference between the experimental chemical shift (C^{exp}) and the random coil chemical shift (C^{coil}). Tabulated random coil values were used for folded proteins. Both Schwarzingler and Wishart random coil values were used to derive the secondary chemical shifts [57, 58]. Secondary chemical shifts ($\Delta\delta C$) for C_α , C_β and C' were calculated as the difference between measured ^{13}C chemical shift (δC_{exp}) and random coil ^{13}C chemical shift (δC_{coil}) [59].

$$\Delta\delta C = \delta C_{exp} - \delta C_{coil}. \quad (2.1)$$

To obtain secondary structure information, the secondary chemical shifts were calculated using the following formula:

$$\Delta\delta(C_\alpha C_\beta C') = \Delta\delta C_\alpha - \Delta\delta C_\beta + \Delta\delta C' \quad (2.2)$$

or by adding $\Delta\delta C_\alpha$ and $\Delta\delta C'$ when $\Delta\delta C_\beta$ is not available.

2.2.5 Residual dipolar couplings (RDCs)

Alignment media for the measurement of RDCs for individual protein is discussed in the Material part (refer section 2.1). RDCs were back-calculated from X-ray or NMR structures using mostly the order matrix analysis using the singular value decomposition (SVD) method as implemented in PALES [49]. A non linear fit using MATHEMATICA was also used to back calculate RDCs from two or more structures simultaneously. This method was used for the analysis for RDCS of citrate free CitAP.

A description of this method is given below. The RDC analysis described here closely follows the description of the RDCs given in chapter 1, section 1.3. Some of the equations are reproduced here for clarity.

Residual dipolar couplings between spins m and n can be expressed as in equation 2.3

$$D^{mn} = D_{\max}^{mn} \sum_{ij=\{x,y,z\}} S_{ij} \cos \phi_i^{nm} \cos \phi_j^{nm} \quad (2.3)$$

where ϕ_i^{nm} is the angle of internuclear vector connecting nuclei n and m relative to the i_{th} molecular axis. S_{ij} are elements of a 3×3 , symmetric and traceless order matrix. Equation 2.3 suggests that if the directional cosines of the internuclear vectors in an arbitrary molecular frame are known, it is possible to determine the order parameters, and hence the molecular orientational properties from just five independent measurements.

Carrying out an order matrix determination mean setting up and solving a system of linear equations of the form $AX = D$.

$$\begin{pmatrix} \cos^2 \phi_y^1 - \cos^2 \phi_x^1 & \cos^2 \phi_z^1 - \cos^2 \phi_x^1 & 2 \cos \phi_x^1 \cos \phi_y^1 & 2 \cos \phi_x^1 \cos \phi_z^1 & 2 \cos \phi_y^1 \cos \phi_z^1 \\ \cos^2 \phi_y^2 - \cos^2 \phi_x^2 & \cos^2 \phi_z^2 - \cos^2 \phi_x^2 & 2 \cos \phi_x^2 \cos \phi_y^2 & 2 \cos \phi_x^2 \cos \phi_z^2 & 2 \cos \phi_y^2 \cos \phi_z^2 \\ \cos^2 \phi_y^3 - \cos^2 \phi_x^3 & \cos^2 \phi_z^3 - \cos^2 \phi_x^3 & 2 \cos \phi_x^3 \cos \phi_y^3 & 2 \cos \phi_x^3 \cos \phi_z^3 & 2 \cos \phi_y^3 \cos \phi_z^3 \\ \cos^2 \phi_y^4 - \cos^2 \phi_x^4 & \cos^2 \phi_z^4 - \cos^2 \phi_x^4 & 2 \cos \phi_x^4 \cos \phi_y^4 & 2 \cos \phi_x^4 \cos \phi_z^4 & 2 \cos \phi_y^4 \cos \phi_z^4 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} S_{yy} \\ S_{zz} \\ S_{xy} \\ S_{xz} \\ S_{yz} \end{pmatrix} = \begin{pmatrix} D_{red}^1 \\ D_{red}^2 \\ D_{red}^3 \\ D_{red}^4 \\ \vdots \\ \vdots \end{pmatrix} \quad (2.4)$$

Here A would be the matrix containing the directional cosines obtained from the protein pdb file, D is the column matrix containing the measured RDCs and X would be column matrix containing the elements of the order matrix. Usually the S_{ij} values are obtained by SVD method. This can also be obtained by a fitting program that minimizes the error function $AX - D$ by looking for the best fit parameters for the elements of the order matrix. This is a versatile method used for a simultaneous fitting of two or more structures with their population weighting factor to a measured set of residual dipolar couplings. Here the equation to be solved for the order tensor would be

$$w_1 A^1 X + w_2 A^2 X + \dots + \left(1 - \sum_i^{n-1} w_i\right) A^n X = D \quad (2.5)$$

Here A^1, A^2 to A^n are the directional cosines derived from the 1,2 and n th pdb files and w_1, w_2 etc are their population weighting factors. Equation 2.5 was used for the fitting of measured RDCs of citrate free CitAP with different pdb files of CitAP. The fitting procedure was implemented in MATHEMATICA and is given in the appendix. The correlation coefficient and the quality or Q-factor were used to evaluate the agreement between a structure and the observed RDCs. The Q-factor was determined as $\text{rms}({}^1D_{HN}^{measured} - {}^1D_{HN}^{calculated}) / \text{rms}({}^1D_{HN}^{measured})$ [60]. A brief description of the type of RDCs measured and the pulse sequence used are given below.

^1H - ^{15}N RDCs

^{15}N - ^1H splittings were measured under isotropic and anisotropic conditions using 2D IPAP- ^{15}N - ^1H HSQC experiments [61]. ^1H - ^{15}N RDCs were extracted by subtraction of the $^1J_{NH}$ scalar coupling measured for the isotropic sample [33]. Alternatively, ^1H - ^{15}N RDCs were obtained from the ^{15}N - ^1H splittings of a modified 3D TROSY-HNCO experiment [62].

C' - ^{15}N and C' - $^{13}\text{C}_\alpha$ RDCs

The quantitative J correlation method was used to determine one-bond dipolar $^{13}\text{C}'$ - ^{15}N couplings. Usually this was measured together with the ^1H - ^{15}N RDCs (See Chapter 2)[63]. $^{13}\text{C}'$ - $^{13}\text{C}_\alpha$ -RDCs were measured from the splitting in the $^{13}\text{C}'$ dimension of a $^{13}\text{C}_\alpha$ coupled 3D HNCO experiment.

2.2.6 NMR relaxation data

Characteristic time for a brownian rotation of a molecule is usually called the molecular tumbling time or rotational correlation time (τ_c). For a spherical molecule of radius r rotating in a solution of viscosity η , τ_c is given by

$$\tau_c = \frac{4\pi r^3 \eta}{3kT} = \frac{V\eta}{kT}. \quad (2.6)$$

where V is the volume of the protein. As V is proportional to molecular mass ($V = \rho_p \times M$, where ρ_p is the partial specific density of the protein written in terms of l/g. ρ_p is usually 1.39×10^{-3} l/g for a protein). Taking this value in equation 2.6 and substituting the values for $\eta = 0.8872 \times 10^{-3}$ Ns/m² for water, and $k = 1.3807 \times 10^{-23}$ m²kg s²K⁻¹ and taking the temperature to be 298 K, we have

$$\tau_c/s = M/kDa \times 0.4459 \times 10^{-9}. \quad (2.7)$$

. where M is the molecular weight of the protein in Kilo-Dalton (kDa). As a rule of thumb the τ_c (in ns) is equal to half the molecular weight of a globular protein expressed in kDa .

An approximate value of τ_c can be derived from the equation of $T_{1\rho}$ relaxation time for an N-H dipolar coupled spins (assuming slow motion limit).

$$T_{1\rho}^{-1} = \frac{1}{8} \left(\frac{\mu_0}{4\pi} \right)^2 \frac{\gamma_N^2 \gamma_H^2 \hbar^2}{r^6} \{4J(0)\} \quad (2.8)$$

where the spectral density is given by equation 2.12.

1D $T_{1\rho}$ experiments with a spin-lock power of 2.5 kHz was used for the measurement. Two 1D $T_{1\rho}$ experiments were acquired. One with a spin-lock pulse duration of 2 ms (Δ_A) and a second with a spin-lock pulse duration Δ_B which corresponds to ~ 1.3 times the relaxation time calculated from the theoretical molecular weight. The average relaxation time of all backbone nitrogens was calculated using the following formula:

$$T_{1\rho} = \frac{\Delta_A - \Delta_B}{\ln(I_B/I_A)} \quad (2.9)$$

2.2.7 ^{15}N - ^1H steady state NOE

An important relaxation parameter that is usually measured for biomolecule is the Heteronuclear Nuclear Overhauser Effect (Het-NOE). This is measured by saturating the proton (^1H) signal and observing ^{15}N signal. The rate at which this occurs is the heteronuclear cross relaxation rate $R_{HN}(\text{H}_z < - > \text{N}_z)$; for long proton saturation, it reaches the steady state $\text{NOE}(\text{H}_z < - > \text{N}_z)$ value. The Het-NOEs of ^{15}N nuclei in NH groups were measured using the Bruker standard pulse program `invinoef3gpsi`. ^{15}N - ^1H Het-NOE values were calculated as the intensity ratios of the ^{15}N - ^1H correlation peaks from pairs of interleaved spectra acquired with and without ^1H presaturation during the recycle time of 5s.

2.2.8 Measurement of rotational correlation time (τ_c)

Rotational correlation time for citrate free CitAP was calculated from the cross correlated relaxation term using the NMR relaxation experiment described below. The cross correlated relaxation rate does not have conformational exchange term associated with it. Hence a true value of τ_c can be derived using this experiment. The experiment used here was presented previously for determining the CSA (chemical shift anisotropy) of ^{15}N spins [64]. Here, the same experiment was used to measure the τ_c for individual residue assuming a constant CSA for ^{15}N (160ppm). A brief description of the experiment is given below. The pulse sequence used here is the same as in reference [64].

The pulse sequence 2.1 is essentially an ^{15}N - ^1H HSQC with a relaxation period 2Δ inserted before the ^{15}N evolution period. The experiment is performed twice, once with

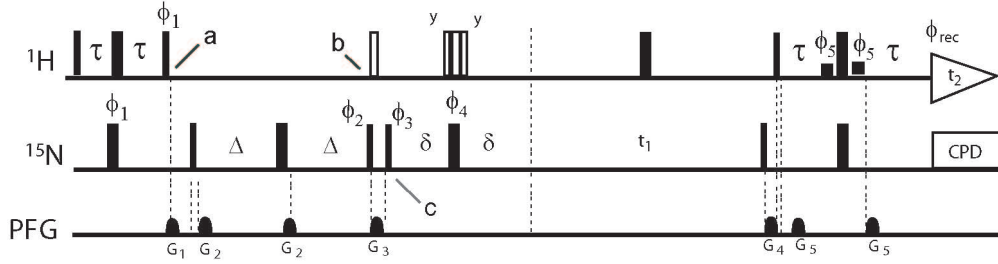


Figure 2.1: Pulse scheme for quantitative measurement of cross correlation. In the reference experiment (scheme B), the open ^1H 90° and composite $(90_y-220_x-90_y)$ 180° pulses are not applied, whereas they are applied in scheme (A), where all resonances are the result of cross correlation effects during the period 2Δ . Narrow and wide pulses correspond to flip angles of 90° and 180° , respectively. The two low power pulses immediately preceding and following the last non-selective ^1H 180° pulse have a width of 1 ms each and correspond to flip angles of 90° . With the carrier positioned on the H_2O resonance, they are part of the WATERGATE water suppression scheme. The radio-frequency phase of all pulses is assumed x , unless indicated. Delay durations: $\tau \approx 2.4$ ms; $\delta = 2.67$ ms, $T_2/8 < \Delta < T_2/2$. Phase cycling: $\phi_1 = y, -y$; $\phi_2 = x, x, -x, -x$; $\phi_4 = 4(x), 4(y), 4(-x), 4(-y)$; $\phi_5 = -x$; Receiver = $x, 2(-x), x, -x, 2(x), -x$. Quadrature detection in the t_1 dimension is accomplished by incrementing ϕ_3 in the States-TPPI manner. All gradients are sine-bell shaped, with an amplitude of 25 G/cm at their center. Durations: $G_{1,2,3,4,5} = 2.75, 2, 1, 1.5,$ and 0.4 ms.

applying the shaded ^1H 90° and composite ^1H 180° pulses, and once without applying these pulses. In the first experiment signal is proportional to $(\epsilon^+ - \epsilon^-)/2$. In the second experiment signal is proportional to $(\epsilon^+ + \epsilon^-)/2$. Here $\epsilon^\pm = \exp[-2\Delta(\lambda \pm \eta)]$. λ is the auto correlation rate and η the cross correlated rate of transverse relaxation of the ^{15}N spin. The ratio of signal intensities obtained in the two experiment will then be

$$I_A/I_B = (\epsilon^- - \epsilon^+)/(\epsilon^- + \epsilon^+) = \tanh(2\Delta\eta) \quad (2.10)$$

where

$$\eta = 2\alpha d \{4J(0) + 3J(\omega_N)\} (3 \cos^2 \theta - 1)/2 \quad (2.11)$$

with $\alpha = \frac{\mu_0 \gamma_I \gamma_S \hbar}{16\sqrt{2}\pi^2 r^3}$ and $d = \frac{\gamma_S B_0 \Delta \sigma}{3\sqrt{2}}$

For a spherical rigid molecule

$$J(\omega) = \frac{2}{5} \frac{\tau_c}{(1 + \omega^2 \tau_c^2)} \quad (2.12)$$

For slow motion limit ($\omega\tau_c \gg 1$), the term $J(\omega)$ can be neglected. Hence only $J(0)$ contribute to the cross relaxation term. Substituting α , d and taking only the $J(0)$ term for the spectral density Equation 2.11 becomes

$$\eta = \frac{\mu_0 \gamma_I \gamma_S^2 h B_0 (\Delta\sigma) \tau_c (3 \cos^2 \theta - 1)}{240 \pi^2 r^3} \quad (2.13)$$

Substitution of equation 2.13 to equation 2.10 and fitting that equation to signal ratios measured at varying Δ gives a value for τ_c . Values of the physical constants are compiled in the table 2.6.

Symbol	Values
$\gamma(^1\text{H})$	$26.7519 \cdot 10^7 \text{ s A kg}^{-1}$
$\gamma(^{15}\text{N})$	$-2.7519 \cdot 10^7 \text{ s A kg}^{-1}$
h (plancks constant)	$6.623 \cdot 10^{-34} \text{ J s}$
μ_0 (permeability of vaccum)	$4 \cdot \pi \cdot 10^{-7} \text{ m kg s}^{-2} \text{ A}^{-2}$
r (distance of N-H bond)	$1.04 \cdot 10^{-10} \text{ m}$
B (field strength)	800.25 MHz

Table 2.6: Compilation of physical constants

For citrate free CitAP, ratios of intensities were obtained for two delays ($\Delta = 16$ and 48ms). These ratios were fitted against the equation in 2.10. The slow motion limit and spectral density for rigid spherical rotor was used.

The value of τ_c obtained using this method does not take into consideration the local motions associated with the spins. Hence a lower bound for the τ_c estimate is obtained using the spectral density function based on rigid spherical model. More general spectral density function has been proposed by Lipari and Szabo [65, 66, 67] which is the sum of two Lorentian shapes.

$$J(\omega) = \frac{2}{5} \left[S^2 \frac{\tau_m}{1 + (\omega\tau_m)^2} + (1 - S^2) \frac{\tau}{1 + (\omega\tau)^2} \right] \quad (2.14)$$

with $\tau^{-1} = \tau_m^{-1} + \tau_e^{-1}$. Here the relaxation rate depends on one global parameter (τ_m) which is same for all residues and two local parameters, τ_e , a correlation time for internal motions and S^2 , which is a measure of amplitude of internal motion, known as the order parameter. Usually τ_e is very small (order of ps) and hence the second term in the equation 2.14 can be neglected. This truncated spectral density function is scaled by a factor of S^2

with respect to the spectral density used for the derivation of τ_c . Common values of order parameter for a globular medium size proteins range from 0.79-0.85. Hence the estimation of the τ_c using spectral density in equation 2.12 underestimates the values by $\sim 20\%$. But for residues showing considerable low values of τ_c , the average would then indicate considerable contribution from internal motions (order parameter low or $S^2 \ll 0.8$).

2.2.9 Chemical shift mapping of binding surfaces

An easy and rapid way to gain qualitative information about the interaction between two molecules is to study the change in chemical shift values. Perturbation in the local environment of the nucleus have a great effects on the chemical shift of that nucleus. The chemical shifts of ^1H and ^{15}N are particularly sensitive to changes of the chemical environment. Therefore ^{15}N and ^1H chemical shift changes recorded via ^{15}N - ^1H -HSQC spectra throughout an NMR titration are commonly combined in order to map a binding site on a protein according to the following equation

$$\Delta\delta_{HN} = \sqrt{\frac{(\Delta\delta_N/5)^2 + (\Delta\delta_H)^2}{2}} \quad (2.15)$$

$\Delta\delta_{HN}$: average amide chemical shift change.

$\Delta\delta_N$: amide nitrogen chemical shift change.

$\Delta\delta_H$: amide proton chemical shift change.

The binding interface is defined by the residues exhibiting the largest ^{15}N and ^1H chemical shift changes upon the NMR titration.

Chapter 3

Simultaneous measurement of protein one-bond residual dipolar couplings without increased resonance overlap

3.1 Abstract

A NMR strategy designed to measure simultaneously and without increased resonance overlap scalar and dipolar couplings (RDCs) in ^{13}C - ^{15}N -labeled proteins is presented. Contrary to common schemes for simultaneous measurement of RDCs, a single reference experiment is used for the extraction of more than one type of coupling, thereby reducing the required measurement time. This is accomplished by a common reference spectrum followed by a series of interleaved experiments, in which a particular coupling dependent parameter is varied according to the quantitative J -correlation method or using accordion spectroscopy. To illustrate this idea, we have modified the 3D TROSY-HNCO and the 3D CBCA(CO)NH experiment allowing efficient measurement of one-bond $^1\text{D}_{\text{NH}}$, $^1\text{D}_{\text{C}\alpha\text{H}\alpha}$, $^1\text{D}_{\text{C}'\text{N}}$, $^1\text{D}_{\text{C}\beta\text{H}\beta}$ and $^1\text{D}_{\text{C}\alpha\text{C}'}$ couplings in small to medium sized proteins. In addition, the experiments are expected to be useful for largely unfolded proteins, which show strong resonance overlap but have very favorable relaxation properties. Measurement of RDCs is demonstrated on uniformly ^{15}N - ^{13}C -labeled ubiquitin and on the sensory domain of the membrane two-component fumarate sensor DcuS of *Escherichia coli* (17kDa). DcuS was found to be unstable and to precipitate in one to two weeks. RDCs obtained from these experiments are in good agreement with the 1.8 Å X-ray structure of ubiquitin.

3.2 Introduction

Residual dipolar couplings (RDCs) can be observed in solution when a molecule is aligned with the magnetic field, either as a result of its own magnetic susceptibility anisotropy [68, 69] or caused by an anisotropic environment such as an oriented liquid crystalline phase [70] or an anisotropically compressed gel [39, 71]. When an alignment can be kept sufficiently weak, the NMR spectra retain the simplicity normally observed in regular isotropic solution, while allowing quantitative measurement of a wide variety of RDCs, even in macromolecules [33, 32]. In proteins, the most commonly measured dipolar couplings are the one bond ${}^1D_{NH}$, ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C'N}$, and ${}^1D_{C\alpha C'}$ RDCs. These couplings are sufficiently large to be measured with good accuracy, allow definition of the orientation of a peptide plane and play a key role in RDC-based structure determination [46, 72, 73, 74].

However, it is observed that certain proteins are unstable under the orienting media condition. If the magnitude and/or the orientation of the alignment tensor are changing slowly over time due to instability of the sample (protein or liquid crystal), the different sets of dipolar couplings will not be consistent with a unique alignment tensor/structure. This will in turn bias the protein structure calculation process. Therefore, it is desirable to measure as many types of RDCs in the committed time in which the protein/liquid crystal is stable. Moreover, much focus is put on dipolar couplings for rapid structure determination, with the aim to establish NMR spectroscopy as an important tool in Structural Genomics [75]. Therefore, there is the need to measure a large set of dipolar couplings in an as short time as possible.

Instead of measuring one type of coupling per NMR experiment, a number of schemes have been proposed recently that allow simultaneous measurement of different types of dipolar couplings in a single spectrum: ${}^1D_{HN}/{}^1D_{C\alpha C'}$ or ${}^1D_{HN}/{}^1D_{C\alpha H\alpha}$ from frequency differences using the IPAP-strategy [61, 76], ${}^1D_{C\alpha C'}/{}^3D_{HNC\alpha}$ or ${}^1D_{NC\alpha}/{}^2D_{NC\alpha}/{}^2D_{HNC\alpha}/{}^3D_{HNC\alpha}$ from 3D HNC α (α/β -C'C α - J)-TROSY and HNC α (α/β -NC α - J)-TROSY [77], ${}^1D_{NC\alpha}/{}^2D_{HNC\alpha}$ from ${}^{13}C\alpha$ coupled or J -modulated [${}^{15}N$, 1H]-HSQC spectra [78], ${}^1D_{C'N}/{}^1D_{NC\alpha}/{}^2D_{NC\alpha}$ from a quantitative J -correlation 3D TROSY-HNC experiment, in which cross and reference axial

peaks are observed in a single spectrum [79], ${}^1D_{C\alpha H\alpha}/{}^2D_{NH\alpha}$ from HNCA-TROSY [80], and ${}^1D_{C'N}/{}^2D_{HNC'}$ from an E.COSY-type HSQC experiment [81]. Approaches in which multiple couplings are allowed to evolve in a single experiment, have the disadvantage that each coupling results in a splitting that reduces the signal intensity by one half. Therefore, the overall sensitivity of these experiments is not higher than when the couplings are measured sequentially (potentially even lower if not multiple values for each $J+D$ value are obtained in the same experiment). In addition, if the multiplet components are not separated in different spectra this will result in increased overlap. Increased overlap is also a problem for the quantitative J -correlation TROSY-HNCO experiment, as reference peaks are located in a single plane of the 3D spectrum and both $C^\alpha(i)$ and $C^\alpha(i-1)$ peaks are present [79].

Here we present a different strategy for simultaneous measurement of dipolar couplings that is based on the quantitative J -correlation method and accordion spectroscopy. Only a single coupling is active during a single experiment, but for different couplings the same reference spectrum is used. This can reduce the total required measurement time without increasing resonance overlap. In addition, the experiments, in which different couplings are active, are recorded interleaved, such that measured RDCs correspond to a unique alignment tensor. In order to illustrate this concept, we have modified the 3D experiments TROSY-HNCO and CBCA(CO)NH for efficient measurement of ${}^1D_{NH}$, ${}^1D_{C\alpha H\alpha}$, ${}^1D_{CN}$, ${}^1D_{C\beta H\beta}$ and ${}^1D_{C\alpha C}$ couplings in small to medium sized, globular or larger, but intrinsically unstructured proteins.

3.3 Description of the pulse sequences

All presented pulse sequences can be used for measurement of one-bond isotropic scalar couplings. Dipolar couplings are obtained by performing the experiments both in an isotropic and a partially aligned phase followed by subtraction of the two measured values.

3.3.1 TROSY-HNCO experiment

The pulse sequence of the 3D TROSY-HNCO experiment for simultaneous measurement of ${}^1D_{CN}$ and ${}^1D_{NH}$ couplings is shown in Figure 3.1. It consists of three interleaved experiments that are based on the 3D TROSY-HNCO [51].

The backbone ${}^1J_{C'N}$ coupling is obtained by the quantitative J -correlation approach [62]. A reference experiment is recorded in which peak intensities are not strongly affected by the coupling of interest. This is achieved with the 180° pulse at position (a) and with a conventional constant time ${}^{15}\text{N}$ evolution period as depicted in panel (i). The ${}^{15}\text{N}$ - ${}^{13}\text{C}'$ dephasing delay is adjusted to $2T = 66.7$ ms and the separation of the 180° pulses amounts to $\Delta = 16.7$ ms. Thus, the effective dephasing time due to ${}^{15}\text{N}$ - ${}^{13}\text{C}'$ couplings is $2T - 2\Delta = 33.3\text{ms}$, i.e., close to $1/(2^1J_{C'N})$. To attenuate cross peak intensities according to the ${}^1J_{C'N}$ coupling, the 180° inversion pulse on C' is shifted to position (b) and the ${}^1J_{C'N}$ coupling is active for the entire interval. ${}^1J_{C'N}$ can then be calculated from the intensity ratios of the reference, I_{ref} , and attenuated resonances, I_{att} , according to

$$\frac{I_{att}}{I_{ref}} = \frac{\sin(2\pi^1J_{C'N}T)}{\sin[2\pi^1J_{C'N}(T - \Delta)]} \quad (3.1)$$

This previously described approach for measurement of ${}^1J_{C'N}$ [62] is extended by addition of a third experiment, which is recorded in an interleaved manner with the two described above, and allows simultaneous measurement of ${}^1J_{NH}$ couplings. The 180° inversion pulse on the carbonyls is kept at position (a), in order to maximize the ${}^{15}\text{N}$ - $\{{}^{13}\text{C}'\}$ transfer amplitude. Opposite to the reference spectrum, however, an additional delay is inserted at the end of the constant time ${}^{15}\text{N}$ evolution period that is incremented together with the ${}^{15}\text{N}$ evolution time (Figure 3.1 panel (ii)) [82]. During this delay, the ${}^{15}\text{N}$ - $\{{}^1\text{H}\}$ coupling evolves with respect to its directly attached proton while chemical shift evolution is refocused. Therefore, the peaks appear at $\omega N - (\kappa + 1)\pi^1J_{NH}$ rather than at $\omega N - \pi^1J_{NH}$ as in the reference spectrum with ${}^1J_{NH} < 0$. Typically κ is set to 2 so that a particular cross peak is shifted between the spectra by ${}^1J_{NH}$ Hz in the ${}^{15}\text{N}$ dimension. In this way, a common reference experiment (Figure 3.1; a, (i)) can be used for extraction of ${}^1J_{C'N}$ (Figure 3.1; b, (i)) and ${}^1J_{NH}$ (Figure 3.1; a, (ii)). Note that the two 180° pulses on ${}^1\text{H}$ and ${}^{15}\text{N}$, which are necessary for measurement of the

${}^1J_{NH}$ coupling, are also present in the reference and the ${}^1J_{C'N}$ -attenuated experiment, to have the same number and type of pulses in all experiments.

3.3.2 CBCA(CO)NH experiment

One-bond ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C\beta H\beta}$, and ${}^1D_{C\alpha C'}$ dipolar couplings are measured with a modified version of the 3D CBCA(CO)NH pulse sequence, which has been enhanced by the use of adiabatic pulses [53, 83]. The experiment is performed four times in an interleaved manner (Figure 3.2). All couplings are obtained using the quantitative J -correlation technique. The ${}^1J_{C\alpha H\alpha}$ and ${}^1J_{C\beta H\beta}$ couplings are obtained from the first three interleaved experiments, in which the 1H 180° pulse during the t_1 evolution period is switched between positions a, b and c, respectively [84]. The rephasing of the ${}^1J_{CH}$ coupling of C_β and C_α at the end of the carbon constant time period depends on the duration of Δ_1 . This is reflected in the spectra by a modulation of signal intensities according to

$$S(\Delta_1) \sim \sin \left[\pi \sum J_{CH} (2\Delta_1 + \Delta_{eff}) \right] \quad (3.2)$$

where Δ_{eff} accounts for the effect of ${}^1J_{CH}$ dephasing during the adiabatic ${}^{13}C$ inversion pulse, during the ${}^{13}C$ 90° pulses, and during short switching delays, and from which $\sum J_{CH}$ can be derived by a SIMPLEX minimization routine.

${}^1J_{C\alpha C'}$ couplings are obtained from the first experiment (this is also used as a reference spectrum for the measurement of ${}^1J_{C\alpha H\alpha}$ and ${}^1J_{C\beta H\beta}$ couplings) and an additional fourth, interleaved experiment. In order to attenuate signal intensities according to the one-bond ${}^1J_{C\alpha C'}$ values, the C' to ${}^{15}N$ transfer step is used (Figure 3.2). In all four interleaved experiments, the total dephasing delay is set to $2(n + \zeta) = 22ms$. In the regular CBCA(CO)NH experiment (i.e. the reference experiment) a 180° inversion pulse is applied a duration ζ after the 90° excitation pulse on C' , such that the ${}^1J_{C\alpha C'}$ coupling is active for $2\zeta = 9ms$, i.e., close to $1/(2{}^1J_{C'N})$. In the fourth interleaved experiment, however, the open 180° pulse on C_α is shifted by a duration Δ closer to the 180° C' pulse making ${}^1J_{C\alpha C'}$ active for $2(\zeta + \Delta) = 18.7ms$. ${}^1J_{C\alpha C'}$ couplings are almost rephased at the end of the C' to ${}^{15}N$ transfer step, yielding weak resonance intensity. The ratio of signal intensities from the reference and

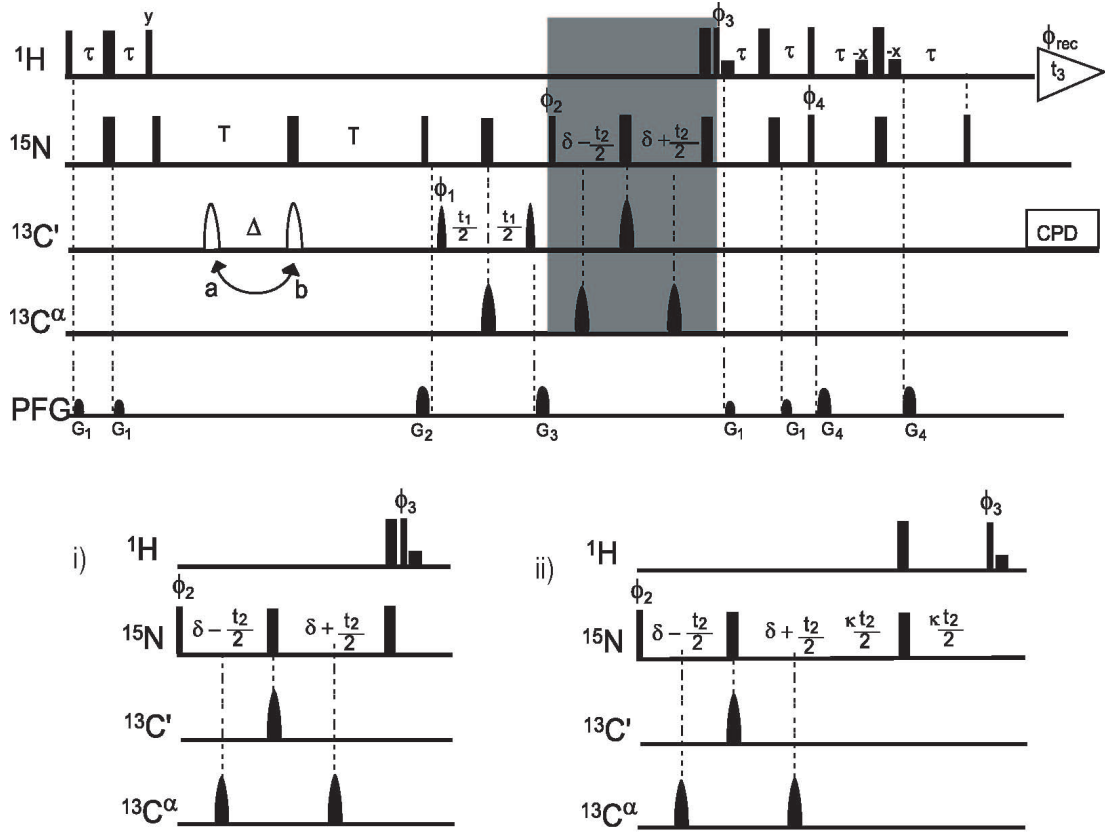


Figure 3.1: Pulse scheme of the modified 3D TROSY-HNCO experiment for simultaneous measurement of ${}^1J_{C'N}$ and ${}^1J_{NH}$ couplings. Narrow and wide pulses correspond to flip angles of 90° and 180° respectively. All pulse phases are x , unless specified otherwise. ${}^{13}\text{C}'$ pulses bracketing the t_1 evolution period are 90° and have the shape of the center lobe of a $\sin x/x$ function, and durations of $120\mu\text{s}$ (at $176\text{ MHz } {}^{13}\text{C}$ frequency). The other two ${}^{13}\text{C}'$ pulses are hyperbolic secant shaped (1 ms) and correspond to 180° . Composite pulse ${}^{13}\text{C}'$ decoupling (CPD) was used during acquisition. All three 180° ${}^{13}\text{C}_\alpha$ pulses are sine-bell shaped and have durations of $150\mu\text{s}$. Delay durations: $\tau=2.5\text{ms}$; $T=33.7\text{ms}$; $\delta=16\text{ms}$; $\Delta=16.6\text{ms}$. For the reference and the ${}^1J_{NH}$ -shifted spectrum, the first 180° ${}^{13}\text{C}'$ pulse (open shape) is applied in position a; for the rephased, ${}^1J_{C'N}$ -attenuated spectrum in position b. The shaded portion of the pulse scheme is given in detail in (i) and (ii). (i) Used for the first and second interleaved spectrum and (ii) used for the third interleaved spectrum for measurement of ${}^1J_{NH}$ couplings. In the shifted spectrum (ii) κ is typically set to 2, to obtain a shift of ${}^1J_{NH}$ Hz in the ${}^{15}\text{N}$ dimension. Two FIDs are acquired and stored separately for obtaining quadrature selection of the TROSY component in the t_2 dimension, with phases $\phi_2 = y, x, \phi_3 = -y, \phi_4 = -y$ and with phases $\phi_2 = y, -x, \phi_3 = y, \phi_4 = y$ [85]. In both cases, $\phi_1 = x, x, -x, -x$, and Receiver = $x, y, -x, -y$. States-TPPI quadrature selection is used in the t_1 dimension. Sine-bell shaped pulsed field gradients (1 ms each) have strength (G/cm) and axis: G1: 3, x ; G2: 18, y ; G3: 18, x ; G4: 18, z .

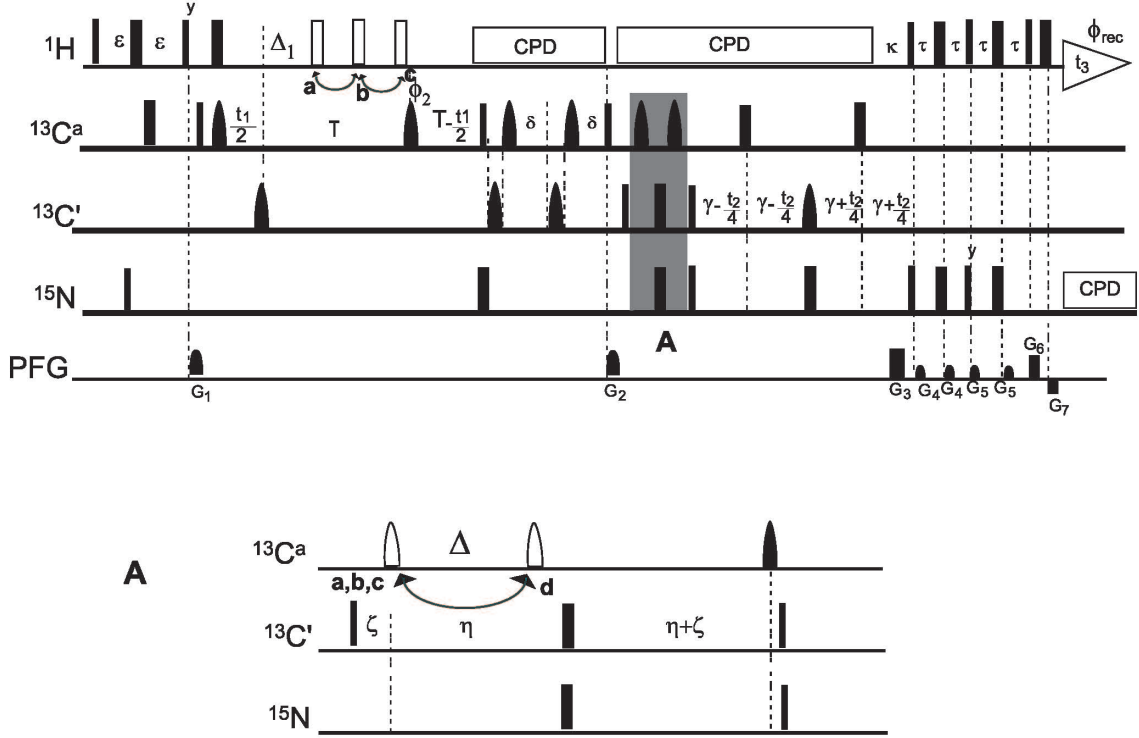


Figure 3.2: Pulse scheme of the 3D CBCA(CO)NH quantitative J_{CH} and $J_{C\alpha C'}$ experiment. Narrow and wide pulses correspond to flip angles of 90° and 180° respectively. All pulse phases are x , unless specified otherwise. The first three $^{13}C'$ 180° pulses have rSNOB profile [86], and a duration of $170\mu s$ (at 150 MHz ^{13}C frequency). All other shaped $^{13}C'$ and $^{13}C_{\alpha/\beta}$ pulses (46 ppm) are of the hyperbolic secant adiabatic inversion type ($500\mu s$), with a squareness level, μ , of 3 [87]. The open 1H pulses, applied at time point a, b, or c, is composite ($90^\circ_x - 180^\circ_y - 90^\circ_x$). Delay durations: $\epsilon = 1.3$ ms; $\Delta_1 = 0.77, 1.67$ and 3.42 ms in the three interleaved experiments; $T = 3.4$ ms; $\delta = 3.4$ ms; $\eta = 6.5$ ms; $\gamma = 6.8$ ms; $\kappa = 5.4$ ms; $\tau = 2.7$ ms. For the reference spectrum, the first 180° 1H pulse (open shape) is applied at position a, which gives a total J_{CH} -modulation duration of $2(\Delta_1 + \Delta_{eff})$ 1.93 ms [84]. In the $J_{C\alpha C'}$ -attenuated experiment (the fourth interleaved spectrum) the open ^{13}C 180° pulse is applied at point d, otherwise at position a/b/c. Delay durations: $\Delta = 4.85$ ms; $\xi = 4.5$ ms. Note that the second 180° pulse on C_α compensates the Block - Siegert shift, and should be applied with a reverse phase profile in order to refocus the coupling evolution during the open ^{13}C 180° pulse [88]. Phase cycling: $\phi_1 = y, -y$; $\phi_2 = x, x, y, y$; Receiver = $x, -x, -x, x$. Rance-Kay t_2 quadrature detection is used, [26] alternating the phase of the first ^{15}N 90° pulse after gradient G_3 between x and $-x$, in concert with the polarity of gradient G_3 . States-TPPI quadrature selection is used in the t_1 dimension. Pulsed field gradients $G_{1,2,4,5}$ are sine-bell shaped (18 G/cm for $G_{1,2}$, 28 G/cm for $G_{4,5}$), and $G_{3,6,7}$ are rectangular (30 G/cm). Durations: $G_{1,2,3,4,5,6,7} = 2, 1.4, 2.705, 0.4, 0.4, 0.2, 0.074$ ms, with respective gradient axes: xy, y, z, x, y, z, z .

attenuated spectrum follow a similar relationship as Eq. [1], from which $^1J_{C\alpha C'}$ couplings can be extracted by a one dimensional grid search or a nonlinear fit. Note that the C_β resonances are also modulated by the $^1J_{C\alpha C'}$ coupling. Thus, if there is signal overlap in the C_α region of the spectrum one can still obtain the couplings from the C_β resonances, provided the particular residues do not have overlapping C_β frequencies.

3D CBCA(CO)NH and TROSY-HNCO are among the most sensitive triple resonance experiments for medium sized proteins and the method is applicable to any protein which yields TROSY-HNCO and CBCA(CO)NH spectra with high S/N ($\geq 20:1$ and $\geq 40:1$, respectively) in a reasonable amount of time.

3.4 Data collection

The method is demonstrated for uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled human ubiquitin (VLI Research Inc., Philadelphia) and on the uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled sensory domain of the membraneous two-component fumarate sensor DcuS of *E. coli* (17 kD) [44]. One-bond couplings were measured for an isotropic sample, containing 0.8 mM ubiquitin in 95% H₂O, 5% D₂O, 50 mM phosphate buffer, pH 6.5 at 35 °C, or containing 0.8 mM DcuS in 95% H₂O, 5% D₂O, 50 mM phosphate buffer, pH 6.5 at 30 °C, and a second sample containing additionally either 5% C12E5/n-hexanol ($r = 0.96$) (in case of ubiquitin) [38] or 10 mg/ml bacteriophage Pf1 (in case of DcuS) [37, 89].

Spectra were recorded on BRUKER AVANCE 600 and 700 MHz spectrometers. TROSY-HNCO experiments were acquired as $24^* \times 40^* \times 512^*$ data matrices with acquisition times of 16.8 ms ($t_1, ^{13}\text{C}'$), 18.7 ms ($t_2, ^{15}\text{N}$), and 50 ms ($t_3, ^1\text{H}$), using 4 scans per fid and a total measuring time of 19 hours for all three interleaved spectra. In case of the CBCA(CO)NH, each interleaved spectrum was recorded as a $46^* \times 25^* \times 512^*$ data matrix with acquisition times of 4.9 ms ($t_1, ^{13}\text{C}^{\alpha/\beta}$), 17.5 ms ($t_2, ^{15}\text{N}$), and 50 ms ($t_3, ^1\text{H}$) using 8 scans per fid and a total measuring time of 48 hours for all four interleaved spectra. Data were processed and analyzed using NMR-PIPE/NMRDRAW [48].

3.5 Results and discussion

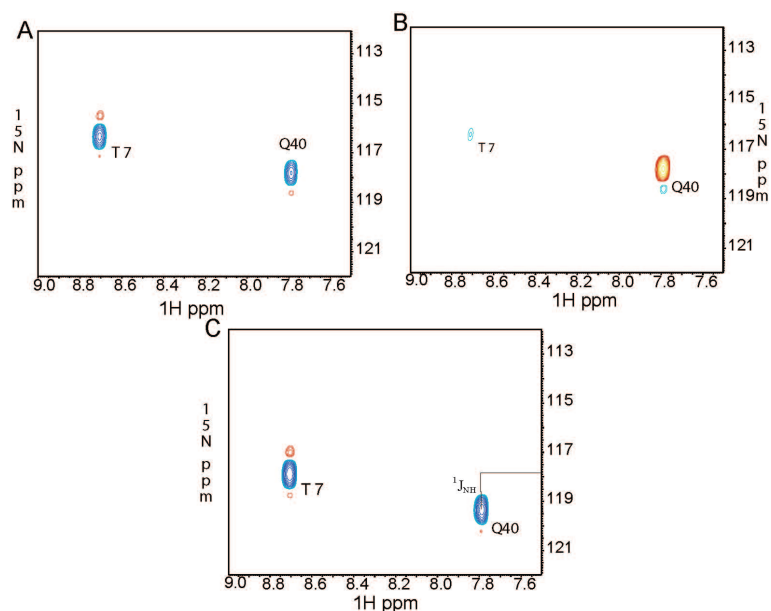


Figure 3.3: Selected regions of the modified TROSY-HNCO spectrum of ubiquitin in isotropic solution. A) reference spectrum, B) $^1J_{C'N}$ -attenuated spectrum and C) $^1J_{NH}$ -shifted spectrum. Dashed contours denote negative intensities. Selected regions have a C' frequency of 171.1 ppm.

Selected regions of the spectra measured with these pulse sequences for ubiquitin in isotropic solution are shown in Figures. 3.3 and 3.4. In Figure 3.3A the first interleaved TROSY-HNCO spectrum is shown, which is used as reference for extraction of both $^1J_{C'N}$ and $^1J_{NH}$ couplings. The second interleaved spectrum is shown in Figure 3.3B, in which signal intensities are modulated according to the $^1J_{C'N}$ coupling. In Figure 3.3C resonances are not attenuated compared to Figure 3.3A, but shifted in the ^{15}N dimension by $^1J_{NH}$ Hz from the reference spectrum of Figure 3.3A. Figure 3.4 shows selected regions of the interleaved CBCA(CO)NH experiments. In Figure 3.4A the reference spectrum is shown and in Figure 3.4B and C signal intensities are modulated by the $^1J_{C\alpha H\alpha}$ and $^1J_{C\beta H\beta}$ couplings. Figure 3.4D encodes the $^1J_{C\alpha C'}$ coupling.

3.5.1 Sensitivity consideration for CBCA(CO)NH experiment

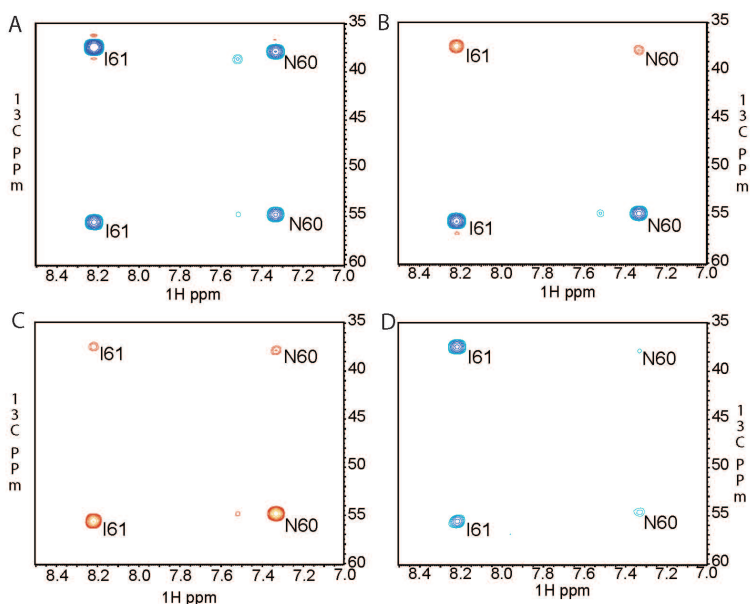


Figure 3.4: Selected regions of the modified CBCA(CO)NH spectrum of ubiquitin in isotropic solution. A is the reference spectrum. B and C are the attenuated spectra for measuring $C_{\alpha}-H_{\alpha}$ and $C_{\alpha}-H_{\beta}$ couplings. D) shows the spectrum for measuring $C_{\alpha}-C'$ couplings. Selected regions have a ^{15}N frequency of 116.2 ppm.

A CBCA(CO)NH experiment is much less sensitive than a conventional HNCO due to the larger number of transfer steps and the fast relaxation of transverse C_{α} and C_{β} magnetization. For a 11 kDa protein such as ribonuclease T1 the CBCA(CO)NH is a factor of three less sensitive than a HNCO experiment (Bruker Avance 3D/Triple Resonance User Manual). For larger proteins the relative sensitivity of the CBCA(CO)NH is even further reduced. However, following considerations have to be taken into account for measurement of $^1\text{D}_{C_{\alpha}C'}$ couplings. The sensitivity of the HNCO($\alpha/\beta-C'C_{\alpha}-J$) experiment proposed by Permi et al.[15], for example, is reduced by a factor of 0.9 due to the insertion of a filter element with a duration of 9 ms and due to an extended ^{13}C evolution time of about 30 ms, which is required to obtain sufficient resolution for accurate measurement of splittings (assuming a C' transverse relaxation time of 117 ms). Moreover, the intensity of a single resonance is split into two reducing the sensitivity by 45% overall. For the approach proposed here, on the other

hand, only a single additional, conventional CBCA(CO)NH experiment is required as the reference experiment is already available either from the measurement of ${}^1D_{C\alpha H\alpha}/{}^1D_{C\beta H\beta}$ couplings or from conventional backbone assignment. No additional delays or pulses had to be introduced into the CBCA(CO)NH pulse sequence avoiding increased relaxation losses or a higher sensitivity towards pulse imperfections. The accuracy of the ${}^1J_{C\alpha C'}$ couplings obtained from the CBCA(CO)NH is further improved, as splittings can be obtained from both the C_α and the C_β resonance. The corresponding values can be averaged, such that the sensitivity of the CBCA(CO)NH (in terms of measurement of couplings) is effectively increased by a factor of $\sqrt{2}$. Therefore, for small to medium sized proteins such as ribonuclease T1, measurement of ${}^1D_{C\alpha C'}$ couplings simultaneously with ${}^1D_{C\alpha H\alpha}/{}^1D_{C\beta H\beta}$ values has a similar efficiency as separate measurement of ${}^1D_{C\alpha C'}$ couplings with a HNC(O)(α/β -C' C_α - J) experiment.

Measurement of ${}^1J_{C\alpha C'}$ couplings as proposed herein has the following favorable properties: (i) Opposite to α/β -filtered spin-state selective experiments J -mismatch is not a problem in quantitative J -correlation experiments. This is particularly important for measurement of RDCs where the filter period no longer matches the $J+D$ value. (ii) Despite the lower resolution of a CBCA(CO)NH compared a HNC(O) experiment, a very complete set of ${}^1J_{C\alpha C'}$ couplings can be obtained as in case of C_α overlap the coupling might be extracted from the C_β resonance. (iii) For extraction of ${}^1J_{C\alpha C'}$ couplings no additional experiment has to be processed, peak picked and analysed. ${}^1J_{C\alpha C'}$ couplings are obtained in parallel with the ${}^1J_{C\alpha H\alpha}$ and ${}^1J_{C\beta H\beta}$ couplings, minimizing the number of errors in extracted dipolar coupling values. (iv) The proposed pulse sequences allow simultaneous measurement of RDCs such that they will be in agreement with a unique effective alignment tensor even if the protein degrades/aggregates rapidly. With the CBCA(CO)NH this is possible for ${}^1J_{C\alpha C'}$, ${}^1J_{C\alpha H\alpha}$, and ${}^1J_{C\beta H\beta}$ couplings without increased resonance overlap and employing a common reference experiment.

3.5.2 Sensitivity consideration for TROSY-HNC(O) experiment

In case of the 3D TROSY-HNC(O), the incremented delay that was introduced for the extraction of the ${}^1J_{NH}$ couplings (Figure 3.1; a, (ii)) slightly lowers the sensitivity of the third

interleaved TROSY-HNCO spectrum compared to the reference experiment. In addition, during this additional delay transverse relaxation proceeds with an effective rate that is given by the average of the individual ^{15}N multiplet components. Therefore, for large molecules a compromise between separation of the upfield and downfield component (which increases with κ) and the sensitivity of this experiment has to be made. For large molecules it might be advantageous to reduce κ . At the same time, however, the 180° pulses on ^{15}N and $^{13}\text{C}'$ could be applied simultaneously after a time $T = 16\text{ms}$, such that $2T$ would be reduced to 32 ms as in a conventional TROSY-HNCO.

$^1\text{J}_{\text{C}'\text{N}}$ couplings can be measured separately using a non-decoupled TROSY-HNCO, in which $^1\text{J}_{\text{C}'\text{N}}$ values are obtained from antiphase splittings in the ^{15}N dimension [32]. In order to resolve the small $^1\text{J}_{\text{C}'\text{N}}$ couplings a rather long ^{15}N acquisition time is required, which is best implemented in a semi-constant time fashion and set to a maximum value, t_2^{max} , roughly equal to the transverse ^{15}N relaxation time t_2 [90]. For example, for deuterated EIN (transverse relaxation time of the downfield component at 600 and 800 MHz: 118 and 131 ms, respectively) a ^{15}N acquisition time of 115 ms was employed. Comparison of the relaxation losses in the quantitative J -correlation experiment, in which the ^{15}N magnetization is 33 ms longer in the transverse plane than in a regular TROSY-HNCO, with those in a non-decoupled TROSY-HNCO, with a semi-constant time evolution $t_2^{\text{max}} = T_2(^{15}\text{N})$, indicates that the quantitative J -correlation experiment is more sensitive for $T_2(^{15}\text{N}) > 100\text{ms}$ [20]. $T_2(^{15}\text{N})$ values larger than 100 ms are commonly observed at room temperature for globular proteins with a molecular weight < 16 kDa, for the downfield component of deuterated proteins up to 30 kDa, such as EIN, or for large protonated, but natively unfolded proteins. For example, for ubiquitin ($T_2(^{15}\text{N}, 25^\circ\text{C}) \sim 160\text{ms}$; $T_2(^{15}\text{N}, 7^\circ\text{C}) \sim 105\text{ms}$) the quantitative J -correlation experiment is 15% and 2% more sensitive than the non-decoupled experiment at 25°C and 7°C , respectively. On the other hand, for a relaxation time of 74 ms ($\sim t_2(^{15}\text{N})$ of protonated EIN) the non-decoupled HNCO would be more sensitive by 20%. Therefore, for transverse relaxation times larger than 100 ms simultaneous measurement of $^1\text{J}_{\text{NH}}$ and $^1\text{J}_{\text{C}'\text{N}}$ couplings as proposed here achieves a time saving that corresponds to about one conventional TROSY-HNCO. Note, that this does not even take into account that the reference

experiment (Figure 3.1; a, (i)) or the shifted experiment (Figure 3.1; a, (ii)) can be used for backbone assignment, which is difficult with the non-decoupled TROSY-HNCO. A further drawback of the non-decoupled TROSY-HNCO is that the overlap is more pronounced.

3.5.3 Error estimation

The random error in coupling values obtained by quantitative J -correlation methods mainly depends on the intensity of the reference spectrum (abbreviated as) I_{ref} , and the random noise (denoted as) σ in the attenuated spectra. This random error in coupling values got by quantitative J -correlation method is approximated as, $\sigma/[2\pi(\zeta + \Delta)I_{ref}]$ where $2(\zeta + \Delta)$ is the total duration in which that particular coupling is active in the attenuated spectrum. So for a signal to noise ratio of 50 in the reference spectrum of the CBCA(CO)NH experiment, the random error in $^1J_{C\alpha C'}$ [$2(\zeta + \Delta) = 18ms$] will be 0.35 Hz ($\Delta = 4.85$ mS, $\zeta = 4.5$ mS). Except for couplings between $^{13}C'$ and carbons resonating in the C_α region, all other couplings attenuate the reference and attenuated spectrum in the same way, and therefore do not affect the derived coupling. The inter-residue $^2J_{C\alpha C'}$ coupling, on the other hand, affect the attenuated and the reference spectrum differentially. The magnitude of $^2J_{C\alpha C'}$ with the next residue being in *cis* position is ≤ 3 Hz [91]. This attenuates the reference spectrum by less than 0.4% and the attenuated spectra by 1.5%, and the I_{att}/I_{ref} ratio is decreased by 1.2%, resulting in a coupling which is at most 1.2% closer to 55 Hz than its real value. When the bonds are *trans*, as it might be the case for residues preceding proline, the coupling could be as large as 6 Hz [91] resulting in a more pronounced overestimation of the derived coupling. However, these residues are any way not detected in CBCA(CO)NH experiments. Due to the fact that the selectivity of the C_α inversion pulse is not perfect, other long-range couplings that may affect the derived $^1J_{C\alpha C'}$ value can include $^2J_{C'C\beta}$ and $^3J_{C'C\gamma}$. These couplings, however, are comparatively smaller. Assuming $^2J_{C'C\beta}$ and $^3J_{C'C\gamma}$ values of 3 Hz and 4 Hz, respectively, the cumulative effect of all the three couplings would be to make the coupling closer to 55 Hz by less than 4.6%. So a coupling of 52.0 Hz will appear as 52.1 Hz and for an actual coupling of 54 Hz the derived coupling would be 54.05 Hz, which typically is considerably smaller than the random error.

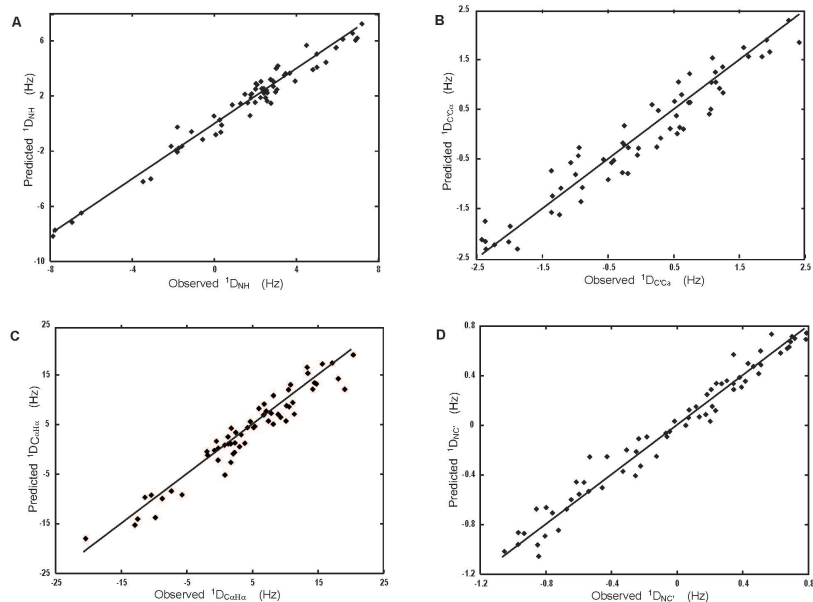


Figure 3.5: Correlation between observed dipolar couplings and values back-calculated from the crystal structure of ubiquitin (PDB code: 1ubq). The correlation coefficients are 0.98, 0.96, 0.98 and 0.96 for ${}^1D_{NH}$, ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C'N}$ and ${}^1D_{C\alpha C'}$ couplings, respectively

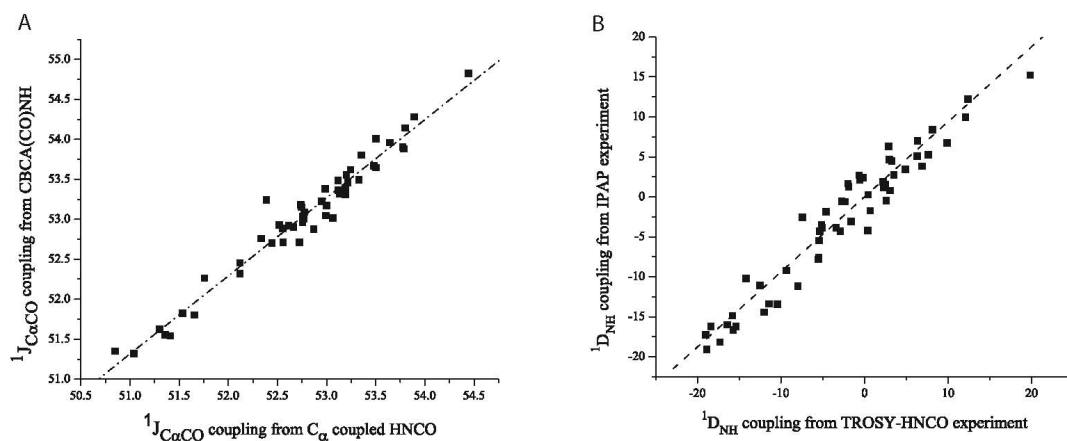


Figure 3.6: **A** Correlation between ${}^1J_{C\alpha C'}$ measured using the CBCA(CO)NH experiment and the same coupling measured using the C_α coupled HNC0 experiment for protein G. **B** Correlation between measured ${}^1D_{NH}$ dipolar couplings using TROSY-HNC0 and IPAP HSQC for DcuS-PD. The correlation coefficients were 0.98 and 0.97 respectively.

Hence, extracted coupling values do not contain any large systematic errors. This is also evident from isotropic ${}^1J_{C\alpha C'}$ couplings measured with the pulse sequence presented here, when compared to those obtained from a C_α -coupled HNCO experiment recorded on the third immunoglobulin-binding domain of streptococcal protein G (56 residues). The coupling values from these two measurements have a correlation of 0.98 (Figure 3.6A). Comparison of ${}^1D_{NH}$ coupling measured for DcuS-PD with the 3D TROSY-HNCO experiment presented here and those obtained from a 2D IPAP-HSQC have a correlation of 0.97 (Figure 3.6B)

3.5.4 Correlation of measured RDCs with structure

Figure 3.5 compares experimental dipolar couplings with those back-calculated from the 1.8 Å X-ray structure of ubiquitin [92] using singular-value decomposition as implemented in the program PALES [49]. Correlation factors of 0.98, 0.96, 0.98 and 0.96 are obtained for ${}^1D_{NH}$, ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C'N}$ and ${}^1D_{C\alpha C'}$ couplings, respectively. This demonstrates that the pulse sequences presented here allow accurate measurement of J and dipolar couplings in medium sized proteins. Residual dipolar couplings measured in DcuS (17 kDa)⁴ were used for the structure determination of this protein [44].

3.6 Conclusions

We have presented a simple strategy for simultaneous measurement of scalar and dipolar couplings in ${}^{13}\text{C}$, ${}^{15}\text{N}$ -labeled proteins. More than one type of coupling is extracted from a single reference experiment, thereby reducing the required measurement time for small to medium sized globular, as well as larger, unfolded proteins. The couplings are obtained either using the quantitative J -correlation approach or accordion spectroscopy; a second spectrum is recorded in which the signal intensities are modulated or signals are shifted according to the coupling of interest. Hence, simultaneous measurement of dipolar couplings does not result in an increase in resonance overlap compared to when the couplings are measured sequentially. The strategy presented is general and can be applied to various 3D experiments.

The interleaved 3D TROSY-HNCO and CBCA(CO)NH experiments presented here allow measurement of four backbone RDCs and ${}^1D_{C\beta H\beta}$, ${}^1D_{NH}$, ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C'N}$ and ${}^1D_{C\alpha C'}$ can be

used in a conventional structure calculation, enabling accurate determination of peptide plane orientations [72], or protein fold determination even in the absence of NOEs [46, 73, 74]. The combination of ${}^1D_{C\alpha H\alpha}$, ${}^1D_{C\beta H\beta}$, and ${}^1D_{C\alpha C'}$ provides χ_1 rotamer information, even in the absence of a backbone structure [84]. Moreover, the reference spectra, which are used here for extraction of couplings, can be used for traditional backbone assignment, further reducing the total required measurement time in a NMR structure determination project. In the broader sense of speeding up structure determination by NMR (e.g. as required for Structural Genomics applications), our strategy also has the advantage that peak picking, refinement of picked peaks and reassignment of resonances only has to be done for the reference spectrum, as long as the quantitative J -correlation approach is employed. For extraction of the ${}^1D_{C\alpha C'}$ coupling no additional experiment has to be processed, peak picked and analyzed. ${}^1D_{C\alpha C'}$ couplings are obtained in parallel with the ${}^1D_{C\alpha H\alpha}$ and ${}^1D_{C\beta H\beta}$ couplings. In addition, if the anisotropic medium is unstable or the protein degrades rapidly, as it is the case for DcuS, measured residual dipolar couplings will be in agreement with a unique effective alignment tensor.

Chapter 4

NMR structural studies on the periplasmic domain of DcuS and CitA

4.1 Introduction

Cells respond to changes in their physical and chemical environment by generating intracellular signals that lead to the modification of their structure, metabolism and movement. Most of these signals originate in one side of the membrane and require the sensing, and then the transfer of the signal to the other side of the membrane facilitating the regulatory responses [93]. Signal transduction across biological membrane is very important for the ability of cells to integrate and process environmental information. In most prokaryotes and a few eukaryotes this is usually performed by an evolutionarily conserved class of proteins, called the two component system [94, 95, 96]

The two-component system in their simplest form consist of two modular proteins, a sensor kinase and a response regulator [97, 98, 99]. In prokaryotes, the sensor protein mostly belongs to a larger family of protein kinases called Histidine kinases, which has the conserved histidine residue in the cytoplasmic domain that upon stimulus functions as a site for autophosphorylation. Autophosphorylation is carried out with the help of other subunit of sensor kinase within a homodimer [100, 101, 102, 103]. The response regulator catalyzes phosphoryl transfer from the phospho-histidine of the Histidine kinase to a conserved aspartate in its own regulatory domain, and controls the expression of target genes by transcriptional activation or repression [104]. The majority of response regulators are transcription factors with DNA-binding effector domains [99]. Refer Figure 4.1 for the schematic representation of the working of a typical two component system.

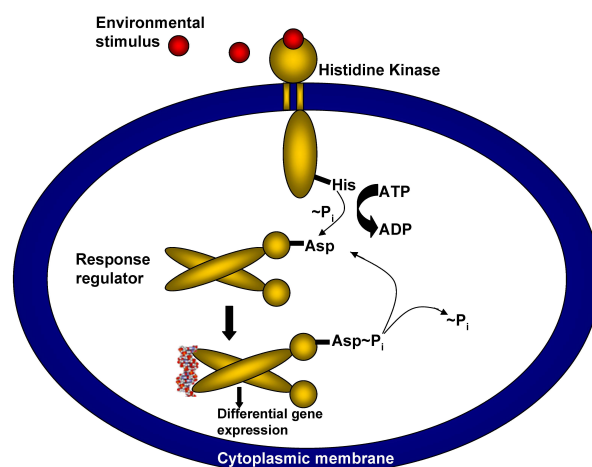


Figure 4.1: Schematic representation of a two component system: Transmembrane sensor histidine kinase act as the sensor domain, which upon stimulus undergoes autophosphorylation at the conserved histidine in the kinase core with an ATP getting converted to an ADP. The phosphoryl group is then transferred to the response regulator which in turn binds to the DNA starting the transcription of genes.

4.1.1 Histidine Kinase (HK)

In a typical two component system, a sensor histidine kinase monitor external stimuli and transmit this information to the response regulator by a phosphorylation event.

The main structural feature of the Histidine Kinase family include the diverse sensing domain and the characteristic kinase core composed of dimerization domain and an ATP/ADP-binding phosphotransfer or catalytic domain [99, 105]. In trans membrane histidine kinase (see Figure 4.2) the sensing domain is connected to the kinase core by a linker domain.

Kinase Core

All protein in the HK family have a characteristic kinase core composed of a dimerization domain and an ATP/ADP binding phospho-transfer or catalytic domain. [106] X-ray structure of C terminal half of *T.maritima* CheA protein [107] and the NMR solution structure of

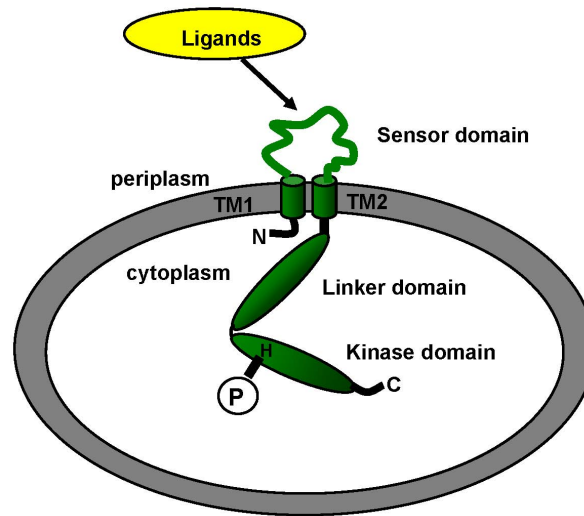


Figure 4.2: Schematic representation of sensory trans-membrane histidine kinase with the sensor domain present outside the cytoplasm. Ligands are sensed in the the periplasm with the help of the sensory domain. The linker domain connects the periplasmic domain to the kinase core which has the conserved histidine where upon binding of ligand, autophosphorylation take place.

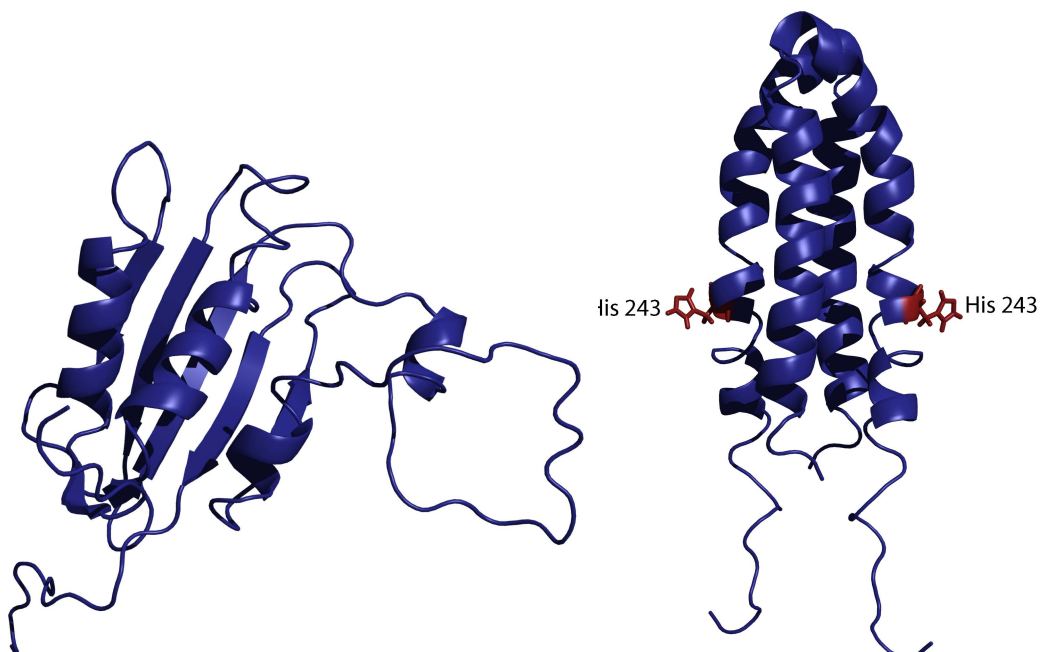


Figure 4.3: The kinase core of histidine protein kinases. The NMR structure of the (left) catalytic and (right) dimerization domains of *E. coli* EnvZ. The dimerization domain of EnvZ (right), like that of most histidine protein kinases, contains a conserved Histidine (blue) that is the site of phosphorylation.

catalytic domain [108], of *E.coli* EnvZ revealed an α/β sandwich fold consisting of five anti parallel β strands and 3 α helices (Figure 4.3 left). The dimerization domains of both EnvZ [109] and CheA form anti parallel four-helix bundles. The dimerization domain of EnvZ, like those of the majority of HKs, houses the conserved histidine, which is positioned midway along the exposed face of first helix. (Figure 4.3 right).

Linker domain

Its the least understood domain of HKs. But studies indicate that these linker regions are critical for proper signal transduction [110, 111, 112]. This domain may promote intra molecular association or as suggested by mutational analysis of the EnvZ linker region, correct structural alignment of monomers within the HK dimer [113].

Sensing domain

Signal transduction cascades begin with signal detection by the sensor domain. These domain senses wide range of ligands, redox potential and ions. These diverse sensing domains share little primary sequence similarity, indicating that they are designed for specific ligand/stimulus interactions. One of the best known sensing module incorporated into the sensing domain is the versatile PAS domains [114, 115].

PAS (PER-ARNT-SIM) domains are a family of sensor protein domains involved in signal transduction in a wide range of organism [116, 117]. The PAS module was first identified in the *Drosophila* clock protein PER and the basic helix-loop-helix containing transcription factors ARNT (aryl-hydrocarbon receptor nuclear translocator) in mammals and SIM (single-minded proteins) in insects [118]. A typical representative of the PAS domain is photoactive yellow protein(PYP) [119], a phtoreceptor to be involved in a photoactive response of the bacterium *Ectothiorhodospira halophila* to intense blue light [120].

Its structure has an α/β fold characterized by a central, twisted, 6 stranded anti parallel β sheet flanked on both sides by loops and helices. Helices $\alpha3$ and $\alpha4$, the short strand $\beta3$, the loop connecting the $\beta3$ strand to the helix $\alpha5$, the helix $\alpha5$, span across the other side of β sheet forms the binding pocket for the light sensitive chromophore p-coumeric

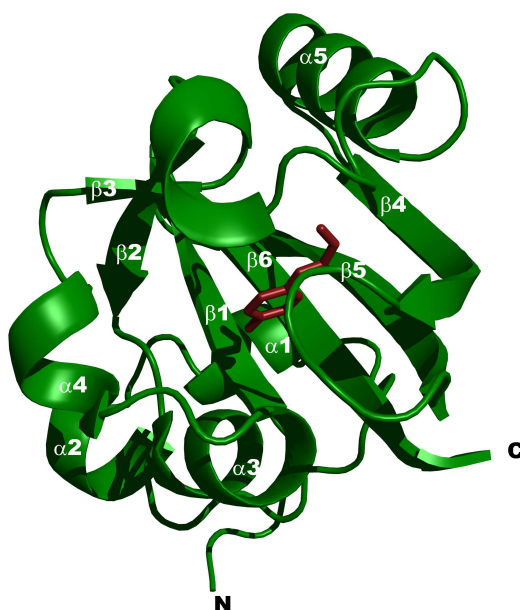


Figure 4.4: X-ray Structure of photo active yellow protein(PYP) revealing a PAS fold . The ligand, 4-hydroxycinnamyl chromophore position shown in sticks (red).

acid (see Figure 4.4). Although the PAS domains have little amino acid similarity, their three-dimensional structures appear to be conserved [121].

In the past two decades, hundreds of two component systems have been identified. However, many aspects of their molecular function remain obscure. In particular, for the ligand-binding sensor kinases little is known about such fundamental processes as the stereochemistry of ligand-binding, the nature of the associated conformational changes, and the mechanism of signal transduction to the cytoplasmic kinase domain.

The aim behind the study was to understand the signal perception and transduction of the signal from the sensor domain to the cytosolic domain. As a first step we describe the NMR studies on the soluble periplasmic domain of CitA and DcuS with and without their ligand. Here we describe the NMR solution structure of the ligand free DcuS sensory periplasmic domain and the X-ray structure of ligand bound and free structure of sensory domain of

CitA. These structures reveal a PAS fold, and are the first of the few determined structures having this fold, present outside the cytoplasm [116].

4.2 Periplasmic domain of the sensory domain of the two component fumarate sensor DcuS

The fumarate sensor DcuS is a prototype for a two component sensory histidine kinase with signal perception in the periplasm, trans membrane signal transfer [122, 123] , and autophosphorylation of a histidine residue in the kinase domain in the cytoplasm [124]. DcuS belongs to the CitA family of sensors that are specific for sensing di- and tricarboxylates [122, 123, 125, 126]. The periplasmic domain of the histidine autokinase CitA works as a highly specific citrate receptor, whereas DcuS uses any type of C4-dicarboxylate, like fumarate, succinate, and malate, as a stimulus [122, 125, 126, 127]. DcuS is predicted to consist of two transmembrane helices and of a periplasmic sensory domain enclosed by the trans membrane helices. The second trans membrane helix is followed by a cytoplasmic PAS domain of unknown function and the kinase with the consensus histidine residue for autophosphorylation. Previous studies suggest that fumarate sensing occurs in the periplasmic domain [122, 125, 127]. After autophosphorylation by DcuS the response regulator DcuR of the DcuSR system (Figure 4.5) activates the expression of target genes like *dcuB* and *frdABCD* encoding an anaerobic fumarate carrier DcuB and fumarate reductase [125, 126].

Biochemical studies on DcuS reconstituted in liposomes indicated that in the presence of fumarate and other dicarboxylates, kinase activity of DcuS is stimulated [122]. But the binding site, or the exact nature of the stimulus were yet to be fully characterized. Moreover, there was a lack of three dimensional structural information for any domain of DcuS or that of its homologues. This study was initiated to address these issues. Hence as a first step to understand the signal perception, NMR structural studies on the periplasmic domain of DcuS were performed. The structure of DcuS-PD derived using NOE was refined using RDCs for better convergence of the structure. The structure of DcuS-PD could be solved only without bound C4-dicarboxylates which precludes direct identification of the effector binding site. Titration experiments were necessary to establish the sensing mode of C4 dicarboxylates by DcuS-PD.

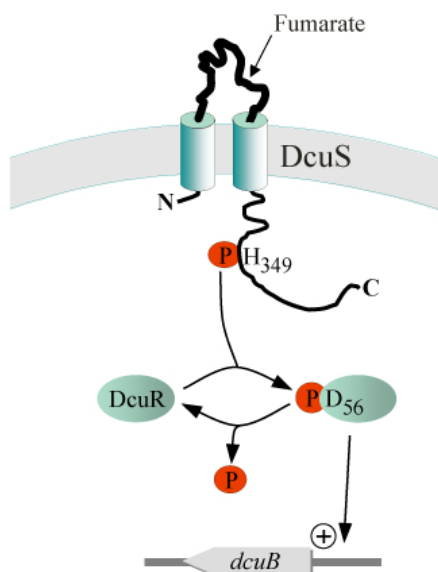


Figure 4.5: Schematic representation of DcuS-DcuR system. After the autophosphorylation of DcuS, phosphoryl group on histidine 349 in DcuS is transferred to aspartate 56 in DcuR. DcuR then binds to the *dcuB* codon, triggering the expression of the genes for fumarate uptake and respiration.

4.2.1 Material and Methods

The periplasmic domain of DcuS was produced in the lab of Prof. Gotfried Unden in University of Mainz. DcuS-PD was overproduced in *E. coli* strain carrying the DcuS_{45–180} expression plasmid [44]. The assignments were taken from the previously published assignment of histidine tagged DcuS-PD [128]. The assignment of DcuS-PD without the His-tag differs only slightly from the one with the His-tag. NMR measurements were carried out with the ¹⁵N-¹³C labeled protein received from Prof. Unden's group in Mainz. The wild type protein containing samples were usually around 1 mM, in a 50 mM sodium phosphate buffer at pH 7.0, including 200 mM NaCl, 0.8 mM CHAPS, 50 mM Glycine, 50 pM of PefablocSC, 0.01 % NaN₃ and H₂O/D₂O. 90/10%. NMR samples were stable for a week after which it degraded gradually.

NOESY experiments were measured and analysed by Dr. Lucia Pappalardo and Dr. Wolfgang Peti. Structural calculations were done by Dr. Jochen Junker in our lab.

Heteronuclear NOE experiment was measured with 5 second recycling delay

RDC measurements

To refine the structure, several sets of RDCs were measured. Protein was oriented in bacterial filamentous phages. The concentration of the phage in the protein sample was 10 mg/ml giving deuterium quadrupolar splitting of 9 Hz. A set of $^1D_{NH}$, $^1D_{NC'}$, $^1D_{CC\alpha}$ and $^1D_{C\alpha H\alpha}$ RDCs of DcuS-PD were measured using the pulse sequence described in chapter 3. The magnitude of the alignment tensor (Da) obtained from histogram of measured dipolar couplings is 10.7 Hz for $^1D_{NH}$ and the corresponding rhombicity (R) is 0.63. RDCs were applied in the structural calculation with different weight factors 1.0($^1D_{NH}$), 0.4($^1D_{C\alpha H\alpha}$), 4.4($^1D_{NC'}$) and 2.5 ($^1D_{CC\alpha}$). The weighting factor used here reflect the gyromagnetic ratios of the nuclei involved in the coupling and also the bond distances.

Effector titration of DcuS-PD

To study the effector binding, wild type DcuS-PD and the mutant forms, DcuS(F120M)-PD, DcuS(R147A)-PD, and DcuS(H110A)-PD were dissolved in a NMR buffer at pH 7.0 containing 45 mM sodium/potassium phosphate, 10 % D₂O, 200 mM NaCl, 0.01 % NaN₃, 50 uM Pefabloc, 50 mM glycine, and 4.5 mM imidazole. Titration with the effectors were conducted by stepwise addition of properly buffered fumarate and tartrate solutions (pH 7.0 and upto 300 mM final concentration) to a sample of 1.2 mM of DcuS-PD. ¹⁵N-¹H HSQC experiments were measured for each step of titration.

Experiments were performed in Bruker 800 DRX, 700 Avance and 600 DRX MHz spectrometers at 30°C. The data sets were processed and analyzed using the NMRPipe/NMRDraw software package.

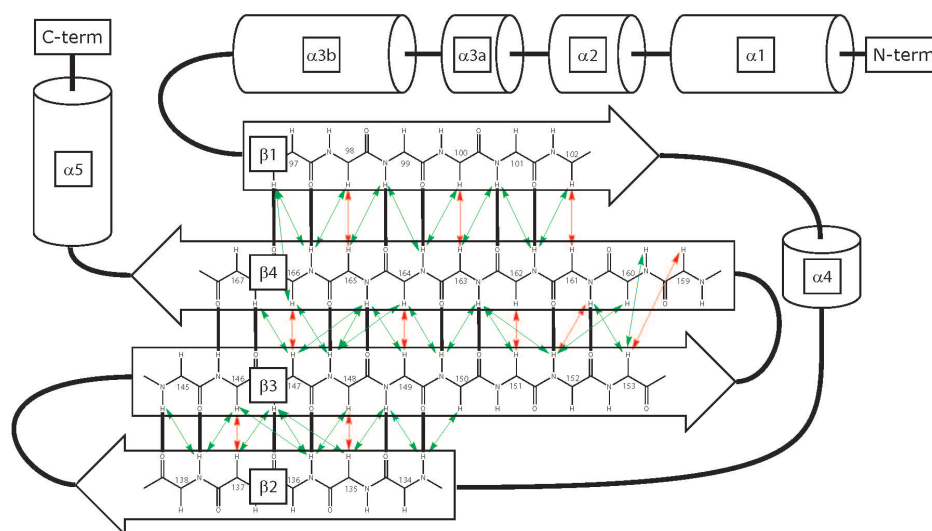


Figure 4.6: Secondary structure of DcuS-PD together with the intra-strand NOEs from which the topology of the β -strands were derived.

4.2.2 Results and discussions

The NMR solution structure of DcuS-PD without bound C4 dicarboxylates was determined. For the final refinement of the structure, four types of RDCs were measured and included in the structural calculations. Detail description of the structure follows.

4.2.3 Structure of DcuS-PD

Secondary structure of DcuS-PD

The secondary structure of DcuS-PD (Figure 4.6) consists of a long N-terminal α -helix (α_1) ranging from amino acid 46 to 64 with a continuation from 68 to 72 (α_2). After a short loop there is another α -helix (α_{3a} : 77-79) and (α_{3b} : 83-92) that is connected to the first β -strand (β_1 : 97-102) of the four stranded anti parallel β -sheet. β_1 is connected via an α -helix (α_4 : 126-128) and a long loop to the second β -strand (β_2 : 134-138), which is connected by a short loop to the third strand (β_3 : 145-153). Yet, another turn connects to the fourth β -strand (β_4 : 159-167). From this strand the C-terminal helix follows after

a short helix (α_5 : 174-179). The secondary structural elements have been established by secondary chemical shift and characteristic sequential NOEs and connectivity of the four stranded anti parallel β -sheet by H^N , H_α and H_α , H_α cross strand NOEs (Figure 4.6).

Structure determination and discussion

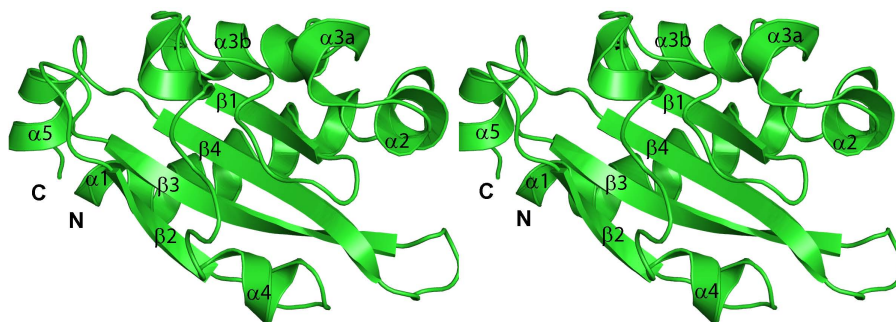


Figure 4.7: Stereo view of the NMR solution structure of periplasmic domain of DcuS. It reveals a PAS fold with 4 β strands and 5 main α helix. The structure is a α,β -fold where both sides of the large β -sheet form hydrophobic cores with α -helices and the long connector between helix α_4 and strand β_2 . The C-terminal helix is not fixed by NOEs and shows only small dipolar couplings, in agreement with a flexible helix.

Final refinement of the structure was done with 382 RDCs. These include 107 ^{15}N - $^1\text{H}^N$, 95 $^1\text{H}_\alpha$ - $^{13}\text{C}_\alpha$, 114 ^{15}N - $^{13}\text{C}'$ and 66 $^{13}\text{C}'$ - $^{13}\text{C}_\alpha$ coupling. $^1\text{J}_{\text{C}_\alpha\text{H}_\alpha}$ couplings were also used to define ϕ angles. ϕ is negative for $^1\text{J}_{\text{C}_\alpha\text{H}_\alpha} > 137$ Hz and in the α -helix range for $^1\text{J}(\text{C}_\alpha, \text{H}_\alpha) > 145$ Hz [129]. The structure is well ordered in the region 46-168, while the first few residues in the N and C terminal are not well ordered. The RDCs in these region are also small compared to the rest of the residues in the structure. The Het-NOE values in these regions of the protein are also relatively small, indicating dis-orderness in these parts of the protein. Figure 4.7 shows the mean structure derived from the 10 structures with the lowest energy.

The structure is a novel α,β -fold completely dissimilar of the four helix bundle structure of the aspartate sensor [130, 131]. Consistent with gel shift data, the average $T_{1\rho}$ of the wild type protein is around 110 mS corresponding to a monomer. This is dissimilar to the aspartate sensor which was a dimer in solution. In a DALI [45] search the closest match is photoactive yellow protein (PYP) from *Halorhodospira halophila* [120], which also shows an α,β -fold, however, with 5 instead of 4 β -strands. The topology of strands β_3 , β_4 and β_5 of PYP is

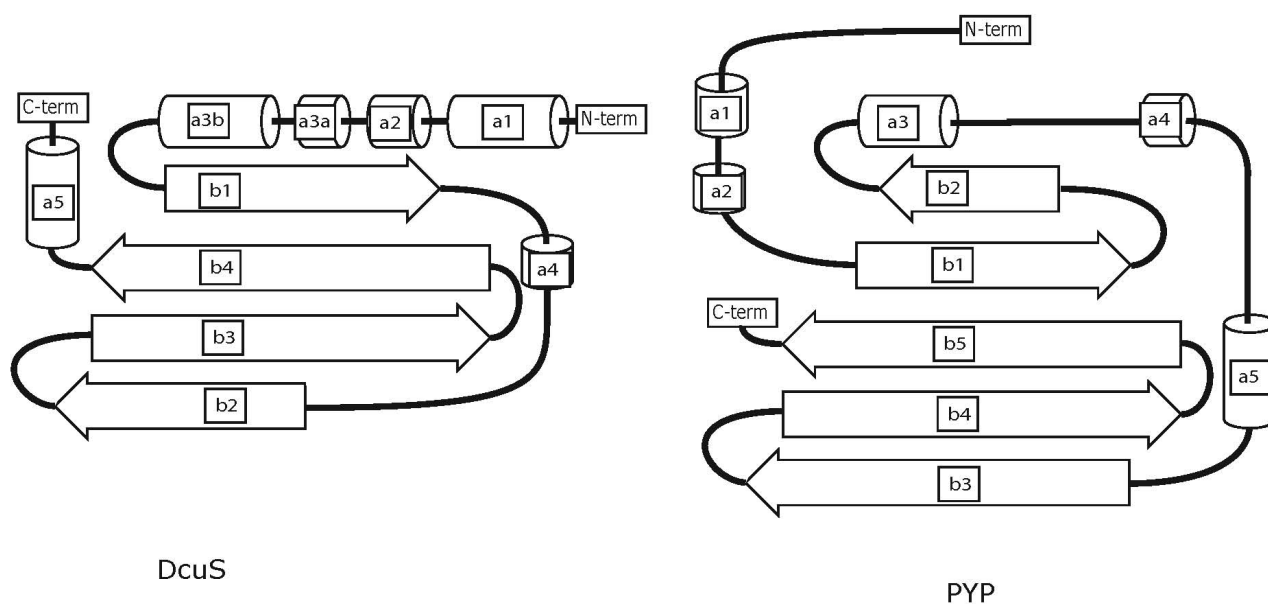


Figure 4.8: Comparison of secondary structure of periplasmic domain of DcuS with its structure homologue, photo active yellow protein (PYP).

similar to the strands β_2 , β_3 and β_4 of DcuS-PD (see Figure 4.8 for comparison of secondary structural element). However, the rest of the secondary structure is quite dissimilar. While in PYP the PAS core domain connects the strands β_2 and β_3 by crossing the whole β -sheet in a diagonal manner, there is no PAS core domain in DcuS-PD and the connection between sheets β_1 and β_2 is achieved on one side of the β -sheet. Similar to PYP, there are two hydrophobic cores on both sides of the β -sheet formed. Helices α_1 and α_{3b} bind to the bottom side of the β -sheet, while helix α_4 and the connector attach to the upper half of the β -sheet. In PYP the chromophore binding site is formed by the PAS core domain. Dissimilar to PYP, in DcuS-PD residues located in the β -sheet (Arg147) as well as in the connector (Arg107 and His110) across the β -sheet contribute to the putative binding site of fumarate.

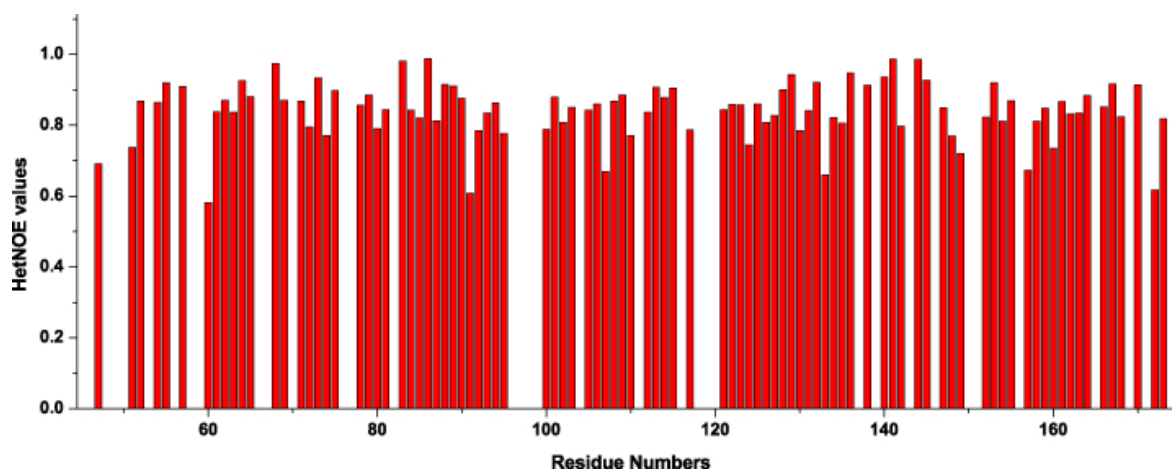


Figure 4.9: ^{15}N - ^1H Heteronuclear NOE values for DcuS-PD are plotted against residue number. Most of the values are between 0.6-0.9 indicating highly rigid structures. Het-NOE values for a number of residues in the N and C terminals were unable to be determined due to poor sensitivity in the reference spectra.

4.2.4 Binding studies on DcuS-PD

DcuS senses a number of C4-dicarboxylates like fumarate (Kd 5 mM), succinate etc [122]. The isolated DcuS-PD was used for the binding studies of the effectors to the domain by NMR spectroscopy. During the addition of fumarate, sharpening of peaks were observed for some of the amino acids of the periplasmic domain in ^{15}N - ^1H HSQC spectra (Figure 4.10). The affected residues cluster in a well defined region in the structure of the periplasmic domain of DcuS involving amino acid residues 107-168 (Figure 4.12 Left panel). This region in the structure correspond to the binding pocket of citrate in the periplasmic domain of CitA. There were no chemical shift changes observed during the titration. Sharpening of peaks were not observed when the protein sample contained a detergent (CHAPS) used for the titration with fumarate. This suggest that the detergent might have shielded the binding pocket of DcuS-PD.

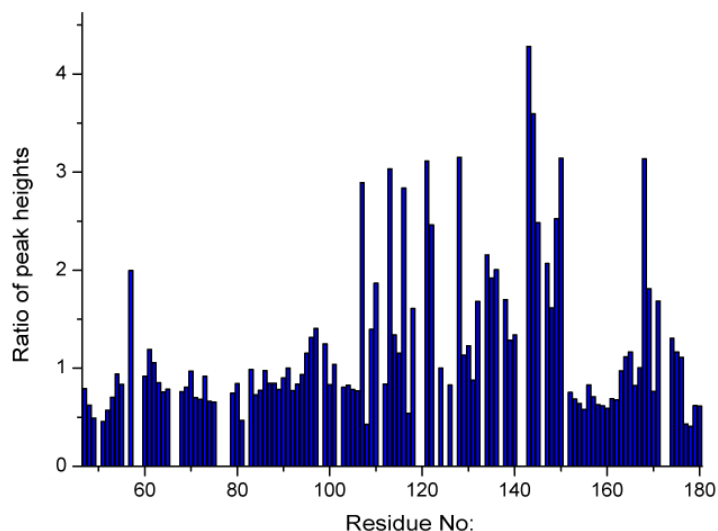


Figure 4.10: Ratio of peak heights in ^{15}N - ^1H HSQC spectrum after the addition of 30-fold excess of fumarate to DcuS-PD. Titrations were carried out with addition of properly buffered sodium fumarate to a solution of protein at 1.2 mM. Residue that shows the maximum peak intensity is Q144

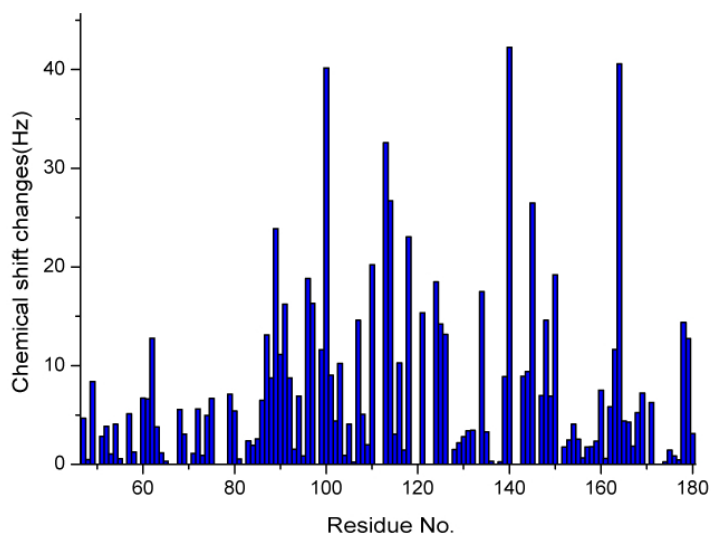


Figure 4.11: Plot of chemical shift changes in ^{15}N - ^1H HSQC when D-tartrate (50-fold excess) is added to periplasmic domain of DcuS (1.2 mM). The chemical shift plotted here is the sum of ^{15}N and $^1\text{H}^N$ chemical shift, properly scaled using the method given in chapter 2, section 2.2.9. Residue that shows the maximum chemical shift change is G140.

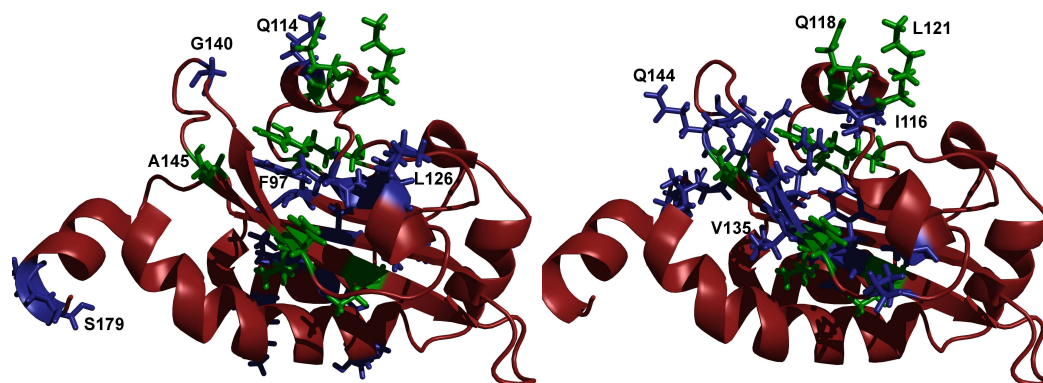


Figure 4.12: (Right) Structure of DcuS periplasmic domain showing the residues most affected by fumarate titration. The residues, represented as sticks in the structure are R107, A113, I116, K121, A128, N134, A136, A143, Q144, A145, R147, F149 and T150. (Left) Structure of DcuS periplasmic domain showing the residues most affected by the tartrate titration. The residues shown as sticks on the structure are V89, K91, L96, F97, V100, H110, A113, Q114, Q118, K121, D124, N134, G140, A145, and T150. Common residues that are affected both by fumarate and tartrate titration are represented in green colour and the residues individually affected by the ligand titration are shown in blue sticks.

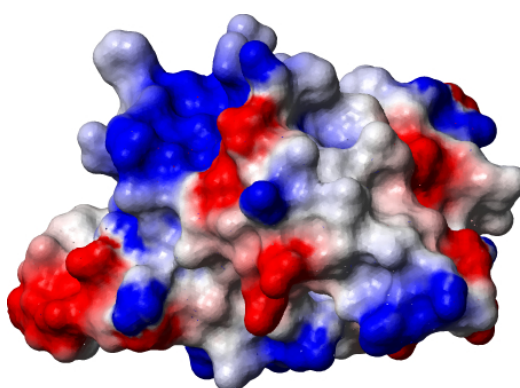


Figure 4.13: Electrostatic surface potential of DcuS-PD with positive and negative potentials coloured blue and red respectively. This reveals a positive charged surface region in the putative binding site of DcuS-PD. Residues of this region of the protein are mainly affected by the ligand titration.

D-tartrate is regarded as a non-physiological stimuli of DcuS [132], but is structurally closely related to other C4 di-carboxylates and binds to DcuS with a higher affinity than fumarate (apparent K_d is 0.5 mM [132]). Chemical shift changes were observed for a number of residues when DcuS-PD was titrated with D-tartrate (Figure 4.11). Most of these residues belong to the same region that was affected by fumarate binding. Such chemical shift changes were observed only for D-tartrate and not for L-tartrate. There was a set of common residues that were affected both by the fumarate and tartrate titration (Ala113, Lys121, Asn134, Ala145 and Thr150-refer Figure 4.12 : Residues coloured in green). Other residues were specifically affected by fumarate (Arg107, Ile116, Ala128, Ala136, Ala143, Gln144, Arg147 and Phe149-refer Figure 4.12: Residues coloured in red in the right panel) or D-tartrate addition (Val89, Lys91, Leu96, Phe97, Val100, His110, Gln114, Gln118, Asp124 and Gly140-refer Figure 4.12: Residues coloured in blue in the left panel). A comparison of the affected residues reveal that fumarate and tartrate bind to the same positively charged binding pocket in the DcuS-PD domain (Figure 4.13). In addition, some residues located outside this defined binding pocket are specifically affected by D-tartrate and more non-polar residues are affected by fumarate than D-tartrate, the latter containing two additional hydroxyl groups. Titration experiments were also performed with nitro-propionate (known to be sensed by DcuS at higher affinity than fumarate, K_d 0.4 mM [132]) as ligand, but unfortunately no visible effects were observed in the spectrum during the titration.

Alignment of the amino acid sequence of the periplasmic domain of DcuS from *E. Coli* with those from the citrate sensor CitA from *Klebsiella pneumoniae*[127, 133] and other DcuS and CitA proteins revealed conserved residues in the members of the CitA/DcuS family (Figure 4.14). Most of the conserved residues are hydrophobic and might be structurally relevant. Charged residues are conserved in both types of proteins (D102, R107, H110 and R147).

Mutation of the charged residue like R147, H110 to a non polar residue like Alanine completely repress *dcuB'*-*lacZ* reporter gene fusion, indicating loss of effector binding. Another

set of mutation for amino acid residues that are specifically conserved in DcuS-type proteins (M103, F120, F149 and Q159) also impair DcuB⁻lacZ expression [132].

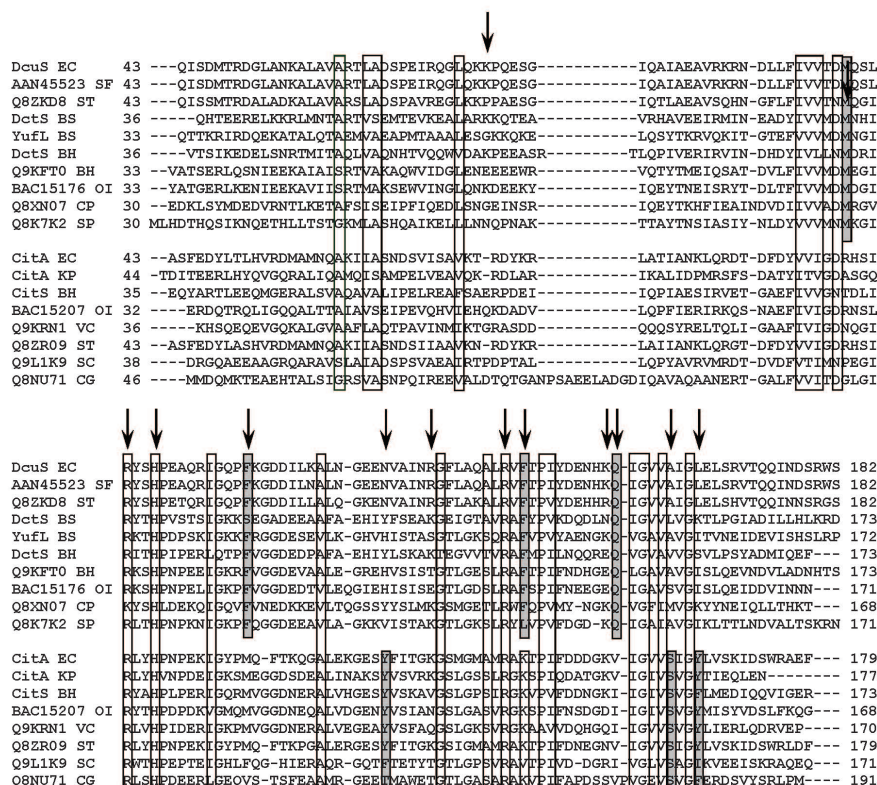


Figure 4.14: Comparison of the amino acid sequences of the periplasmic sensor domains of C4-dicarboxylate or citrate sensory histidine kinases. Amino acid residues conserved in both types of proteins are boxed; amino acids conserved in only one of the types are boxed and shaded. Amino acid residues of *E. coli* DcuS that were changed by site-directed mutagenesis are indicated by arrows. BS, *Bacillus subtilis*; BH, *Bacillus halodurans*; CG, *Corynebacterium glutamicum*; CP, *Clostridium perfringens*; EC, *Escherichia coli*; KP, *Klebsiella pneumoniae*; OI, *Oceanobacillus ihejensis*; SC, *Streptomyces coelicolor*; SF, *Shigella flexneri*; SP, *Streptococcus pyogenes*; ST, *Salmonella typhimurium*; VC, *Vibrio cholerae*. This figure has been reproduced from [132].

The wild type periplasmic domain used in the study was well folded (See Figure 4.15 peaks in blue) and the residues R107, H110, and R147, found from in silico studies to be essential for fumarate binding appear to be in close proximity in the structure, suggesting that the binding motif is retained in the periplasmic domain. In order to observe the influence of mutation on the effector binding on DcuS-PD, ¹⁵N-¹H HSQC spectra of the mutants were measured. ¹⁵N-¹H HSQC spectra of the mutants H110A and F120M showed very small

chemical shift dispersion in proton dimension which are characteristic of unfolded proteins (Figure 4.15 panel B and C). For R147A, severe line broadening of resonances were observed denoting a high oligomerization state of the protein(Figure 4.15 panel D). Moreover R147A-PD was highly unstable and precipitated in the NMR tube very quickly. Unfolding and aggregation of these mutants indicate a loss of structure and stability in these mutants. This study illustrates that these residues are very important for the fold of DcuS-PD. The decrease in the binding affinity of ligands in the biological assay could partly be due to these point mutations which are related to the loss of the three dimensional structure of mutated sensory domain.

4.2.5 Conclusion

Solution structure of the ligand free periplasmic domain of DcuS was determined. Structure reveals a PAS fold with the putative binding site of the ligand formed by the beta sheet and the inter strand loops. Titration experiments were carried out with dicarboxylates namely, fumarate and tartrate. The affected residues cluster around the putative binding pocket of DcuS-PD. Upon fumarate titration only sharpening of peaks were observed indicating a stabilization of the structure with the binding of fumarate. Upon titration with the non-physiological stimuli tartrate, small chemical shift changes were observed for a subset of peaks in the HSQC spectrum. The presence of two hydroxyl group in tartrate makes this C4 di-carboxylate to bind to DcuS-PD with much higher affinity. Because of the presence of two more charges, more polar residues are affected upon tartrate titration. However, even with tartrate, no major conformational changes could be observed in the C-terminus of the protein to indicate signal transduction.

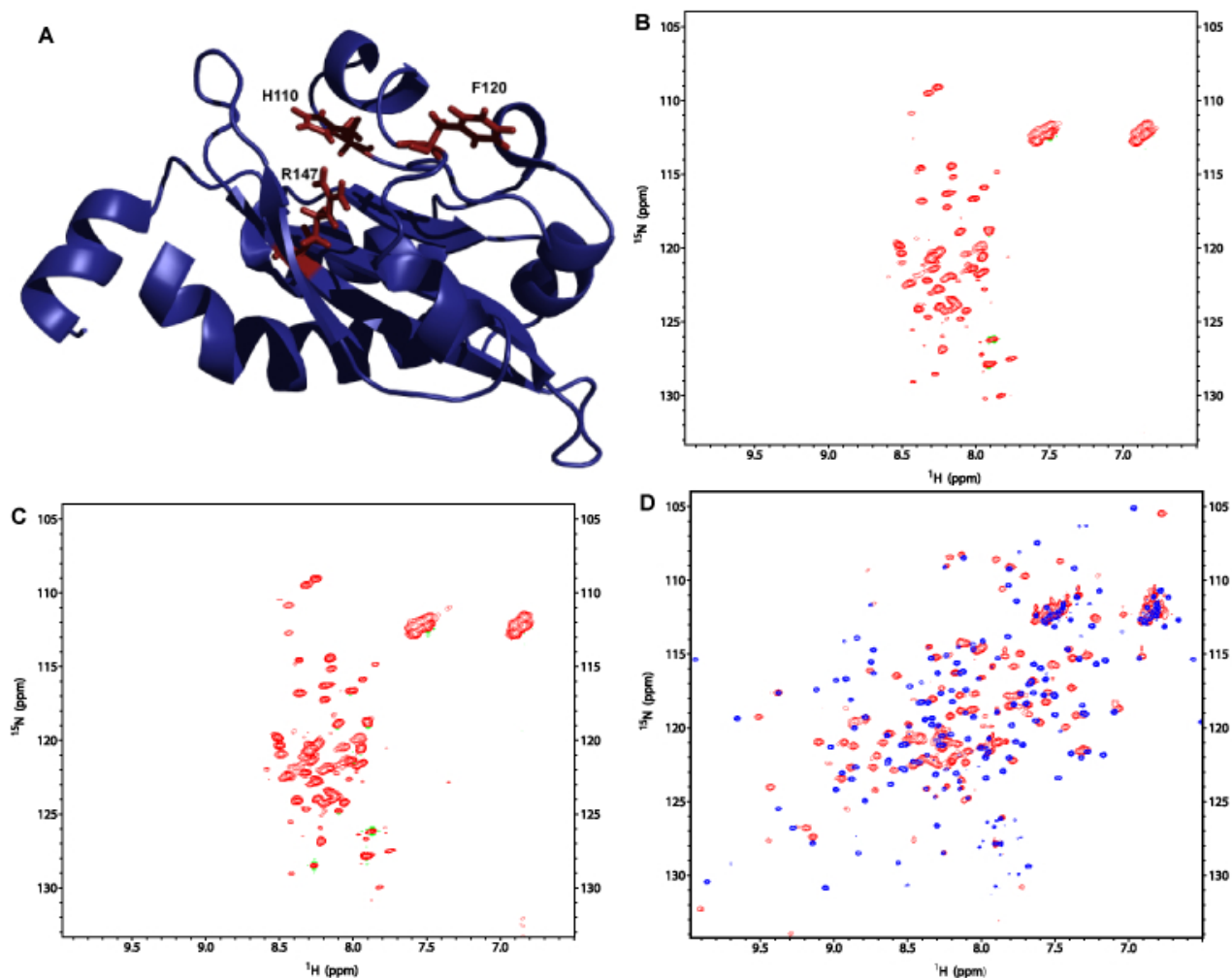


Figure 4.15: **(A)** DcuS periplasmic domain with the residues that are mutated are indicated by sticks. The residues shown are H110, F120 and R147. **(B)** ^{15}N - ^1H HSQC of DcuS-PD mutant H110A. The protein was unfolded with small dispersion in chemical shift in the proton dimension. **(C)** ^{15}N - ^1H HSQC of DcuS-PD mutant F120A. This protein is also unfolded with small dispersion in chemical shift in the proton dimension. **(D)** Overlay of ^{15}N - ^1H HSQC of wild type DcuS periplasmic domain (blue) with R147A mutant of DcuS-PD (red). Severe line broadening is observed for the peaks of the mutant in the HSQC. Large number of peaks were missing when compared to peaks in the HSQC of the wild type DcuS-PD.

4.3 Periplasmic domain of the sensory domain of the two component citrate sensor CitA

4.3.1 Introduction

Several species of enterobacteria are able to use citrate as the sole carbon as energy source in both aerobic and anaerobic conditions. The most extensively studied organism with respect to citrate metabolism is *Klebsiella pneumoniae*. During aerobic growth, citrate is taken up in symport with protons and then metabolized via the tri-carboxylate cycle. The transport is catalysed by the a protein encoded by the *citH* gene, which is expressed constitutively in the presence of oxygen [134, 135, 136]. During anaerobic growth, uptake of citrate is a Na^+ - dependent process [137] catalysed by the CitS protein. Within the cell, citrate is metabolized by the citrate specific fermentation enzymes like Citrate lyase, oxaloacetate decarboxylase etc.

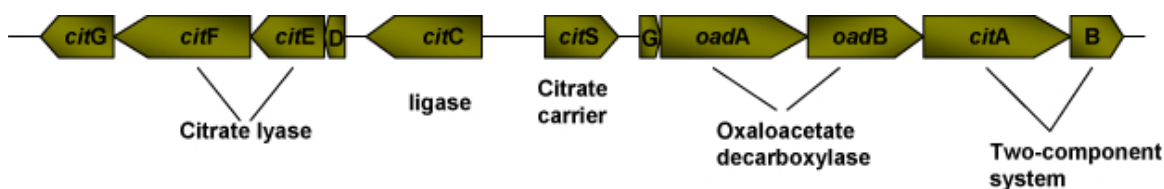


Figure 4.16: The *cit* regulon of *Klebsiella pneumoniae*. The regulator CitB binds between *citC* and *citS*, where it induces the transcription of the genes required for citrate fermentation.

The cluster of genes which confer the property for bacteria to grow anaerobically is believed to form a regulon [138] (See Figure 4.16). These include genes for the citrate carrier protein CitS [139], that is flanked downstream by the genes for oxaloacetate decarboxylase(*oadGAB*) [140, 141, 142] and upstream by the *citCDEFG* genes coding for citrate lyase ligase (CitC), citrate lyase (CitDEF) and a protein of yet unknown function (CitG) [143]. The two component regulatory system *citAB* located downstream of the *oadB* gene codes for the sensor histidine kinase (CitA) and its response regulator (CitB) [138]. Mutational analysis suggest that the CitAB system is essential for the expression of citrate specific fermentation enzymes. CitB null mutants of *Klebsiella pneumoniae* were unable to grow

anaerobically with citrate as the sole carbon source and none of the citrate specific fermentation enzymes were synthesized [138]. This showed that *citS*, *oadGAB* and *citDEF* required the CitB protein for expression and therefore are part of a regulon. From analogy to other two component systems [144, 124, 145], the signal trigger should then come from the sensor histidine kinase CitA [127]. In the presence of specific signal molecule(citrate), CitA is first autophosphorylated with ATP as substrate. CitA delivers the phosphoryl moiety to CitB. CitB then binds to a specific target within the *cit* regulon (Figure 4.16), thereby inducing the transcription of its genes.

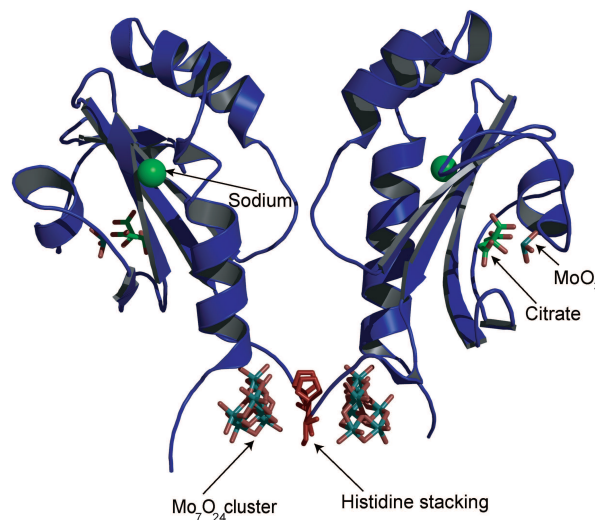


Figure 4.17: X-ray structure of GJ dimer of CitAP in complex with citrate and molybdate (pdb id 1P0Z). The citrate-binding pocket is on the right. Citrate-molybdate (right) and isopolymolybdate (bottom) groups are shown as stick figures. A putative sodium ion is shown as a green sphere

CitA, like DcuS is a membrane bound sensor kinase consisting of a periplasmic domain flanked by two trans-membrane helices, a linker domain and the conserved kinase domain. The structural morphology of CitA is also very similar to DcuS. As mentioned above, the signal detection happens in the periplasmic domain of CitA (hence forth called CitAP). CitAP is very specific for citrate [127] and bio-chemical studies indicate that citrate binds to CitAP with a K_d of $5 \mu\text{M}$ [133]. The aim of the study was to investigate the structure of CitAP with and without the ligand and to determine the structural changes that may cause the signal transduction across the membrane into the cytoplasmic domain.

X-ray structure of CitAP in complex with citrate and molybdate was already determined [146] (pdb id: 1P0Z, Figure 4.17). The structure reveals that CitAP adopts a PAS domain fold with the ligand binding pocket formed by the β sheets, a short helix and the flanking random coil residues (Figure 4.17). Two arginines, a lysine and a histidine were found to interact with citrate in the binding pocket. Due to the crystallization condition, there was also a molybdate(MoO_3) moiety in the binding pocket which was in complex with citrate. The presence of molybdate in the binding pocket was unexpected as citrate is transported into the cell without any metal complex formation. Two dimeric forms of CitAP were observed in the crystal lattice. One of it might be artificially induced by the stacking of the C-terminal histidine tags, used for the purification of the protein, as seen in the Figure 4.17. A large molybdate cluster(Mo_7O_{24}) was also found interacting with the N- and C-terminal region. The citrate free structure of CitAP was still undetermined and hence the structural change accompanying the ligand binding was poorly understood.

4.3.2 Materials and Methods

Sample Optimization

Initially, CitAP protein was produced in the lab of Prof. Bott in Jülich. The protein was first produced with a C-terminal 6 membered histidine tag. For testing the stability of the protein, two fragments of the protein were made, a longer fragment consisting of amino acids 39-185(Cit₃₉₋₁₈₅) and a shorter one consisting of amino acids 45-176(Cit₄₅₋₁₇₆). The longer protein fragment included a small part of the trans-membrane helix in the N- and C-terminal. ^{15}N - ^1H HSQC spectrum of the ^{15}N labeled protein produced from the two fragments are shown in the Figure 4.18. The longer fragment (panel **(A)** Figure 4.18) was unstable and precipitated very quickly. The HSQC spectrum of the longer fragment had more peaks than the one from the shorter fragment (panel **(B)** Figure 4.18) but most of the peaks were very broad. This could be due to the presence of hydrophobic residues in the longer construct (more specifically tryptophan, leucine etc) resulting in an increased propensity to aggregate. The shorter fragment (more stable) was used for all further studies as the peaks in its ^{15}N - ^1H HSQC spectrum had more favorable line width for three dimensional experiments.

Longer fragment -39-MSWYLTMDITEERLHYQVGQRALIQAMQISAMPELVEAVQKRDLARIKALIDPMRSFSDATYITVG DASGQR
 LYHVNP DEIGKSMEGGDSDEALINAKSYVSVRKGSLGSSLRGKSPIQDATGKVGIVSVGYTIEQLEWLSLQIS-185+6H

Shorter fragment -----45-MDITEERLHYQVGQRALIQAMQISAMPELVEAVQKRDLARIKALIDPMRSFSDATYITVG DASGQR
 LYHVNP DEIGKSMEGGDSDEALINAKSYVSVRKGSLGSSLRGKSPIQDATGKVGIVSVGYTIEQLE-176+6H-----

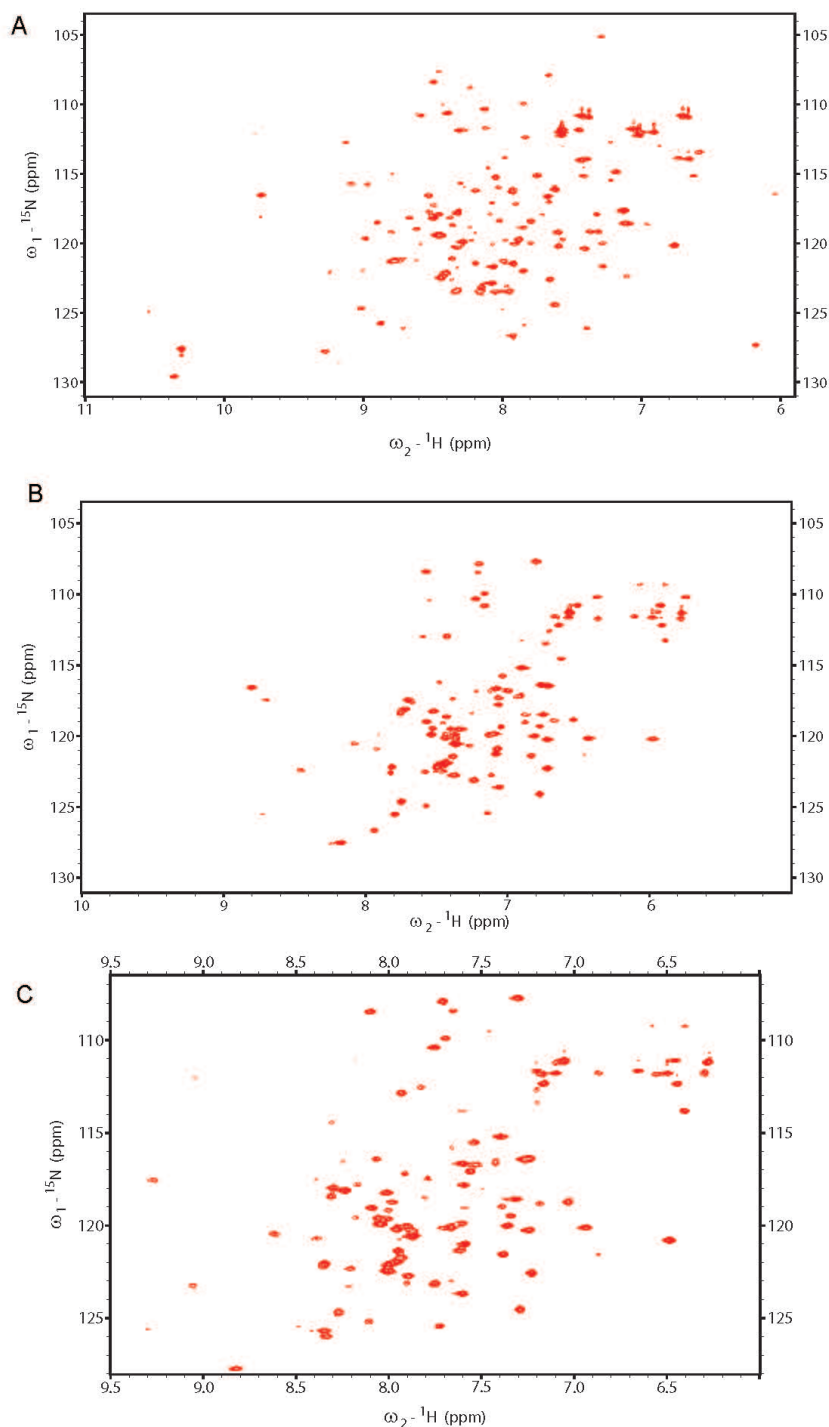


Figure 4.18: Amino acid residues for the short and long fragment of the protein prepared are given on the top. **(A)** ^{15}N - ^1H HSQC of longer construct. The spectrum was measured on the 600DRX spectrometer at 310K. **(B)** ^{15}N - ^1H HSQC spectrum of the shorter construct measured on the 600 DRX spectrometer at 310K. Only 62% of the expected peaks are seen in the spectrum. **(C)** ^{15}N - ^1H HSQC of CitAP of the shorter construct at higher salt concentration(300 mM). Fewer peaks are seen in this spectrum compared to the HSQC spectrum in **(B)**

Nevertheless the shorter construct had only 60% of the expected peaks. Hence to further improve the spectrum of the smaller construct, different conditions for experiments were tried.

1) Change of pH of the buffer from 7 to 6 did not improve the spectrum. Lowering the pH reduces the amide-proton and water chemical exchange. This confers better relaxation property for NMR measurements. Hence all later experiments were measured at pH 6.00

2) Salt titration was carried out with step-wise addition of properly buffered NaCl solution (panel (C) Figure 4.18). Again no improvement in the spectra were observed (0 to 100 mM). On addition of more salt, some resonances start to disappear (100 to 300 mM). All later experiments were measured with no additional salt other than the 50 mM sodium present in the phosphate buffer.

3) To determine the optimal temperature for measurements, HSQC experiments were measured by varying temperature from 290 K to 310 K. Better spectra were observed at higher temperatures.

The final buffer condition were 50 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ at pH 6.00. The spectrum also improved when the histidine tag on the protein was cleaved off. The conditions optimized for the histidine tagged protein were later used for all experiments measured for preparations without the histidine tag. All three dimensional experiments were measured at 298 K using the protein prepared with the shorter construct and without the histidine tag. At higher temperatures the protein started to precipitate in the course of longer experiment. The triple resonance experiments measured for the backbone assignment are listed in the table 2.2 and 2.3. The citrate bound form of CitAP was obtained by the step wise addition of sodium citrate to the citrate free sample of CitAP. Three dimensional experiments for the backbone assignment were repeated for the citrate bound form of CitAP. These samples were stable for a week and then degraded very quickly.

Later, ^{15}N , ^{13}C labeled protein samples were produced in our department in the molecular biology laboratory, supervised by Dr. Stefan Becker. Samples were very stable and did not degrade even after keeping them at 4°C for months. These samples were used to re-measure the three dimensional experiments acquired with the previous sample. The measurements

were also repeated for the citrate bound form of CitAP. Because of the higher stability of the samples all three dimensional experiments were measured at 310 K. These experiments were used for the protein backbone assignments of citrate free and citrate bound form of CitAP.

In order to measure RDCs, citrate free and bound form of CitAP samples were oriented in bacterial filamentous phages. The concentration of the phages in the sample were 12 mg/ml in the citrate free and 15 mg/ml in the citrate bound CitAP. $^1D_{NH}$ couplings were measured for the citrate bound form of CitAP. A set of $^1D_{NH}$, $^1D_{NC}$ and $^1D_{CC\alpha}$ were measured for citrate free form of CitAP. The $^1D_{NCO}$ and $^1D_{NH}$ couplings were measured using the TROSY-HNCO sequence described in chapter 3. $^1D_{C'C\alpha}$ couplings were measured with HNCO experiment by removing $C\alpha$ decoupling in the C' dimension. The measured residual dipolar couplings of CitA of the free and bound form are given in the Table B.6 and B.5 respectively.

Simultaneous ^{15}N and ^{13}C NOESY experiment was measured with a mixing time of 120 ms. Heteronuclear NOE experiments were measured with a recycling delay of 5 seconds. The average rotational correlation time (τ_c) of CitAP was determined using the NMR experiment given in reference [64]. A detailed description of the calculations of (τ_c) is given in the methods part (Chapter 2, section 2.2.8).

4.3.3 Results and Discussion

Titration results

As indicated earlier, the available structure of CitAP(1P0Z) was obtained by crystallizing a histidine tagged protein. The structure was found to contain citrate in complex with molybdate.

In order to determine the influence of molybdate on the binding of citrate to His-tagged CitAP, titration experiments were carried out with citrate and molybdate using histidine tagged ^{15}N labeled CitAP. ^{15}N - ^1H HSQC experiment was measured for each titration step. Titrations were carried out by step-wise addition of sodium citrate to the protein. Sharpening

of peaks and chemical shifts changes were observed in the ^{15}N - ^1H HSQC with an addition of one equivalent of citrate to the protein, indicating the binding of citrate to CitAP (Figure 4.19 peaks in blue). A second sample was used to add sodium molybdate to the protein. Broadening of signals were observed indicating a higher aggregation state of the protein. Most of the signals disappeared upon addition of 20 fold molar excess of molybdate (Figure 4.19 peaks in green). On addition of citrate to the molybdate bound CitAP solution, line broadening decreased and peaks started to re-appear. No major changes were obtained after addition of 20 fold excess of citrate to the molybdate bound CitAP solution (Figure 4.19 peaks in red). This corresponds to the published crystallization condition of citrate bound CitAP (1P0Z). The difference in the HSQC spectra of citrate bound CitAP and citrate and molybdate bound-CitAP, demonstrates that there might be a difference in the protein structure with citrate alone and citrate and molybdate together bound to CitAP. From the titration experiments it is clear that the citrate binding to CitAP does not require the presence of molybdate. Since the importance of molybdate in signal transduction is less understood, further experiments for backbone assignments were obtained without the influence of molybdate.

Therefore the new study of CitAP will be based on the backbone assignment of the CitAP shorter fragment, without the histidine tag, in a buffer which does not contain any molybdate.

Extent of assignment

Usual three dimensional experiments were acquired for the backbone assignment of the citrate free and citrate bound CitAP. These are listed in Table 2.4. Only 85 peaks were observed out of 129 peaks expected in the ^{15}N - ^1H HSQC spectrum of citrate free CitAP. The use of TROSY based NMR experiments did not give any additional peaks. Out of the 80 peaks for which $C\alpha$ connectivities were present, 72 peaks were assigned. The Backbone chemical shifts of the assigned residues are given in Table B.1. The ^{15}N - ^1H HSQC spectrum of the citrate bound form of CitAP gave 109 peaks out of the possible 129 peaks expected from the amino acid sequence. Since the binding of citrate to CitAP was in slow exchange, peaks could not be followed during the titration. Hence the three dimensional experiments for

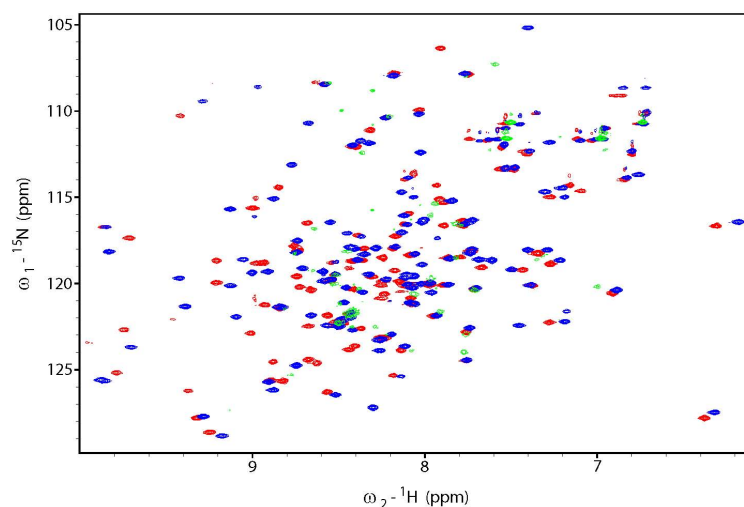


Figure 4.19: Overlay of ^{15}N - ^1H HSQC spectra of molybdate titration using the His-tagged CitAP shorter construct. Peaks of the spectrum with 20 fold excess of molybdate is in green. Extreme line broadening is observed for most of the peaks. Peaks in red were obtained from a spectrum with 20 fold excess of citrate and molybdate. Peaks in blue were obtained with 5 fold excess of citrate in the protein sample. Large changes in chemical shifts are seen between the spectrum of molybdate and citrate, and citrate alone.

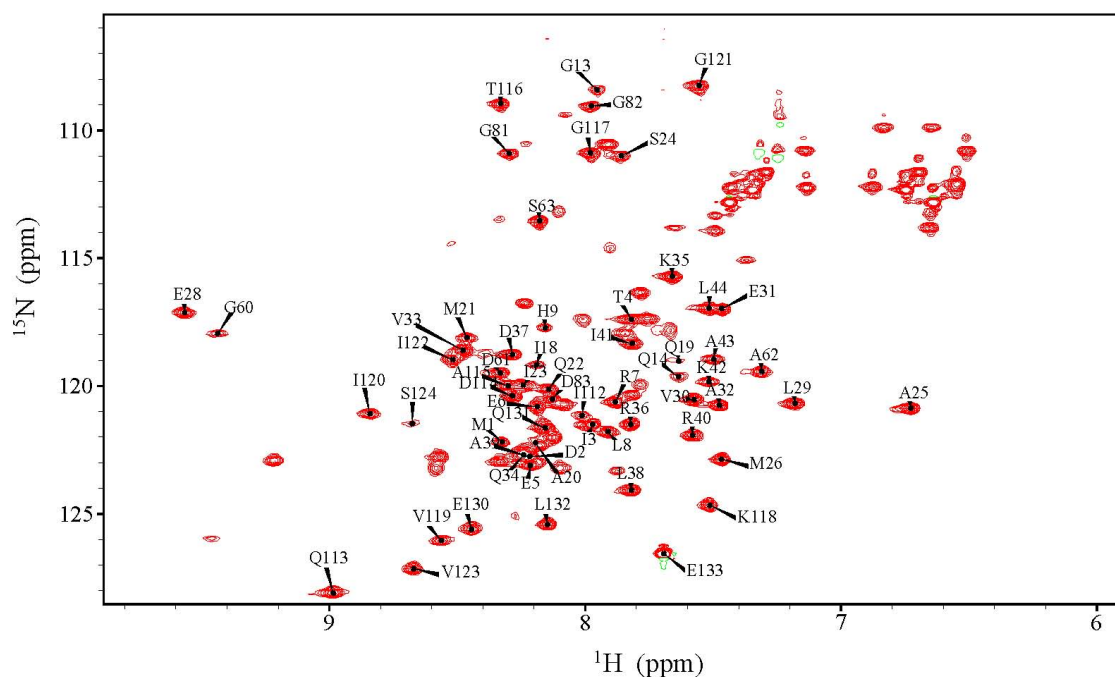


Figure 4.20: ^{15}N - ^1H HSQC spectrum of citrate free CitAP with the assignment of amino acid given on top of the peaks. The spectrum was measured at 900 MHz at 310 K.

backbone assignments had to be re-measured for the citrate bound form of CitAP (Table 2.5). Assignment was obtained for 100 out of 109 peaks seen in the ^{15}N - ^1H HSQC spectrum. Chemical shifts of the assigned peaks are listed in Table B.2. ^{15}N - ^1H HSQC of the citrate free and citrate bound forms of CitAP including their assignments are given in the Figure 4.20 and Figure 4.21 respectively.

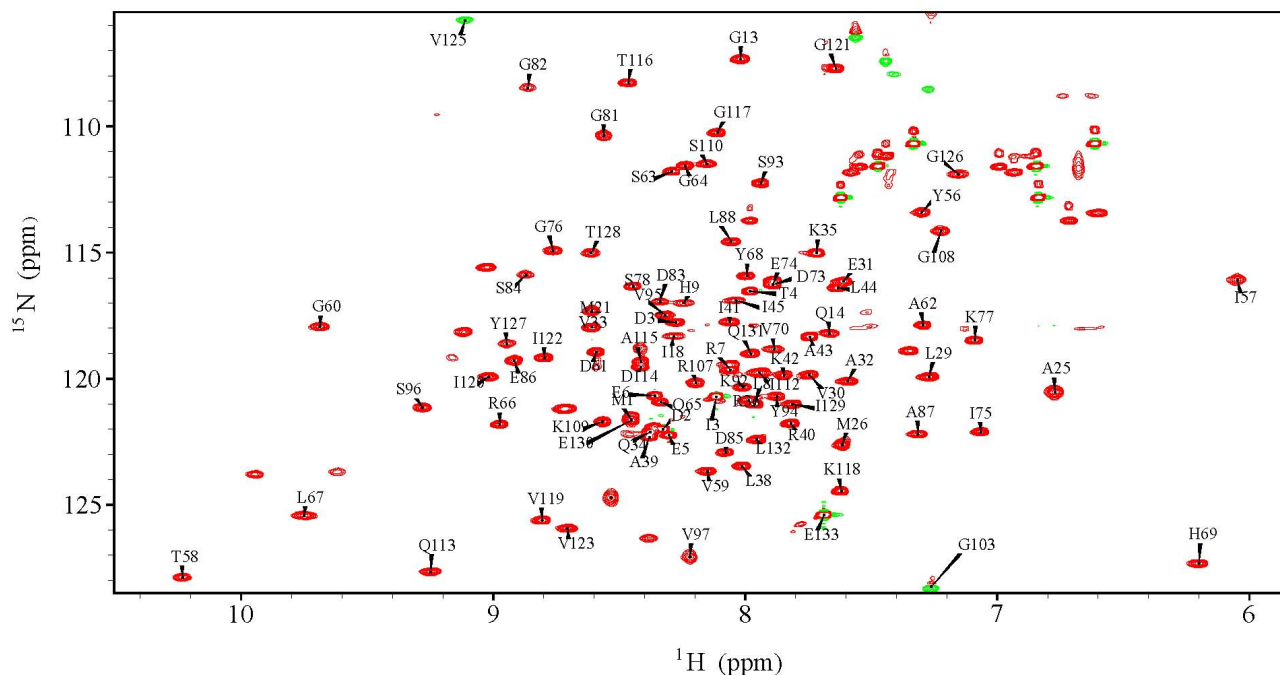


Figure 4.21: ^{15}N - ^1H HSQC spectrum of citrate bound CitAP with the assignment of amino acid given on top of the peaks. Ten fold excess of citrate was present in the solution. The spectrum was measured at 900 MHz at 310 K.

4.3.4 X-ray structure of citrate bound form of CitAP without molybdate

Due to lack of complete backbone assignments of the citrate free and bound form of CitAP, the NMR solution structure could not be determined. X-ray crystal trials were simultaneously performed using the protein made with the shorter construct and without histidine tag as well as without molybdate in the crystal screening buffer. Crystallization trials were carried out by Dr. Stefan Becker in our department. He obtained crystals of CitAP in the citrate bound form and solved the structure in collaboration with the group of Prof. George M. Sheldrick at the university of Göttingen.

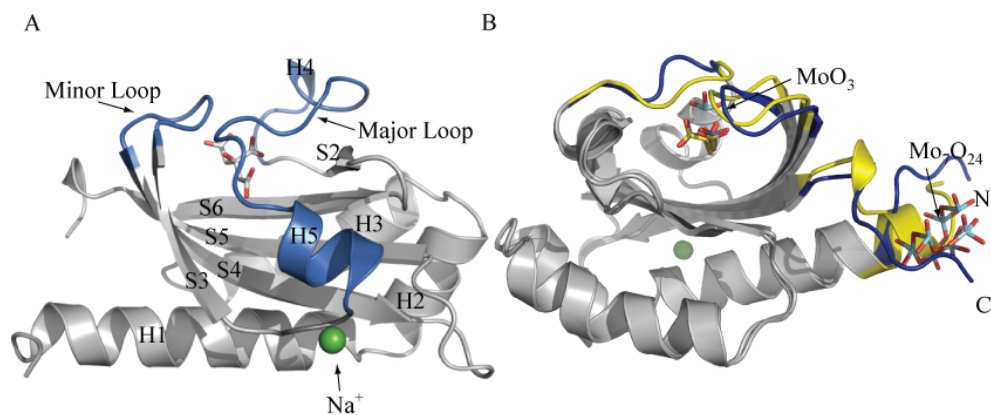


Figure 4.22: Citrate-bound structures of CitAP. **(A)** The secondary structural elements, major loop and minor loop residues highlighted in blue, bound Na⁺ in green and the bound citrate in stick representation. **(B)** Superposition of 1P0Z (differences shown in dark blue) on the new citrate-bound structure (differences shown in yellow). Citrate is shown in stick representation (yellow: new structure, blue: 1P0Z). MoO₃ near the active site and bound Mo₇O₂₄ near the N- and C-terminus of 1P0Z are shown in stick representation.

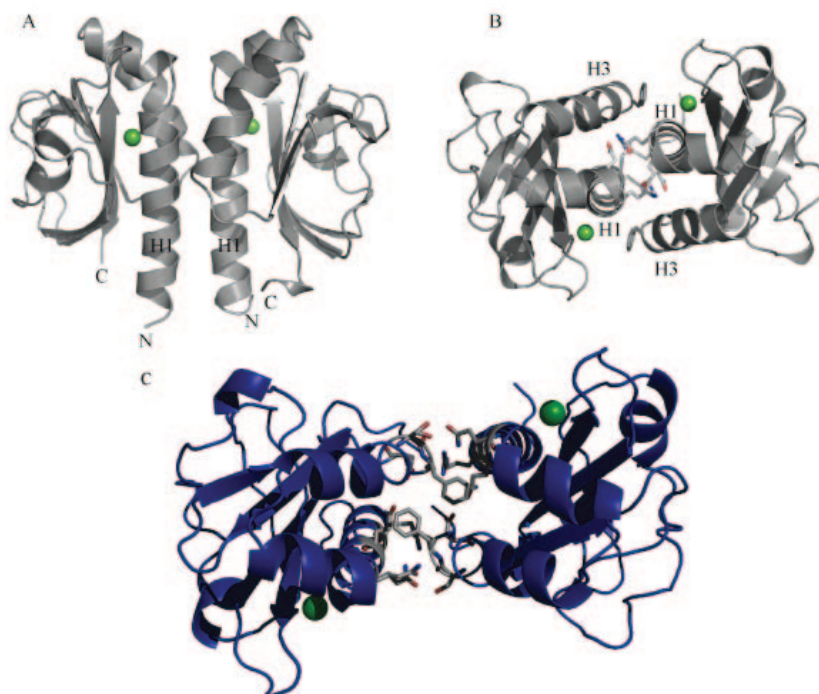


Figure 4.23: **(A)** Dimer in the asymmetric unit of our citrate-bound structure showing bound Na⁺ (green spheres). **(B)** Top view of the dimer in the asymmetric unit of our citrate-bound structure showing the interface residues.

The crystallization condition was 0.1 M HEPES, pH 7.5, 1.6 M $(\text{NH}_4)_2\text{SO}_4$ and 5 mM sodium citrate.

Structure discussion

The amino acid numbers used in the discussion of the X-ray and NMR part is the same as in the published CitAP structure. These residue numbers have an offset of 46 from the amino acid number in the CitA protein sequence.

There are two monomers in the asymmetric unit (citrate-bound-A: residues 4-132, citrate-bound-B: residues 6-129), which form a dimer. Structurally (Figure 4.22 panel A) each monomer has a PAS domain fold, highly similar to the monomers found in the asymmetric unit of 1P0Z (Figure 4.22 panel B) with an r.m.s.d. of only 1.2 Å. The bound citrate and the Na^+ in both structures superimpose nearly perfectly and their binding modes are very similar (Figure 4.22 panel B). The binding pocket is made up of the minor and major loop at the top and the β sheet in the bottom. In this structure the backbone of the major loop (residues 63-92) is twisted between residues 78 and 84 by about 20 degrees in direction of the β -sheet, occupying the space where MoO_3 is located in 1P0Z. Additionally the minor loop (residues 96-106) is relocated by about 1 Å towards the major loop in our structure. Overall these conformational differences in minor and major loop result in a much tighter closure of the citrate binding pocket in this structure compared to 1P0Z. The extended conformations of the N- and C-termini in 1P0Z are likely to reflect both domain truncation and tight packing against the Mo_7O_{24} cluster (Figure 4.17). In the new structure the N-terminal region has a helical conformation from residue 7, and in the C-terminal region there is a 3_{10} helix from residues 129-131 (monomer A). Several lattice contacts are made by the N-terminal residues. Therefore the specific structure of this region must also be considered with caution. NMR chemical shifts, on the other hand, support a significant helical propensity of residues 129-131 in solution (Figure 4.36). In either case, the conformation of both termini may also be further altered by attachment to the trans membrane domains TM1 and TM2.

Dimer interface

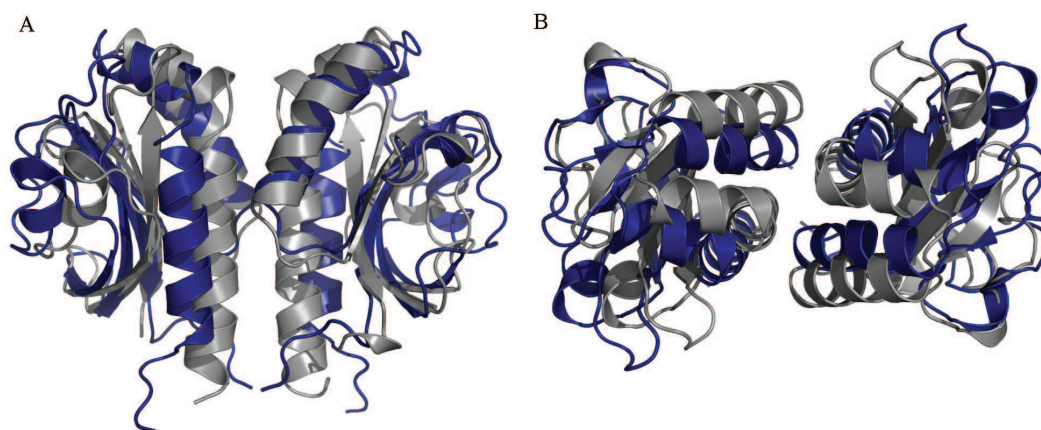


Figure 4.24: Superposition of the GJ type dimer of 1P0Z (dark blue) and the dimer of the new citrate-bound CitAP structure in the asymmetric unit. (A) side view (B) top view.

The overall shape of the dimer found in the asymmetric unit of the new citrate-bound CitAP structure (Figure 4.23 panel A) resembles the GJ-type dimer of 1P0Z (Figure 4.17). In 1P0Z the dominant feature of this dimer interface is the pairwise parallel association of the N-terminal helix H1 from one monomer with the H3 helix of the other (Figure 4.23 panel C). In the new structure, helices H1 form a central parallel bundle involving residues 11-27, with helices H3 (residues 48-51) packed against them on either side (residues 21-25) (Figure 4.23 panel A and B). In the new structure dimerization mode results in a tightly packed dimer with an interface covering 2123 Å² of accessible surface area. The GJ-type dimer in 1P0Z is by far less tightly packed, with an interface of only 1415 Å². The superposition of their backbones (Figure 4.24) results in an r.m.s.d. of 3.6 Å, compared to only 1.2 Å for the monomers, emphasizing the distinct dimer packing interfaces. Both structures differ also strongly in the hydrophobicity of their interface. In the GJ-type dimer of 1P0Z the Phe51 residues form a central hydrophobic patch (Figure 4.23 panel C) that is extended via hydrophobic interactions into the hydrophobic cores of both monomers. In the new structure the interface between the monomers is in the region between helices H1 by far less hydrophobic. Thus, despite high structural similarity between the monomers, the dimerization modes in the new citrate-bound CitAP structure and in the GJ-type dimer from 1P0Z are very different in

size and nature. Given the more extensive interface observed, the new citrate-bound dimer is more likely to represent the physiological association.

4.3.5 X-ray structure of citrate free form of CitA

Structure of citrate free CitAP was yet to be determined. Hence simultaneously with the crystallization of the citrate bound form of CitAP, Xray crystal trials were done on the free form as well. The citrate free CitAP was able to be crystallized in a buffer containing 20 mM HEPES, pH 7.5, 0.63 M NaH_2PO_4 and 0.63 M KH_2PO_4 . The structure of citrate-free CitAP was determined to a resolution of 2.0 Å . There are two monomers in the asymmetric unit (citrate-free-A and citrate-free-B) (Figure 4.25). In citrate-free-A residues 1 to 131 were traced, except for a part of the major loop (76-85). In citrate-free-B, residues 3 to 128, including the complete main chain of the major loop residues (63-92), were traced (yellow in Figure 4.25). For a number of these residues the refinement resulted in decreased occupancies. This suggests mobility of the major loop, in agreement with the fact that part of the major loop residues could only be traced in one of the monomers. Pairwise r.m.s.d of backbone residues of the two monomer structures show large deviations in the major loop region (Figure 4.35, top panel). Only 12 common hydrogen bonds were found between the two monomer structures. This is a hint toward a very flexible structure for the citrate free form of CitAP.

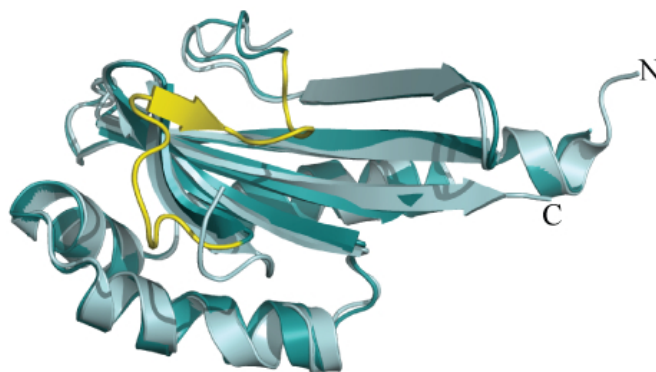


Figure 4.25: Structure of Citrate-free CitAP with superposition of both monomers in the asymmetric unit. The major loop that could be traced in one of the monomers is colored in yellow

4.3.6 NMR studies on the citrate free form of CitAP

Only 62% of the possible peaks were observed in the ^{15}N - ^1H HSQC spectrum of the citrate free form of CitAP. The average rotational correlation time (τ_c) measured for the protein was 7.54 ns, indicating a monomeric CitAP in solution. No improvement and no additional peaks were observed with TROSY-HSQC spectrum. These results suggest that line broadening due to conformational exchange might be the cause of the disappearance of around 40% of the peaks.

Usually, three dimensional experiments for protein backbone assignments are acquired with amide proton detection in the direct dimension. These amide protons are the most sensitive to line broadening due to chemical exchange. In order to limit line broadening due to conformational exchange, several non-standard NMR experiments were acquired. To reduce amide proton chemical exchange broadening, CPMG-HSQC experiments were measured. CPMG sequence was applied during the INEPT (refer to Chapter 1 for description of INEPT) transfer period in the normal HSQC. Relaxation due to chemical exchange is reduced in the CPMG sequence. No additional peaks were obtained in the CPMG-HSQC spectra.

Rather than using amide detection in the direct dimension, H_α detected experiments were also tried. Back and forth HACACO experiment with H_α detection in the direct dimension did not give any additional peaks (refer to Chapter 2, section 2.2 for more details). Only 80 peaks were observed in the spectrum (Figure 4.26 panel **A**). This implies that line broadening due to chemical exchange is not restricted to amide group only.

To determine the presence of conformational exchange in the side chain resonances in the protein, ^{13}C - ^1H HSQC experiment was measured. Methyl groups have the most favorable T_2 relaxation time due to rapid rotation of its carbon-carbon single bonds. The ^{13}C - ^1H HSQC spectrum of the methyl region of the citrate free CitAP shows only 62 peaks (Figure 4.26 panel **B**). 85 possible peaks were expected from the protein sequence. The missing number of methyl resonances fit well with the number of methyl peaks that were expected from the unassigned amino acids in the protein sequence. This proves that line broadening due to

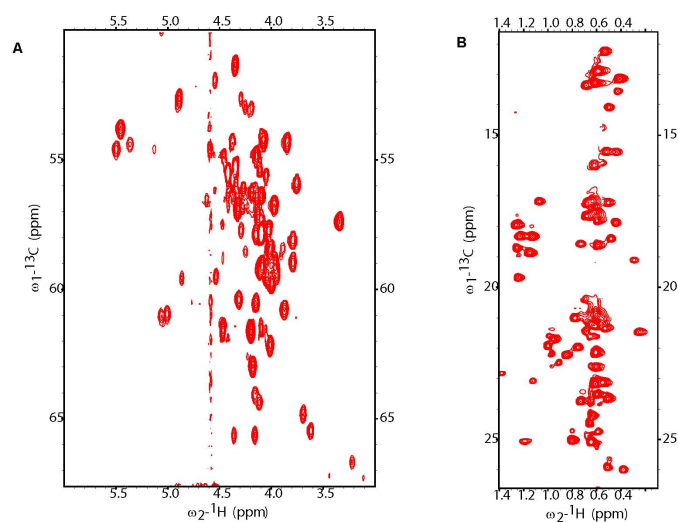


Figure 4.26: **(A)** $^{13}\text{C}_\alpha\text{-}^1\text{H}_\alpha$ projection of three dimensional HACACO experiment. Only 80 peaks are observed in this spectrum. **(B)** $^{13}\text{C}\text{-}^1\text{H}$ HSQC spectrum of the methyl region. Only 62 peaks out of 85 peaks are seen in the spectrum.

conformational interchange affect the side chain resonances as well. These results indicate that there must be different conformations of CitAP in solution. This is in accordance with the identification of two different structures for the monomers of CitAP in the asymmetric unit of the citrate free CitAP crystals.

Nevertheless, the backbone assignment of 72 of the 80 peaks observed in the HSQC spectrum were possible. The unassigned residues of CitAP is plotted onto the monomer B of the structure of citrate free form of CitAP (Figure 4.39). It is interesting to observe that most of the unassigned residues in the sequence are those residues which show large backbone r.m.s.d deviations between the two forms of the monomers of the citrate free structure of CitAP. These residues are mostly in the two β sheets and the parts of the major loop which also defines the binding pocket for citrate in the citrate bound structure. The three residues that are assigned in the major loop region have considerably high flexibility associated with

it (refer to discussion below), this could have reduced the relaxation process which leads to the disappearance of the other peaks in the major loop.

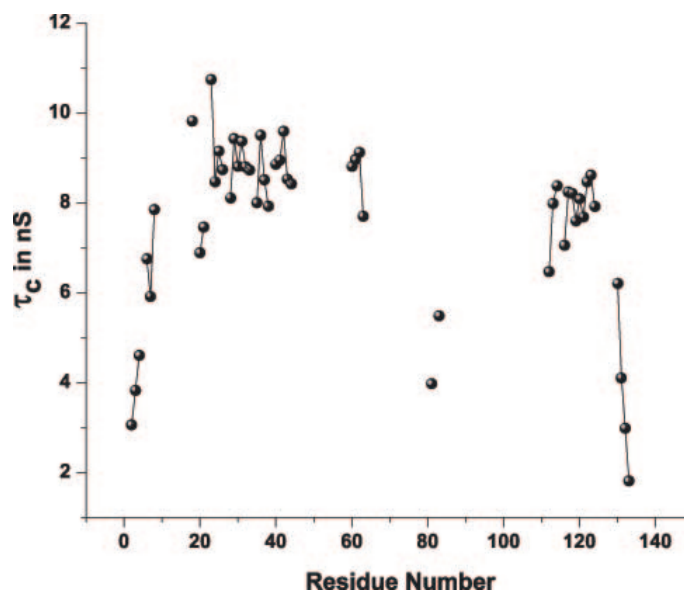


Figure 4.27: Plot of measured rotational correlation time against individual residue number of citrate free-CitAP. The plot gives a measure of flexibility of residues. Residues at the N and C terminal and the two residues that was assigned in the major loop region show low τ_c values indicating major flexibility in these region.

Usual T_1 , T_2 methods for estimation of τ_c have R_{ex} (relaxation rate due to conformational exchange) term associated with it. For proteins for which R_{ex} is the dominant relaxation rate, these methods give a very large value for τ_c . Hence cross correlated relaxation rate was used to determine the τ_c values. The τ_c values obtained using this method does not have contribution from chemical exchange. As indicated earlier, the average τ_c of CitAP determined using this method (described in 2.2.8) is 7.54 ns. Spectral density function of a rigid body is used in the calculation for τ_c s. Hence a lower bound for the values of τ_c s are obtained using this method. τ_c for individual amino acids are determined. As described in chapter 2, section 2.2.8, the τ_c values are a function of the amplitude of local motion associated with that amino acid residue (S^2). Residues with lower τ_c values than the average value have significant flexibility term associated with it. Most of the residues showing low τ_c values are in the N and C terminals. The τ_c for the amino acids assigned in the major loop

region also have lower than average value, indicating high flexibility and deviation from rigid body approximation (Figure 4.27). Heteronuclear NOE experiment was also measured for the citrate free CitAP. The amplitude of motions in the pico- to nano-second time-scale is investigated with the steady state ^{15}N - ^1H NOE. ^{15}N - ^1H Het-NOE values therefore give a measure of residue flexibility. The three residues for which assignment was possible in the major loop (residues 81,83 and 84), have lower Het-NOE value (lower than 0.6) indicating high flexibility in this region (Figure 4.28). This correlates well with the lower τ_c values of these residues (residues 81 and 83).

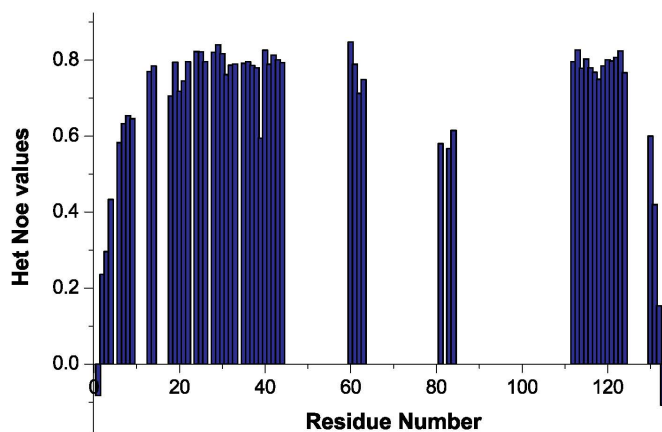


Figure 4.28: ^{15}N - ^1H NOE values measured for citrate free-CitAP are plotted against residue number. Residues near the N and C terminal and the residue in the main loop region have lower Het-NOE values indicating high flexibility.

4.3.7 NMR studies on the citrate bound form of CitAP

Citrate bound form of CitAP was obtained by the addition of one equivalent of citrate to the protein sample. Binding of citrate to CitAP is in the slow exchange limit. 107 peaks were observed in the ^{15}N - ^1H HSQC spectrum of the citrate bound form of CitAP. Backbone assignment was obtained for hundred of these peaks. The unassigned residues are mainly located in the N-terminal helix (H1: residues 9–12, 15–19, and 22–24), linker between the H3 helix and the first β strand (residues 48–54) and in the citrate binding area (mainly in the minor loop, residues 97–102) (refer Figure 4.30). These residues might be broadened beyond detection because of chemical exchange. As discussed for the crystal structure of citrate

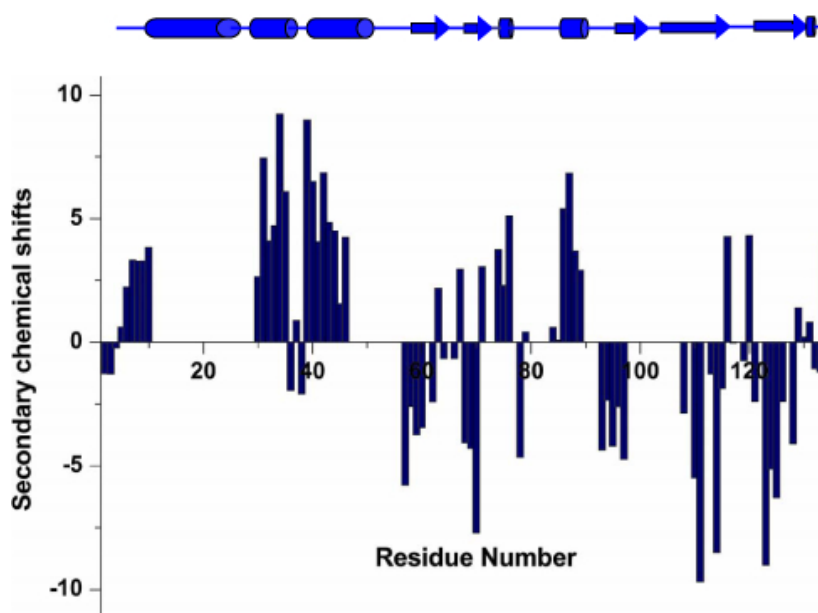


Figure 4.29: Secondary structure of citrate bound CitAP. Secondary chemical shifts of $\Delta C_{\alpha} - \Delta C_{\beta} + \Delta C'$. Secondary structural element from the new citrate bound crystal structure of CitAP is given at the top of the figure.

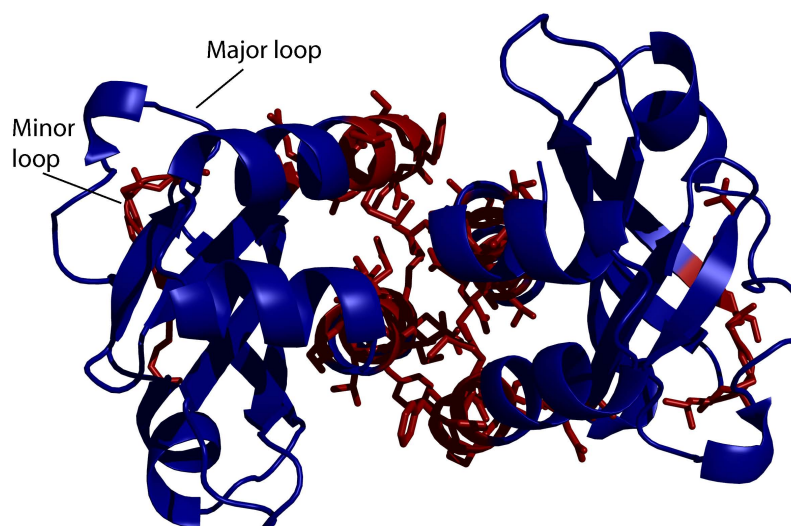


Figure 4.30: The unassigned residues in citrate bound form of CitAP are plotted on to the structure of citrate bound dimer. The unassigned residues (red sticks) are mainly located in the dimer interface and also in the minor loop of CitAP.

bound CitAP, the N terminal helix (namely Helix1) is involved in the dimer interface. This might be the reason for the chemical exchange broadening of the peaks due to a possible dimer-monomer exchange (residues in the H1 and H3 helix are involved in dimer interface), though previous studies indicate the presence of only 10% dimer in the solution of citrate bound form of CitAP.

The secondary structure element was determined from the deviation of the C_{α} , C_{β} and C' chemical shifts from the random coil values and are shown in Figure 4.29. Positive values indicate helical propensity and negative value denote β strand propensity. The plot of the secondary chemical shifts of citrate bound CitAP correlate well with the secondary structural elements from X-ray structure of the citrate bound form of CitAP (Figure 4.29, top panel). ^{15}N - ^1H HetNOE values were also measured. Heteronuclear amide NOE values larger than 0.7 for most of the residues of bound form of CitAP imply rigidity of the structure. Amide NOE values are slightly decreased in the N and C terminal regions of the protein indicating flexibility in these regions.

4.3.8 Sodium binding to CitAP

Early microbiological evidence had indicated that anaerobic growth of *K.pneumoniae* on citrate was Na^+ dependent [147, 148, 149]. It was also previously shown that induction of the target genes by CitA/CitB requires not only citrate, but also Na^+ [138]. This feature is physiologically meaningful as both citrate transport and metabolism are also strictly dependent on Na^+ [150, 151]. The X-ray structures seem to suggest a functional connection between citrate and Na^+ binding to CitAP, as Na^+ was identified only in the citrate-bound CitAP structure. Although valence bond calculations strongly support this choice [152], the identity of the Na^+ modeled in this structure is not absolutely certain. Similarly, the fact that we could not identify an ordered Na^+ in the citrate-free form does not strictly exclude the possibility that Na^+ may be present but disordered. However NMR studies corroborate the Na^+ binding to the citrate bound form of CitAP.

Significant chemical shift changes and sharpening of peaks are the characteristic changes observed in the ^{15}N - ^1H HSQC when citrate is bound to CitAP. These changes are observed with a one fold addition of sodium citrate to CitAP. On further addition of sodium citrate (10 fold to 20 fold) in excess, small chemical shift changes were observed (Figure 4.31 left). Specific subset of residues are affected in the titration. Those residues which show additional changes are in a well defined region in the CitAP structure, which corresponds to the Na^+ binding site in the CitAP structure (Figure 4.31 right). Hence as evident from the Na^+ titration, citrate bound form of CitAP binds sodium in the fast exchange time scale.

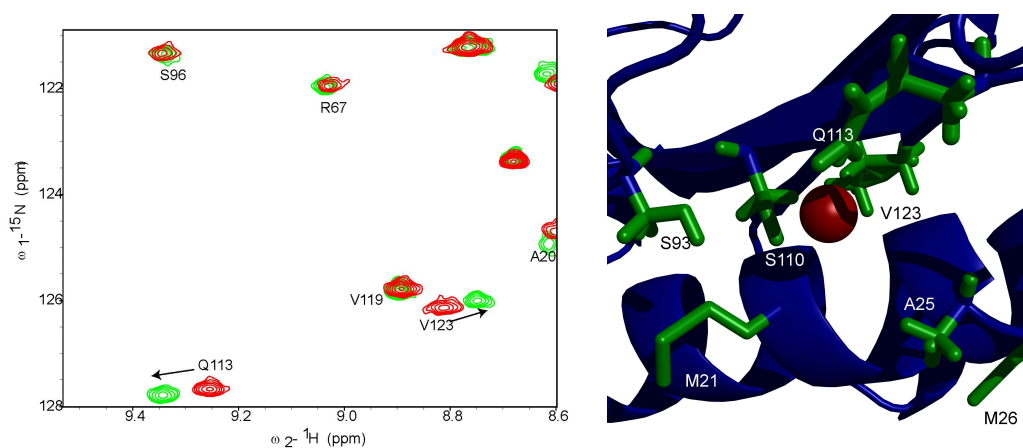


Figure 4.31: (left) Effect of Na^+ on citrate bound CitAP. Overlay of the portion of HSQC showing few peaks affected by Na^+ titration. Peaks in red are from spectrum measured for one equivalent of sodium citrate in the protein sample and peaks in green are measured at 20 fold excess of sodium citrate. (right) Citrate bound-CitAP structure: with residues most affected by Na^+ titration in stick representation. Na^+ ion is given in red sphere. The residue shown are M21, A25, M26, S93, S110, Q113 and V123.

4.3.9 Residual dipolar coupling analysis

Residual dipolar couplings (RDCs) analysis is an efficient method to determine whether the X-ray structure of a protein is retained in solution (refer to chapter 1, section 1.3). To check for the presence of the X-ray structure of CitAP in solution, residual dipolar couplings were measured. The citrate free and bound form of CitAP were aligned in bacterial filamentous phages. Quadrupolar deuterium splitting was 9 and 7 Hz respectively. $^1\text{D}_{\text{NH}}$ RDCs were measured for citrate bound CitAP while a set of $^1\text{D}_{\text{NH}}$, $^1\text{D}_{\text{NC}}$ and $^1\text{D}_{\text{CC}\alpha}$ RDCs were measured for citrate free CitAP.

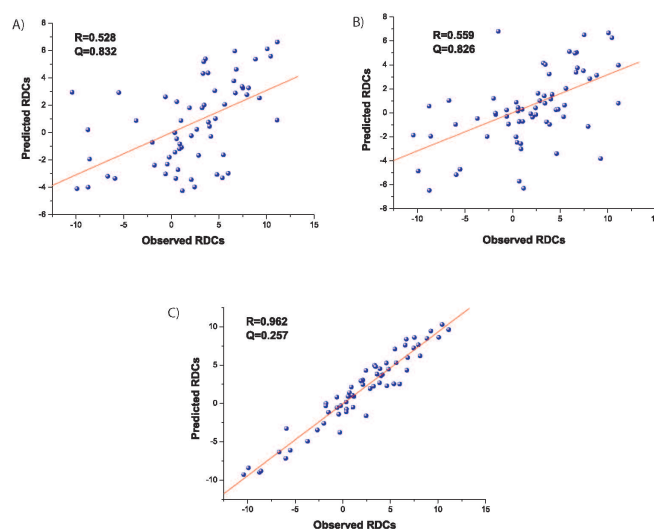


Figure 4.32: Correlation of measured H-N RDCs of citrate bound CitAP with back calculated RDCs from different crystal structures of CitAP. (A) Correlation between measured HN RDCs with back calculated RDCs from monomer A (A), monomer B (B) of citrate free-CitAP and the new citrate bound structure of CitAP (C). The correlation coefficients were 0.528, 0.559 and 0.962 respectively.

RDCs of citrate bound CitAP

PALES software was used to fit the measured residual dipolar couplings to the structures using the SVD (Single Value Decomposition) method. 75 well resolved NH couplings were obtained for the citrate bound form of CitAP. Measured RDCs gave a correlation of 0.962 with the new citrate bound structure of CitAP (refer Figure 4.32 panel C). As expected the same RDCs gave a very bad correlation with the two free structures of CitAP (refer Figure 4.32 panel A and B). Good correlation of the measured citrate bound RDCs with the bound structure suggest that in solution structure of the bound form of CitAP is very similar to crystal structure of the bound form. Bad correlation of the bound RDCs with the citrate free structures indicate that there are hardly any contribution from the free structures to the bound RDCs and that the citrate bound structure would be very different from the citrate free structure.

RDCs of citrate free CitAP

66 well resolved N-H and N-C' couplings were obtained for the citrate free form of CitAP. RDCs of the free form of CitAP fit with a correlation of 0.825 with the monomer A of the crystal structure of citrate free CitAP (Figure 4.33 panel A). The measured RDCs also fit to 0.763 with monomer B of the citrate free form of CitAP (Figure 4.33 panel B). Surprisingly the same RDCs gave a correlation of 0.93 with the new citrate bound structure of CitAP (monomer A). (Figure 4.33 panel C). It is possible that the measured RDCs could have contribution from both the monomer structures in the citrate free form. In order to analyze this, simultaneous fit of the measured RDCs with the two monomer structures of CitAP were carried out. A non linear fit implemented in MATHEMATICA using a single alignment tensor for both the monomer structure of the citrate free CitAP yielded a correlation of 0.836 (Figure 4.33 panel D). The fitting procedure is explained in more detail in Chapter 2, section 2.2.5. These results can be explained on the basis of multiple conformation of protein in solution. The measured RDCs will have contributions from all of the conformers of the protein in solution. Only slight increase in the correlation was observed with the simultaneous fit of the RDCs with the two citrate free monomer structures of CitAP. This suggests that there may be more possible conformers in solution. It could also be that the X-ray structures of citrate free CitAP are not well defined and the increase in the correlation with the simultaneous fit could just be the averaging of structural noise in these two structures. Hence these results have to be taken with caution. Nevertheless taking these results and severe line broadening observed for the peaks in the ^{15}N - ^1H HSQC it is reasonable to assume the presence of multiple conformers of CitAP in solution. The correlation of measured RDCs with those derived from monomer A of citrate free CitAP and the new bound structure of CitAP was 0.936 (Figure 4.33 panel E). The fitted weighting factors were 0.18 for the monomer B and 0.82 for the bound citrate-CitAP structure. Same correlation was obtained by using the two structures of the citrate free CitAP and the new citrate bound CitAP (Figure 4.33 panel G). The fit converged only when the weighting factor for monomer B of citrate free CitAP approached zero. This could be explained by a badly defined X-ray structure or an improbable conformer in solution. So the outcome of the RDC analysis is that in the citrate free solution, the major

population of the protein is in the citrate bound conformer (about 80%) and a part in the citrate free monomer A form (About 20%). Better correlation of the measured RDCs with the citrate bound structure suggest that the bound structure is one of the conformer present in the citrate free CitAP solution giving a significant contribution to the measured RDCs of citrate free CitAP.

4.3.10 Comparison of citrate free CitAP and citrate bound CitAP using X-ray and NMR

The major difference in the structures of citrate free and citrate bound CitAP are in the N and C terminals and in the major and minor loops. In both citrate-free-A and citrate-free-B the short helices H4 (71-74) and H5 (86-90), which are found in the major loop in the new citrate-bound CitAP structure, are unwound (Figure 4.34 A). Instead, a short strand is formed by residues 78-81 in citrate-free-B. A superposition of citrate-free-B with citrate-bound-A (Figure 4.34 B), shows that a part of the major loop fills the space in the citrate-free form that is occupied by citrate in the citrate-bound form. Although the overall structure of citrate-free CitAP is similar to the new citrate-bound structure, the central five-stranded anti-parallel β -sheet is less bent in the citrate-free state. As a result, the tip of the minor loop has moved by 13.51 Å away from its position in citrate-bound CitAP (Figure 4.34 A). This wide open conformation of the minor loop (formed partly by strands S3 and S4) causes the strand S5 to extend up to the C-terminus. A typical β -sheet-type backbone hydrogen-bonding pattern exists between residues of the C-terminal region and residues of the neighboring strand S4 (green in Figure 4.34b). The helix H1 is slightly extended in the citrate free form of CitAP compared to bound form of CitAP. In the C-terminal part, a 3_{10} helix is found in the bound form of CitAP, while in the citrate free-CitAP, the strand S5 extend up to C terminus.

NMR studies

The back-bone structure of CitAP becomes more rigid with the binding of citrate to CitAP. This is indicated by the positive values obtained from the difference in Het-NOE values of citrate bound and citrate free form for most of the residues (Figure 4.36 middle panel). Major

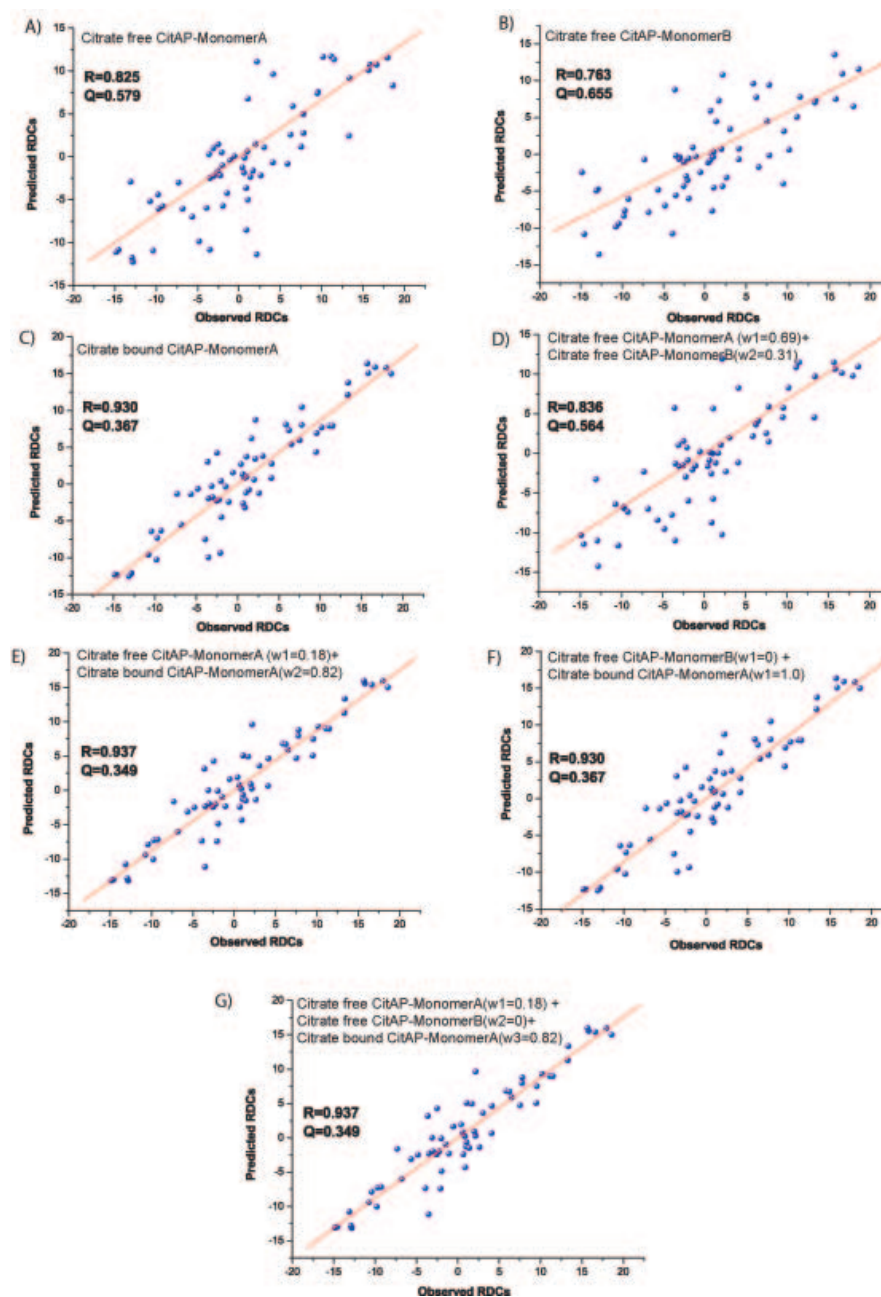


Figure 4.33: Correlation between measured RDCs of citrate free-CitAP with back calculated RDCs from different Xray structures of CitAP and their combinations. Correlation between the measured N-H and N-C' RDCs with back calculated RDCs from single alignment tensors derived from X ray structures: monomer A of citrate free-CitAP (A), monomer B of citrate free-CitAP (B) and citrate bound CitAP structure (C). Correlation of the measured RDCs with back calculated RDCs using a single alignment tensor and a population weighting factor derived by campaigning two or more X ray structures: monomer A and B of citrate free-CitAP (D): the population factors were 0.69 and 0.31 respectively, monomer A of citrate free CitAP and citrate bound CitAP (E): population factors were 0.18 for monomer A and 0.82 for citrate bound structure, monomer B of citrate free CitAP and citrate bound CitAP (F): population factor was zero for monomer B and 1.00 for citrate bound structure, monomer A and B of citrate free CitAP and the citrate bound CitAP (G): the population factor was zero for monomer B of citrate-free CitAP and 0.18 for monomer A of citrate-free CitAP, the population factor was 0.82 for the citrate bound form.

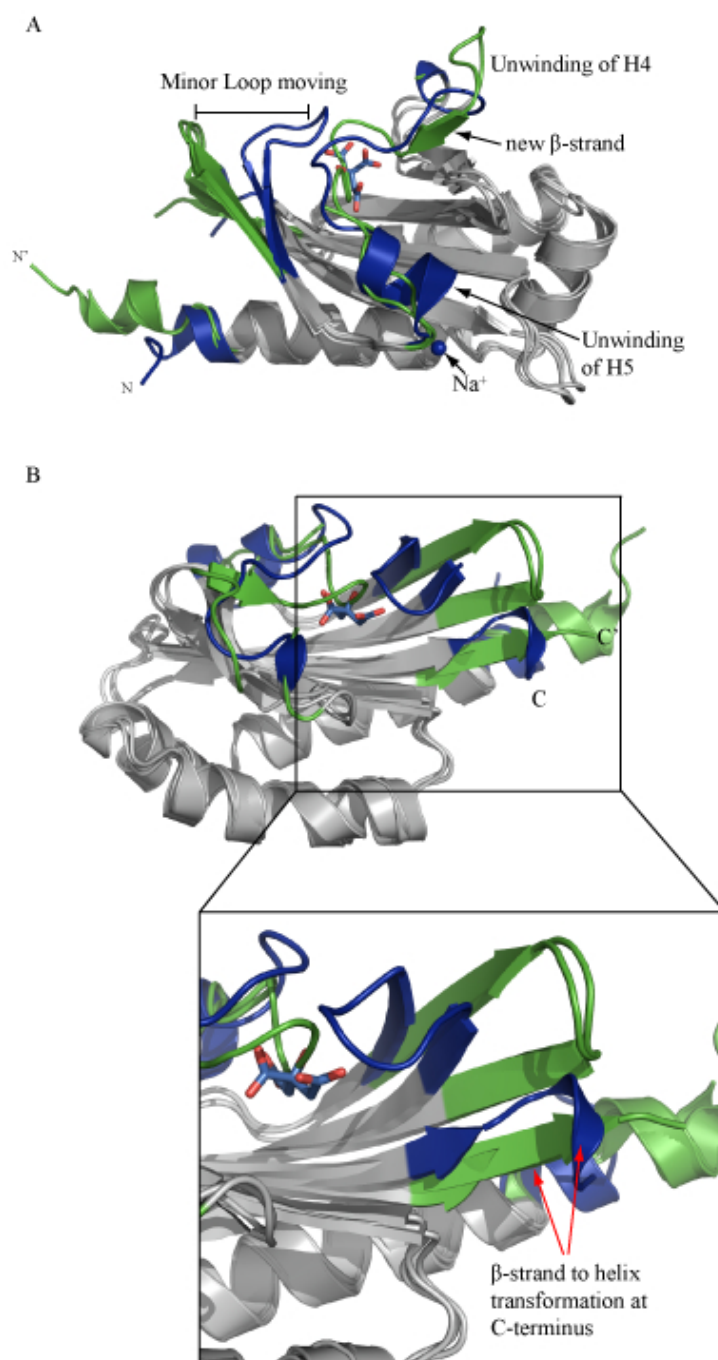


Figure 4.34: Conformational differences between citrate-bound and citrate-free CitAP. (A) Side view of superposition of citrate-bound structure (differences highlighted in dark blue) on citrate-free structure (both monomers) (differences highlighted in green) showing the minor loop movement of $\sim 45^\circ$. (B) View from opposite side of the superposition of citrate-bound CitAP (differences highlighted in dark blue) on the citrate-free structure (both monomers) (differences highlighted in green) showing the displacement of the major loop upon citrate binding, the β -strand in the major loop of the citrate-free structure and the conformational difference in the C-terminal region (inset). The displacement of the C-terminus is indicated by the location of residue Gln131.

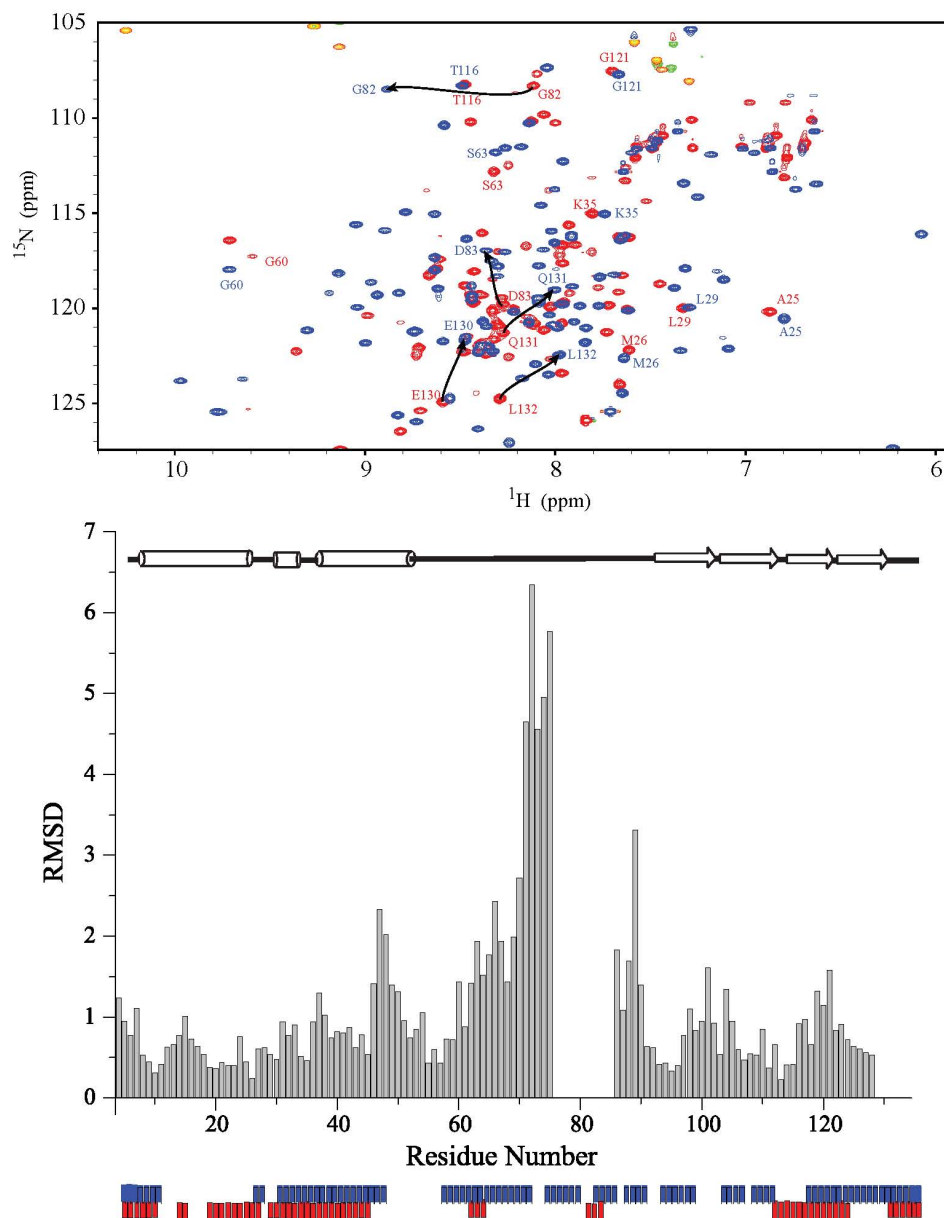


Figure 4.35: Top panel: Overlay of HSQC spectra showing the free (red) and citrate bound (blue) resonances for CitA. Arrows follows those residues which show a large chemical shift change upon citrate binding. Bottom panel: Backbone r.m.s.d plot (grey) for the two free monomer structures of CitA. Red and blue bars indicate respectively those residues of CitA for which NMR resonance assignments were obtained for the citrate-free and citrate-bound form. Secondary structural elements from the citrate free structure is given on top of the r.m.s.d plot.

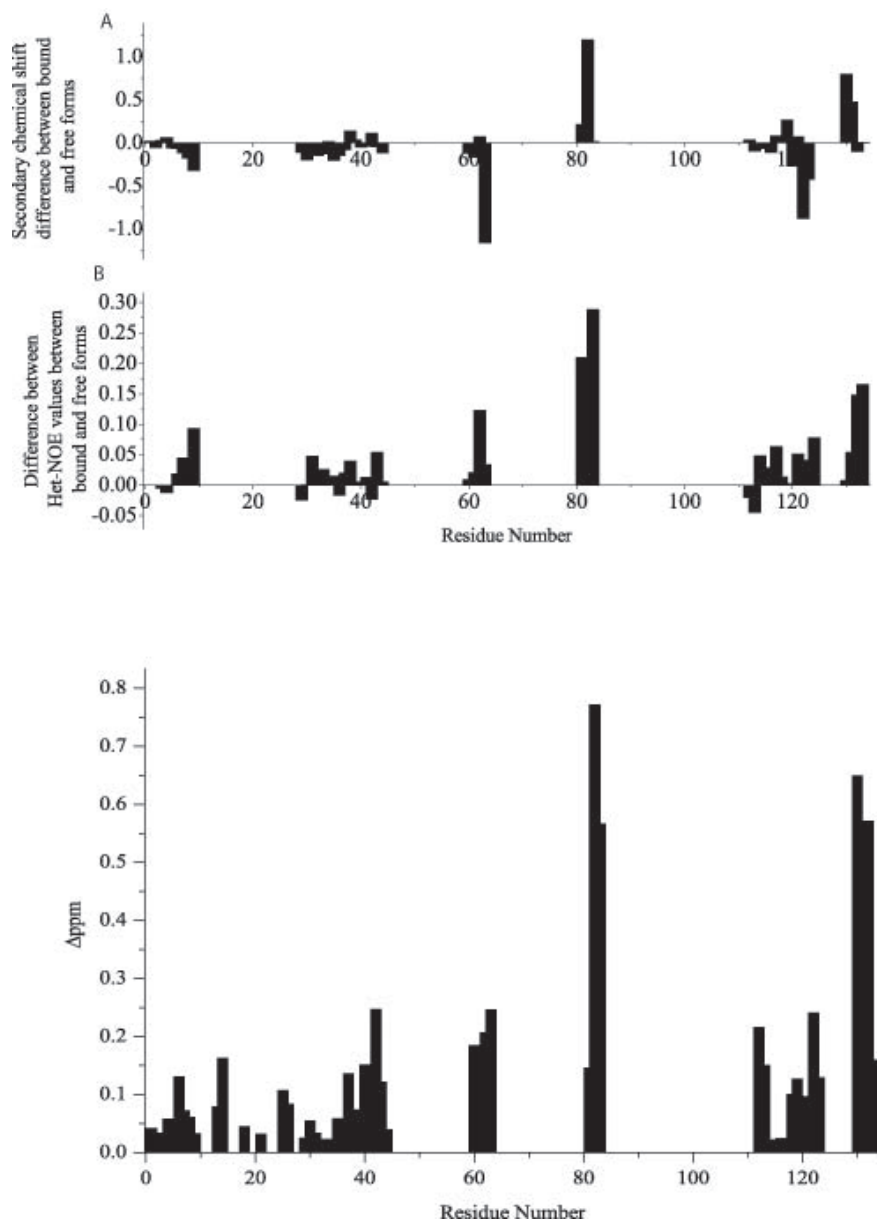


Figure 4.36: Structural and dynamic changes in CitAP upon binding to citrate as monitored by multidimensional NMR spectroscopy. Top panel: Difference between secondary C_{α}/C chemical shifts (secondary chemical shifts = difference between experimental chemical shifts and random coil values) in the citrate-bound and the citrate-free state. Note the significant increase in helical propensity for residues 129-130. Middle panel: Difference between ^1H - ^{15}N NOE values between the citrate-bound and citrate-free state. Overall the ^1H - ^{15}N NOE values were slightly increased indicating a global increase in the rigidity of the backbone upon binding to citrate. The strongest increase was observed for residues neighboring the major loop and for residues 129-130. Bottom panel: Difference in $^1\text{H}/^{15}\text{N}$ chemical shifts between citrate-bound and citrate-free state.

changes were observed in the C terminal region and the part of the major loop suggesting an increased stability of the bound form in these regions of the protein. Large chemical shift changes were also observed in the major loop regions (residues 81-84) and the C terminal (residues 130-133) as seen in Figure 4.36 bottom panel (also see the HSQC spectrum in Figure 4.35). Difference in secondary chemical shift of the assigned residues of the citrate bound and free form also indicate changes in the major loop region and the C terminus region of the protein (Figure 4.36 top panel). Significant increase in helical propensity was observed for residues 129-131 corresponding to the 3_{10} helix in the citrate bound crystal structure. Large $^1\text{H}/^{15}\text{N}$ chemical shift differences are also observed in the major loop region and C terminal region of the protein (Figure 4.36 bottom panel). These observations suggest structural and dynamic changes in the major loop region and the C terminal region when citrate is bound to CitAP.

4.3.11 Mechanism of Signal Transduction

The structural differences found in this study between the citrate-bound and citrate-free structures of CitAP give for the first time a more detailed view of potential signaling mechanisms for this histidine kinase. The open minor loop and the extended S4-S5 structure seen in the citrate-free structure is formed in a region that experiences only weak, water mediated lattice contacts, and the multiple main-chain interactions stabilizing the extended structure suggest that it is likely to reflect the pre-signaling conformation of the receptor. In the citrate-bound molecules, no lattice contacts are found around the C-terminus, and only a few distinct crystal contacts are formed around this region in the citrate-free molecules. Therefore the fundamental observation is that in the citrate-bound PAS domain, the end of the β - sheet is pulled in towards the citrate-binding site, forming tight interactions with the activating ligand. This disrupts the strand-to-strand interactions between S4 and S5 and simultaneously pulls the C-terminus upward, away from the membrane. Regardless of the exact geometry of the C-terminal attachment, these changes are likely to be communicated directly to the TM2 helix (Figure 4.34B and 4.37a), causing an immediate transduction of ligand binding into a conformational change at the transmembrane junction. At the N-terminus the first helical turn of helix H1 in the citrate-free CitAP structure unwinds to

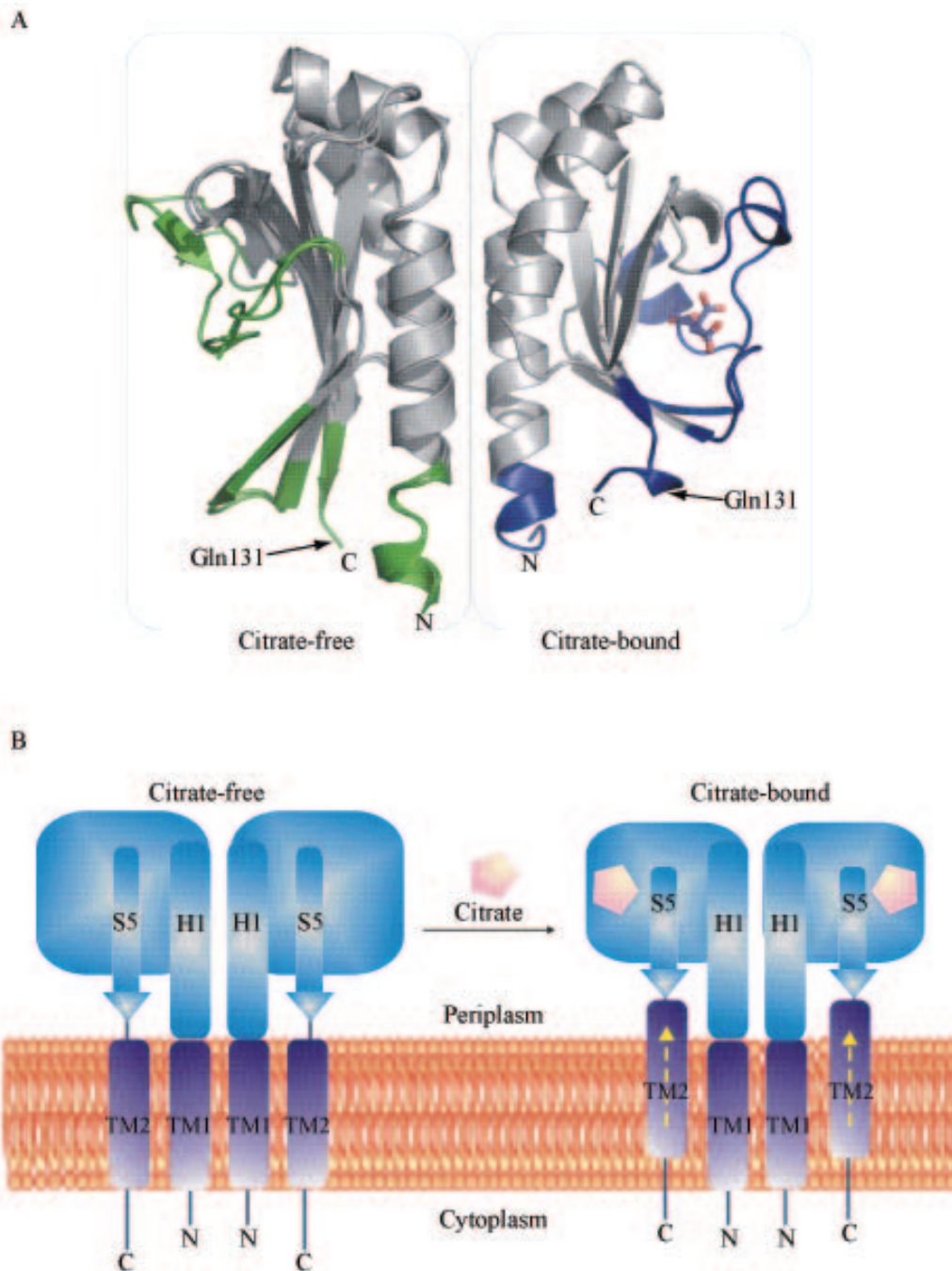


Figure 4.37: Signal transduction mechanism in CitA. (a) Comparison of the citrate-free monomer (superposed on the Citrate-bound-B) with the Citrate-bound-A monomer. (b) Model of the ligand-induced switch triggering a vectorial movement of TM2.

an extended loop in our citrate-bound CitAP structure. An immediate connection of this structural change to citrate binding is not obvious. It is rather probable that the unwinding is due to the missing connection of helix H1 to TM1, and to the formation of extensive lattice contacts involving residues 4-12 in the citrate-free structure. Attempts to express CitAP with a more extended N-terminus were not successful. We assume that the N-terminal helix as an extension of TM1 most probably serves as an anchor keeping the periplasmic sensor domain in a stable position regarding its distance to the membrane surface. The contraction of CitAP upon ligand binding might then result in a vectorial movement of TM2 towards the sensory domain, constituting the trigger of the kinase activity (Figure 4.37a and b). This piston-type movement of TM2 relative to TM1 has been suggested as a possible motion type of helices in membranes [153, 154]. Experimental evidence for the piston model was found for the trans membrane signaling of the aspartate receptor [155].

4.3.12 Comparison of structures of DcuS-PD and CitAP

DcuS of *E.Coli* and CitA of *Klebsilla Pneumonia* are homologous proteins. DcuS senses a wide range of C4 dicarboxylates while CitA is highly specific for citrate. Their periplasmic domains have 25% sequence identity and 70% sequence homology. The structures of DcuS-PD and CitAP are similar and reveal a PAS fold. The backbone r.m.s.d for DcuS-PD structure with citrate bound and citrate free monomer A and B are 3.5 Å, 4.09 Å and 4.38 Å respectively (see Figure 4.38). Except for the small beta strand (S2) in citrate bound CitAP, secondary structural elements of citrate bound CitAP and DcuS-PD are conserved. In DcuS-PD, the N terminal helix is slightly bent inward with the beta sheets. Binding site for ligands in DcuS-PD and CitAP are conserved. As seen from the Figure 4.38, the binding site is more open in DcuS-PD than citrate bound structure of CitAP, but is less so from the citrate free CitAP. Consequently the β sheets are less bent in the DcuS-PD structure compared to citrate bound CitAP. The residues of DcuS-PD affected by fumarate binding are located in the 4 main β strands and in the major and minor loops. In CitAP the residues having direct hydrogen bonds with citrate are also located in the four β strands and the major and minor loops. The residue showing major peak intensity increase on binding of fumarate is Q144 which is in the minor loop region of DcuS-PD. The largest chemical shift

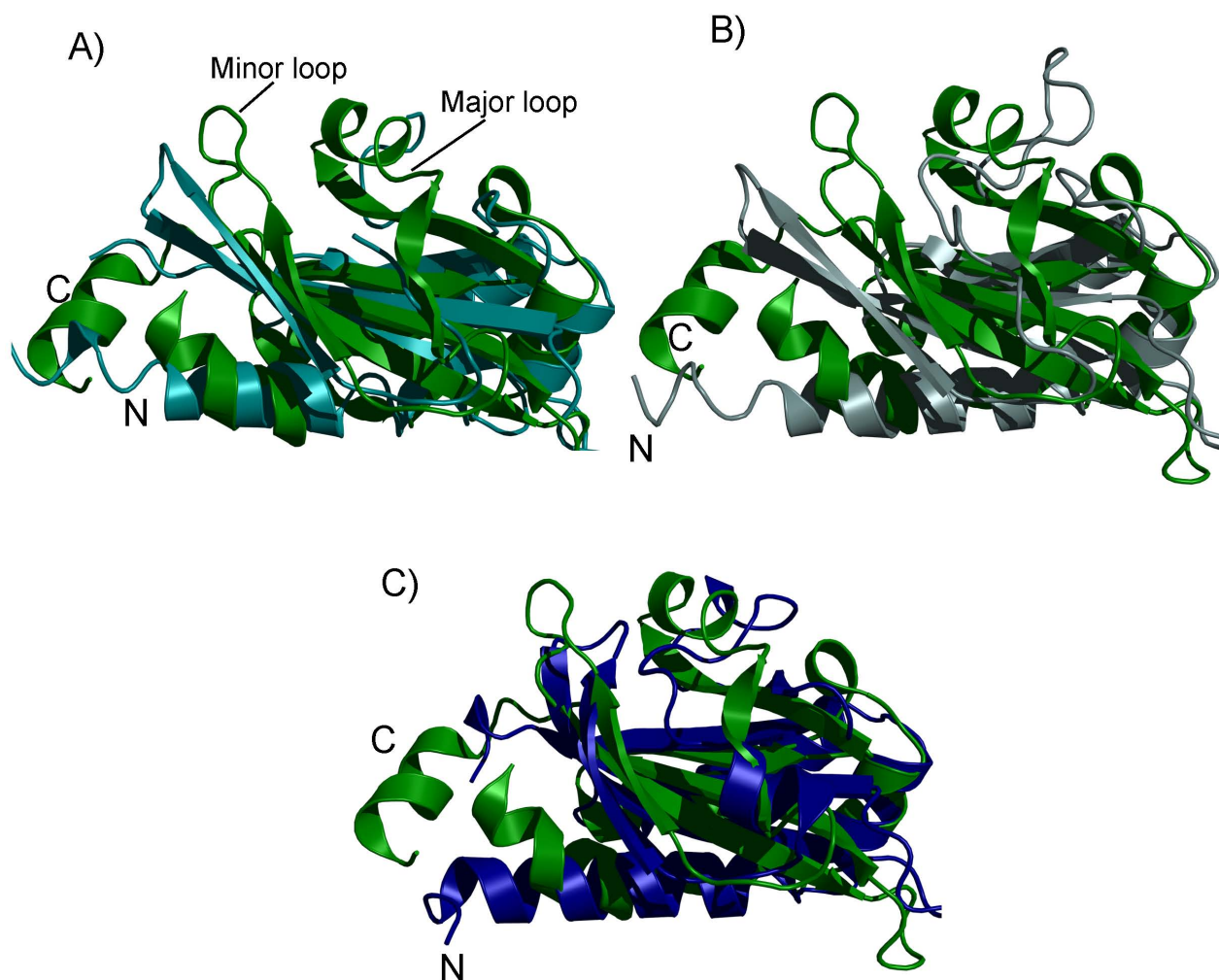


Figure 4.38: Comparison of the DcuS-PD structure with citrate-free monomer and citrate-bound monomer structures. A) Superposition of the DcuS-PD structure (green) with the citrate free monomer A structure of CitAP (cyan), r.m.s.d : 4.09 Å . B) Superposition of the DcuS-PD structure (green) with the citrate free monomer B structure of CitAP (pale cyan), r.m.s.d : 4.38 Å . C) Superposition of the DcuS-PD structure (green) with the citrate bound structure of CitAP (blue), r.m.s.d : 3.5 Å. Notable differences are in the minor loop and the N and C terminal regions. The N terminal helix of DcuS-PD is curved in with the beta sheets and helix at the C terminal region is well formed in the case of DcuS-PD. The structure of DcuS-PD resembles the citrate bound structure of CitAP, with the binding pocket slightly more open than the citrate bound structure of CitAP.

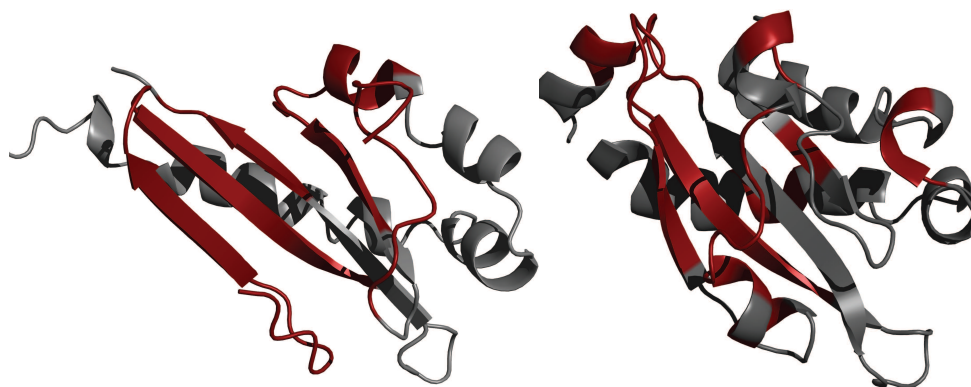


Figure 4.39: Citrate free CitAP is characterized by major line broadening due to conformational exchange. The unassigned residues are plotted on the structure of citrate free monomer A of CitAP(left). DcuS-PD also show line broadening. Residues of DcuS-PD showing very low peak intensity in the HSQC are plotted on the NMR structure of DcuS-PD (right).

change upon tartrate addition was observed for residue G140 which again is in the minor loop region in the structure of DcuS-PD. This is in accordance with the large conformational differences in the minor loop region between the citrate-free and citrate-bound CitAP X-ray structure.

Characteristic of citrate free CitAP in solution is the severe line broadening (missing peaks) in the ^{15}N - ^1H HSQC spectrum due to conformational exchange. These missing peaks are mostly from residues in the first three β strands and the part of the major loop region and helix H3 (residues from 64-111). Dcus-PD also showed chemical exchange broadening. ^{15}N - ^1H HSQC spectrum of DcuS-PD showed very small peak intensities for a number of residues. Het-NOE values of these residue were in the 0.7 to 0.8 range, very similar to rest of the residues in the sequence, suggesting a rigid structure in these parts of the protein. Residues corresponding to the peaks showing small intensities are mapped onto the structure of DcuS-PD. This region in the structure of DcuS-PD is similar to the unassigned regions in citrate free CitAP (Figure 4.39) .

4.3.13 Conclusion

In this section, the NMR solution studies on the periplasmic domain of CitA (CitAP) with and without the ligand are presented. In the course of this study the X-ray crystallographic structures of CitAP with and without the ligand were determined. The description of these structures is also presented.

The X-ray structures reveal a PAS fold similar to DcuS-PD and other PAS domains. The citrate bound structure was determined without the non physiological molybdate which was present in the previously determined X-ray structure.

There were two conformers of CitAP (monomer A and monomer B) in the asymmetric unit of the citrate free CitAP crystal structure. A part of the major loop could not be traced in monomer A and the C terminal residues were missing in monomer B. This suggests that the citrate-free CitAP is flexible in the C terminal and also in some parts the major loop region. But fortunately specific conformational changes were observed between the structures of the citrate-free and -bound form. Based on this, a possible mechanism for signal transduction is discussed.

By NMR titration it was shown that in solution molybdate might induce a different structural response than with only citrate bound to CitAP. Molybdate also induces aggregation of the protein. The newly determined bound structure did indeed differ in the dimer interface and slight change in the binding of citrate to CitAP from previously published X-ray structure. Previous studies suggest that CitAP might also sense sodium ion. A sodium ion was tentatively localized in the citrate bound structure of CitAP. With NMR titrations, this is found to be true in solution as well.

NMR backbone assignments for 90% of the peaks seen in the ^{15}N - ^1H HSQC were obtained for both the bound and free form of CitA. The NMR solution structures of CitAP could not be determined because of the large number of missing peaks in NMR spectra due to chemical/conformational exchange broadening. Nevertheless, with the available assignments, the chemical shift difference and the secondary structural propensity difference between the free and the bound forms of the CitAP protein were determined. These values indicate significant difference in chemical shifts and secondary structural propensities in the C-terminal regions

of the protein. The model proposed here for signal transduction is based on the structures of the citrate free and bound CitAP. This is also consistent with the present NMR data and the previous knowledge of signal transduction mechanisms.

Correlation co-efficients from residual dipolar coupling analysis of citrate free CitAP suggest that monomer A might be relatively more closer to the solution structure than monomer B. The correlation coefficient of RDCs of citrate free CitAP fits even better with the citrate bound structure of CitAP. RDCs of citrate-free CitAP therefore have a major contribution from the citrate-bound CitAP structure (RDC analysis indicate that the population of citrate bound structure would be 80% and that of the monomer A of citrate free form to be 20%). This suggest that the binding pocket of citrate is already formed in the citrate free solution of CitAP. CitAP might require the two transmembrane helices to keep the N- and C-terminal parts of the protein together, and without the transmembrane part the protein is not confined to one conformational space. Smaller Het-NOE values in the N- and C-terminal regions indicate greater flexibility in these regions of the protein, hence providing greater freedom for the protein to sample different conformational space.

Chapter 5

Summary and outlook

The ability of organisms to adapt to changing environmental conditions is essential for their long-term survival and requires a set of receptors that can recognize external stimuli and activate the appropriate biochemical responses. In bacteria, two-component regulatory systems play an important role in these processes. In particular, membrane-bound sensor histidine kinases allow the detection of ligands outside of the cytoplasmic membrane, so that the cellular response does not require transport or diffusion of these ligands into the cell. However, despite their key role in bacterial adaptation, many aspects of the molecular mechanisms of these proteins are unknown. In this thesis a detailed NMR studies, complemented by crystallographic results on two sensory periplasmic domain of histidine kinases, DcuS and CitA, are presented. DcuS of *E.Coli* and CitA of *Klebsiella pneumoniae* are homologous proteins involved in sensing different types of C4 di-carboxylates. While CitA is highly specific for citrate, DcuS senses a wide range of C4 dicarboxylates with much lower affinity.

NMR solution structure of the ligand free periplasmic domain of DcuS was determined. The structure of the periplasmic domain was refined using RDCs. In order to measure all types of RDCs in the committed time in which the protein was stable, a simple strategy for simultaneous measurement of different types of RDCs with minimum resonance overlap is also presented. ^{15}N - ^1H HSQC based titrations with fumarate and tartrate were performed to determine the binding pocket in DcuS-PD. Binding of fumarate resulted in sharpening of peaks and tartrate binding resulted in small chemical shift changes in NMR spectrum for residues in the binding pocket of the structure. However no such changes were observed for residues outside the binding region as it would have been expected for a signal transfer from the periplasmic domain to transmembrane part.

NMR studies on periplasmic domain of CitA was motivated by the fact that citrate binds to CitAP with a higher affinity than any of the C4 di carboxylates binding to DcuS-PD. The higher affinity and the higher specificity of CitAP for citrate could lead to a more clearer difference in the conformation of ligand free and bound form of the periplasmic domain. So far the X-ray structure of citrate bound CitAP was known. At first, a citrate free CitAP structure would be required to obtain a possible conformational changes which would allow the description of the signal transduction process. The first aim of this study focuses on obtaining the structure of citrate free CitAP. The X-ray structure of the citrate bound form that existed prior to this study was obtained with the exogenous histidine tag still attached to the protein. Moreover, two types of molybdate moieties were found bound to the protein in the citrate binding pocket and at the protein dimer interface. To minimize the risk of misleading interpretation, this work includes the study of the influence of the histidine tag and molybdate in CitAP. There after major part of the work is dedicated to the structural and dynamic behavior of citrate free and bound CitAP without histidine tag and molybdate.

Due to conformational broadening only about 62% of the peaks were observed in the ^{15}N - ^1H HSQC spectrum of citrate free CitAP and 83% peaks were present in the citrate bound CitAP HSQC spectrum. Because of the large number of missing peaks, the NMR solution structure of the citrate free and bound form of CitAP could not be determined. However the X-ray structure of the citrate free and bound form of CitAP could be determined. The structure has a PAS fold, very similar to DcuS-PD structure. The binding pocket is formed by the β sheet of the PAS fold and the inter-strand minor and major loops. The binding pocket is very similar to the C4 di-carboxylate binding pocket observed in DcuS-PD. In the citrate bound structure a Na^+ ion was tentatively localized between N terminal helix and the β sheets. This was also confirmed by NMR titrations. Hence CitAP may be involved in sensing both citrate and Na^+ ion in solution.

There are two conformers of citrate free monomer structures in the asymmetric unit of CitAP. Specific conformational change can be observed between the citrate free and citrate bound structures of CitAP. The major conformational differences are seen in the binding

pocket and also in the C-terminal region of the protein. The minor loop and the major loop (which defines the binding pocket of citrate) are in an open conformation in the citrate free structure while it is more closed in the citrate bound structure. The C terminal 3_{10} helix in the bound structure is absent in the citrate free structure. In solution, definite helical propensity is observed in the C-terminal region in the citrate bound-CitAP. Large chemical shift changes are observed in the C-terminal and major loop region of the protein. A model of signal transfer is described using the conformational differences seen in the crystal structures. A vectorial movement of the trans-membrane helix attached to the C-terminus of the protein could be used for the signal transduction cascade. This model is consistent with the model of signal transduction described for aspartate sensor.

A part of the major loop region and also the C-terminal part of the protein could not be traced in one of the two monomer structures. In solution, the citrate free form of CitAP is a monomer. C-terminal residues and the three residues that could be assigned in the major loop region showed lower Het-NOE values, indicating higher flexibility. Surprisingly the RDCs measured for citrate free CitAP fit better with citrate bound structure of CitAP. RDC analysis also indicate that in citrate free CitAP solution, larger contribution to RDCs come from the citrate bound conformer (around 80%), compared to citrate free monomer A (20%) conformer. .

The first step in the signal transduction mechanism would require the signal from the binding pocket to be transferred to the C-terminal part of the periplasmic domain. Hence there should exist a conserved motion of residues between the binding pocket and the C-terminal part of the protein which allows for the signal transfer to occur. This would mean that a different conformational state of protein in either end of the signal transfer process will have a complementary effect on the other end. Since the transmembrane part of the protein are absent in the present study, the N- and C-terminal part of the periplasmic domain are not fixed onto the membrane. This makes the N and C-terminal residues highly flexible (as seen from lower Het-NOE values). The flexibility of the N- and C-terminal region therefore provides more conformational freedom to the protein. On the other hand if the protein is

fixed on the membrane, there would be limited mobility in the N and C terminal region and that would make the protein sample a narrow conformational space. In the cases studied here, the conformational freedom due to absence of transmembrane part of the protein could explain the severe line broadening observed in the NMR spectra of citrate free form of CitAP. From RDC analysis it seems that the major conformer in the citrate free CitAP solution would be the bound form of CitAP. Stability in the binding pocket also leads to stability in the N and C terminal region of the protein. Addition of citrate populates the citrate bound conformer which was already present in solution and makes the overall structure more stable (indicated by sharpening of peaks and Het-NOE values of 0.8 for most of the residues). The sharpening of peaks as seen in the fumarate titration of DcuS-PD also arises as a result of selectively populating one of the conformers in the DcuS-PD fumarate free solution. In such systems the bound conformers seems to be more populated as it is less dynamic and more stable.

A better picture for signal transduction mechanism could be obtained if in the free form of the protein, the N- and C-terminal regions are made less mobile. Cysteine mutation with a disulphide linkage between the N and C-terminal parts of the protein may stabilize the protein in the ligand free form. It would be interesting to see the difference in the structure of such a system with ligand free and bound structure described here. Like in aspartate sensor, the disulphide linkage would most likely resemble the ligand bound structure than the ligand free structure. The solid state NMR studies of DcuS reconstituted in liposomes with their transmembrane part intact would be very useful to corroborate the signal transduction mechanism described here.

References

- [1] F. Bloch. Nuclear induction. *Physical Review*, **70**(7-8):460–474, 1946.
- [2] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Physical Review*, **69**(1-2):37–38, 1946.
- [3] N. F. Ramsey and E. M. Purcell. Interactions between nuclear spins in molecules. *Physical Review*, **85**(1):143–144, 1952.
- [4] M. P. Williamson, T. F. Havel, and K. Wuthrich. Solution conformation of proteinase Inhibitor-IIa from bull seminal plasma by H¹ nuclear magnetic-resonance and distance geometry. *Journal of Molecular Biology*, **182**(2):295–315, 1985.
- [5] R. R. Ernst and W. A. Anderson. Sensitivity enhancement in magnetic resonance .2. Investigation of intermediate passage conditions. *Review of Scientific Instruments*, **36**(12):1696, 1965.
- [6] W. P. Aue, E. Bartholdi, and R. R. Ernst. 2-Dimensional spectroscopy - Application to nuclear magnetic-resonance. *Journal of Chemical Physics*, **64**(5):2229–2246, 1976.
- [7] R. H. Griffey and A. G. Redfield. Proton-detected heteronuclear edited and correlated nuclear-magnetic-resonance and nuclear overhauser effect in solution. *Quarterly Reviews of Biophysics*, **19**(1-2):51–82, 1987.
- [8] A. P. Hansen, A. M. Petros, A. P. Mazar, T. M. Pederson, A. Rueter, and S. W. Fesik. A practical method for uniform isotopic labeling of recombinant proteins in mammalian-cells. *Biochemistry*, **31**(51):12713–12718, 1992.
- [9] T. Kigawa, Y. Muto, and S. Yokoyama. Cell-free synthesis and amino acid-selective stable-isotope labeling of proteins for NMR analysis. *Journal of Biomolecular NMR*, **6**(2):129–134, 1995.

- [10] J. W. Lustbader, S. Birken, S. Pollak, A. Pound, B. T. Chait, U. A. Mirza, S. Ramnarain, R. E. Canfield, and J. M. Brown. Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: An advance toward rapid determination of glycoprotein structures. *Journal of Biomolecular NMR*, **7**(4):295–304, 1996.
- [11] L. P. McIntosh, E. Brun, and L. E. Kay. Stereospecific assignment of the NH₂ resonances from the primary amides of asparagine and glutamine side chains in isotopically labeled proteins. *Journal of Biomolecular NMR*, **9**(3):306–312, 1997.
- [12] L. P. McIntosh, F. W. Dahlquist, and A. G. Redfield. Proton NMR and NOE structural and dynamic studies of larger proteins and nucleic-acids aided by isotope labels - T4 Lysozyme. *Journal of Biomolecular Structure and Dynamics*, **5**(1):21–34, 1987.
- [13] D. C. Muchmore, L. P. McIntosh, C. B. Russell, D. E. Anderson, and F. W. Dahlquist. Expression and ¹⁵N labeling of proteins for proton and ¹⁵N nuclear-magnetic-resonance. *Methods in Enzymology*, **177**:44–73, 1989.
- [14] B. J. Stockman, M. D. Reily, W. M. Westler, E. L. Ulrich, and J. L. Markley. Concerted two-dimensional NMR Approaches to ¹H, ¹³C, and ¹⁵N resonance assignments in proteins. *Biochemistry*, **28**(1):230–236, 1989.
- [15] R. A. Venters, W. J. Metzler, L. D. Spicer, L. Mueller, and B. T. Farmer. Use of ¹H^N-¹H^N NOEs to determine protein global folds in perdeuterated proteins. *Journal of the American Chemical Society*, **117**(37):9592–9593, 1995.
- [16] W. M. Westler, B. J. Stockman, and J. L. Markley. Correlation of ¹³C and ¹⁵N chemical-shifts in selectively and uniformly labeled proteins by heteronuclear two-dimensional NMR-spectroscopy. *Journal of the American Chemical Society*, **110**(18):6256–6258, 1988.
- [17] C. Griesinger, O. W. Sorensen, and R. R. Ernst. A practical approach to 3-dimensional NMR-spectroscopy. *Journal of Magnetic Resonance*, **73**(3):574–579, 1987.

-
- [18] C. Griesinger, O. W. Sorensen, and R. R. Ernst. 3-Dimensional fourier spectroscopy - application to high-resolution NMR. *Journal of Magnetic Resonance*, **84**(1):14–63, 1989.
- [19] L. E. Kay, G. M. Clore, A. Bax, and A. M. Gronenborn. 4-Dimensional heteronuclear triple-resonance NMR-spectroscopy of Interleukin-1- β in solution. *Science*, **249**(4967):411–414, 1990.
- [20] M. Sattler, J. Schleucher, and C. Griesinger. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **34**(2):93–158, 1999.
- [21] C. Griesinger, M. Hennig, J. P. Marino, B. Reif, C. Richter, and H. Schwalbe. *Modern techniques in protein NMR series*, **16**. Kluwer Academics, New York 1998, 1999.
- [22] J. H. Prestegard. New techniques in structural NMR - Anisotropic interactions. *Nature Structural Biology*, **5**:517–522, 1998.
- [23] G. A. Morris and R. Freeman. Enhancement of nuclear magnetic-resonance signals by polarization transfer. *Journal of the American Chemical Society*, **101**(3):760–762, 1979.
- [24] O. W. Sorensen, G. W. Eich, M. H. Levitt, G. Bodenhausen, and R. R. Ernst. Product operator-formalism for the description of NMR pulse experiments. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **16**:163–192, 1983.
- [25] J. Schleucher, M. Sattler, and C. Griesinger. Coherence selection by gradients without signal attenuation - Application to the 3-Dimensional HNC0 experiment. *Angewandte Chemie-International Edition in English*, **32**(10):1489–1491, 1993.
- [26] L. E. Kay, P. Keifer, and T. Saarinen. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the american chemical society*, **114**(10):10663–10665, 1992.

- [27] M. Sattler, M. G. Schwendinger, J. Schleucher, and C. Griesinger. Novel Strategies for sensitivity enhancement in heteronuclear multidimensional NMR experiments employing pulsed-field gradients. *Journal of Biomolecular NMR*, **6**(1):11–22, 1995.
- [28] M. Sattler, P. Schmidt, J. Schleucher, O. Schedletzky, S. J. Glaser, and C. Griesinger. Novel pulse sequences with sensitivity enhancement for in-phase coherence transfer employing pulsed-field gradients. *Journal of Magnetic Resonance Series B*, **108**(3):235–242, 1995.
- [29] K. Wuthrich. *NMR of protein and nucleic acids*. Wiley Interscience, New York, 1986.
- [30] Sharon J. Archer, Mitsuhiro Ikura, Dennis A. Torchia, and Ad Bax. An alternative 3D NMR technique for correlating backbone ^{15}N with side chain $\text{H}\beta$ resonances in larger proteins. *Journal of Magnetic Resonance (1969)*, **95**(3):636–641, 1991.
- [31] J. Cavanagh, A. G. Palmer, P. E. Wright, and M. Rance. Sensitivity improvement in proton-detected 2-Dimensional heteronuclear relay spectroscopy. *Journal of Magnetic Resonance*, **91**(2):429–436, 1991.
- [32] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard. Nuclear magnetic dipole interactions in field-oriented proteins - Information for structure determination in solution. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(20):9279–9283, 1995.
- [33] N. Tjandra and A. Bax. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, **278**(5340):1111–1114, 1997.
- [34] N. Tjandra. Establishing a degree of order: Obtaining high-resolution NMR structures from molecular alignment. *Structure with Folding and Design*, **7**(9):R205–R211, 1999.
- [35] C. R. Sanders and J. H. Prestegard. Magnetically orientable phospholipid-bilayers containing small amounts of a bile-salt analog, CHAPSO. *Biophysical Journal*, **58**(2):447–460, 1990.

-
- [36] C. R. Sanders and J. P. Schwonek. Characterization of magnetically orientable bilayers in mixtures of Dihexanoylphosphatidylcholine and Dimyristoylphosphatidylcholine by solid-state NMR. *Biochemistry*, **31**(37):8898–8905, 1992.
- [37] M. R. Hansen, L. Mueller, and A. Pardi. Tunable alignment of macromolecules by filamentous Phage yields dipolar coupling interactions. *Nature Structural Biology*, **5**(12):1065–1074, 1998.
- [38] M. Ruckert and G. Otting. Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *Journal of the American Chemical Society*, **122**(32):7793–7797, 2000.
- [39] R. Tycko, F. J. Blanco, and Y. Ishii. Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. *Journal of the American Chemical Society*, **122**(38):9340–9341, 2000.
- [40] M. P. Nieh, C. J. Glinka, S. Krueger, R. S. Prosser, and J. Katsaras. Sans study of the structural phases of magnetically alignable Lanthanide-Doped phospholipid mixtures. *Langmuir*, **17**(9):2629–2638, 2001.
- [41] A. Bax, G. Kontaxis, and N. Tjandra. Dipolar couplings in macromolecular structure determination. *Method Enzymology*, **339**:127–174, 2001.
- [42] A. Saupe. Recent results in field of liquid crystals. *Angewandte Chemie-International Edition*, **7**(2):97, 1968.
- [43] J. A. Losonczi, M. Andrec, M. W. F. Fischer, and J. H. Prestegard. Order matrix analysis of residual dipolar couplings using singular value decomposition. *Journal of Magnetic Resonance*, **138**(2):334–342, 1999.
- [44] L. Pappalardo, I. G. Janausch, V. Vijayan, E. Zientz, J. Junker, W. Peti, M. Zweckstetter, G. Unden, and C. Griesinger. The NMR structure of the sensory domain of the membranous two-component Fumarate sensor (histidine protein kinase) DcuS of *Escherichia Coli*. *Journal of Biological Chemistry*, **278**(40):39185–39188, 2003.

- [45] L. Holm and C. Sander. Protein-structure comparison by alignment of distance matrices. *Journal of Molecular Biology*, **233**(1):123–138, 1993.
- [46] Y. S. Jung, M. Sharma, and M. Zweckstetter. Simultaneous assignment and structure determination of protein backbones by using NMR dipolar couplings. *Angewandte Chemie-International Edition*, **43**(26):3479–3481, 2004.
- [47] R. Koradi, M. Billeter, and K. Wuthrich. MolMol: A program for display and analysis of macromolecular structures. *Journal of Molecular Graphics*, **14**(1):51–55, 1996.
- [48] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax. NmrPipe - a multidimensional spectral processing system based on unix pipes. *Journal of Biomolecular NMR*, **6**(3):277–293, 1995.
- [49] M. Zweckstetter and A. Bax. Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *Journal of the American Chemical Society*, **122**(15):3791–3792, 2000.
- [50] A. Bax and S. Grzesiek. Methodological advances in protein NMR. *Accounts of Chemical Research*, **26**(4):131–138, 1993.
- [51] M. Salzmann, K. Pervushin, G. Wider, H. Senn, and K. Wuthrich. TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(23):13585–13590, 1998.
- [52] W. Peti, C. Griesinger, and W. Bermel. Adiabatic TOCSY for C,C and H,H J-transfer. *Journal of Biomolecular NMR*, **18**(3):199–205, 2000.
- [53] M. Zweckstetter and T. A. Holak. Robust refocusing of $^{13}\text{C}_\alpha$ magnetization in multidimensional NMR experiments by adiabatic fast passage pulses. *Journal of Biomolecular NMR*, **15**(4):331–334, 1999.
- [54] M. Sattler, M. Maurer, J. Schleucher, and C. Griesinger. A simultaneous ^{15}N , ^1H -HSQC and ^{13}C , ^1H -HSQC with sensitivity enhancement and a heteronuclear gradient-echo. *Journal of Biomolecular NMR*, **5**(1):97–102, 1995.

-
- [55] Y. L. Xia, A. Yee, C. H. Arrowsmith, and X. L. Gao. $^1\text{H}_C$ and $^1\text{H}^N$ Total NOE correlations in a single 3D NMR experiment. ^{15}N and ^{13}C time-sharing in T(1) and T(2) dimensions for simultaneous data acquisition. *Journal of Biomolecular NMR*, **27**(3):193–203, 2003.
- [56] F. A. A. Mulder, C. A. E. M. Spronk, M. Slijper, R. Kaptein, and R. Boelens. Improved HSQC experiments for the observation of exchange broadened signals. *Journal of Biomolecular NMR*, **8**(2):223–228, 1996.
- [57] S. Schwarzingler, G. J. A. Kroon, T. R. Foss, J. Chung, P. E. Wright, and H. J. Dyson. Sequence-dependent correction of random coil NMR chemical shifts. *Journal of the American Chemical Society*, **123**(13):2970–2978, 2001.
- [58] D. S. Wishart and B. D. Sykes. Chemical-shifts as a tool for structure determination. *Methods in Enzymology*, **239**:363–392, 1994.
- [59] W. J. Metzler, K. L. Constantine, M. S. Friedrichs, A. J. Bell, E. G. Ernst, T. B. Lavoie, and L. Mueller. Characterization of the 3-Dimensional solution structure of human Profilin – ^1H , oneC and ^{15}N NMR assignments and global folding pattern. *Biochemistry*, **32**(50):13818–13829, 1993.
- [60] G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax. Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *Journal of the American Chemical Society*, **120**(27):6836–6837, 1998.
- [61] M. Ottiger, F. Delaglio, and A. Bax. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *Journal of Magnetic Resonance*, **131**(2):373–378, 1998.
- [62] J. J. Chou, F. Delaglio, and A. Bax. Measurement of one-bond ^{15}N – ^{13}C dipolar couplings in medium sized proteins. *Journal of Biomolecular NMR*, **18**(2):101–105, 2000.

- [63] V. Vijayan and M. Zweckstetter. Simultaneous measurement of protein one-bond residual dipolar couplings without increased resonance overlap. *Journal of Magnetic Resonance*, **174**(2):245–253, 2005.
- [64] N. Tjandra, A. Szabo, and A. Bax. Protein backbone dynamics and ^{15}N chemical shift anisotropy from quantitative measurement of relaxation interference effects. *Journal of the American Chemical Society*, **118**(29):6986–6991, 1996.
- [65] G. Lipari, A. Szabo, and R. M. Levy. Protein dynamics and NMR Relaxation - comparison of simulations with experiment. *Nature*, **300**(5888):197–198, 1982.
- [66] G. Lipari and A. Szabo. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules .1. Theory and range of validity. *Journal of the American Chemical Society*, **104**(17):4546–4559, 1982.
- [67] G. Lipari and A. Szabo. Model-Free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules .2. Analysis of experimental results. *Journal of the American Chemical Society*, **104**(17):4559–4570, 1982.
- [68] A. A. Bothnerby. Encyclopedia of nuclear magnetic resonance. pages 2932–2938. Wiley, Chichester, 1996.
- [69] C. Gayathri, A. A. Bothnerby, P. C. M. Vanzijl, and C. Maclean. Dipolar magnetic-field effects in NMR-spectra of liquids. *Chemical Physics Letters*, **87**(2):192–196, 1982.
- [70] A. Saupe and G. Englert. High-resolution nuclear magnetic resonance spectra of oriented molecules. *Physical Review Letters*, **11**(10):462–464, 1963.
- [71] J. Sass, F. Cordier, A. Hoffmann, A. Cousin, J. G. Omichinski, H. Lowen, and S. Grzesiek. Purple membrane induced alignment of biological macromolecules in the magnetic field. *Journal of the American Chemical Society*, **121**(10):2047–2055, 1999.
- [72] J. J. Chou, S. P. Li, and A. Bax. Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. *Journal of Biomolecular NMR*, **18**(3):217–227, 2000.

-
- [73] F. Delaglio, G. Kontaxis, and A. Bax. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *Journal of the American Chemical Society*, **122**(9):2142–2143, 2000.
- [74] J. C. Hus, D. Marion, and M. Blackledge. Determination of protein backbone structure using only residual dipolar couplings. *Journal of the American Chemical Society*, **123**(7):1541–1542, 2001.
- [75] J. H. Prestegard, H. Valafar, J. Glushka, and F. Tian. Nuclear magnetic resonance in the era of structural genomics. *Biochemistry*, **40**(30):8677–8685, 2001.
- [76] E. de Alba, M. Suzuki, and N. Tjandra. Simple multidimensional NMR experiments to obtain different types of one-bond dipolar couplings simultaneously. *Journal of Biomolecular NMR*, **19**(1):63–67, 2001.
- [77] P. Permi, P. R. Rosevear, and A. Annala. A set of HNC(O)-based experiments for measurement of residual dipolar couplings in ^{15}N , ^{13}C , (^2H)-labeled proteins. *Journal of Biomolecular NMR*, **17**(1):43–54, 2000.
- [78] S. Heikkinen, P. Permi, and I. Kilpelainen. Methods for the measurement of $^1\text{J}_{\text{N}\alpha}$ and $^2\text{J}_{\text{N}\alpha}$ from a simplified $^{13}\text{C}_\alpha$ -coupled ^{15}N Se-HSQC spectrum. *Journal of Magnetic Resonance*, **148**(1):53–60, 2001.
- [79] H. L. J. Wienk, M. M. Martinez, G. N. Yalloway, J. M. Schmidt, C. Perez, H. Ruterjans, and F. Lohr. Simultaneous measurement of protein one-bond and two-bond nitrogen-carbon coupling constants using an internally referenced quantitative J-correlated [^{15}N , H^1]-TROSY-HNC(O) experiment. *Journal of Biomolecular NMR*, **25**(2):133–145, 2003.
- [80] P. Permi. Measurement of residual dipolar couplings from $^1\text{H}_\alpha$ to $^{13}\text{C}_\alpha$ and ^{15}N using a simple HNCA-based experiment. *Journal of Biomolecular NMR*, **27**(4):341–349, 2003.
- [81] K. Y. Ding and A. M. Gronenborn. Simultaneous and accurate determination of one-bond ^{15}N - $^{13}\text{C}'$ and two-bond $^1\text{H}^{\text{N}}$ - $^{13}\text{C}'$ dipolar couplings. *Journal of the American Chemical Society*, **125**(38):11504–11505, 2003.

- [82] D. W. Yang, R. A. Venters, G. A. Mueller, W. Y. Choy, and L. E. Kay. TROSY-based HNCO pulse sequences for the measurement of $^1\text{H}^N\text{-}^{15}\text{N}$, $^{15}\text{N}\text{-}^{13}\text{C}'$, $^1\text{H}^N\text{-}^{13}\text{C}'$, $^{13}\text{C}'\text{-}^{13}\text{C}_\alpha$ and $^1\text{H}^N\text{-}^{13}\text{C}_\alpha$ dipolar couplings in ^{15}N , ^{13}C , H^2 -labeled proteins. *Journal of Biomolecular NMR*, **14**(4):333–343, 1999.
- [83] S. Grzesiek and A. Bax. Correlating backbone amide and side-chain resonances in larger proteins by multiple relayed triple resonance NMR. *Journal of the American Chemical Society*, **114**(16):6291–6293, 1992.
- [84] J. J. Chou and A. Bax. Protein side-chain rotamers from dipolar couplings in a liquid crystalline phase. *Journal of the American Chemical Society*, **123**(16):3844–3845, 2001.
- [85] K. V. Pervushin, G. Wider, and K. Wuthrich. Single transition-to-single transition polarization transfer (ST2-PT) in $^{15}\text{N}, ^1\text{H}$ -TROSY. *Journal of Biomolecular NMR*, **12**(2):345–348, 1998.
- [86] E. Kupce, J. Boyd, and I. D. Campbell. Short selective pulses for biochemical applications. *Journal of Magnetic Resonance Series B*, **106**(3):300–303, 1995.
- [87] M. S. Silver, R. I. Joseph, and D. I. Hoult. Highly selective $\pi/2$ and π -pulse generation. *Journal of Magnetic Resonance*, **59**(2):347–351, 1984.
- [88] E. Kupce and R. Freeman. Compensation for spin-spin coupling effects during adiabatic pulses. *Journal of Magnetic Resonance*, **127**(1):36–48, 1997.
- [89] M. Zweckstetter and A. Bax. Characterization of molecular alignment in aqueous suspensions of Pf1 Bacteriophage. *Journal of Biomolecular NMR*, **20**(4):365–377, 2001.
- [90] G. Kontaxis, G. M. Clore, and A. Bax. Evaluation of cross-correlation effects and measurement of one-bond couplings in proteins with short transverse relaxation times. *Journal of Magnetic Resonance*, **143**(1):184–196, 2000.
- [91] R. H. Contreras and J. E. Peralta. Angular dependence of spin-spin coupling constants. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **37**(4):321–425, 2000.

-
- [92] S. Vijaykumar, C. E. Bugg, and W. J. Cook. Structure of Ubiquitin refined at 1.8 Å resolution. *Journal of Molecular Biology*, **194**(3):531–544, 1987.
- [93] S. Aizawa, C. S. Harwood, and R. J. Kadner. Signaling components in bacterial locomotion and sensory reception. *Journal of Bacteriology*, **182**(6):1459–1471, 2000.
- [94] B. T. Nixon, C. W. Ronson, and F. M. Ausubel. 2-Component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. *Proceedings of the National Academy of Sciences of the United States of America*, **83**(20):7850–7854, 1986.
- [95] A. J. Ninfa and B. Magasanik. Covalent modification of the *glnG* product, NR_I , by the *glnL* product, NR_{II} , regulates the transcription of the *glnALG* operon in *Escherichia-Coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **83**(16):5909–5913, 1986.
- [96] J. F. Hess, K. Oosawa, N. Kaplan, and M. I. Simon. Phosphorylation of 3 proteins in the signaling pathway of bacterial chemotaxis. *Cell*, **53**(1):79–87, 1988.
- [97] J. A. Hoch. Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology*, **3**(2):165–170, 2000.
- [98] L. A. Egger, H. Park, and M. Inouye. Signal transduction via the histidyl-aspartyl phosphorelay. *Genes to Cells*, **2**(3):167–184, 1997.
- [99] A. M. Stock, V. L. Robinson, and P. N. Goudreau. Two-component signal transduction. *Annual Review of Biochemistry*, **69**:183–215, 2000.
- [100] M. Levit, Y. Liu, M. Surette, and J. Stock. Active site interference and asymmetric activation in the chemotaxis protein histidine kinase *CheA*. *Journal of Biological Chemistry*, **271**(50):32057–32063, 1996.
- [101] Y. Yang and M. Inouye. Intermolecular complementation between 2 defective mutant signal-transducing receptors of *Escherichia-Coli*. *Proceedings of the National Academy of Sciences of the United States of America*, **88**(24):11057–11061, 1991.

- [102] L. Qin, R. Dutta, H. Kurokawa, M. Ikura, and M. Inouye. A monomeric histidine kinase derived from EnvZ, an Escherichia Coli osmosensor. *Molecular Microbiology*, **36**(1):24–32, 2000.
- [103] S. J. Cai and M. Inouye. Spontaneous subunit exchange and biochemical evidence for trans-autophosphorylation in a dimer of Escherichia Coli histidine kinase (EnvZ). *Journal of Molecular Biology*, **329**(3):495–503, 2003.
- [104] V. L. Robinson, D. R. Buckler, and A. M. Stock. A tale of two components: A novel kinase and a regulatory switch. *Nature Structural Biology*, **7**(8):626–633, 2000.
- [105] M.Y. Galperin. Bacterial signal transduction network in a genomic perspective. *Environmental Microbiology*, **6**(6):552–567, 2004.
- [106] J. Stock. Signal transduction: gyrating protein kinases. *Current Biology*, **9**(10):R364–R367, 1999.
- [107] A. M. Bilwes, L. A. Alex, B. R. Crane, and M. I. Simon. Structure of CheA, a signal-transducing histidine kinase. *Cell*, **96**(1):131–141, 1999.
- [108] T. Tanaka, S. K. Saha, C. Tomomori, R. Ishima, D. J. Liu, K. I. Tong, H. Park, R. Dutta, L. Qin, M. B. Swindells, T. Yamazaki, A. M. Ono, M. Kainosho, M. Inouye, and M. Ikura. NMR Structure of the histidine kinase domain of the E-Coli osmosensor EnvZ. *Nature*, **396**(6706):88–92, 1998.
- [109] C. Tomomori, T. Tanaka, R. Dutta, H. Y. Park, S. K. Saha, Y. Zhu, R. Ishima, D. J. Liu, K. I. Tong, H. Kurokawa, H. Qian, M. Inouye, and M. Ikura. Solution structure of the homodimeric core domain of Escherichia Coli histidine kinase EnvZ. *Nature Structural Biology*, **6**(8):729–734, 1999.
- [110] M. R. Atkinson and A. J. Ninfa. Characterization of Escherichia-Coli glnL Mutations affecting nitrogen regulation. *Journal of Bacteriology*, **174**(14):4538–4548, 1992.
- [111] L. A. Collins, S. M. Egan, and V. Stewart. Mutational analysis reveals functional similarity between NARX, a nitrate sensor in Escherichia-Coli K-12, and the methyl-accepting chemotaxis proteins. *Journal of Bacteriology*, **174**(11):3667–3675, 1992.

-
- [112] J. S. Fassler, W. M. Gray, C. L. Malone, W. Tao, H. Lin, and R. J. Deschenes. Activated alleles of Yeast SLN1 Increase Mcm1-dependent reporter gene expression and diminish signaling through the Hog1 osmosensing pathway. *Journal of Biological Chemistry*, **272**(20):13365–13371, 1997.
- [113] H. Park and M. Inouye. Mutational analysis of the linker region of EnvZ, an Osmosensor in Escherichia Coli. *Journal of Bacteriology*, **179**(13):4382–4390, 1997.
- [114] C. P. Ponting and L. Aravind. PAS: A multifunctional domain family comes to light. *Current Biology*, **7**(11):R674–R677, 1997.
- [115] I. B. Zhulin and B. L. Taylor. Correlation of PAS domains with electron transport-associated proteins in completely sequenced microbial genomes. *Molecular Microbiology*, **29**(6):1522–1523, 1998.
- [116] B. L. Taylor and I. B. Zhulin. PAS Domains: Internal sensors of oxygen, redox potential, and light. *Microbiology and Molecular Biology Reviews*, **63**(2):479–+, 1999.
- [117] I. B. Zhulin and B. L. Taylor. PAS domain S-boxes in archaea, bacteria and sensors for oxygen and redox. *Trends in Biochemical Sciences*, **22**(9):331–333, 1997.
- [118] J. R. Nambu, J. O. Lewis, K. A. Wharton, and S. T. Crews. The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell*, **67**(6):1157–1167, 1991.
- [119] J. L. Pellequer, K. A. Wager-Smith, S. A. Kay, and E. D. Getzoff. Photoactive Yellow Protein: A structural prototype for the three-dimensional fold of the PAS domain superfamily. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(11):5884–5890, 1998.
- [120] G. E. O. Borgstahl, D. R. Williams, and E. D. Getzoff. 1.4 Angstrom structure of Photoactive Yellow Protein, a cytosolic photoreceptor - unusual fold, active-site, and chromophore. *Biochemistry*, **34**(19):6278–6287, 1995.

- [121] M. H. Hefti, K. J. Francoijs, S. C. de Vries, R. Dixon, and J. Vervoort. The PAS fold - A redefinition of the PAS domain based upon structural prediction. *European Journal of Biochemistry*, **271**(6):1198–1208, 2004.
- [122] I. G. Janausch, E. Zientz, Q. H. Tran, A. Kroger, and G. Unden. C-4-Dicarboxylate carriers and sensors in bacteria. *Biochimica Et Biophysica Acta-Bioenergetics*, **1553**(1-2):39–56, 2002.
- [123] I. G. Janausch, I. Garcia-Moreno, and G. Unden. Function of DcuS from Escherichia Coli as a Fumarate-stimulated histidine protein kinase in vitro. *Journal of Biological Chemistry*, **277**(42):39809–39814, 2002.
- [124] J. S. Parkinson and E. C. Kofoed. Communication modules in bacterial signaling proteins. *Annual Review of Genetics*, **26**:71–112, 1992.
- [125] E. Zientz, J. Bongaerts, and G. Unden. Fumarate regulation of gene expression in Escherichia Coli by the DcuSR (Dcusr Genes) Two-component regulatory system. *Journal of Bacteriology*, **180**(20):5421–5425, 1998.
- [126] P. Golby, S. Davies, D. J. Kelly, J. R. Guest, and S. C. Andrews. Identification and characterization of a two-component sensor-kinase and response-regulator system (DcuS-DcuR) controlling gene expression in response to C4-dicarboxylates in Escherichia Coli. *Journal of Bacteriology*, **181**(4):1238–1248, 1999.
- [127] S. Kaspar, R. Perozzo, S. Reinelt, M. Meyer, K. Pfister, L. Scapozza, and M. Bott. The periplasmic domain of the histidine autokinase CitA functions as a highly specific citrate receptor. *Molecular Microbiology*, **33**(4):858–872, 1999.
- [128] T. N. Parac, B. Coligaev, E. Zientz, G. Unden, W. Peti, and C. Griesinger. Assignment of ^1H , ^{13}C and ^{15}N resonances to the sensory domain of the membraneous two-component fumarate sensor (histidine protein kinase) DcuS of Escherichia Coli. *Journal of Biomolecular NMR*, **19**(1):91–92, 2001.

-
- [129] G. W. Vuister, F. Delaglio, and A. Bax. An empirical correlation between $^1J_{C\alpha H\alpha}$ and protein backbone conformation. *Journal of the American Chemical Society*, **114**(24):9674–9675, 1992.
- [130] M. V. Milburn, G. G. Prive, D. L. Milligan, W. G. Scott, J. Yeh, J. Jancarik, D. E. Koshland, and S. H. Kim. 3-Dimensional structures of the ligand-binding domain of the bacterial aspartate receptor with and without a ligand. *Science*, **254**(5036):1342–1347, 1991.
- [131] J. I. Yeh, H. P. Biemann, J. Pandit, D. E. Koshland, and S. H. Kim. The 3-dimensional structure of the ligand-binding domain of a wild-type bacterial chemotaxis receptor - structural comparison to the cross-linked mutant forms and conformational-changes upon ligand-binding. *Journal of Biological Chemistry*, **268**(13):9787–9792, 1993.
- [132] H. Kneuper, I. G. Janausch, V. Vijayan, M. Zweckstetter, V. Bock, C. Griesinger, and G. Uden. The nature of the stimulus and of the fumarate binding site of the fumarate sensor DcuS of Escherichia Coli. *Journal of Biological Chemistry*, **280**(21):20596–20603, 2005.
- [133] T. Gerharz, S. Reinelt, S. Kaspar, L. Scapozza, and M. Bott. Identification of basic amino acid residues important for citrate binding by the periplasmic receptor domain of the sensor kinase CitA. *Biochemistry*, **42**(19):5917–5924, 2003.
- [134] E. Schwarz and D. Oesterhelt. Cloning and expression of Klebsiella-Pneumoniae genes-coding for citrate transport and fermentation. *Embo Journal*, **4**(6):1599–1603, 1985.
- [135] M. E. Vanderrest, E. Schwarz, D. Oesterhelt, and W. N. Konings. DNA-sequence of a citrate carrier of Klebsiella-Pneumoniae. *European Journal of Biochemistry*, **189**(2):401–407, 1990.
- [136] M. E. Vanderrest, T. Abee, D. Molenaar, and W. N. Konings. Mechanism and energetics of a citrate-transport system of Klebsiella-Pneumoniae. *European Journal of Biochemistry*, **195**(1):71–77, 1991.

- [137] P. Dimroth and A. Thomer. Citrate transport in *Klebsiella-Pneumoniae*. *Biological Chemistry Hoppe-Seyler*, **367**(8):813–823, 1986.
- [138] M. Bott, M. Meyer, and P. Dimroth. Regulation of anaerobic citrate metabolism in *Klebsiella-Pneumoniae*. *Molecular Microbiology*, **18**(3):533–546, 1995.
- [139] M. E. Vanderrest, R. M. Siewe, T. Abee, E. Schwarz, D. Oesterhelt, and W. N. Konings. Nucleotide-sequence and functional-properties of a sodium-dependent citrate transport-system from *Klebsiella-Pneumoniae*. *Journal of Biological Chemistry*, **267**(13):8971–8976, 1992.
- [140] E. Laussermair, E. Schwarz, D. Oesterhelt, H. Reinke, K. Beyreuther, and P. Dimroth. The sodium-ion translocating oxaloacetate decarboxylase of *Klebsiella-Pneumoniae* - Sequence of the integral membrane-bound subunit-beta and gamma. *Journal of Biological Chemistry*, **264**(25):14710–14715, 1989.
- [141] E. Schwarz, D. Oesterhelt, H. Reinke, K. Beyreuther, and P. Dimroth. The Sodium-ion translocating oxalacetate decarboxylase of *Klebsiella-Pneumoniae* - Sequence of the biotin-containing alpha-subunit and relationship to other biotin-containing enzymes. *Journal of Biological Chemistry*, **263**(20):9640–9645, 1988.
- [142] G. Woehlke, E. Laussermair, E. Schwarz, D. Oesterhelt, H. Reinke, K. Beyreuther, and P. Dimroth. Sequence of the beta-subunit of oxaloacetate decarboxylase from *Klebsiella-Pneumoniae* - A correction of the C-terminal part. *Journal of Biological Chemistry*, **267**(32):22804–22805, 1992.
- [143] M. Bott and P. Dimroth. *Klebsiella-Pneumoniae* genes for Citrate Lyase and Citrate Lyase Ligase - Localization, sequencing, and expression. *Molecular Microbiology*, **14**(2):347–356, 1994.
- [144] J. A. Hoch and T.J. Silhavy. *Two-component signal transduction*. ASM press, Washington, DC, 1995.

-
- [145] R. V. Swanson, L. A. Alex, and M. I. Simon. Histidine and aspartate phosphorylation - 2-component systems and the limits of homolog. *Trends in Biochemical Sciences*, **19**(11):485–490, 1994.
- [146] S. Reinelt, E. Hofmann, T. Gerharz, M. Bott, and D. R. Madden. The Structure of the periplasmic ligand-binding domain of the sensor Kinase CitA reveals the first extracellular PAS domain. *Journal of Biological Chemistry*, **278**(40):39189–39196, 2003.
- [147] Brien R. W. O and J. R. Stern. Role of sodium in determining alternate pathways of aerobic citrate catabolism in aerobacter-aerogenes EnZ alpha keto glutarate dehydrogenase Enz citritase EnZ oxalacetate decarboxylase. *Journal of Bacteriology*, **99**(2):389–394, 1969.
- [148] R. W. O'Brien. Induction by sodium of citrate fermentation enzymes in Klebsiella-Aerogenes. *Febs Letters*, **53**(1):61–63, 1975.
- [149] R. W. O'Brien. Effect of aeration and sodium on metabolism of citrate by Klebsiella-Aerogenes. *Journal of Bacteriology*, **122**(2):468–473, 1975.
- [150] P. Dimroth, P. Jockel, and M. Schmid. Coupling mechanism of the oxaloacetate decarboxylase Na⁺ pump. *Biochimica Et Biophysica Acta-Bioenergetics*, **1505**(1):1–14, 2001.
- [151] K. M. Pos and P. Dimroth. Functional properties of the purified Na⁺ -dependent citrate carrier of Klebsiella Pneumoniae: Evidence for asymmetric orientation of the carrier protein in proteoliposomes. *Biochemistry*, **35**(3):1018–1026, 1996.
- [152] P. Muller, S. Kopke, and G. M. Sheldrick. Is the bond-valence method able to identify metal atoms in protein structures? *Acta Crystallographica Section D-Biological Crystallography*, **59**:32–37, 2003.
- [153] M. Hulko, F. Berndt, M. Gruber, J. U. Linder, V. Truffault, A. Schultz, J. Martin, J. E. Schultz, A. N. Lupas, and M. Coles. The HAMP domain structure implies helix rotation in transmembrane signaling. *Cell*, **126**(5):929–940, 2006.

- [154] E. E. Matthews, M. Zoonens, and D. M. Engelman. Dynamic helix interactions in transmembrane signaling. *Cell*, **127**(3):447–450, 2006.
- [155] K. M. Ottemann, W. Z. Xiao, Y. K. Shin, and D. E. Koshland. A piston model for transmembrane signaling of the aspartate receptor. *Science*, **285**(5434):1751–1754, 1999.
- [156] O. W. Sorensen, G. W. Eich, M. H. Levit, G. Bodenhausen, and R. R. Ernst. Product operator formalism for the description of NMR pulse experiments. *Progress in nuclear magnetic resonance*, **16**:163–192, 1984.

Appendix A

Time evolution of product operators

The discussion given in the following section is based on the density operator formalism [156], which permits the most convenient description of quantum mechanical system dynamics. The density operator is a mathematical description of spin system, and is given by the sum of all spins and spin states weighted by time dependent factors. $b_s(t)$.

$$\sigma(t) = \sum_S b_S(t) B_S \quad (\text{A.1})$$

where B_S are the set of basis vectors describing the different spin states, and are given by

$$B_S = 2^{(q-1)} \prod_{k=1}^N (I_{kv})^{a_{sk}} \quad (\text{A.2})$$

where N is the total number of spins half in the system; k and v are respectively, the index of the spin and x,y, or z; q is the number of single spin operators I_x, I_y or I_z in the product; for these spin a_{sk} is 1.

The complete basis set (B_S) for a system with N spin 1/2 consists of 4^N product operators (B_S). As an example in table A.1 , the complete set of 16 product operators (B_S) for a two spin system with $I = S = 1/2$ are listed.

The time evolution of the density operator $\sigma(t)$ is given by the Liouville-von Neumann equation

$$\frac{d}{dt} \sigma(t) = -i [H(t), \sigma(t)]. \quad (\text{A.3})$$

where H is the hamiltonian under which the spin density operator $\sigma(t)$ (defined by A.1) evolves.

For H being constant during the time interval 0 to t , equation A.3 integrates to

$$\sigma(t) = \exp(-iHt) \sigma(0) \exp(iHt) \quad (\text{A.4})$$

q	Product or basis operator (B_S)
0	unity operator $\frac{1}{2}E$
1 single spin operators	I_x, I_y, I_z, S_x, S_y and S_z
2 two-spin operators	$2I_x S_x, 2I_x S_y, 2I_x S_z$ $2I_y S_x, 2I_y S_y, 2I_y S_z$ $2I_z S_x, 2I_z S_y, 2I_z S_z$

Table A.1: Basis operators for $q = 0, 1$ or 2 for an IS system of spin $1/2$

This type of exponential operator function can be evaluated using the Baker-Campbell-Hausdorff(BCH) formula. The BCH formula consolidates the series expansion of exponential operator to evaluate such operator evolution.

$$\exp(-iHt) \sigma(0) \exp(iHt) = \sigma(0) - (it) [H, \sigma(0)] + \frac{(it)^2}{2!} [H, [H, \sigma(0)]] - \frac{(it)^3}{3!} [H, [H, [H, \sigma(0)]]] + \dots \quad (\text{A.5})$$

Now collecting the terms corresponding to the cosine and sine terms, we get the equation for the evolution of $\sigma(0)$. A NMR problem has a well defined, finite dimensional space associated with it and so the number of independent commutators is limited and a series expansion quickly leads to an operator recursion.

Evolution of different initial density operators with some common spin Hamiltonians are given in the table A.2

Table A.2: Time evolution of common product operators for a two spin systems under different spin hamiltonians

Hamiltonians	Time evolution
Under pulse rotation of angle β and phase φ $H = I_x \cos \varphi + I_y \sin \varphi$	$I_z \xrightarrow{I_x \cos \varphi + I_y \sin \varphi} I_z \cos \beta + I_x \sin \beta \sin \varphi - I_y \sin \beta \cos \varphi$ $I_x \xrightarrow{I_x \cos \varphi + I_y \sin \varphi} -I_z \sin \beta \sin \varphi + I_x (\cos \beta \sin^2 \varphi + \cos^2 \varphi) + I_y \sin^2 \left(\frac{\beta}{2}\right) \sin 2\varphi$ $I_y \xrightarrow{I_x \cos \varphi + I_y \sin \varphi} I_z \sin \beta \cos \varphi + I_x \sin^2 \left(\frac{\beta}{2}\right) \sin 2\varphi + I_y (\cos \beta \cos^2 \varphi + \sin^2 \varphi)$
Under chemical shift $\Omega t I_z$	$I_x \xrightarrow{\Omega t I_z} I_x \cos \Omega t + I_y \sin \Omega t$ $I_y \xrightarrow{\Omega t I_z} I_y \cos \Omega t - I_x \sin \Omega t$ $I_x S_z \xrightarrow{\Omega t I_z} I_x S_z \cos \Omega t + I_y S_z \sin \Omega t$

Continued on next page

Table A.2 – continued from previous page

Hamiltonians	Time evolution
	$I^\pm \xrightarrow{\Omega I_z} I^\pm e^{\mp \Omega t}$
<p style="text-align: center;">Under weak coupling J_{IS}</p> <p style="text-align: center;">$\pi J_{IS} t 2I_z S_z$</p>	$I_x \xrightarrow{\pi J_{IS} t 2I_z S_z} I_x \cos(\pi J_{IS} t) + 2I_y S_z \sin(\pi J_{IS} t)$ $I_y \xrightarrow{\pi J_{IS} t 2I_z S_z} I_y \cos(\pi J_{IS} t) - 2I_x S_z \sin(\pi J_{IS} t)$ $2I_x S_z \xrightarrow{\pi J_{IS} t 2I_z S_z} 2I_x S_z \cos(\pi J_{IS} t) + I_y \sin(\pi J_{IS} t)$ $2I_y S_z \xrightarrow{\pi J_{IS} t 2I_z S_z} 2I_y S_z \cos(\pi J_{IS} t) - I_x \sin(\pi J_{IS} t)$

Appendix B

Chemical shift assignments and residual dipolar couplings.

In the following section, the backbone assignments of CitAP with and without citrate are given. These chemical shifts were not deposited in the BMRB.

Residual dipolar couplings for different proteins measured during this thesis are also summarized.

B.1 Chemical shift Assignments

Table B.1: Backbone chemical shifts of citrate free CitAP at 310K.

Residue No.	Residue name	C'	C $_{\alpha}$	C $_{\beta}$	H N	N
1	MET	175.4	55.41	32.97	8.46	121.5
2	ASP	176.3	54.42	41.32	8.355	122.1
3	ILE	176.2	61.67	38.62	8.109	120.8
4	THR	175.3	62.98	69.36	7.961	116.7
5	GLU	177.2	58.35	29.98	8.359	122.4
6	GLU	178	58.54	29.73	8.329	120.1
7	ARG	177.7	58.21	30.1	8.023	119.9
8	LEU	178.2	57.21	41.88	8.058	121.1
9	HIS	177	58.04	29.73	8.299	117
13	GLY	-	47.82	-	8.095	107.7
14	GLN	-	58.47	28.65	7.774	118.9
18	ILE	178.8	65.14	37.5	8.329	118.5
19	GLN	177.8	59.37	29.59	7.793	118.3
20	ALA	180.1	55.85	17.21	8.322	121.6
21	MET	178.9	59.32	33.56	8.607	117.4
22	GLN	179.1	59.09	28.55	8.282	119.4
23	ILE	177.5	65.5	38.69	8.391	119.3
24	SER	174.5	61.15	63.39	7.998	110.3

Continued on next page

Table B.1 – continued from previous page

Residue No.	Residue name	C'	C _α	C _β	H ^N	N
25	ALA	178.2	51.25	19.73	6.869	120.2
26	MET	175.6	55.42	33.56	7.612	122.2
28	GLU	179.1	59.57	29.97	9.71	116.4
29	LEU	177.5	56.87	42.02	7.323	120
30	VAL	177.7	67.58	31.86	7.719	119.8
31	GLU	177.9	59.44	29.96	7.609	116.3
32	ALA	179.8	55.52	19.02	7.619	120.1
33	VAL	179.4	67.45	31.05	8.622	117.9
34	GLN	178.3	59.78	28.87	8.355	122.1
35	LYS	175.1	56.5	32.91	7.8	115
36	ARG	174.2	57.58	28.18	7.966	120.8
37	ASP	175.9	52.32	40.16	8.423	118.1
38	LEU	179.9	59.04	40.35	7.964	123.4
39	ALA	180.5	54.98	18.07	8.355	122.1
40	ARG	178	58.03	30.09	7.728	121.3
41	ILE	176.8	66.78	38.22	7.961	117.6
42	LYS	177.1	59.85	32.34	7.666	119.2
43	ALA	179.8	54.07	18.36	7.646	118.3
44	LEU	178	56.64	43.44	7.66	116.2
60	GLY	172.7	43.94	–	9.592	117.3
61	ASP	177.3	52.59	41.07	8.471	118.8
62	ALA	177.2	54.17	18.37	7.449	118.7
63	SER	177.2	57.91	64.36	8.32	112.8
81	GLY	175.1	45.38	–	8.443	110.2
82	GLY	173.8	45.71		8.113	108.3
83	ASP	176.5	54.22	40.75	8.273	119.8
112	ILE	175.3	60.62	39.58	8.147	120.4
113	GLN	175.3	53.83	34.06	9.129	127.4
114	ASP	177.9	51.99	40.93	8.429	119.7
115	ALA	179	55.01	18.76	8.442	119.3
116	THR	175.3	61.81	70.18	8.473	108.2
117	GLY	173.8	44.66	–	8.124	110.2
118	LYS	176.2	56.44	32.8	7.658	124
119	VAL	176.8	63.92	31.46	8.707	125.4
120	ILE	174.6	60.7	39.65	8.985	120.4
121	GLY	170.7	45.99	–	7.697	107.5
122	ILE	173	59.61	41.97	8.661	118.3
123	VAL	174.1	61.11	34.18	8.812	126.5
124	SER	-	55.84	65.04	8.812	120.8
130	GLU	175.8	56.79	30.54	8.587	124.9
131	GLN	175.2	55.6	29.91	8.296	120.9
132	LEU	176.2	55.37	42.6	8.289	124.8

Continued on next page

Table B.1 – continued from previous page

Residue No.	Residue name	C'	C _α	C _β	H ^N	N
133	GLU	180.8	57.87	31.32	7.835	125.9

Table B.2: Backbone chemical shifts of citrate bound CitAP at 310K.

Residue No.	Residue name	C'	C _α	C _β	H ^N	N
1	MET	175.4	55.45	32.98	8.474	121.7
2	ASP	176.3	54.31	41.45	8.349	122
3	ILE	176.3	61.63	38.62	8.137	120.7
4	THR	175.4	63.01	69.69	8.003	116.5
5	GLU	177.3	58.13	29.99	8.327	122.3
6	GLU	177.9	58.51	29.68	8.381	120.7
7	ARG	177.9	57.78	30.15	8.084	119.7
8	LEU	178.1	56.95	41.99	8.012	120.9
9	HIS	177	57.4	29.4	8.266	117
13	GLY	174.8	65.61	–	8.043	107.4
14	GLN	178.1	63.96	28.51	7.69	118.2
18	ILE	178.7	65	–	8.307	118.3
20	ALA	178.6	55.9	–	8.555	124.7
21	MET	178.8	59.18	33.54	8.633	117.3
25	ALA	178.1	51.11	19.79	6.797	120.6
26	MET	–	55.25	33.36	7.639	122.6
29	LEU	177.4	56.75	41.98	7.297	120
30	VAL	177.7	67.19	31.9	7.77	119.9
31	GLU	178	59.11	29.96	7.636	116.2
32	ALA	179.7	55.32	19.18	7.618	120.1
33	VAL	179.4	67.19	31.08	8.634	118
34	GLN	178.4	59.71	28.89	8.349	122
35	LYS	175	56.19	33	7.741	115
36	ARG	174	57.48	28.41	7.988	121
37	ASP	175.9	52.16	40.45	8.3	117.8
38	LEU	180	59.22	40.59	8.035	123.5
39	ALA	180.6	54.95	18.17	8.404	122.3
40	ARG	178.2	57.73	30.14	7.842	121.8
41	ILE	176.9	66.6	38.38	8.086	117.8
42	LYS	177	60.18	32.53	7.87	119.9
43	ALA	179.7	54.08	18.47	7.767	118.4
44	LEU	178.2	56.21	43.6	7.657	116.4
45	ILE	177.8	60.42	35.49	8.063	116.9
56	TYR	172.2	56.39	40.2	7.325	113.4
57	ILE	176.5	61.3	42.85	6.075	116.1

Continued on next page

Table B.2 – continued from previous page

Residue No.	Residue name	C'	C _α	C _β	H ^N	N
58	THR	173.6	61.88	72.54	10.26	127.9
59	VAL	175	60.87	34.31	8.174	123.7
60	GLY	172.5	43.9	-	9.713	118
61	ASP	177.2	52.36	40.96	8.616	119
62	ALA	176.9	54.61	18.56	7.317	117.9
63	SER	175.4	57.38	64.77	8.264	111.6
64	GLY	172.8	45.64	-	8.264	111.6
65	GLN	177.2	55.28	29.52	8.361	120.9
66	ARG	177	59.13	30.83	8.998	121.8
67	LEU	175	54.5	43.93	9.771	125.4
68	TYR	172.7	59.94	42.21	8.02	115.9
69	HIS	172.8	55.21	35.06	6.225	127.3
70	VAL	175.7	64.23	31.21	7.909	118.9
73	ASP	176.7	55.79	40.34	7.914	116.3
74	GLU	178.3	55.18	27.93	7.914	116.3
75	ILE	176.4	64.66	37.22	7.09	122.1
76	GLY	173.4	44.92	-	8.786	114.9
77	LYS	176	54.49	34.67	7.114	118.5
78	SER	174.2	59.57	63.86	8.466	116.4
81	GLY	175.8	45.11	-	8.584	110.4
82	GLY	175.2	46.71	-	8.884	108.5
83	ASP	178.3	52.45	39.98	8.363	117
84	SER	173.4	60.92	64.31	8.896	115.9
85	ASP	177.9	58.19	40.32	8.103	122.9
86	GLU	178.6	59.89	28.32	8.935	119.3
87	ALA	178	54.8	18.77	7.341	122.2
88	LEU	178	57.32	42.8	8.075	114.6
92	LYS	174.6	55.86	34.63	8.029	120.3
93	SER	174.8	57.66	65.44	7.96	112.3
94	TYR	172.2	56.45	38.53	7.901	120.7
95	VAL	176.1	60.21	33.88	8.339	117.5
96	SER	172.7	56.31	64.87	9.304	121.2
97	VAL	-	61.17	33.69	8.242	127.1
103	GLY	-	43.26	-	7.253	122.6
107	ARG	175.4	56.17	32.73	8.218	120.2
108	GLY	171.4	44.22	-	7.251	114.2
109	LYS	175	53.88	34.26	8.59	121.7
110	SER	172.5	54.28	66.53	8.178	111.5
112	ILE	175.5	60.49	39.66	7.967	119.8
113	GLN	175.2	53.74	34.67	9.267	127.7
114	ASP	178	51.84	40.94	8.439	119.6
115	ALA	179.1	54.78	18.75	8.436	119.3

Continued on next page

Table B.2 – continued from previous page

Residue No.	Residue name	C'	C _α	C _β	H ^N	N
116	THR	175.3	61.59	70.11	8.488	108.3
117	GLY	174	44.63	–	8.136	110.3
118	LYS	176.3	56.37	32.89	7.646	124.5
119	VAL	176.9	64.36	31.78	8.828	125.6
120	ILE	174.3	60.46	39.96	9.04	120
121	GLY	170.7	46.14	–	7.666	107.7
122	ILE	171.2	59.65	41.51	8.822	119.2
123	VAL	173.2	61.16	34.5	8.73	126
124	SER	174	54.65	66.18	–	–
125	VAL	174.7	61.52	33.1	9.135	128.8
126	GLY	171.9	44.28	–	7.18	111.9
127	TYR	176.1	56.38	41.7	8.968	118.6
128	THR	175.2	62.24	68.72	8.633	115
129	ILE	176.7	61.56	38.81	7.839	121
130	GLU	176.6	57.59	29.93	8.474	121.7
131	GLN	175.7	56.05	29.53	8.003	119
132	LEU	176.2	55.17	42.62	7.979	122.4
133	GLU	180.8	57.88	31.36	7.71	125.4

B.2 Residual dipolar coupling

Table B.3: Residual dipolar couplings measured for DcuS-PD in phages

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
24	MET	H	24	MET	N	-7.272
25	THR	H	25	THR	N	6.874
26	ARG	H	26	ARG	N	-5.494
27	ASP	H	27	ASP	N	-7.022
28	GLY	H	28	GLY	N	-5.500
29	LEU	H	29	LEU	N	-2.338
30	ALA	H	30	ALA	N	-9.714
31	APN	H	31	APN	N	-16.478
32	LYS	H	32	LYS	N	-8.552
33	ALA	H	33	ALA	N	6.582
34	LEU	H	34	LEU	N	-11.398
35	ALA	H	35	ALA	N	-19.018
36	VAL	H	36	VAL	N	-7.022
37	ALA	H	37	ALA	N	-13.298
38	ARG	H	38	ARG	N	-21.694
39	THR	H	39	THR	N	-15.794
40	LEU	H	40	LEU	N	-9.366
41	ALA	H	41	ALA	N	-19.112
42	ASP	H	42	ASP	N	-21.046
46	ILE	H	46	ILE	N	7.626
47	ARG	H	47	ARG	N	8.266
48	GLN	H	48	GLN	N	2.606
49	GLY	H	49	GLY	N	-5.478
50	LEU	H	50	LEU	N	8.634
51	GLN	H	51	GLN	N	12.094
52	LYS	H	52	LYS	N	-15.724
53	LYS	H	53	LYS	N	-5.336
56	ELU	H	56	ELU	N	1.194
57	SER	H	57	SER	N	14.250
60	GLN	H	60	GLN	N	-10.462
61	ALA	H	61	ALA	N	-17.306
62	ILE	H	62	ILE	N	-7.982
63	ALA	H	63	ALA	N	-10.458
64	GLU	H	64	GLU	N	-18.378
65	ALA	H	65	ALA	N	-18.870
66	VAL	H	66	VAL	N	-15.392
67	ARG	H	67	ARG	N	-16.418
68	LYS	H	68	LYS	N	-16.696
70	APN	H	70	APN	N	-12.502

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
71	ASP	H	71	ASP	N	-5.140
72	LEU	H	72	LEU	N	-3.780
74	PHE	H	74	PHE	N	-0.572
76	VAL	H	76	VAL	N	-7.428
77	VAL	H	77	VAL	N	2.332
78	THR	H	78	THR	N	2.972
79	ASP	H	79	ASP	N	14.192
80	MET	H	80	MET	N	-14.188
81	GLN	H	81	GLN	N	3.190
82	SER	H	82	SER	N	15.838
83	LEU	H	83	LEU	N	-5.574
84	ARG	H	84	ARG	N	19.856
85	TYR	H	85	TYR	N	-11.966
87	HIS	H	87	HIS	N	6.322
90	ALA	H	90	ALA	N	-5.450
91	GLN	H	91	GLN	N	-12.010
92	ARG	H	92	ARG	N	-3.392
93	ILE	H	93	ILE	N	-2.746
94	GLY	H	94	GLY	N	8.126
95	GLN	H	95	GLN	N	-3.454
97	PHE	H	97	PHE	N	-1.800
98	LYS	H	98	LYS	N	1.610
99	GLY	H	99	GLY	N	-1.602
101	ASP	H	101	ASP	N	-4.604
102	ILE	H	102	ILE	N	12.356
103	LEU	H	103	LEU	N	2.232
104	LYS	H	104	LYS	N	-7.958
105	ALA	H	105	ALA	N	6.346
106	LEU	H	106	LEU	N	2.624
108	GLY	H	108	GLY	N	-1.972
109	GLU	H	109	GLU	N	15.184
110	GLU	H	110	GLU	N	5.218
112	VAL	H	112	VAL	N	3.264
113	ALA	H	113	ALA	N	4.866
114	ILE	H	114	ILE	N	6.614
115	APN	H	115	APN	N	9.890
116	ARG	H	116	ARG	N	-8.614
121	GLN	H	121	GLN	N	-1.332
122	ALA	H	122	ALA	N	0.396
123	LEU	H	123	LEU	N	9.168
125	VAL	H	125	VAL	N	-1.902
126	PHE	H	126	PHE	N	-2.088

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
127	THR	H	127	THR	N	2.014
129	ILE	H	129	ILE	N	16.536
130	TYR	H	130	TYR	N	3.082
131	ASP	H	131	ASP	N	10.472
132	GLU	H	132	GLU	N	-2.890
133	APN	H	133	APN	N	13.834
134	HIS	H	134	HIS	N	2.788
136	GLN	H	136	GLN	N	3.522
137	ILE	H	137	ILE	N	0.682
138	GLY	H	138	GLY	N	2.878
139	VAL	H	139	VAL	N	-0.630
140	VAL	H	140	VAL	N	-0.220
141	ALA	H	141	ALA	N	-5.122
144	LEU	H	144	LEU	N	1.302
145	GLU	H	145	GLU	N	10.844
146	LEU	H	146	LEU	N	7.608
148	ARG	H	148	ARG	N	-17.148
149	VAL	H	149	VAL	N	6.244
150	THR	H	150	THR	N	-1.920
151	GLN	H	151	GLN	N	-4.052
152	GLN	H	152	GLN	N	-3.960
153	ILE	H	153	ILE	N	2.510
154	APN	H	154	APN	N	-0.164
155	ASP	H	155	ASP	N	-2.700
156	SER	H	156	SER	N	0.432
157	ARG	H	157	ARG	N	0.444
23	ASP	C	24	MET	N	-0.622
24	MET	C	25	THR	N	-0.382
25	THR	C	26	ARG	N	1.683
26	ARG	C	27	ASP	N	0.225
27	ASP	C	28	GLY	N	-0.765
28	GLY	C	29	LEU	N	1.606
29	LEU	C	30	ALA	N	0.628
30	ALA	C	31	APN	N	0.739
31	APN	C	32	LYS	N	-1.178
32	LYS	C	33	ALA	N	0.647
33	ALA	C	34	LEU	N	0.132
34	LEU	C	35	ALA	N	0.282
35	ALA	C	36	VAL	N	0.273
36	VAL	C	37	ALA	N	1.602
37	ALA	C	38	ARG	N	0.292
38	ARG	C	39	THR	N	-0.614

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
39	THR	C	40	LEU	N	1.937
40	LEU	C	41	ALA	N	0.523
41	ALA	C	42	ASP	N	1.477
42	ASP	C	43	SER	N	0.302
44	PRO	C	45	GLU	N	-1.591
45	GLU	C	46	ILE	N	-1.031
46	ILE	C	47	ARG	N	0.908
47	ARG	C	48	GLN	N	0.745
48	GLN	C	49	GLY	N	-1.598
49	GLY	C	50	LEU	N	0.169
50	LEU	C	51	GLN	N	0.533
51	GLN	C	52	LYS	N	-0.893
52	LYS	C	53	LYS	N	0.647
55	GLN	C	56	ELU	N	-0.638
56	ELU	C	57	SER	N	-1.319
57	SER	C	58	GLY	N	-0.173
59	ILE	C	60	GLN	N	1.274
60	GLN	C	61	ALA	N	0.411
61	ALA	C	62	ILE	N	-0.394
62	ILE	C	63	ALA	N	1.708
63	ALA	C	64	GLU	N	-0.252
64	GLU	C	65	ALA	N	1.128
65	ALA	C	66	VAL	N	-0.432
66	VAL	C	67	ARG	N	5.007
67	ARG	C	68	LYS	N	0.094
68	LYS	C	69	ARG	N	0.864
69	ARG	C	70	APN	N	1.093
70	APN	C	71	ASP	N	0.385
71	ASP	C	72	LEU	N	-1.628
72	LEU	C	73	LEU	N	1.268
73	LEU	C	74	PHE	N	-0.606
75	ILE	C	76	VAL	N	0.123
76	VAL	C	77	VAL	N	-1.327
77	VAL	C	78	THR	N	-0.364
78	THR	C	79	ASP	N	-0.934
79	ASP	C	80	MET	N	-2.152
80	MET	C	81	GLN	N	0.252
81	GLN	C	82	SER	N	-0.175
82	SER	C	83	LEU	N	-1.468
83	LEU	C	84	ARG	N	-0.553
84	ARG	C	85	TYR	N	1.928
85	TYR	C	86	SER	N	-1.422

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
86	SER	C	87	HIS	N	-0.992
89	GLU	C	90	ALA	N	-0.226
90	ALA	C	91	GLN	N	1.530
91	GLN	C	92	ARG	N	-0.327
92	ARG	C	93	ILE	N	1.138
93	ILE	C	94	GLY	N	-0.688
94	GLY	C	95	GLN	N	-1.783
96	PRO	C	97	PHE	N	0.400
97	PHE	C	98	LYS	N	-2.131
98	LYS	C	99	GLY	N	0.299
99	GLY	C	100	ASP	N	0.000
100	ASP	C	101	ASP	N	-0.138
101	ASP	C	102	ILE	N	-1.360
102	ILE	C	103	LEU	N	1.556
103	LEU	C	104	LYS	N	-1.478
104	LYS	C	105	ALA	N	0.042
105	ALA	C	106	LEU	N	1.219
107	APN	C	108	GLY	N	0.458
108	GLY	C	109	GLU	N	-0.550
109	GLU	C	110	GLU	N	-0.978
111	APN	C	112	VAL	N	-1.237
112	VAL	C	113	ALA	N	0.376
113	ALA	C	114	ILE	N	0.831
114	ILE	C	115	APN	N	-0.356
115	APN	C	116	ARG	N	0.576
116	ARG	C	117	GLY	N	0.686
120	ALA	C	121	GLN	N	-0.590
121	GLN	C	122	ALA	N	-0.735
122	ALA	C	123	LEU	N	1.189
124	ARG	C	125	VAL	N	1.133
125	VAL	C	126	PHE	N	-0.352
126	PHE	C	127	THR	N	-0.905
128	PRO	C	129	ILE	N	-1.748
129	ILE	C	130	TYR	N	1.793
130	TYR	C	131	ASP	N	-2.142
131	ASP	C	132	GLU	N	-0.205
132	GLU	C	133	APN	N	1.344
133	APN	C	134	HIS	N	-2.017
135	LYS	C	136	GLN	N	1.132
136	GLN	C	137	ILE	N	0.177
137	ILE	C	138	GLY	N	-2.278
138	GLY	C	139	VAL	N	-1.086

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
139	VAL	C	140	VAL	N	-0.730
140	VAL	C	141	ALA	N	0.568
143	GLY	C	144	LEU	N	-5.000
144	LEU	C	145	GLU	N	-0.540
145	GLU	C	146	LEU	N	1.181
147	SER	C	148	ARG	N	-0.060
148	ARG	C	149	VAL	N	-0.015
149	VAL	C	150	THR	N	1.012
150	THR	C	151	GLN	N	-0.742
151	GLN	C	152	GLN	N	0.264
152	GLN	C	153	ILE	N	0.571
153	ILE	C	154	APN	N	-1.519
154	APN	C	155	ASP	N	1.224
155	ASP	C	156	SER	N	0.118
156	SER	C	157	ARG	N	0.261
23	ASP	CA	23	ASP	C	-1.081
24	MET	CA	24	MET	C	0.277
25	THR	CA	25	THR	C	0.764
26	ARG	CA	26	ARG	C	2.235
27	ASP	CA	27	ASP	C	0.431
28	GLY	CA	28	GLY	C	0.090
29	LEU	CA	29	LEU	C	-1.835
30	ALA	CA	30	ALA	C	3.015
31	APN	CA	31	APN	C	-0.675
32	LYS	CA	32	LYS	C	-8.427
33	ALA	CA	33	ALA	C	-4.001
34	LEU	CA	34	LEU	C	3.442
35	ALA	CA	35	ALA	C	-2.338
36	VAL	CA	36	VAL	C	-0.189
37	ALA	CA	37	ALA	C	5.921
38	ARG	CA	38	ARG	C	3.584
40	LEU	CA	40	LEU	C	-2.709
41	ALA	CA	41	ALA	C	3.417
42	ASP	CA	42	ASP	C	-4.112
45	GLU	CA	45	GLU	C	-1.326
46	ILE	CA	46	ILE	C	5.741
47	ARG	CA	47	ARG	C	-3.139
48	GLN	CA	48	GLN	C	6.411
49	GLY	CA	49	GLY	C	4.842
50	LEU	CA	50	LEU	C	2.117
51	GLN	CA	51	GLN	C	-1.069
52	LYS	CA	52	LYS	C	2.474

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
56	ELU	CA	56	ELU	C	2.229
58	GLY	CA	58	GLY	C	-6.173
59	ILE	CA	59	ILE	C	3.399
60	GLN	CA	60	GLN	C	-4.454
61	ALA	CA	61	ALA	C	0.686
62	ILE	CA	62	ILE	C	-0.121
63	ALA	CA	63	ALA	C	-3.641
64	GLU	CA	64	GLU	C	3.267
65	ALA	CA	65	ALA	C	-1.189
66	VAL	CA	66	VAL	C	-4.916
67	ARG	CA	67	ARG	C	-4.333
68	LYS	CA	68	LYS	C	1.372
69	ARG	CA	69	ARG	C	-1.871
70	APN	CA	70	APN	C	2.475
71	ASP	CA	71	ASP	C	0.367
72	LEU	CA	72	LEU	C	0.732
73	LEU	CA	73	LEU	C	-6.924
74	PHE	CA	74	PHE	C	4.530
77	VAL	CA	77	VAL	C	2.571
79	ASP	CA	79	ASP	C	-5.637
81	GLN	CA	81	GLN	C	-2.988
82	SER	CA	82	SER	C	-0.767
83	LEU	CA	83	LEU	C	1.164
85	TYR	CA	85	TYR	C	0.023
86	SER	CA	86	SER	C	-3.298
90	ALA	CA	90	ALA	C	7.086
91	GLN	CA	91	GLN	C	-1.244
92	ARG	CA	92	ARG	C	0.217
94	GLY	CA	94	GLY	C	2.992
101	ASP	CA	101	ASP	C	-2.307
102	ILE	CA	102	ILE	C	51.908
103	LEU	CA	103	LEU	C	2.348
104	LYS	CA	104	LYS	C	-1.997
105	ALA	CA	105	ALA	C	-3.310
106	LEU	CA	106	LEU	C	-0.756
107	APN	CA	107	APN	C	3.349
108	GLY	CA	108	GLY	C	-1.693
109	GLU	CA	109	GLU	C	-2.021
110	GLU	CA	110	GLU	C	4.691
111	APN	CA	111	APN	C	-0.212
113	ALA	CA	113	ALA	C	-4.221
114	ILE	CA	114	ILE	C	-0.802

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
120	ALA	CA	120	ALA	C	1.901
122	ALA	CA	122	ALA	C	-0.460
124	ARG	CA	124	ARG	C	-3.375
125	VAL	CA	125	VAL	C	-1.010
126	PHE	CA	126	PHE	C	-1.275
130	TYR	CA	130	TYR	C	4.437
131	ASP	CA	131	ASP	C	4.063
132	GLU	CA	132	GLU	C	-0.263
133	APN	CA	133	APN	C	-1.012
135	LYS	CA	135	LYS	C	1.766
137	ILE	CA	137	ILE	C	-1.675
138	GLY	CA	138	GLY	C	-4.745
139	VAL	CA	139	VAL	C	0.584
140	VAL	CA	140	VAL	C	-5.503
141	ALA	CA	141	ALA	C	-7.929
142	ILE	CA	142	ILE	C	-1.596
143	GLY	CA	143	GLY	C	2.066
144	LEU	CA	144	LEU	C	0.871
148	ARG	CA	148	ARG	C	-0.034
150	THR	CA	150	THR	C	1.342
151	GLN	CA	151	GLN	C	0.621
152	GLN	CA	152	GLN	C	-0.630
153	ILE	CA	153	ILE	C	-1.318
154	APN	CA	154	APN	C	-0.776
155	ASP	CA	155	ASP	C	-0.395
156	SER	CA	156	SER	C	-0.362
23	ASP	CA	23	ASP	HA	-1.992
24	MET	CA	24	MET	HA	-16.375
25	THR	CA	25	THR	HA	-25.682
26	ARG	CA	26	ARG	HA	-9.467
27	ASP	CA	27	ASP	HA	-14.372
28	GLY	CA	28	GLY	HA	-6.572
31	APN	CA	31	APN	HA	-21.858
34	LEU	CA	34	LEU	HA	-10.503
35	ALA	CA	35	ALA	HA	-28.208
36	VAL	CA	36	VAL	HA	-1.588
37	ALA	CA	37	ALA	HA	-6.806
38	ARG	CA	38	ARG	HA	-13.595
40	LEU	CA	40	LEU	HA	10.515
41	ALA	CA	41	ALA	HA	-4.692
42	ASP	CA	42	ASP	HA	-23.383
45	GLU	CA	45	GLU	HA	-22.265

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
47	ARG	CA	47	ARG	HA	23.448
48	GLN	CA	48	GLN	HA	-35.860
50	LEU	CA	50	LEU	HA	20.066
51	GLN	CA	51	GLN	HA	9.372
52	LYS	CA	52	LYS	HA	-39.761
58	GLY	CA	58	GLY	HA	33.644
59	ILE	CA	59	ILE	HA	-22.077
60	GLN	CA	60	GLN	HA	-7.230
61	ALA	CA	61	ALA	HA	-14.769
62	ILE	CA	62	ILE	HA	22.705
63	ALA	CA	63	ALA	HA	4.796
64	GLU	CA	64	GLU	HA	-11.438
65	ALA	CA	65	ALA	HA	-9.486
66	VAL	CA	66	VAL	HA	28.202
67	ARG	CA	67	ARG	HA	0.779
68	LYS	CA	68	LYS	HA	-1.144
69	ARG	CA	69	ARG	HA	-15.202
70	APN	CA	70	APN	HA	7.626
71	ASP	CA	71	ASP	HA	-2.654
72	LEU	CA	72	LEU	HA	-15.433
73	LEU	CA	73	LEU	HA	-11.133
74	PHE	CA	74	PHE	HA	-35.971
77	VAL	CA	77	VAL	HA	-18.338
79	ASP	CA	79	ASP	HA	33.309
81	GLN	CA	81	GLN	HA	11.727
82	SER	CA	82	SER	HA	14.098
83	LEU	CA	83	LEU	HA	20.346
85	TYR	CA	85	TYR	HA	-11.966
86	SER	CA	86	SER	HA	-1.462
90	ALA	CA	90	ALA	HA	13.911
91	GLN	CA	91	GLN	HA	-27.534
92	ARG	CA	92	ARG	HA	31.162
94	GLY	CA	94	GLY	HA	-11.802
103	LEU	CA	103	LEU	HA	-39.398
104	LYS	CA	104	LYS	HA	-7.715
105	ALA	CA	105	ALA	HA	15.062
106	LEU	CA	106	LEU	HA	4.954
108	GLY	CA	108	GLY	HA	-4.875
109	GLU	CA	109	GLU	HA	-3.847
110	GLU	CA	110	GLU	HA	3.242
111	APN	CA	111	APN	HA	-8.163
113	ALA	CA	113	ALA	HA	-10.839

Continued on next page

Table B.3 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
114	ILE	CA	114	ILE	HA	24.258
120	ALA	CA	120	ALA	HA	30.245
122	ALA	CA	122	ALA	HA	-15.533
124	ARG	CA	124	ARG	HA	23.059
125	VAL	CA	125	VAL	HA	12.949
130	TYR	CA	130	TYR	HA	30.152
131	ASP	CA	131	ASP	HA	-3.947
132	GLU	CA	132	GLU	HA	-28.927
133	APN	CA	133	APN	HA	35.768
135	LYS	CA	135	LYS	HA	1.344
137	ILE	CA	137	ILE	HA	8.414
138	GLY	CA	138	GLY	HA	25.908
140	VAL	CA	140	VAL	HA	5.031
141	ALA	CA	141	ALA	HA	-15.067
142	ILE	CA	142	ILE	HA	-7.068
143	GLY	CA	143	GLY	HA	20.430
148	ARG	CA	148	ARG	HA	-23.074
150	THR	CA	150	THR	HA	-0.293
151	GLN	CA	151	GLN	HA	-5.697
152	GLN	CA	152	GLN	HA	-4.922
153	ILE	CA	153	ILE	HA	9.942
154	APN	CA	154	APN	HA	-0.997
155	ASP	CA	155	ASP	HA	-0.381
156	SER	CA	156	SER	HA	-3.383

Table B.4: Residual dipolar couplings measured for Ubiquitine in Otting phase

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
3	ILE	HN	3	ILE	N	2.987
4	PHE	HN	4	PHE	N	2.739
5	VAL	HN	5	VAL	N	1.794
6	LYS	HN	6	LYS	N	2.311
7	THR	HN	7	THR	N	-2.102
8	LEU	HN	8	LEU	N	1.744
10	GLY	HN	10	GLY	N	1.635
11	LYS	HN	11	LYS	N	-6.481
12	THR	HN	12	THR	N	0.310
13	ILE	HN	13	ILE	N	0.076
14	THR	HN	14	THR	N	2.885
15	LEU	HN	15	LEU	N	3.678

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
16	GLU	HN	16	GLU	N	-0.570
17	VAL	HN	17	VAL	N	-1.804
18	GLU	HN	18	GLU	N	-7.763
20	SER	HN	20	SER	N	-3.474
21	ASP	HN	21	ASP	N	-1.734
22	THR	HN	22	THR	N	5.953
23	ILE	HN	23	ILE	N	2.506
25	APN	HN	25	APN	N	1.278
26	VAL	HN	26	VAL	N	2.015
27	LYS	HN	27	LYS	N	2.420
28	ALA	HN	28	ALA	N	2.000
29	LYS	HN	29	LYS	N	2.551
30	THR	HN	30	THR	N	1.777
31	GLN	HN	31	GLN	N	2.870
32	ASP	HN	32	ASP	N	1.493
33	LYS	HN	33	LYS	N	2.092
34	GLU	HN	34	GLU	N	2.579
35	GLY	HN	35	GLY	N	-1.577
36	ILE	HN	36	ILE	N	-7.862
39	ASP	HN	39	ASP	N	4.810
40	GLN	HN	40	GLN	N	-0.011
41	GLN	HN	41	GLN	N	6.743
42	ARG	HN	42	ARG	N	3.941
43	LEU	HN	43	LEU	N	3.421
44	ILE	HN	44	ILE	N	2.577
45	PHE	HN	45	PHE	N	-1.870
46	ALA	HN	46	ALA	N	0.263
47	GLY	HN	47	GLY	N	2.258
49	GLN	HN	49	GLN	N	-1.110
50	LEU	HN	50	LEU	N	2.277
51	GLU	HN	51	GLU	N	-3.101
52	ASP	HN	52	ASP	N	-6.940
54	ARG	HN	54	ARG	N	0.358
55	THR	HN	55	THR	N	2.420
56	LEU	HN	56	LEU	N	5.443
57	SER	HN	57	SER	N	6.892
58	ASP	HN	58	ASP	N	3.056
59	TYR	HN	59	TYR	N	3.019
60	APN	HN	60	APN	N	6.967
61	ILE	HN	61	ILE	N	3.106
62	GLN	HN	62	GLN	N	4.501
63	LYS	HN	63	LYS	N	-1.799

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
64	GLU	HN	64	GLU	N	5.003
65	SER	HN	65	SER	N	6.272
67	LEU	HN	67	LEU	N	2.036
68	HIS	HN	68	HIS	N	1.821
69	LEU	HN	69	LEU	N	0.880
70	VAL	HN	70	VAL	N	5.007
71	LEU	HN	71	LEU	N	5.932
72	ARG	HN	72	ARG	N	3.477
74	ARG	HN	74	ARG	N	2.760
2	GLN	C	3	ILE	N	-1.053
3	ILE	C	4	PHE	N	0.497
4	PHE	C	5	VAL	N	-0.330
5	VAL	C	6	LYS	N	0.344
6	LYS	C	7	THR	N	0.202
7	THR	C	8	LEU	N	-0.844
9	THR	C	10	GLY	N	-0.222
10	GLY	C	11	LYS	N	0.509
11	LYS	C	12	THR	N	-0.851
12	THR	C	13	ILE	N	0.432
13	ILE	C	14	THR	N	-0.254
14	THR	C	15	LEU	N	0.208
15	LEU	C	16	GLU	N	0.788
16	GLU	C	17	VAL	N	0.182
17	VAL	C	18	GLU	N	0.345
19	PRO	C	20	SER	N	0.785
20	SER	C	21	ASP	N	0.416
21	ASP	C	22	THR	N	-0.645
22	THR	C	23	ILE	N	-0.062
24	GLU	C	25	ASN	N	-0.805
25	ASN	C	26	VAL	N	0.691
26	VAL	C	27	LYS	N	-0.968
27	LYS	C	28	ALA	N	0.683
28	ALA	C	29	LYS	N	-0.674
29	LYS	C	30	THR	N	0.115
30	THR	C	31	GLN	N	0.069
31	GLN	C	32	ASP	N	-0.533
32	ASP	C	33	LYS	N	0.510
33	LYS	C	34	PHE	N	-0.457
34	PHE	C	35	GLY	N	0.672
35	GLY	C	36	ILE	N	0.381
38	PRO	C	39	ASP	N	-0.933
39	ASP	C	40	GLN	N	0.175

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
40	GLN	C	41	GLN	N	-0.568
41	GLN	C	42	ARG	N	0.717
42	ARG	C	43	LEU	N	-0.126
43	LEU	C	44	ILE	N	0.303
44	ILE	C	45	PHE	N	0.135
45	PHE	C	46	ALA	N	0.392
46	ALA	C	47	GLY	N	-0.615
47	GLY	C	48	LYS	N	0.631
48	LYS	C	49	GLN	N	-0.760
49	GLN	C	50	LEU	N	0.235
50	LEU	C	51	GLU	N	0.466
51	GLU	C	52	ASP	N	-0.184
53	GLY	C	54	ARG	N	0.575
54	ARG	C	55	THR	N	-0.969
55	THR	C	56	LEU	N	-0.250
56	LEU	C	57	SER	N	0.214
57	SER	C	58	ASP	N	-0.724
58	ASP	C	59	TYR	N	0.271
59	TYR	C	60	APN	N	-0.539
60	APN	C	61	ILE	N	0.239
61	ILE	C	62	GLN	N	-0.045
62	GLN	C	63	LYS	N	-0.798
63	LYS	C	64	GLU	N	-0.016
64	GLU	C	65	SER	N	-0.235
66	THR	C	67	LEU	N	0.052
67	LEU	C	68	HIS	N	-0.858
68	HIS	C	69	LEU	N	0.073
69	LEU	C	70	VAL	N	-0.598
70	VAL	C	71	LEU	N	-0.428
71	LEU	C	72	ARG	N	0.343
72	ARG	C	73	LEU	N	-0.069
75	GLY	C	76	GLY	N	-0.308
1	MET	CA	1	MET	C	0.300
2	GLN	CA	2	GLN	C	0.266
3	ILE	CA	3	ILE	C	0.538
4	PHE	CA	4	PHE	C	-2.016
5	VAL	CA	5	VAL	C	1.135
6	LYS	CA	6	LYS	C	-2.363
7	THR	CA	7	THR	C	1.965
8	LEU	CA	8	LEU	C	0.728
9	THR	CA	9	THR	C	-2.232
10	GLY	CA	10	GLY	C	0.749

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
11	LYS	CA	11	LYS	C	1.065
12	THR	CA	12	THR	C	1.844
13	ILE	CA	13	ILE	C	-0.195
14	THR	CA	14	THR	C	1.641
15	LEU	CA	15	LEU	C	-0.997
16	GLU	CA	16	GLU	C	0.552
17	VAL	CA	17	VAL	C	-0.054
19	PRO	CA	19	PRO	C	-0.253
20	SER	CA	20	SER	C	-0.916
21	ASP	CA	21	ASP	C	-0.498
22	THR	CA	22	THR	C	0.571
24	GLU	CA	24	GLU	C	-0.955
25	ASN	CA	25	ASN	C	1.198
26	VAL	CA	26	VAL	C	-1.366
27	LYS	CA	27	LYS	C	1.140
28	ALA	CA	28	ALA	C	-1.350
29	LYS	CA	29	LYS	C	1.245
30	ILE	CA	30	ILE	C	-0.283
31	GLN	CA	31	GLN	C	0.647
32	ASP	CA	32	ASP	C	0.446
34	GLU	CA	34	GLU	C	-2.371
35	GLY	CA	35	GLY	C	-0.449
38	PRO	CA	38	PRO	C	-0.198
39	ASP	CA	39	ASP	C	2.411
40	GLN	CA	40	GLN	C	-0.894
41	GLN	CA	41	GLN	C	0.167
42	ARG	CA	42	ARG	C	-2.423
43	LEU	CA	43	LEU	C	1.252
44	ILE	CA	44	ILE	C	-1.881
45	PHE	CA	45	PHE	C	0.619
46	ALA	CA	46	ALA	C	-2.375
47	GLY	CA	47	GLY	C	-0.252
49	GLN	CA	49	GLN	C	1.090
50	LEU	CA	50	LEU	C	-0.571
51	GLU	CA	51	GLU	C	0.743
53	GLY	CA	53	GLY	C	-1.362
54	ARG	CA	54	ARG	C	-1.221
55	THR	CA	55	THR	C	1.073
56	LEU	CA	56	LEU	C	-0.944
58	ASP	CA	58	ASP	C	1.915
60	ASN	CA	60	ASN	C	0.514
61	ILE	CA	61	ILE	C	-1.242

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
62	GLN	CA	62	GLN	C	1.043
63	LYS	CA	63	LYS	C	1.568
64	GLU	CA	64	GLU	C	-0.037
67	LEU	CA	67	LEU	C	0.592
69	LEU	CA	69	LEU	C	-0.414
70	VAL	CA	70	VAL	C	0.242
71	LEU	CA	71	LEU	C	-1.991
72	ARG	CA	72	ARG	C	-0.276
74	ARG	CA	74	ARG	C	-1.069
2	GLN	CA	2	GLN	HA	-1.938
3	ILE	CA	3	ILE	HA	0.761
4	PHE	CA	4	PHE	HA	10.574
5	VAL	CA	5	VAL	HA	10.108
6	LYS	CA	6	LYS	HA	-0.207
7	THR	CA	7	THR	HA	2.327
8	LEU	CA	8	LEU	HA	1.712
9	THR	CA	9	THR	HA	7.423
10	GLY	CA	10	GLY	HA	17.939
11	LYS	CA	11	LYS	HA	-5.744
12	THR	CA	12	THR	HA	-10.440
13	ILE	CA	13	ILE	HA	2.252
14	THR	CA	14	THR	HA	10.185
15	LEU	CA	15	LEU	HA	9.290
16	GLU	CA	16	GLU	HA	5.156
17	VAL	CA	17	VAL	HA	8.181
19	PRO	CA	19	PRO	HA	11.370
20	SER	CA	20	SER	HA	1.758
21	ASP	CA	21	ASP	HA	17.148
25	ASN	CA	25	ASN	HA	8.169
26	VAL	CA	26	VAL	HA	-12.486
27	LYS	CA	27	LYS	HA	13.326
28	ALA	CA	28	ALA	HA	-20.446
29	LYS	CA	29	LYS	HA	10.793
30	ILE	CA	30	ILE	HA	-8.764
31	GLN	CA	31	GLN	HA	5.350
32	ASP	CA	32	ASP	HA	-0.250
34	GLU	CA	34	GLU	HA	3.817
35	GLY	CA	35	GLY	HA	24.092
38	PRO	CA	38	PRO	HA	-7.371
39	ASP	CA	39	ASP	HA	2.435
40	GLN	CA	40	GLN	HA	-9.823
41	GLN	CA	41	GLN	HA	14.735

Continued on next page

Table B.4 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
42	ARG	CA	42	ARG	HA	19.047
43	LEU	CA	43	LEU	HA	8.906
44	ILE	CA	44	ILE	HA	1.254
45	PHE	CA	45	PHE	HA	4.671
46	ALA	CA	46	ALA	HA	14.170
47	GLY	CA	47	GLY	HA	12.662
48	LYS	CA	48	LYS	HA	-11.402
49	GLN	CA	49	GLN	HA	-0.511
50	LEU	CA	50	LEU	HA	-0.798
51	GLU	CA	51	GLU	HA	-12.923
53	GLY	CA	53	GLY	HA	8.942
54	ARG	CA	54	ARG	HA	7.065
55	THR	CA	55	THR	HA	4.232
56	LEU	CA	56	LEU	HA	1.702
57	SER	CA	57	SER	HA	-1.862
58	ASP	CA	58	ASP	HA	1.291
59	TYR	CA	59	TYR	HA	3.033
60	ASN	CA	60	ASN	HA	2.074
61	ILE	CA	61	ILE	HA	20.333
62	GLN	CA	62	GLN	HA	15.636
63	LYS	CA	63	LYS	HA	0.779
64	GLU	CA	64	GLU	HA	3.271
65	SER	CA	65	SER	HA	18.062
66	THR	CA	66	THR	HA	10.475
67	LEU	CA	67	LEU	HA	6.716
69	LEU	CA	69	LEU	HA	5.977
70	VAL	CA	70	VAL	HA	11.118
71	LEU	CA	71	LEU	HA	14.458
72	ARG	CA	72	ARG	HA	13.446
73	LEU	CA	73	LEU	HA	6.784
74	ARG	CA	74	ARG	HA	7.838
75	GLY	CA	75	GLY	HA	10.529

Table B.5: Residual dipolar couplings measured for citrate bound form of CitAP in phages

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
3	ILE	H	3	ILE	N	0.584
4	THR	H	4	THR	N	-1.763
6	GLU	H	6	GLU	N	2.450
7	ARG	H	7	ARG	N	-0.618

Continued on next page

Table B.5 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
9	HIS	H	9	HIS	N	3.238
13	GLY	H	13	GLY	N	5.361
14	GLN	H	14	GLN	N	1.070
21	MET	H	21	MET	N	2.672
25	ALA	H	25	ALA	N	-0.422
26	MET	H	26	MET	N	-8.741
29	LEU	H	29	LEU	N	6.818
30	VAL	H	30	VAL	N	0.889
31	GLU	H	31	GLU	N	0.361
32	ALA	H	32	ALA	N	3.582
33	VAL	H	33	VAL	N	4.160
34	GLN	H	34	GLN	N	0.858
35	LYS	H	35	LYS	N	0.562
36	ARG	H	36	ARG	N	8.836
37	ASP	H	37	ASP	N	3.856
38	LEU	H	38	LEU	N	4.581
39	ALA	H	39	ALA	N	2.103
40	ARG	H	40	ARG	N	-0.600
41	ILE	H	41	ILE	N	2.867
42	LYS	H	42	LYS	N	5.602
43	ALA	H	43	ALA	N	0.696
45	ILE	H	45	ILE	N	3.331
56	TYR	H	56	TYR	N	7.518
59	VAL	H	59	VAL	N	3.405
60	GLY	H	60	GLY	N	4.796
61	ASP	H	61	ASP	N	-5.918
62	ALA	H	62	ALA	N	3.451
63	SER	H	63	SER	N	6.668
64	GLY	H	64	GLY	N	-5.512
65	GLN	H	65	GLN	N	-9.921
66	ARG	H	66	ARG	N	-6.680
68	TYR	H	68	TYR	N	7.935
69	HIS	H	69	HIS	N	9.256
70	VAL	H	70	VAL	N	11.106
73	ASP	H	73	ASP	N	-3.705
75	ILE	H	75	ILE	N	1.163
76	GLY	H	76	GLY	N	0.704
77	LYS	H	77	LYS	N	-1.498
78	SER	H	78	SER	N	-2.658
81	GLY	H	81	GLY	N	-1.790
83	ASP	H	83	ASP	N	2.420
85	ASP	H	85	ASP	N	-5.986

Continued on next page

Table B.5 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
86	GLU	H	86	GLU	N	-8.760
87	ALA	H	87	ALA	N	-10.412
88	LEU	H	88	LEU	N	-8.587
92	LYS	H	92	LYS	N	-1.988
93	SER	H	93	SER	N	4.620
94	TYR	H	94	TYR	N	10.087
96	SER	H	96	SER	N	6.719
109	LYS	H	109	LYS	N	11.118
110	SER	H	110	SER	N	7.454
114	ASP	H	114	ASP	N	0.392
115	ALA	H	115	ALA	N	10.445
117	GLY	H	117	GLY	N	3.891
119	VAL	H	119	VAL	N	1.904
120	ILE	H	120	ILE	N	1.010
121	GLY	H	121	GLY	N	8.121
122	ILE	H	122	ILE	N	6.566
123	VAL	H	123	VAL	N	5.976
127	TYR	H	127	TYR	N	4.012
128	THR	H	128	THR	N	5.460
129	ILE	H	129	ILE	N	-0.205
130	GLU	H	130	GLU	N	2.096
131	GLN	H	131	GLN	N	-0.308
132	LEU	H	132	LEU	N	0.361
133	GLU	H	133	GLU	N	2.116

Table B.6: Residual dipolar couplings measured for citrate free form of CitAP in phages

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
18	ILE	H	18	ILE	N	-10.411
22	GLN	H	22	GLN	N	-12.835
24	SER	H	24	SER	N	2.122
25	ALA	H	25	ALA	N	-9.721
26	MET	H	26	MET	N	-15.662
28	GLU	H	28	GLU	N	0.839
29	LEU	H	29	LEU	N	11.105
30	VAL	H	30	VAL	N	0.391
31	GLU	H	31	GLU	N	0.997
32	ALA	H	32	ALA	N	4.118
33	VAL	H	33	VAL	N	7.505
34	GLN	H	34	GLN	N	0.602

Continued on next page

Table B.6 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
35	LYS	H	35	LYS	N	2.608
37	ASP	H	37	ASP	N	11.477
38	LEU	H	38	LEU	N	6.496
39	ALA	H	39	ALA	N	-4.833
40	ARG	H	40	ARG	N	-1.974
41	ILE	H	41	ILE	N	-1.458
42	LYS	H	42	LYS	N	-2.527
43	ALA	H	43	ALA	N	0.884
44	LEU	H	44	LEU	N	-3.026
61	ASP	H	61	ASP	N	-5.673
63	SER	H	63	SER	N	15.812
112	ILE	H	112	ILE	N	-2.552
114	ASP	H	114	ASP	N	1.370
115	ALA	H	115	ALA	N	18.573
117	GLY	H	117	GLY	N	7.779
118	LYS	H	118	LYS	N	-2.091
119	VAL	H	119	VAL	N	9.551
120	ILE	H	120	ILE	N	-3.127
121	GLY	H	121	GLY	N	13.372
122	ILE	H	122	ILE	N	16.623
131	GLN	H	131	GLN	N	2.057
132	LEU	H	132	LEU	N	3.035
17	LEU	C	18	ILE	N	-0.889
19	GLN	C	20	ALA	N	-0.371
21	MET	C	22	GLN	N	-0.131
23	ILE	C	24	SER	N	0.710
24	SER	C	25	ALA	N	-0.243
25	ALA	C	26	MET	N	0.508
27	PRO	C	28	GLU	N	0.081
28	GLU	C	29	LEU	N	0.129
29	LEU	C	30	VAL	N	0.128
30	VAL	C	31	GLU	N	0.367
31	GLU	C	32	ALA	N	-1.570
32	ALA	C	33	VAL	N	1.899
33	VAL	C	34	GLN	N	-1.769
34	GLN	C	35	LYS	N	0.755
35	LYS	C	36	ARG	N	-0.275
36	ARG	C	37	ASP	N	-0.426
37	ASP	C	38	LEU	N	-1.125
38	LEU	C	39	ALA	N	1.612
39	ALA	C	40	ARG	N	-1.189
40	ARG	C	41	ILE	N	0.242

Continued on next page

Table B.6 – continued from previous page

Res: Number1	Res: name1	Atom1	Res: Number2	Res: name2	Atom2	D
41	ILE	C	42	LYS	N	0.262
42	LYS	C	43	ALA	N	-0.825
43	ALA	C	44	LEU	N	0.941
60	GLY	C	61	ASP	N	0.495
61	ASP	C	62	ALA	N	-1.803
62	ALA	C	63	SER	N	1.150
113	GLN	C	114	ASP	N	-0.477
114	ASP	C	115	ALA	N	-0.430
116	THR	C	117	GLY	N	-1.592
117	GLY	C	118	LYS	N	1.231
118	LYS	C	119	VAL	N	-1.306
119	VAL	C	120	ILE	N	0.204
120	ILE	C	121	GLY	N	-0.069
132	LEU	C	133	GLU	N	-0.170

Appendix C

Bruker Pulseprogramme and MATHEMATICA scripts

Bruker pulse-programme for CBCA(CO)NH experiment used for simultaneous measurement of $^1J_{C\alpha H\alpha}$, $^1J_{C\beta H\beta}$ and $^1J_{C\alpha C'}$ couplings:

```
;3D CBCACONH-J(CH)
;Takahisa Ikegami      15Feb2002
;Chou and Bax          JACS 2001, 123, 3844-3845
;Grzesiek and Bax      JACS 1992, 114, 6291-6293
;Markus Zweckstetter  04Jan2002
;vinesh vijayan       12Dec2002

#include "bits.mz"
#include <Avance.incl>
#include <Grad.incl>

#define CARBON
#define NITROGEN
;#define INTERLEAVE      ;undefine this for just the reference spectrum
;#define EXE1D

define loopcounter NLOOP
define loopcounter CLOOP
"NLOOP=30" ;IN = 220u max = 30
"CLOOP=60" ;IN = 40u max = 84

;All parameters for 600MHz spectrometer

;H = hydrogen frequency on water
;p1 = high power proton 90 at p11
;p30= low power proton 90 at p130 (~15db) CPD decoupling cpds1

;C1 = carbon frequency at 46 ppm, C2 = 177 ppm
;for hsec.3, Q=|w(eff)/(d(theta)/dt)|=3 BW: 20kHz (600MHz)
; p16:sp16 -> 180 hsec22k500u at 12.5kHz (700MHz)
; pL3, p3 -> 90 Cab
; p14:sp14 -> 180 Rsnob from Cab to Co 150u (700MHz)
; pL15,p15 -> 90 Co
; p13:sp13 -> 180 hsec.3 from Co to Ca
; pL6, p6 -> 180 Co

;N = nitrogen frequency (~118 ppm)
;p7 = high power nitrogen 90 ~50us at p17
;p31 = low power 90 pulse ~200u at p131 waltz16 cpd2

;***** N evolution *****
;in22=in24=1/(4sw) for N
;MAX NLOOP = 6.8m/in22
```

```

;***** CA/CB evolution *****
;in16=in18=1/(2sw) for C
;MAX LOOP = 3.4m/in18

;***** pulses *****
"p2=p1*2"
"p8=p7*2"

"17=1"

;***** delays *****
"d3=2.7m-p22"
"d9=5.4m-p26"
"d6=0.26m-p24"
"d7=0.26m-p25-10u"
"d11=50m"
"d13=p16*0.5-p1"
"d14=p7-p6*0.5"
"d15=p3*4/3.14159"
"d20=1.8m-p1*4-6u"
"d26=p7-p1"
"d27=p16*0.5-p1*2-3u"

;**Increment or Decrement delays**
#ifdef EXE1D
  "d16=3u"
#else
  "d16=in16*0.5"
#endif
"d18=3.4m-d16"

"d22=6.8m"
"d24=d22-5.4m"
"d25=d22"

;---The following must be optimized for different machine---
;"p26=2.705m          ;gp26=+50%z, gp27=-50%z (rectangular)
;"p24=0.2m           ;gp24=+50%z          (rectangular)
;"p25=0.070m        ;gp25=-50%z          (rectangular)
;***** end declaration *****

ze
d11
2  10u
   10u do:N
   10u do:H
   10u do:C1
d12
3  10u
   d12*3
4  d12
   d12*3
6  d12
   d12*3
7  d12
   d12 do:H
   d12 do:N
   d12 fq2:C1 ;->CaCb
   200u p11:H
   200u p17:N
d1
(p1 ph10):H
1.3m
(d13 p2 ph10):H (p16:sp2 ph10):C1
1m

```

```

300u UNBLKGRAD
    (p1 ph7):H
    10u
    2mp:gp1
    100u
    300u p13:C1
;*****      inept to c aliphatic      *****
    (p3 ph4):C1
    6u
    (p16:sp2 ph10):C1 (d27 p1 ph10 3u p2 ph11 3u p1 ph10):H
    6u
    d15                      ;p3*4/Pi
    d16                      ;d16 increment delay
    (p14:sp14 ph10):C1      ;sinc1.0 C' refocusing (~150u)
100u
    if "l7==1 or 4" goto 71 ;1.93m  length of effective J-modulation
    if "l7==2" goto 72      ;3.73m  ...
    if "l7==3" goto 73      ;7.22m  ...
71
    600u
    (p1 ph10 3u p2 ph11 3u p1 ph10):H
    d20
    900u
    goto 74
72
    1.5m
    (p1 ph10 3u p2 ph11 3u p1 ph10):H
    d20
    goto 74
73
    1.5m
    d20
    (p1 ph10 3u p2 ph11 3u p1 ph10):H
74
    5u
    20u p130:H
    (p16:sp2 ph17):C1
    d18                      ;d18 decrement delay
    (p14:sp14 ph10):C1      ;B.S. compensate
    5u
    20u p13:C1
    (p3 ph10):C1
;*****      inept to ca      *****
    5u
    20u cpds1:H
    (p14:sp14 ph10):C1      ;sinc compensation
    5u
    (p16:sp2 ph10):C1      ;hsec.3 compensation
    30u
    d15
    3.4m
    (p14:sp14 ph10):C1      ;sinc on C'
    5u
    (p16:sp2 ph10):C1      ;500u hsec.3 on Ca
    5u
    3m
    400u p13:C1
    (p3 ph10):C1
    5u
    20u do:H
    1.4mp:gp2
    100u
    100u fq2:C1              ; -> Co
    100u p130:H
    80u p115:C1
    20u cpds1:H
;*****      inept to carbonyl      *****
    (p15:sp11 ph8):C1
    500u
    if "l7==4" goto 76      ;measurement of C-CA coupling
75
    4m                      ;9m length of effective J-modulation
    (p13:sp13 ph10):C1

```

```

5.5m
goto 77
76 8.85m ;18m
(p13:sp13 ph10):C1
0.65m
77 25u
1m p16:C1
(d14 p6 ph10):C1 (p8 ph10):N
11m
(p13:sp3 ph10):C1
5u
20u p15:C1
(p15:sp12 ph10):C1
5u
20u do:H
1.7mp:gp22
100u
280u p130:H
20u cpds1:H
;***** inept to nitrogen *****
(p7 ph27):N
11u
d22
(p13:sp13 ph10):C1
8u
d22 p16:C1
(d14 p6 ph10):C1 (p8 ph27):N
8u
d25
(p13:sp13 ph10):C1
d24
3u do:H
if "l1==1" goto 88
8u do:H
p26:gp26 ;Rance-Kay encoding gradient (+)
d9 p11:H
(p7 ph19):N (d26 p1 ph0):H
goto 89
88 8u do:H
p26:gp27 ;Rance-Kay encoding gradient (-)
d9 p11:H
(p7 ph9):N (d26 p1 ph0):H
;----- end CT evolution on N -----
;----- Rance Kay Transfer back to H -----
89 2u
p22:gp3
d3
(d26 p1*2 ph0):H (p7*2 ph27):N
2u
p22:gp3
d3
(p1 ph15):H (p7 ph28):N
2u
p22:gp4
d3
(d26 p1*2 ph0):H (p7*2 ph27):N
2u
p22:gp4
d3
(d26 p1 ph0):H
4u
p24:gp24 ;Rance-Kay decoding gradient (+)
d6
(p1*2 ph0):H
2u
p25:gp25 ;Rance-Kay decoding gradient (-)
d7 p131:N
10u BLKGRAD

```

```

        (2u ph0)
        go=2 ph31 cpd3:N
        5u do:N
#ifdef EXE1D
        d11 wr #0
#else
        d11 wr #0 if #0 zd

#ifdef INTERLEAVE
        d12 iu7
        lo to 3 times 4
        d12 ru7
#endif

#ifdef NITROGEN
        10u iu1
        lo to 3 times 2
        d12 dd22
        10u ru1
        d12 id24
        d12 id25
        lo to 4 times NLOOP ;46 ;change back to 46
        d12 rd22
        d12 rd24
        d12 rd25
#endif

#ifdef CARBON
        d12 ip4
        lo to 6 times 2
        d12 id16
        d12 dd18
        d12 ip31
        d12 ip31
        lo to 7 times CLOOP
        d12 rd16
        d12 rd18
#endif

#endif

        d12 BLKGRAD
        d12 do:N
        d12 do:C1
        d12 do:H

exit

ph0=0
ph4=0
ph6=0
ph7=1 3
ph8=0 0 0 0 2 2 2 2
ph9=2 ; or 0
ph10=0
ph11=1
ph15=1
ph17= 0 0 1 1 2 2 3 3
ph19=0
ph27=0
ph28=1
ph31=0 2 2 0 2 0 0 2
;

```

Bruker pulse-programme for TROSY-HNCO experiment for simultaneous measurement of $^1J_{C'N}$ and $^1J_{NH}$ couplings.

```
#include "bits.mz"
```

```

#include <Avance.incl>
#include <Grad.incl>

;3D TROSY-HNCOJ
;Modified from the original pulse scheme by vivi, for including shorter NCo transfer delay for the 3rd interleaved experiment.
;3D TROSY-HNCOJ

;-----
; NOTE: C' on C1; CA on C1 using shaped pulses with offset
;-----

;--RF Pulses-----
;p1 = 90 deg (10us) 1H pulse @p11
;p2 = 90 deg @p12 1H; should be ~1ms
;p3 = 90 hard on C'
;p6 = sp0 90 sinc1.0 on C' (~100 us at 800 MHz), center lobe sinx/x
;p7 = 90 deg (~50us) 15N pulse @p17
;p8 = sp3 sinc1.0 on CA (~100 us at 800 MHz)
;p16= sp2 180 hsec.3 on C' (1000 us at 800 MHz), hyperbolic secant

;--Gradients-----
;p21=1m, gp1=5%x      sine.50      ;sine bell shaped gradient
;p20=1m, gp0=30%y    sine.50
;p20=1m, gp2=30%x    sine.50
;p21=1m, gp3=30%z    sine.50

;--Options-----
;--Options-----
#define NITRO
#define CARBON
#define INTERLEAVE

;--Evolutions-----
define loopcounter NLOOP
define loopcounter CLOOP

"NLOOP=36"                ;max = 62 for 500u increment in 15N dim
;;"CLOOP=32"              ;number of complex C' increments
"CLOOP=18"                ;number of complex C' increments

;
;15N evolution   in20=in21=in22=in24   Total evolution = 4*in20*NLOOP
;C' evolution    in0                    Total evolution = 2*in0*CLOOP

;--Delays-----
"d0=in0*0.5-p8*0.5-p6*0.54"          ;90,-180 phase correction in C' dim
"d3=2.5m-p21"
"d4=2.6m-p21-p2"
;"d5=16.6m-p2-40u"
;"d6=33.1m-p2-40u"
"d7=d3-p2-20u"
"d11=50m"
"d16=p16"
"d19=in21*0.444"                  ;fudge delay for zero ph1 in 15N dim
"d20=7.5m+p16*0.5"
"d21=7.5m"
"d22=7.5m"
"d23=d20-p16-5u"
"d24=2u"                          ;fudge delay for zero ph1 in 15N dim
"d26=p7-p1"
"d27=p16*0.5-p7"
"d28=d24*2"
"d29=p8*0.5-p7"

1      ze
      1m

```



```

2      d11 do:f3
3      3m
7      5u do:f2
      10u pl3:f2
      d1 setnmr3^0 setnmr0^34^32^33
      1m setnmr3|0 setnmr0|34|32|33
(p3 ph0):f2          ;kill C' boltzman magnetization
      10u pl1:f1
      10u pl7:f3
;----- start 90-degree on 1HN -----
      (p1 ph0):f1          ;the first 1H 90
      3u
p21:gp1          ;1m @ 5%x
      d3
      (d26 p1*2 ph0):f1 (p7*2 ph0):f3
p21:gp1          ;1m @ 5%x
      d3
      (p1 ph1):f1
      3u
      p20:gp0          ;scrambler gradient 1m @ 30%y
500u
;----- INEPT to 15N -----
(p7 ph0):f3          ;INEPT to 15N
      if "l2==3" goto 22          ; shorter transfer for the 3rd interleaved experiment.
if "l2==2" goto 20          ;interleaving expt with 34.2 and 67.4msec J(NC') dephasing
16.6m pl6:f2
(p16:sp2 ph0):f2          ;selective C' adiabatic 180
16.6m
goto 21
20 33.1m pl6:f2
      95u
(p16:sp2 ph0):f2
      5u
21 (p7*2 ph0):f3
d16          ;d16 = p16
33.2m
      goto 23
22 16.6m pl6:f2
(d27 p7*2 ph0):f3 (p16:sp2 ph0):f2          ;selective C' adiabatic 180
16.6m
23 (p7 ph1):f3
p20:gp0          ;scrambler gradient 1m @ 30%y
1m pl5:f2
(p6:sp0 ph6):f2          ;selective C' 90 (sinc1.0)
;----- Start C' evolution-----
      d0          ;start C' evolution
      (p8:sp3 ph0):f2 (d29 p7*2 ph0):f3          ;CA and 15N refocusing
      d0
;----- End C' evolution-----
(p6:sp0 ph0):f2          ;selective C' 90 (sinc1.0)
6u
p20:gp2          ;scrambler gradient 1m @ 30%x
1m
      if "l1==1" goto 30          ;select appropriate TROSY phase
      (p7 ph17):f3          ;ph17=1 2  --|
goto 31          ;          |
30 (p7 ph7):f3          ;ph7=1 0  --|
;----- Start CT 15N evolution-----
31 d20          ;decrement d20 = 7.5m
(p8:sp3 ph0):f2
d20          ;decrement d20 = 7.5m
(p7*2 ph0):f3
(p16:sp2 ph0):f2
      d21          ;increment d21
(p8:sp3 ph0):f2
      if "l2==3" goto 25          ;interleaving expt without J(NHN) dephasing
      5u

```

```

        d22                ;increment d21
        d28
        (d26 p1*2 ph9):f1 (p7*2 ph0):f3
        d28
        goto 26

25      5u                ;interleaving expt with J(NHN) dephasing
        d22
        d24
        d24
        (d26 p1*2 ph9):f1 (p7*2 ph0):f3
        d24                ;increment d24 two times
        d24

26      5u
;----- End CT 15N evolution-----
        if "l1==1" goto 60                ;select appropriate TROSY phase
(p1 ph12):f1                ;ph12 = 1  --|
10u pl2:f1
(p2 ph12):f1                ;1m water 90 pulse
goto 61                ;
60 (p1 ph2):f1                ;ph2 = 3  --|
10u pl2:f1
(p2 ph2):f1                ;1m water 90 pulse
61 5u
10u pl1:f1
p21:gp1                ;1m @ 5%x
d7
        (d26 p1*2 ph0):f1 (p7*2 ph0):f3
5u
10u pl2:f1
p21:gp1                ;1m @ 5%x
d7
(p2 ph0):f1                ;1m water 90 pulse
10u pl1:f1
        if "l1==1" goto 70                ;select appropriate TROSY phase
(p1 ph0):f1 (p7 ph4):f3                ;DOUBLE 90 ph14 = 1  --|
goto 71                ;
70 (p1 ph0):f1 (p7 ph14):f3                ;DOUBLE 90 ph4 = 3  --|
71 5u
p21:gp3                ;1m @ 30%z
d4 pl2:f1
(p2 ph15):f1
3u
5u pl1:f1
        (d26 p1*2 ph0):f1 (p7*2 ph0):f3
8u pl2:f1
(p2 ph15):f1
5u
p21:gp3                ;1m @ 30%z
d4 pl30:f2
(p7 ph0):f3
        go=2 ph31 cpd1:f2                ; CPD for C' during acquisition
        10u do:f2
        1m BLKGRAD
        d11 wr #0 if #0 zd
;
#ifdef INTERLEAVE
        10u iu2                ;increment loopcounter 12
        lo to 3 times 3
        10u ru2                ;reset loopcounter 12 to 1
#endif

#ifdef NITRO
0.1m iu1                ;increment loopcounter 11
lo to 3 times 2

```

```

10u dd20
10u id21
10u id22
10u id24
10u id24
10u ru2
10u ru1                                ;reset loopcounter l1 to 1
lo to 3 times NLOOP
10u rd21
10u rd24
10u rd22
10u rd20
#endif
;
#ifdef CARBON
10u ip6
lo to 3 times 2
10u id0
10u ip31                                ;TPPI-States in C' dim
10u ip31
lo to 3 times CLOOP
#endif

;
1m do:f2
1m do:f3
1m
1m
exit

ph0=0
ph1=1
ph2=3
ph12=1
ph4=3
ph14=1
ph15=(360)178
ph6=0 0 2 2
ph7=1 0
ph8=2
ph9=0 2
ph17=1 2
ph10=(360)0
ph11=0
ph31=1 2 3 0                            ;receiver phase

```

C.1 MATHEMATICA scripts

MATHEMATICA scripts for fitting RDCs to a structure.

```

In[1]:=
<<LinearAlgebra`MatrixManipulation`
Needs["Optimization`UnconstrainedProblems`"]

y1[a1_,a2_,a3_,a4_,a5_,x1_,x2_,x3_,x4_,
x5_]:= -21583*((a1*x1+a2*x2+a3*x3+a4*x4+
a5*x5))(*Fitting function for one structures*)

y2[w1_,a1_,a2_,a3_,a4_,a5_,x1_,x2_,x3_,x4_,x5_,x6_,x7_,x8_,x9_,
x10_]:= (-21523.11)*(((w1*(a1*x1+a2*x2+a3*x3+a4*x4+
a5*x5))+((1-(w1))*(a1*x6+a2*x7+a3*x8+a4*x9+
a5*x10)))) (*Fitting function for two structures*)

y5[w1_,w2_,a1_,a2_,a3_,a4_,a5_,x1_,x2_,x3_,x4_,x5_,x6_,x7_,x8_,x9_,x10_,x11_,

```

```

x12_,x13_,x14_,
x15_]:=(-21523.11)*(((w1*(a1*x1+a2*x2+a3*x3+a4*x4+a5*x5)+(w2*(a1*x6+
a2*x7+a3*x8+a4*x9+a5*x10))+((1-(w1+w2))*(a1*x11+a2*x12+
a3*x13+a4*x14+
a5*x15))))>(*Fitting function for three structures*)

(*****5*****)

(*an example of fitting procedure for three structures*)

erfunc5[w1_,w2_,a1_,a2_,a3_,a4_,a5_,
datum_]:=(<y5[w1,w2,a1,a2,a3,a4,a5,datum[[1]],datum[[2]],datum[[3]],
datum[[4]],datum[[5]],datum[[6]],datum[[7]],datum[[8]],datum[[9]],
datum[[10]],datum[[11]],datum[[12]],datum[[13]],datum[[14]],
datum[[15]]]-datum[[16]]^2 (*error function for three structures*)

In[30]:=
chisq5=Map[erfunc5[w1,w2,a1,a2,a3,a4,a5,#]&,data];

soln5=NMinimize[{Apply[Plus,chisq5],1>w1>0&&1>w2>0},{w1,w2,a1,a2,a3,a4,
a5]}>(*fitting function for three structures*)

Out[31]=
{534.878,{a1->-0.000344807,a2->0.000159985,a3->-0.000718622,
a4->0.000152141,a5->-0.000126468,w1->0.184418,w2->0.}}

In[32]:=
ev5abc[w1_,w2_,a1_,a2_,a3_,a4_,a5_,datum_]:=
y5[w1,w2,a1,a2,a3,a4,a5,datum[[1]],datum[[2]],datum[[3]],datum[[4]],
datum[[5]],datum[[6]],datum[[7]],datum[[8]],datum[[9]],datum[[10]],
datum[[11]],datum[[12]],datum[[13]],datum[[14]],datum[[15]]]

In[33]:=
calc5abc={Map[ev5abc[w1,w2,a1,a2,a3,a4,a5,#]&,
data]/.{a1->-0.0003448069828227771',
a2->0.00015998488798546685',a3->-0.0007186215447121178',
a4->0.0001521411659575144',a5->-0.00012646774391763385',
w1->0.18441827328210506',w2->0.'}};

R5abc=Correlation[Flatten[ex],
Flatten[calc5abc]]>(*correlation function for three structures*)

Out[34]=
0.936602

Sqrt[Apply[Plus,(Flatten[ex]-Flatten[calc5abc])^2]/64]/
Sqrt[Apply[Plus,Flatten[ex]^2]/64] (*quality factor for three structures*)

Out[35]=
0.348577

(*****5*****)

```

MATHEMATICA scripts for fitting intensity ratios to obtain $\tau(c)$

```

h=6.6261*10^-34;
Bh=800.25;
uo=4*Pi*10^-7;
Gn=-2.712*10^7;
Gh=2.6752 *10^8;
Gc=6.73 *10^7;
rhn=1.04 10^-10;
sig15n=-155*10^-6 ;(* CSA = S|_ -S|| = Szz - 1/2 (Sxx + Syy) *)
sig1h=15*10^-6;
Bo=Bh*10^6*2*Pi/Gh;
theta=17;

```


Lebenslauf

Name	Vinesh Vijayan
Geburtsdatum	16.Dezember 1977
Geburtsort	Vaikom,Kerala (Indien)
Staatsangehörigkeit	Malayalam (Indien)
Religion	Hindu
Familienstand	Verheiratet

Ausbildung

1994-1996	Gymnasium aus der Mahatma Gandhi Universität in Kottayam, Indien
1996-1999	Bachelor in Chemie an der Mahatma Gandhi Universität in Kottayam, Indien
1996-1999	Stipendium für studentische Exzellenz der staatlichen Bildungsbehörde
2000-2002	Anfertigung einer Master-Arbeit am Indien Institut für Technologie Madras in Chennai, Indien Thema der Master-Arbeit: "Selective Heteronuclear Cross-Polarization in Scalar Coupled Spin Systems"; Betreuer: Prof. Dr. N. Chandrakumar
2001	Stipendium für studentische Exzellenz der staatlichen Bildungsbehörde
2002	Master in Chemie
2002	Rat der wissenschaftlichen Industrieforschung - Nationaler Eignungstest qualifiziert
2002	Graduierte Eignungsprüfung in der Technik qualifiziert
2002-2007	Wissenschaftlicher Mitarbeiter am Institut für Biophysikalische Chemie, Abteilung NMR basierte Strukturbiologie (Prof.Griesinger) in Göttingen
2002-2007	Anfertigung der vorliegenden Doktorarbeit unter Prof. Dr. Christian Griesinger; Thema der Doktorarbeit: "NMR Structural Studies on the Periplasmic Domain of CitA and DcuS".

Göttingen, den 21.03.2007