A NMR INVESTIGATION OF STRUCTURE AND ALLOSTERY IN THE HSP70 CHAPERONE SYSTEM

by

Akash Bhattacharya

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in The University of Michigan
2010

Doctoral Committee:

Professor Erik R P Zuiderweg, Chair
Professor Hashim M Al-Hashimi
Professor Charles L Brooks III
Professor Nils G Walter
Associate Professor Mark A Saper

Karmanye Vadhikaraste Ma Phaleshu Kadachana
Ma Karma Phala Hetur Bhurmatey Sangostva Akarmani

- 2:47 SRIMAD BHAGVAD GITA

To Chinmay, my father, Bharati, my mother and Swati, my sister

ACKNOWLEDGEMENTS

If it were not for the untiring efforts of my parents, I would not be writing this. Chances are, I would be talking to you if you dialed a 1-800 number. That is why, this is for my father who believed that the most beautiful mysteries of the world could be understood by a ten year old brat if he wanted to. And so, I found myself perched with David Bowman orbiting Japetus in *Discovery*. And I was hooked for life. My mother told me probably when I was nine or nearabouts that if I cycle really fast, the cycle would shrink and that people on the sidewalk would grow older, but I wouldn't. I thought she was fibbing. She had studied languages. I was nine. My sister has always been the smart sibling. But if she hadn't been around, I would have probably a) been trampled by a herd of buffalo in 1986, b) not in grad school. To these three people, I owe a quirky, yet wonderful upbringing (have you ever had someone explain exactly where a snake's poison fangs are, or been taught Bengali epic poems while driving across the oven that Northern India turns into in July?).

I have an awesome boss. Prof. Zuiderweg has been the best possible mentor. I am extremely grateful for the training I have received under his supervision. Apart from my deep appreciation of Prof. Zuiderweg's scientific leadership, I have the highest regard for his fortitude under adverse circumstances. His strength of character and resolve stand as examples to me. While on a conference, Prof. Zuiderweg once took us on a detour so that I could spend a couple of extra hours shooting landscapes (I am an avid photographer). Working with him has been a privilege and a pleasure. The Zuiderweg lab is/has been home to several great people. I started learning wetlab work from Dr. Bertelsen, who has the most inspiring (and exhausting) work ethic possible. Dan Weaver has been a trusted friend as we have negotiated the minefields of graduate school and large protein structures together. Dr. Ahmed's stewardship in the last year has been a strong driving force behind much of our work. Ramsey Macdonald is an excellent undergraduate intern who has been working on the DnaK/DnaJ project with Dr. Ahmed and myself. Of course, hardly anything would get done if our spectrometers weren't lovingly maintained by Dr. Alex Kurochkin. Our collaborators in Prof. Jason Gestwicki's lab have helped us set up and run several biochemical assays. I would especially like to thank Lyra Chang for her help and advice.

The last four-and-something years will be special because of certain people. I have been lucky enough to have an amazing group of friends at Michigan. This is probably an incomplete list and someone might get mad at me for being left out, but here goes: Chamaree, the kids (Young Sethu and Young Visha), Rachna, Pradeep-babu, our man Nandi, our man Smith, Courtney, Samopriyo, Young Karthik, Young Krishnan, Bhadra, McDowell, Katherine, Ronnie and Shivani. Friends not at Michigan who have been here if not in person, then sometimes by airwaves include Sutirtha, Padmaparna, Jayanta, Smita, Akash (yes another one), Mili, Kaushik, Bibhuti, Gautam, Tanya, Soumita, Susmita, Jishnu, Debayan, Joydeep, Sudipta and my cousin Malini. All of these people score highly on a scale of general spiffingness.

My home department, Biophysics has been a great place work, with a warm and friendly office staff (thanks for the cookies, Sara). I have benefited from talking to Profs. Gafni, Steel and Meiners in their capacity as Department and Programme Chairs. Finally, I would like to that the members of my thesis committee for their advice and help.

Table of Contents

DEDICATION.....ii

ACKNOWLEDGEMENTS.....iii

LIST OF FIGURES.....xi

LIST OF EQUATIONS.....xvi

LIST OF TABLES.....xviii

ABSTRACT.....xix

CHAPTER 1 AN INTRODUCTION TO THE STRUCTURAL BIOPHYSICS OF Hsp70 CHAPERONE PROTEINS.....1

1.1 INTRODUCTION – PROTEIN FOLDING.....1

1.2 PROTEIN HOMEOSTASIS.....4

1.3 CHAPERONES: A FIRST LOOK.....5

1.4 THE COMPONENTS OF THE CELLULAR CHAPERONE SYSTEM.....6

1.4.1 STAGE ONE – RIBOSOME BINDING CHAPERONES.....7

1.4.2 STAGE TWO – Hsp70 CHAPERONES.....7

1.4.3 STAGE THREE – CHAPERONINS.....8

1.5 Hsp70 CHAPERONES – AN OVERVIEW.....10

1.6 STRUCTURE HIGHLIGHTS OF Hsp70.....11

1.7 STRUCTURE OF THE Hsp70 NUCLEOTIDE BINDING DOMAIN – CRYSTALLOGRAPHY STUDIES.....12

1.7.1 NUCLEOTIDE BINDING SITE.....14

1.8 STRUCTURE OF THE SUBSTRATE BINDING DOMAIN (SBD): INITIAL STRUCTURAL INVESTIGATION.....16

1.8.1 STRUCTURE OF THE SUBSTRATE BINDING DOMAIN (SBD): CRYSTALLOGRAPHY STUDIES.....17

1.8.2 THE QUESTION OF SUBSTRATE BINDING.....19

1.9 ALLOSTERY IN Hsp70s.....20

1.10 OTHER PARTICIPANTS IN THE Hsp70 CYCLE: CO-CHAPERONES.....22

1.10.1 A BRIEF LOOK AT DnaJ.....22

1.10.2 A BRIEF LOOK AT GrpE.....25

1.11 ASSEMBLING THE Hsp70 CYCLE.....28

1.12 THE QUESTIONS ADDRESSED IN THIS DISSERTATION.....32

1.12.1 HOW DOES GrpE (OR ANY NUCLEOTIDE EXCHANGE FACTOR) ACT TO EXCHANGE ADP FOR ATP IN THE NUCLEOTIDE BINDING CLEFT?.....32

1.12.2 HOW IS THE ALLOSTERIC SIGNAL COMMUNICATED FROM THE NUCLEOTIDE BINDING DOMAIN TO THE SUBSTRATE BINDING DOMAIN? 33
1.12.3 HOW DOES DnaJ COMMUNICATE WITH DnaK DURING 'LOADING' OF SUBSTRATE PEPTIDE? 33
1.13 ALLOSTERY AND MOLECULAR COMMUNICATION 33
1.14 METHODS 35
1.15 A BRIEF DISCUSSION OF THE BASIC PRINCIPLES OF NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY 35
1.16 A MORE DETAILED EXPOSITION OF THE NMR METHODS USED IN THIS PROJECT 38
1.16.1 RESIDUAL DIPOLAR COUPLINGS IN STRUCTURAL ANALYSIS 39
1.16.2 PARAMAGNETIC RELAXATION IN STRUCTURAL STUDIES 41
1.17 REFERENCES 42
CHAPTER 2 ALLOSTERY IN THE HSP70 CHAPERONES IS TRANSDUCED BY SUBDOMAIN ROTATIONS. 49
2.1 ABSTRACT 49
2.2 INTRODUCTION 50
2.2.1 CURRENT MODEL FOR THE ALLOSTERIC MECHANISM OF Hsp70s 51
2.2.2 CRYSTAL STRUCTURES OF THE NUCLEOTIDE BINDING DOMAIN DO NOT SHOW ANY CHANGE AS Hsp70 TURNS OVER NUCLEOTIDE 51
2.2.3 CHANGES IN Hsp70 NUCLEOTIDE BINDING DOMAIN CHEMICAL SHIFTS SEEN VIA SOLUTION NMR EXPERIMENTS 54
2.3 MOTIVATING THIS PROBLEM 56
2.4 A BRIEF DESCRIPTION OF THIS PROJECT 57
2.5 RESULTS 57
2.6 ASSIGNING TTh-NBD 58
2.7 USING RESIDUAL DIPOLAR COUPLINGS TO PROBE STRUCTURE 59
2.8 VISUALISING THE DIFFERENCES IN SUBDOMAIN ORIENTATIONS BETWEEN DIFFERENT NUCLEOTIDE STATES OF TTh-NBD: SANSON FLAMSTEED SINUSOIDAL PROJECTIONS 62
2.9 FROM RDCs TO STRUCTURES 65
2.10 ARE THE STRUCTURAL DIFFERENCES IN THE TWO NUCLEOTIDE STATES STATISTICALLY SIGNIFICANT? 66
2.11 DISCUSSION 68
2.11.1 A NEW PARADIGM FOR NUCLEOTIDE EXCHANGE IN THE NBD 72
2.12 SUMMARISING THE RESULTS 74
2.13 INTERPRETING STRUCTURAL DIFFERENCES BETWEEN NUCLEOTIDE STATES IN TERMS OF ALLOSTERY 75
2.13.1 INTERACTION OF THE LINKER WITH THE IA/IIA INTERFACE CLEFT 77
2.14 MATERIALS AND METHODS 78
2.14.1 SAMPLES 78

2.14.2 NMR EXPERIMENTS 78
2.14.3 DATA ANALYSIS: EXTRACTING RDCs 79
2.14.4 DATA ANALYSIS: CULLING THE RDCs 79
2.14.5 DATA ANALYSIS: RDCs TO STRUCTURE 80
2.14.6 DATA ANALYSIS: HOMOLOGY MODEL 81
2.14.7 SELF VALIDATION 82
2.15 MOLECULAR MODELS 86
2.15.1 ENSEMBLE OF RDC STRUCTURES 87
2.15.2 STATISTICS OF THE ENSEMBLE 87
2.16 SUMMARY 88
2.17 REFERENCES 91

CHAPTER 3 MEASUREMENT AND INTERPRETATION OF ^{15}N-^{1}H RESIDUAL DIPOLAR COUPLINGS IN LARGER PROTEINS 96
3.1 ABSTRACT 96
3.2 INTRODUCTION 96
3.3 ABCs OF PROTEIN STRUCTURE DETERMINATION 98
3.4 STRUCTURE DETERMINATION WHEN NOEs ARE NOT AVAILABLE 99
3.4.1 EXTRACTING COUPLINGS VIA AN IP/AP EXPERIMENT 100
3.4.2 WHEN IP/AP HSQC FAILS 103
3.4.3 USING TROSY EFFECTS TO MEASURE COUPLINGS – AN OVERVIEW 106
3.5 A BRIEF DESCRIPTION OF THIS PROJECT 106
3.6 THEORY 107
3.6.1 COMPUTATIONS OF TROSY LINE WIDTHS 107
3.6.2 PULSE SEQUENCE OVERVIEW 108
3.6.3 MATHEMATICAL ANALYSIS OF THE RDC-TROSY (κ-BLOCK) PULSE SEQUENCE USING THE STATES/TRANSITIONS FORMALISM 110
3.6.4 HOW THE κ-BLOCK RDC-TROSY EXPERIMENT SHIFTS CROSSPEAKS 117
3.6.5 EFFECT OF AMIDE PROTON SPIN FLIPS 118
3.7 MATERIALS AND METHODS 119
3.7.1 PROTOCOLS FOR DnaK Tth-NBD 119
3.7.2 PROTOCOLS FOR DnaK Tth-NBD-SBD 120
3.7.3 BACKBONE RESONANCE ASSIGNMENTS 120
3.8 RESULTS AND DISCUSSION 121
3.8.1 SPECTRA FROM RDC-TROSY κ-BLOCK EXPERIMENTS 121
3.8.2 SIMULATIONS: OPTIMUM VALUES OF κ 122
3.8.3 THE EFFECT OF AMIDE PROTON FLIPS 126
3.8.4 SIMULATIONS SUMMARY 128
***3.9* APPLICATION OF RDC-TROSY TO SUBDOMAIN ORIENTATION OF THE NBD OF DnaK *Thermus thermophilus* 129**
3.9.1 EXTRACTING RDCS FOR DnaK Tth-NBD AND DnaK Tth-NBD-SBD 129

3.9.2 PROCEDURE USED TO DETERMINE A SUBSET OF EXPERIMENTAL RDCS OF MAXIMAL MUTUAL CONSISTENCY 130
3.9.3 STATISTICS OF THE COMPUTATION 133
3.9.4 SUBDOMAIN ORIENTATIONS 134
3.10 STATISTICAL VALIDATION 140
3.11 STRUCTURAL INSIGHTS FROM RDC ANALYSIS 142
3.12 3D STRUCTURAL INFORMATION – SOLUTION STRUCTURES OF Tth-NBD AND Tth-NBD-SBD 143
3.13 CONCLUSIONS 147
3.14 REFERENCES 148

CHAPTER 4 UNDERSTANDING THE Hsp70 CYCLE - INTERACTION OF DnaK AND DnaJ 152
4.1 ABSTRACT 152
4.2 INTRODUCTION 153
4.2.1 RECAPITULATING THE Hsp70 FUNCTIONAL CYCLE 153
4.2.2 THE BIOCHEMICAL ROLE OF J-PROTEINS 153
4.2.3 DIFFERENT FAMILIES OF J-PROTEINS 154
4.2.4 BIOPHYSICAL ASSAYS OF J-FUNCTION 156
4.2.5 STRUCTURAL STUDIES ON J-PROTEINS 156
4.2.6 BIOLOGICAL ROLES OF J-PROTEINS 158
4.2.7 STUDYING THE J-DOMAIN ALONE VS. STUDYING FULL LENGTH DnaJ 159
4.2.8 AUXILIN AND THE C-TERMINAL J-DOMAIN 160
4.2.9 SEARCHING FOR CORRELATIONS BASED ON J-DOMAIN POSITION: TOPOLOGY-FUNCTION RELATIONSHIPS 162
4.2.10 PREVIOUS EFFORTS TO STUDY THE STRUCTURE OF Hsp70-J-DOMAIN COMPLEXES: THE HPD LOOP 163
4.2.11 COMPLEMENTARY MUTATIONS IN DnaJ AND DnaK 163
4.3 CRYSTAL STRUCTURE OF A DISULPHIDE LINKED ADDUCT BETWEEN BOVINE Hsc70 NBD AND THE AUXILIN J-DOMAIN 165
4.3.1 CRITICISM OF THE DISULPHIDE LINKED Hsc70 NBD – AUXILIN J DOMAIN STRUCTURE 166
4.4 A BRIEF DESCRIPTION OF THIS PROJECT 168
4.5 THEORY: PARAMAGNETIC LABELING AS A TOOL IN STRUCTURAL BIOLOGY 169
4.6 MATERIALS AND METHODS 171
4.6.1 PREPARATION OF DnaK 171
4.6.2 PREPARATION OF DnaJ 172
4.6.3 RESIDUES CHOSEN FOR CYSTEINE MUTATION 173
4.6.4 MTSL-TAGGING OF DnaJ 173
4.6.5 NMR EXPERIMENTS 174
4.6.6 BLANK TESTS WITH SPIN LABEL AND DnaK 176

4.7 RESULTS **176**
4.8 ASSEMBLING THE COMPLEX STRUCTURE **177**
4.9 ATP HYDROLYSIS ASSAYS **178**
4.10 DISCUSSION **184**
4.11 A POSSIBLE MODEL FOR ALLOSTERY IN THE Hsp70 SYSTEM **186**
4.12 LIMITATIONS OF THIS ANALYSIS **188**
4.13 REFERENCES **189**

CHAPTER 5 EPILOGUE AND FUTURE OUTLOOK **194**
5.1 ABSTRACT **194**
5.2 ALLOSTERY IN THE Hsp70 SYSTEM **194**
5.2.1 INITIAL WORK ON STRUCTURES 194
5.2.2 MOVING STRUCTURES AND ALLOSTERY 195
5.2.3 DETAILED WORK ON ALLOSTERY 196
5.2.4 WORK ON THE LINKER 196
5.2.5 THE FIRST FULL LENGTH DnaK STRUCTURE 197
5.2.6 MORE DETAILED WORK ON ALLOSTERY IN THE NUCLEOTIDE BINDING DOMAIN 197
5.3 THE ATP STATE OF Hsp70 ..... **198**
5.4 SOLUTION NMR STUDIES ON THE ATP STATE OF Hsp70 ..... **199**
5.5 FURTHER STUDIES ON DnaK-DnaJ INTERACTION **200**
5.5.1 DOUBLE MUTANT STUDIES 200
5.5.2 TRUNCATED DnaJ STUDIES 201
5.5.3 FULL LENGTH DnaJ STUDIES 201
5.6 STUDYING INTERFACES BETWEEN LARGE PROTEINS: CHEMICAL SHIFT PERTURBATION NMR **202**
5.7 STUDYING INTERFACES BETWEEN LARGE PROTEINS: HYDROGEN-DEUTERIUM EXCHANGE NMR **203**
5.8 STUDYING INTERFACES BETWEEN LARGE PROTEINS: CROSS SATURATION NMR 204
5.9 SUMMARIZING PROTEIN-PROTEIN INTERFACE ANALYSIS **204**
5.10 A FEW COMMENTS ON CURRENT TRENDS IN STRUCTURAL BIOLOGY **205**
5.11 AFTERWORD **207**
5.12 REFERENCES **208**

List of Figures

Figure 1.1: The protein folding funnel (Ref: [4; 5]) 2
Figure 1.2: The many possible fates of a protein Ref:[18] 4
Figure 1.3: The three stage chaperone machinery deployed by the cell 6
Figure 1.4: Schematic of the GroEL operational cycle 10
Figure 1.5: DnaK sequence schematic 11
Figure 1.6: DnaK structure schematic 11
Figure 1.7: The nucleotide-binding domain of Hsp70 13
Figure 1.8: The topology of the Hsp70 nucleotide-binding domain 13
Figure 1.9: Residues identified in the active site of the Hsp70 NBD 15
Figure 1.10: NMR structure of Hsc70 (rat) SBD 15
Figure 1.11: Crystal structure of the SBD of E.coli DnaK 18
Figure 1.12: Ribbon diagram of the DnaK Substrate Binding Domain showing the Substrate-peptide (NRLLLTG) binding channel {Zhu, 1996 #1523} 18
Figure 1.13: ATP binding driven conformation change leads to substrate release in Hsp70. Ref: [56] 20
Figure 1.14 E.coli DnaJ sequence schematic 23
Figure 1.15: NMR solution structure of the J-domain from E.coli DnaJ [79; 80] 23
Figure 1.16: Sequence schematic for bovine Auxilin 25
Figure 1.17: Crystal structure of the Auxilin J-Domain [88] 25
Figure 1.18: Crystal structure of DnaK NBD in complex with GrpE [90] 26
Figure 1.19: Hydrogen bond between Arg183 of the proximal GrpE β-sheet and Glu28 of the DnaK NBD 27
Figure 1.20: Overlay of the DnaK NBD from the crystal structure in complex with GrpE with the bovine Hsc70 NBD (with ADP.Pi) crystal structure (shown here in light salmon colour) {Harrison, 1997 #10372} 27
Figure 1.21: Schematic of the Hsp70 operational cycle 30
Figure 1.22: NMR solution structure of full length DnaK in the ADP.Pi state [92] 30
Figure 1.23: Crystal structure of Hsp110 (yeast) solved by Hendrickson and co-workers. This is a candidate model for the ATP state of Hsp70 [93]. 32
Figure 1.24: Hsp70 operational cycle with the structures investigated so far 32
Figure 1.25: NMR fundamentals 37
Figure 1.26: Two spin system 40
Figure 2.1: Collision model of the ADP.Pi state of DnaK based on RDC solution NMR spectroscopy 51

Figure 2.2: Stereograph of crystal structures of different nucleotide states of the Hsp70 NBD 52
Figure 2.3: Different crystal structures used to evaluate changes in the NBD 53
Figure 2.4: Backbone alignment RMSD (Angstrom) of NBD residues from different crystal structures 54
Figure 2.5: Chemical shift differences between different nucleotide states of Bovine Hsc70 NBD in solution 55
Figure 2.6: Chemical shift differences between different nucleotide states of Thermus thermophilus DnaK NBD in solution 55
Figure 2.7: ^{15}N-^{1}H TROSY spectrum of TTh-NBD 58
Figure 2.8: Chemical shift differences ΔδNH between the ADP.Pi and the AMPPNP states of Tth-NBD (in phage) 59
Figure 2.9: RDC-TROSY spectrum of TTh-NBD used to extract RDCs 60
Figure 2.10: Sanson Flamsteed projection of subdomain orientation for TTh-NBD in the AMPPNP state 64
Figure 2.11: Sanson Flamsteed projection for subdomain orientation of TTh-ND in the ADP.Pi state 64
Figure 2.12: RDC derived eigenstructures of TTh-NBD in the ADP.Pi and AMPPNP bound states 66
Figure 2.13: Comparing the average ADP.Pi structure of TTh-NBD with the ensemble of AMPPNP structures 67
Figure 2.14: Comparing the average of the AMPPNP structure of TTh-NBD with the ensemble of ADP.Pi structures 67
Figure 2.15: Overlay of the TTh-NBD AMPPNP (blue) eigenstructure with the crystal structure of bovine Hsc70 NBD in ADP.Pi state (3HSC.pdb, red). 71
Figure 2.16: Overlay of the TTh-NBD ADP.Pi eigenstructure (red) with the crystal structure of bovine Hsc70 NBD in complex with ADP.Pi (3HSC.pdb, red). 71
Figure 2.17: Overlay of the bovine Hsc70 NBD in the ADP.Pi state (3HSC.pdb, green) and in complex with BAG-1 (1HX1.pdb, brown) 71
Figure 2.18: Overlay of Hsc70 NBD in complex with BAG-1(brown) with Hsp110 NBD, ATP state (yellow) 73
Figure 2.19: The IA/IIA interface of TTh-NBD in the ADP.Pi state 75
Figure 2.20: The IA/IIA interface of TTh-NBD in the AMPPNP state 76
Figure 2.21: RDC derived eigenstructures of Tth-NBD in ADP.Pi (red) and AMPPNP (blue) states, calculated with a 'culled' set of RDCs 80
Figure 2.22: Reference model for Tth-NBD (purple), overlaid on bovine Hsc70 NBD (3HSC.pdb, shown in green) 82
Figure 2.23: Testing the self-validation procedure 83

Figure 2.24: Sanson Flamsteed sinusoidal projection of subdomain orientation for Tth-NBD in the AMPPNP bound state with 60% data retained. 85
Figure 2.25: Sanson Flamsteed sinusoidal projection of subdomain orientation for Tth-NBD in the ADP.Pi bound state with 60% data retained. 86
Figure 2.26: Conventional model of the exchange factor mechanism of Hsp70: Induced Fit 89
Figure 2.27: Our model of the exchange factor mechanism: Conformation Capture. 89
Figure 2.28: Proposed allosteric mechanism of Hsp70s 91
Figure 3.1: Two spin system 98
Figure 3.2: Fingerprinting a protein: each NH group corresponds to a crosspeak 99
Figure 3.3: HSQC pulse sequence 100
Figure 3.4: One-dimensional projection of cross peaks in a decoupled and coupled HSQC experiment 101
Figure 3.5: IP/AP sum and difference spectra leading to alpha and beta peaks 102
Figure 3.6: Field dependence of the ^{15}N-TROSY effect 102
Figure 3.7: Undecoupled HSQC on a large protein 103
Figure 3.8: Peak height drops as molecular weight increases 104
Figure 3.9: Quartet of peaks in an undecoupled HSQC seen for a large protein 105
Figure 3.10: IPAP spectra run on Tth-NBD 105
Figure 3.11: Comparison of decoupled HSQC and TROSY spectra for Tth-NBD 106
Figure 3.12: the RDC-TROSY (kappa block) pulse sequence 109
Figure 3.13: Energy level diagram for an NH two-spin system 110
Figure 3.14: Detailed view of the RDC-kappa block 112
Figure 3.15: Shifting of crosspeaks in a RDC-TROSY experiment 117
Figure 3.16: RDC-Kappa block spectra for Tth-NBD 122
Figure 3.17: Detailed view of RDC-Kappa block spectra for Tth-NBD 122
Figure 3.18: ^{1}H TROSY effect simulated for a protein tumbling with a correlation time of ~ 22 nsec 125
Figure 3.19: Simulation of RDC-TROSY vs. HSQC peakshapes 126
Figure 3.20: ^{15}N lineshapes for an amide proton flip rate of 3/sec 127
Figure 3.21: ^{15}N lineshapes for an amide proton flip rate of 10/sec 127
Figure 3.22: ^{15}N lineshapes for an amide proton flip rate of 30/sec 127
Figure 3.23: ^{15}N lineshapes for an amide proton flip rate of 100/sec 127
Figure 3.24: Shift of the TROSY line with increasing kappa in an RDC-TROSY experiment 128
Figure 3.25: Shifting spectra to eliminate J-coupling and extract RDCs 130
Figure 3.26: Consistency check plot of experimental RDCs before the first filter 131
Figure 3.27: Consistency check plot of experimental RDCs after the first filter 132

Figure 3.28: Consistency check plot of experimental RDCs after the second filter .132
Figure 3.29: Sanson Flamsteed sinusoidal projection for the AMPPNP bound state of Tth-NBD 138
Figure 3.30: Sanson Flamsteed sinusoidal projection for the ADP.Pi bound state of Tth-NBD 139
Figure 3.31: Sanson Flamsteed sinusoidal projection for the apo state of Tth-NBD 139
Figure 3.32: Sanson Flamsteed sinusoidal projection for the ADP.Pi bound state of Tth-NBD-SBD 140
Figure 3.33: Sanson Flamsteed projection for 40% RDCs rejected solutions for Tth-NBD in the AMPPNP bound state.....141
Figure 3.34: Sanson Flamsteed sinusoidal projection for 40% RDCs rejected solution for Tth-NBD in the ADP.Pi bound state 141
Figure 3.35: Sanson Flamsteed sinusoidal projection of 40% RDCs rejected solution for Tth-NBD in the apo form 141
Figure 3.36: Sanson Flamsteed sinusoidal projection of 40% RDCs rejected solution for Tth-NBD-SBD in the ADP.Pi bound state 142
Figure 3.37: Overlay of Tth-NBD RDC eigenstructures (standard view) 145
Figure 3.38: Overlay of Tth-NBD RDC eigenstructures (IA/IIA interface view) 145
Figure 3.39: Overlay of RDC eigenstructures of nucleotide bound forms of Tth-NBD and Tth-NBD-SBD: standard view 146
Figure 3.40: Overlay of RDC eigenstructures of nucleotide bound forms of Tth-NBD and Tth-NBD-SBD: bottom view of the IA/IIA interface.....147
Figure 4.1: Hsp70 functional cycle 153
Figure 4.2: Sequence schematic for J-proteins 154
Figure 4.3: Representative list of J-proteins 155
Figure 4.4: Sequence alignment comparison of different J-proteins 155
Figure 4.5: J-domain NMR structure solved by Wuthrich and co-workers 157
Figure 4.6: Crystal structure of S.cerevisaie DnaJ homolog J-domain 157
Figure 4.7: Solution structure of the Cys rich region from DnaJ 158
Figure 4.8: Clathrin triskelia and lattice assembly.....161
Figure 4.9: Auxilin 1 and Auxilin 2 sequence schematic 161
Figure 4.10: Cryo-electron microscopy derived low resolution structure of Clathrin heavy chain lattice with the Auxilin J-domain[236] 162
Figure 4.11: Residues on DnaK, which affect binding to DnaJ 165
Figure 4.12: Residues on the J-domain critical for DnaK binding 165
Figure 4.13: Disulphide linked 'complex' of Hsc70 NBD with the Auxilin J – domain 166

Figure 4.14: Disulphide linked structure of bovine Hsc70 NBD and the Auxilin J-domain showing interaction interfaces 168
Figure 4.15: MTSL chemical structure 171
Figure 4.16: DnaJ J-domain showing residues chosen for Cysteine mutations 173
Figure 4.17: The effect of MTSL on neighbouring residues 175
Figure 4.18: Sample spectra (peak A41) for DnaK ADP.Pi complexed with DnaJ K41C-MTSL 176
Figure 4.19: Sample spectra (peak F67) for DnaK ADP.Pi complexed with DnaJ K41C-MTSL 177
Figure 4.20: Model for the complex interaction of DnaK.ADP.Pi with DnaJ J-domain 177
Figure 4.21: Malachite green absorbance calibration curve with potassium phosphate 179
Figure 4.22: Stimulation of DnaK638WT (full length) by DnaJ (WT-FL), modulated by peptide 180
Figure 4.23: Stimulation of ATP hydrolysis rates of DnaK638 (WT-FL) by DnaJ (WT-FL), modulated by peptide 181
Figure 4.24: Stimulation of DnaK638(WT-FL) by untagged J-domain (1-108) mutants modulated by peptide 182
Figure 4.25: Stimulation of DnaK638(Wt-FL) by MTSL tagged J-domain (1-108) mutants modulated by peptide 182
Figure 4.26: Stimulation of ATP hydrolysis rates of DnaK (WT-FL) by J-domain (1-108) mutants with and without MTSL tags in the absence of peptide 183
Figure 4.27: Stimulation of ATP hydrolysis rates of DnaK (WT-FL) by J-domain (1-108) mutants with and without MTSL tags in the presence of peptide 183
Figure 4.28: Annotated model of the disulphide linked 'complex' between Hsc70 NBD and the Auxilin J domain 184
Figure 4.29: DnaK - DnaJ complex interaction model based on paramagnetic labeling NMR 185
Figure 4.30: Peptide backbone overlay of NMR based interaction complex of DnaK-J-domain and the X-ray structure of the disulphide linked adduct of Hsc70 NBD the Auxilin J-domain 186
Figure 4.31: The role of the J-domain in allosteric behaviour in the Hsp70 system 187
Figure 5.1: DnaK in the ADP.Pi state showing Cys mutation points 200

List of Equations

Equation 1.1: Energy stored in a magnetic dipole 36
Equation 1.2: Magnetic dipole energy to frequency of on resonance irradiation 36
Equation 1.3: Boltzmann distribution and spin temperature 36
Equation 1.4: Dipolar coupling Hamiltonian 39
Equation 1.5: Dipolar coupling Hamiltonian under the weak coupling approximation 40
Equation 1.6: The dipolar coupling term 40
Equation 2.1: Jacknifing choices 84
Equation 2.2: T-statistic 87
Equation 2.3: Squared standard deviation 88
Equation 3.1: The NMR Hamiltonian for a two spin system 96
Equation 3.2: The chemical shift Hamiltonian 97
Equation 3.3: The scalar coupling Hamiltonian 97
Equation 3.4: Dipolar Coupling Hamiltonian 97
Equation 3.5: Two spin reduced coupling Hamiltonian 98
Equation 3.6: Spatial averaging of dipolar coupling to zero 98
Equation 3.7: Transverse relaxation due to dipole-dipole coupling 107
Equation 3.8: Transverse relaxation due to ^{15}N CSA 108
Equation 3.9: CSA/dipole-dipole cross-correlated relaxation rate 108
Equation 3.10: Definitions 108
Equation 3.11: The initial INEPT transfer 110
Equation 3.12: Rewriting the output of the initial INEPT block 111
Equation 3.13: Expanding from Cartesian to Ladder operators 111
Equation 3.14: Rewriting the INEPT block output in terms of ^{15}N coherences 111
Equation 3.15: Density matrix elements at the end of the first INEPT block 111
Equation 3.16: Transition rules for single transition operators 113
Equation 3.17: Density matrix elements halfway through the RDC-kappa block 114
Equation 3.18: Interconversion of density matrix elements due to the 180(H) pulse 114
Equation 3.19: Interconversion of density matrix elements due to the 180(N) pulse along x 115
Equation 3.20: Density matrix elements at the end of the ^{15}N chemical shift-labeling period (t_1) 115
Equation 3.21: Density matrix elements after simplification 116
Equation 3.22: Density matrix element, which is extracted by the ST2-PT element and read out 116
Equation 3.23: Reduction to a standard TROSY signal 116

Equation 3.24: Kappa shifted TROSY signal 116
Equation 3.25: Evolution of ^{15}N doublet components due to amide proton spin flips 118
Equation 3.26: Evaulating the effect of the (^{15}N) 180 degree pulse halfway through the RDC-kappa block 119
Equation 3.27: Evaluating the effect of the (^{1}H) 180 degree pulse halfway through the RDC-kappa block 119
Equation 3.28: Effective ^{15}N transverse relaxation rate 123
Equation 4.1: Paramagnetic Relaxation Enhancement of linewidth 170
Equation 4.2: Electron-proton correlation time 170

List of Tables

Table 2.1: Solution statistics for RDC analysis of TTh-NBD subdomains....62
Table 2.2 Computed and experimental radii of gyration of different NBD constructs from different species [163] 70
Table 2.3: Tabulation of data used to verify the self validation procedure....84
Table 3.1: ^{15}N H-alpha effective linewidths showing the limitations of RDC-TROSY vs HSQC experiments....123
Table 3.2: ^{15}N H-alpha effective intensities (peak heights) for RDC-TROSY vs. HSQC experiments 124
Table 3.3: Statistics of RDC computing for TTh-NBD....137
Table 4.1: Different families of J-proteins....154
Table 4.2: SPR results on interaction of various DnaK and DnaJ mutants....164
Table 4.3: SPR results on binding of DnaJ to various fragments of DnaK....164
Table 4.4: ATPase rates of various bovine Hsc70 fragments and stimulation by the Auxilin J-domain....167
Table 4.5: ATPase activity of different Hsc70 fragments with R171C mutation and stimulation by the Auxilin J-domain disulphide linkage....167
Table 4.6: Residues on DnaJ J-domain chosen for Cys mutations 173
Table 4.7: Peak behaviour in NMR experiments....175
Table 4.8: Peaks showing expected signature upon DnaJ-mtsl addition 176
Table 4.9: Equivalent residues on bovine Hsc70 NBD with Auxilin J domain compared to E.coli DnaK and J-domain....184
Table 5.1: Suggested further experiments for DnaK-DnaJ interaction studies 201

ABSTRACT

The Hsp70 Chaperone system is a complex molecular machine, which is responsible for assisting in protein folding/refolding/processing and transport. Hsp70 proteins have been historically associated with the 'stress' or 'heatshock' and response. They function in association with several co-chaperones in a delicately balanced and regulated two stage functional cycle. This system and its component molecules has been the subject of intensive efforts in structural biophysics for the last two decades. In this dissertation we present, for the first time an analysis of the allosteric machinery which operates in this molecular system. Using residual dipolar coupling analysis, a state-of-the-art method in solution nuclear magnetic resonance (NMR) spectroscopy, we have been able to see distinct changes in the *T.th*-DnaK (the thermophilic Hsp70 molecule) as the protein switches between the different stages of its 'two-stroke' functional cycle. Such changes have been observed for the first time in this field. We have refined new methods to detect residual dipolar couplings and scrutinized our results with the most stringent possible statistical analysis.

In the second stage of this dissertation, we have studied the interaction of *E.coli* DnaK (the bacterial Hsp70 molecule) with DnaJ (the equivalent bacterial ch-chaperone) using paramagnetic relaxation enhancement in solution NMR. Based on this analysis, we have come up with a model for the complex interaction of DnaK and DnaJ. This study sheds new light on how the DnaJ co-chaperone system structurally interacts with the primary Hsp70 system and also provides a possible mechanical model for further elucidating how the allosteric machinery operates in the complex.

CHAPTER 1 AN INTRODUCTION TO THE STRUCTURAL BIOPHYSICS OF Hsp70 CHAPERONE PROTEINS

1.1 INTRODUCTION – PROTEIN FOLDING

Proteins are complex molecular machines, which participate in almost all biochemical processes (including, but not restricted to catalysis, regulation, binding and transport of client molecules and molecular signaling). In order for them to carry out specific functions, proteins operate as distinct structural units. It has been a cornerstone of biochemistry that proteins are required to exist in unique three-dimensional structures in order to function correctly. Unsurprisingly, this unique three-dimensional structure is referred to as the 'native' conformation. This naturally leads to the question that if proteins all start out as a linear sequence of amino acids assembled at the ribosome, how do they fold to their native conformations? It had long been believed that the native structure of a protein was determined uniquely by its amino acid sequence [1]. This precept has been enshrined as 'Anfinsen's Dogma'. Subsequent work has led to refinement of this idea in the language of statistical mechanics, leading to the concept of the protein folding 'funnel' [2]. It is now understood that proteins can 'sample' multiple conformations, which correspond to microstates on a phase space diagram [3]. A highly mobile 'floppy' polypeptide chain, or in other words an unfolded protein, can sample many conformational microstates. Thus, it is located on the periphery, or rim of a folding funnel. A correctly folded 'native' protein, on the other hand does not sample many microstates, and is hence found at the very bottom of the folding funnel. In this picture, it is now possible to view

'misfolded', or damaged proteins as situated in the many 'free energy traps' in the folding funnel. The process of protein folding can now be viewed as the transition from the periphery/rim of the funnel (unfolded state) to the bottom (natively folded state).

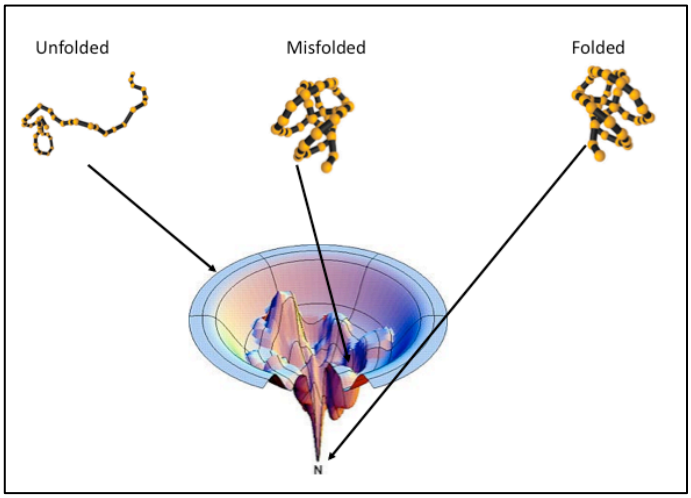

Figure 1.1: The protein folding funnel (Ref: [4; 5])

There are, of course many caveats to this framework: natively unfolded proteins constitute one [6]. Furthermore, the idea that the native structure of a protein, when defined, is a rigid three-dimensional entity has expanded to include the fact that proteins are dynamic entities [7; 8; 9; 10]. Indeed, the current viewpoint is that proteins are required to be somewhat flexible in order to execute their biological functions [11]. However, regardless of the exact nature of a native protein, the question still remains, how is such the native state reached?

The mechanism by which proteins fold after being translated inside the cell, has been, and remains under investigation [12; 13; 14; 15]. Small proteins can fold rapidly through the formation of a hydrophobic core, formation of secondary structure elements and then stabilization of structure by tertiary interactions [16]. Nucleation is centered on small folding units or 'foldons', such as two-helical bundles [17; 18]. The folding pathway includes intermediates such as 'molten globules', which are compact and have defined secondary and tertiary structure elements but which are not yet

completely packed. These intermediates may exist in equilibrium with the denatured/unfolded form of the protein and are eventually converted to the stable, native form. This conversion, is, however, not spontaneous. There are two broad classes of biomolecules, which come into play at this juncture [16][15][14]. Firstly, several enzymes catalyze isomerization steps, which enable molten globule like intermediates to convert to the native state. Protein disulphide isomerase (PDI), for instance catalyzes thiol/disulphide interchanges, thus allowing for rapid cycling of disulphide bonds [16]. In a statistical sense, this is a way for the folding protein to perform a natural 'grid search' over a larger phase space in order to find its native state/free energy minimum. Secondly, 'chaperones' are a class of proteins, which stabilize unfolded or partially folded intermediate states by preventing 'inappropriate' intra and inter molecular interactions. Lattice-based simulations of folding processes lend evidence to support this mechanism [12; 19].

However, large multidomain proteins do not follow such a relatively simple path to the folded, native state. This is because of the formation of partially folded and misfolded intermediates, which are prone to aggregation [20]. Such intermediates can result from 'inappropriate' interactions between residues that are separated in sequence and in the native tertiary structure. Such 'inappropriate' interactions would typically be hydrophobic interactions between residues that would normally be buried in the native form of a protein. Misfolding can thus be viewed as nonnative interactions between foldons, which would then drive the protein into a kinetic trap [18]. Intermediates thus formed can persist long enough to interfere in the normal folding process. Such misfolded intermediates may aggregate into amorphous structures and sometimes ordered amyloid β-fibrils [21]. The crowded environment of the cytosol, combined with 'excluded volume' effects of large molecules only exacerbates the propensity of proteins to misfold, and thence aggregate.

1.2 PROTEIN HOMEOSTASIS

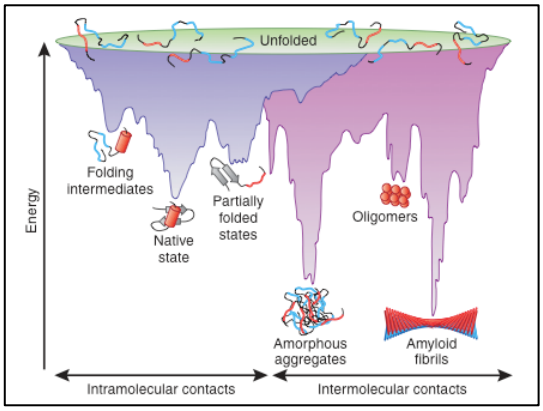

Figure 1.2: The many possible fates of a protein Ref:[18]

The ideas discussed so far lead us inexorably to one conclusion – that protein folding is a complex process which can potentially fail at several points, hence the cell must have evolved a mechanism to prevent such failures. Part of this mechanism has already been touched upon, such as enzyme-catalyzed conversion of folding intermediates to native proteins, and chaperones which stabilize intermediates. But even thinking on very general lines, this is merely part of the problem. Just as nascent polypeptide chains are vulnerable to misfolding under adverse environmental conditions, native proteins can also partially and/or transiently unfold to form 'off path' intermediates. These intermediates will then also have to be processed by the cell. Following in the theme of this discussion, as the connection between structure and function for proteins is well established (albeit a much more complex relationship than was earlier envisioned); the question arises, 'what is the homeostatic mechanism by which proteins are formed and maintained in their functional or native conformations'. This brings us to the topic of cellular chaperones, which is the subject of this work.

Maintenance of protein structure homeostasis is a complex problem, which calls for multiple solutions. If external stress factors induce unfolding of native proteins, the cell has several mechanisms to activate – to begin with, down-regulation of protein translation, followed by deployment of the chaperone machinery, followed, in turn, by targeting damaged proteins for degradation in the endoplasmic reticulum (ERAD). If the

stress conditions persist, thus overwhelming the homeostatic machinery, the result might be apoptosis [21]. Our focus is on the chaperone machinery. The ERAD mechanism can be further subdivided into classes, based on the misfolded domains/motifs that are identified. Although all ERAD processes finally converge on the ubiquitination-proteasome degradation pathway, it is important that a subclass of chaperone proteins is important in targeting misfolded cytosolic domains of proteins for ER degradation. This is distinct from the 'chaperone' functions of these proteins [21].

1.3 CHAPERONES: A FIRST LOOK

Chaperones prevent protein aggregation, both during translation of proteins as well as when proteins are partially unfolded during stress conditions. During translation, yet-unfolded intermediates can aggregate easily. The exit channel of the large ribosome subunit (50S) is about 20 Angstrom wide. This does not allow the egress of ready-made foldons apart from alpha helices [17; 22]. Furthermore, translation of a 300-residue protein takes of the order of a minute. Hence, nascent polypeptide chains must be stabilized by chaperones until a folding-competent domain has emerged from the ribosome [22; 23]. In *Eubacteria*, a 50 kDa Trigger Factor chaperone acts to 'provide a cradle' for the proper folding of nascent proteins [23; 24; 25]. Trigger Factor (TF) is the first of the three pronged chaperone machinery deployed by *Eubacteria*. The second 'prong' is the Hsp70 (DnaK) system, and the third 'prong' is the GroEL system. Mutational studies have shown that deletion of the Trigger Factor can be compensated for by the presence of DnaK and vice versa, but co-deletion of both proteins is lethal to bacteria at 37 C. However, an overexpression of GroEL, the third chaperone system, or lower environmental temperature may still allow bacteria deficient in both Trigger Factor and DnaK to grow [25; 26]. This variance in survivability, based on temperature directly segues into a discussion of thermal stress. While many chaperones are constitutively expressed, their production is stepped up when the cell is exposed to

stress. Such chaperones are classified as 'stress proteins', or 'heatshock proteins' (Hsps). As a historical aside, a 'heatshock response' was noted in Drosophila as a 'puffing' of chromosomes in 1962. Subsequent work led to the first identification of heatshock proteins in 1974 [27; 28]. There are several types of chaperones, each with its own mechanism of action. Nevertheless, some general principles of how a chaperone functions may be discussed. Misfolded proteins are identified by exposed hydrophobic surfaces and unstructured backbones. Chaperone binding prevents such surfaces from mutually interacting, thus staving off aggregation. Damaged proteins may be 're-engineered' to fold into their native states, fuelled by ATP turnover. There are frequently cofactors acting with the heatshock proteins to assist in these tasks [20].

1.4 THE COMPONENTS OF THE CELLULAR CHAPERONE SYSTEM

As already mentioned, cells have a three stage chaperone machinery. We shall briefly discuss their principal features before going on to discuss Hsp70s in more detail.

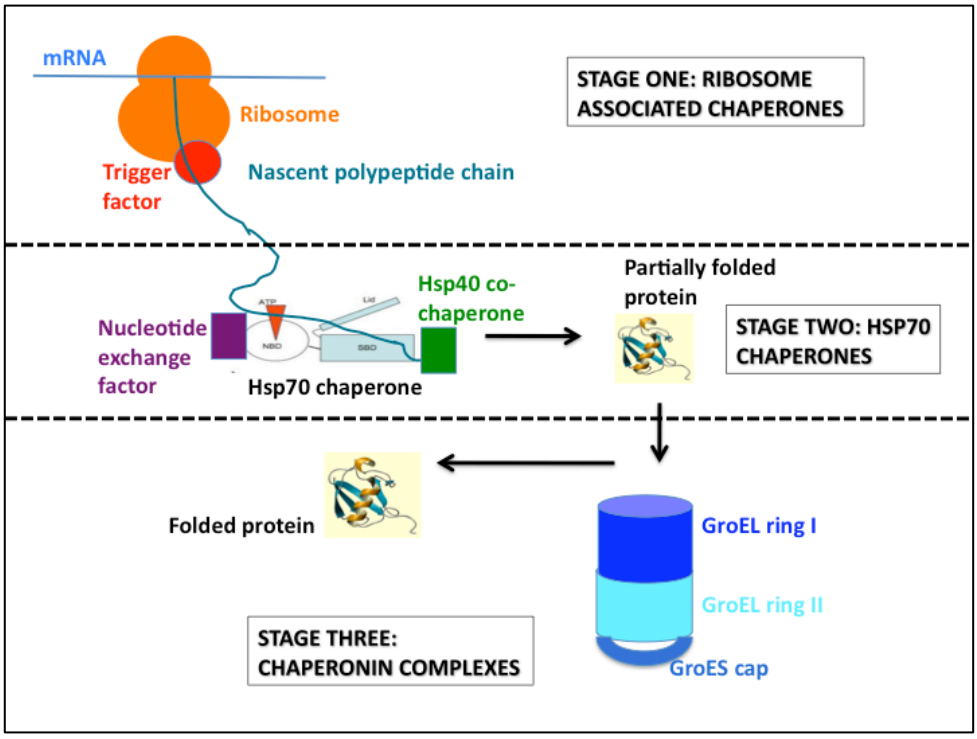

Figure 1.3: The three stage chaperone machinery deployed by the cell

1.4.1 STAGE ONE – RIBOSOME BINDING CHAPERONES

The first is deployed immediately 'after' translation. In Bacteria, Trigger Factor (TF), a 422 residue molecular chaperone that is conserved to about 70% sequence identity across species, plays this role. TF binds to the 50S subunit of the ribosome near the exit channel and is the first chaperone encountered by nascent proteins. This protein possesses a peptidyl-prolyl cis-trans isomerase (PPIase) action that catalyzes proline isomerization in protein folding [25; 29]. Proline isomerization is frequently a rate-limiting step in protein folding [29; 30]. Trigger Factor has three domains: an N-terminal ribosome binding domain, a PPIase domain and a C-terminal domain, where nascent polypeptide binds. TF exists in dimeric form in the cytoplasm and in monomeric form when bound to the ribosome. TF accounts for up to 70% of the nascent protein folding process. In *Archaea*, the equivalent chaperone is the Nascent Chain Associated Complex (NAC) [18; 31]. *Eukarya* fold nascent polypeptides by means of a NAC, which works in conjunction with a Ribosome Assisted Complex (RAC). The RAC is made up of Hsp70 homologues, such as Ssb1/2 in yeast and the Hsp70L1 and Mpp11 in mammals. As mentioned, the defining characteristic of this set of chaperones is that they interact directly with the ribosome as proteins are translated. There is no ATP turnover associated with substrate binding and release.

1.4.2 STAGE TWO – Hsp70 CHAPERONES

Hsp70s are 70 kDa multidomain proteins, which assist in folding polypeptides longer than are handled by the stage-one chaperone machinery. They operate ATP-driven cycles of binding and release of substrate peptide. This system does not interact directly with the ribosome, although it does operate with several 'partner' molecules. Firstly, there are Hsp40 co-chaperones, which are associated with the 'loading' of 'client' peptides onto the Hsp70 system. Later in the Hsp70 operational cycle, Nucleotide Exchange Factors (NEFs) allow for ADP to diffuse out of the Hsp70

molecule and ATP to diffuse in, thus 'resetting' the machinery for another round of substrate binding and release [18; 32]. The combined action of co-chaperones and exchange factors has enabled the Hsp70 system to play many roles in this molecular ballet, from preventing protein aggregation to cross-membrane transport to regulation. Although Hsp70s can be thought of as a stage leading to further protein folding/processing by the Chaperonin system, there is also evidence from reactivation of thermally damaged luciferase that DnaK (a bacterial Hsp70), GrpE (bacterial NEF) and DnaJ (bacterial Hsp40) work in ATP turnover driven cycles to repair damaged proteins [33]. Thus the complete Hsp70 system with its partner molecules is also a protein-refolding machine in its own right.

1.4.3 STAGE THREE – CHAPERONINS

Large protein complexes occupy the third stage. Group I is best exemplified by the bacterial chaperonin GroEL [34], which has been the best characterized of all such protein complexes [35]. In Archaea, the equivalent system is called the Thermosome [36], which is classified in Group II, along with the eukaryotic equivalent: CCT [37; 38]. The structure of GroEL [39], which is the most studied of these protein complexes, is as follows: GroEL is a dual-ring tetradecamer. The monomeric units are three-domain chaperones called Hsp60s. The interior cavity of GroEL is hydrophobic and is the unfolding-refolding 'workshop' for substrate proteins [40; 41]. The top and bottom ring work in opposite 'phase' in the operational cycle [42; 43]. Each ring turns over 7 ATP molecules per cycle.

The operational cycle runs as described below: surface exposed hydrophobic patches characterize unfolded substrate proteins. These substrate hydrophobic patches are recognized by a hydrophobic patch on the rim of the GroEL cavity, and thus bind to the rim. In this process they become partially unfolded. Substrate binding, along with ATP binding induces the binding of a 'lid' domain called GroES capping the

cavity. GroES is a heptameric ring made of 10 kDa subunits. The substrate protein is now trapped inside a cavity, which is about 85 Å long and 80 Å in diameter. The capping of the GroEL cavity by GroES induces a conformational change in the interior of the cavity, which restricts access to the hydrophobic binding site. Now that the cavity is largely hydrophilic, the native conformation of the substrate protein is favored, where hydrophobic residues are buried deep inside, and the partially unfolded substrate protein refolds. Substrate proteins are typically sequestered for ~ 10 sec [44]. ATP hydrolysis and binding of substrate protein to the other tetradecameric ring induces another conformational change where the lid is cast off. Subsequently, the substrate protein is released into the cytosol. Substrate proteins may undergo multiple binding/refolding/release cycles before they are properly folded [45]. The binding stoichiometry, along with the binding geometry of the cavity obviously puts limits on the size of substrate proteins that can be handled by this system. Proteins of up to 60 kDa can be confined within the 'Anfinsen Cage' created by the GroEL/GroES cavity [46; 47]. The eukaryotic equivalent of GroEL; CCT is a system consisting of two 8-mer rings. This molecular chaperone has been characterized in much less detail, and its role in the protein folding in general is not yet so clear.

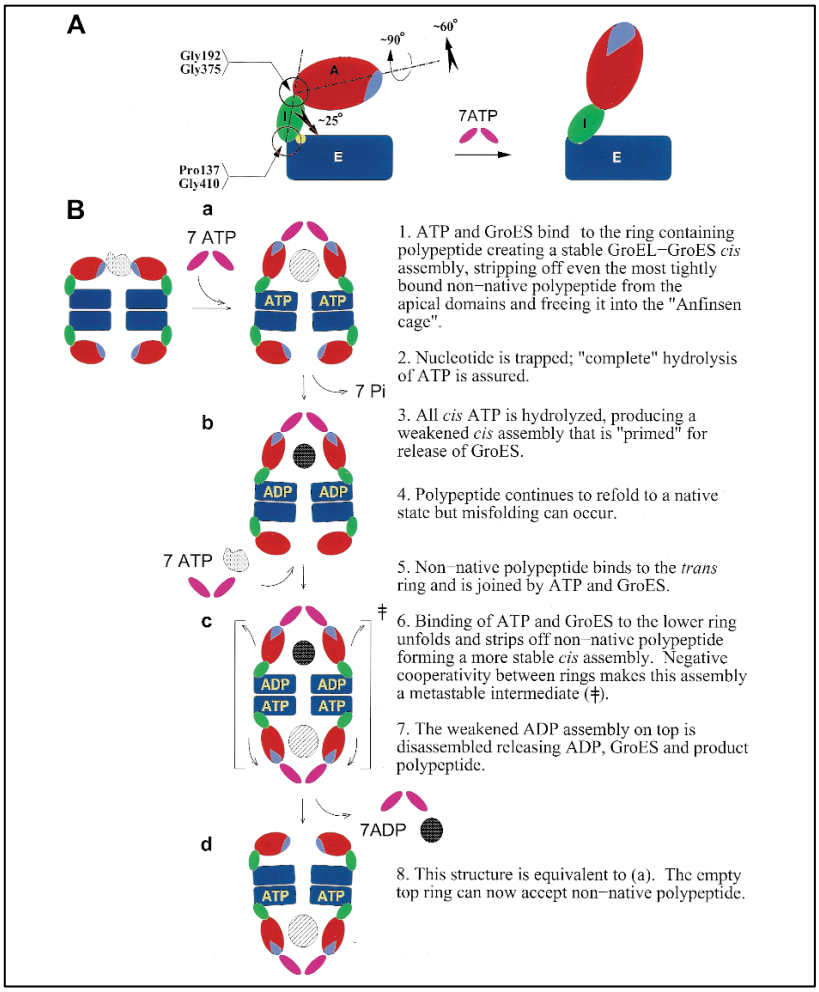

Figure 1.4: Schematic of the GroEL operational cycle

Ref: [45]

1.5 Hsp70 CHAPERONES – AN OVERVIEW

Hsp70 chaperone proteins are 70 kDa proteins, which are both constitutive as well as overexpressed during stress. They bind to and stabilize nascent polypeptide chains; but unlike Trigger Factor, do not associate directly with the ribosome. Exceptions to this behaviour are seen in Hsp70 homologues such as Ssb1/2 and Ssz1 in *S. cerevisiae* and Hsp70L1 in *Mammalia* [18]. Apart from nascent polypeptides, Hsp70s also assist in stabilizing misfolded proteins and preventing aggregation. Substrate binding and release is regulated by ATP turnover. Several co-chaperones

and exchange factors work along with Hsp70s to create a highly 'tunable' response system [32; 48]. Together, they can also elicit a protein refolding mechanism [33].

1.6 STRUCTURE HIGHLIGHTS OF Hsp70

E.coli Hsp70 (DnaK) is 638 residues long and is the one of the best studied of this class of molecules.

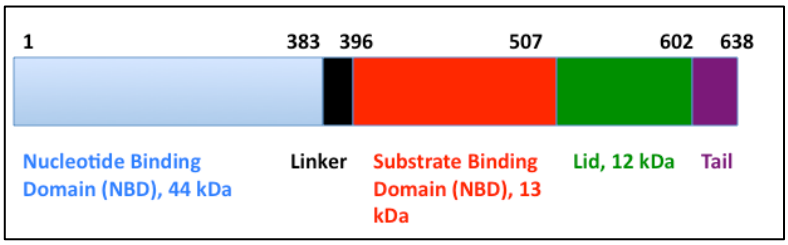

Figure 1.5: DnaK sequence schematic

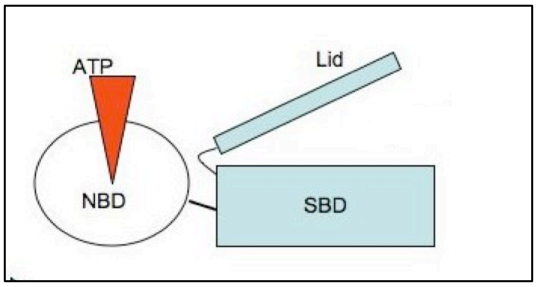

Figure 1.6: DnaK structure schematic

Hsp70s have a N-terminal 45 kDa ATPase domain. This Nucleotide Binding Domain (NBD) is formally divided into four subdomains (IA/IB/IIA/IIB), where IA-IB and IIA-IIB form the two arms of a V-shaped structure [49]. Nucleotide is bound at the centre of the V-shaped cleft. The NBD is connected via a linker to the Substrate Binding Domain (SBD). The linker is 10-12 residues in length and is highly conserved, showing a characteristic LLDV*P hydrophobic sequence.

The Substrate Binding Domain is a β-basket of 13kDa, which is made up of two antiparallel β-sheets of four strands each. The substrate-binding pocket is located between the two β sheets, and is highly hydrophobic, with a high affinity for hydrophobic residues such as Leu.

1.7 STRUCTURE OF THE Hsp70 NUCLEOTIDE BINDING DOMAIN – CRYSTALLOGRAPHY STUDIES

Before we discuss the structure of the NBD, it is worthwhile to dwell briefly on the functional context of the first such structural studies. Apart from the obvious action of preventing protein misfolding and aggregation, Hsp70s are also studied in the context of Clathrin: a protein involved in coating vesicles. Clathrin forms triskelion structures with three heavy and three light chains. These triskelions then form polyhedral lattices, which coat vesicles. During endocytosis, the clathrin coats are disassembled enzymatically. The enzyme implicated in this process is Hsp70. This was the original context of structural studies on Hsp70 [50; 51; 52]. Early studies identified two binding sites on the protein, but somewhat erroneously tried to correlate both sites to nucleotide binding [50]. Subsequent studies however showed that one of the binding sites corresponds to the nucleotide-binding site, and the other site is the peptide binding, or functional site.

The four subdomains which make up the V-shape of the Nucleotide Binding Domain have the following domain boundaries (Bovine Hsc70 count [49]):

IA – 1 to 39, 116 to 188 and 361 to 387

IB – 40 to 115

IIA – 189 to 228 and 307 to 360

IIB – 229 to 306

Linker – 388 to 398

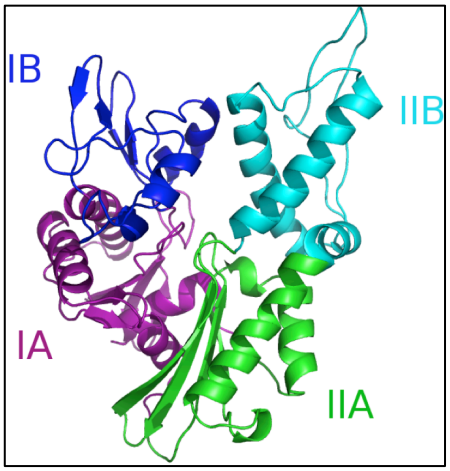

Figure 1.7: The nucleotide-binding domain of Hsp70

McKay and co-workers were the first to solve the structure of the NBD by crystallography. Crystals of Hsp70 NBD were grown in 1mM Mg+2 – ADP, 1M NaCl, 20% PEG-8000 and 50mM buffer in different pH's ranging from 6 to 10 in order to yield a structure refined to 2.5 Å. The four domains within the NBD were first defined in this work. In terms of overall structure, domains IA and IIA have somewhat similar topology βββ..(IB or IIB domain)…αβαβα. There is also weak sequence similarity between IA and IIA. Further, McKay and co-workers also identified topological similarities between IA and IIA and the corresponding domains in hexokinase [49].

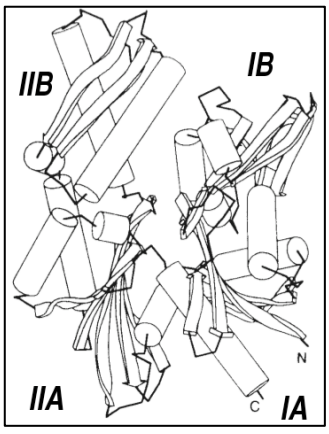

Figure 1.8: The topology of the Hsp70 nucleotide-binding domain

Ref:[49; 50]

It is important, however, to not over-interpret such similarities, especially since

domain IIB is not similar to hexokinase and IB is completely absent. Further, Hsp70 does not display phosphorylation activity.

1.7.1 NUCLEOTIDE BINDING SITE

Four residues were initially identified as active site residues, potentially involved in nucleotide hydrolysis. These are: Asp10 and Asp199 – both identified as interacting with divalent Mg^{+2}, and Glu175 and Asp206, both candidate proton acceptors. Mutational studies (changing the Glutamate/Aspartate to the equivalent Glutamine/Asparagine) were carried out to measure the change in ATP hydrolysis rates with modified active site candidates. Mutant Hsp70 NBD crystal structures were also solved (to about 2.5 Å). ^{32}P based ATP hydrolysis assays indicate that the k_{cat} drops to 0.01 that of the wild type NBD when Asp10 or Asp199 are mutated, and it drops to 0.1 of k^{WT}_{cat} when Glu175/Asp206 are mutated. The K_m is not changed significantly when Asp199/Asp206 are changed, but it increases by a factor of 10-100 when Asp10/Glu175 are changed. Thus, the proposed mechanism for ATP hydrolysis does not depend on just one residue de-protonating an attacking water molecule/stabilizing an attacking hydroxyl ion/stabilizing the transition state, but several residues acting in concert, especially those which interact directly with the divalent magnesium ion [53]. Crystal structures with various mutations and nucleotide states also implicate several residues that interact with inorganic phosphate. Lys71 forms a salt bridge and Thr13 and Thr204 form hydrogen bonds with PO_4^{-3}. Further, several residues such as Thr14 and Thr204 form hydrogen bonds with the β and γ phosphate of ATP [54]. Further work involving mutational studies on Lys71 (K71A/K71M/K71E) indicated that all ATPase activity is lost upon changing the Lys71 to any other amino acid. These mutant proteins were crystallized and yielded structures with a resolution of 1.7 Å. The mutant structures (ATP bound) are all essentially the same as the WT structure (ADP.Pi bound) with backbone aligned RMSDs between 0.15 and 0.32 Å. This study indicates that Lys71 is an essential (and highly conserved) residue, which participates in ATP-

hydrolysis. The proposed mechanism is that ATP binds in a β-γ-bidentate complex with the divalent magnesium ion, which stabilizes the γ PO^{-3}_4 for hydrolysis [55]. A water molecule or hydroxyl ion, which is stabilized by Lys71 then attacks the γ-phosphate and thus hydrolyses ATP [54; 56; 57]. These studies note that a plausible reaction mechanism for the hydrolysis of ATP can be suggested which does not involve substantial conformation changes in the NBD. The authors however offer the caveat that changes in the full-length molecule cannot be ruled out. This is of great import, in terms of allosteric signaling, as we shall shortly see when we discuss the structure of the substrate-binding domain (SBD).

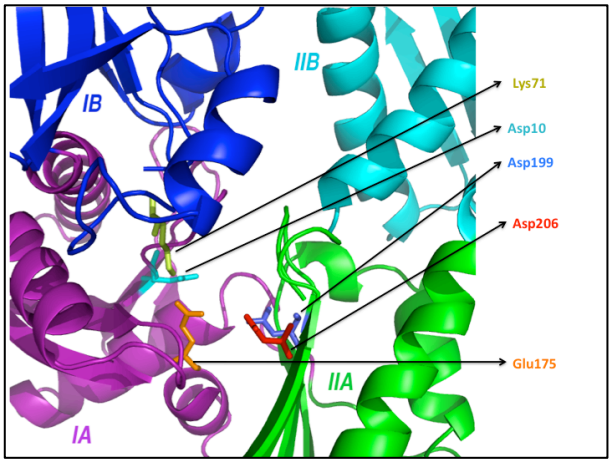

Figure 1.9: Residues identified in the active site of the Hsp70 NBD

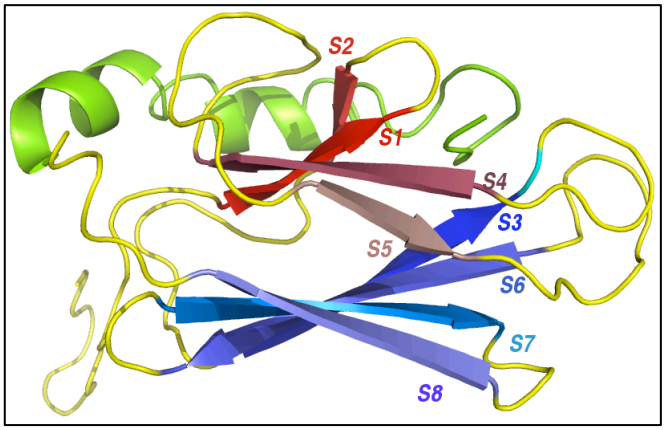

Figure 1.10: NMR structure of Hsc70 (rat) SBD

Sheet I elements (β-strands 1,2,4 and 5) are shown in shades of red-brown. Sheet II elements (β-strands 3,6,7 and 8) are shown in shades of blue. The C-terminal helix is in green [58].

1.8 STRUCTURE OF THE SUBSTRATE BINDING DOMAIN (SBD): INITIAL STRUCTURAL INVESTIGATION

Zuiderweg and co-workers carried out the first studies yielding conformational information on the C-terminal Substrate Binding Domain (SBD). Rat Hsc70 SBD (identical to the human and bovine forms) was expressed in *E.coli* and solution phase NMR experiments were conducted to yield a topological model of the SBD. HSQC and 3 dimensional NMR experiments were used to 'fingerprint' the protein, NOESY experiments were used to find distance constraints and amide exchange was monitored to find solvent exposed residues [59; 60; 61; 62]. The model thus developed indicated two anti-parallel β-sheets (Sheet I: strands 1,2,4 and 5; and Strand II: 3,6,7 and 8). The centre of Sheet II is amphipathic, which points to a protein-solvent boundary. This study also noted the presence of a noncontiguous C-terminal helix [58]. The same system was subjected to a structure determination analysis using NOE derived distance constraints.

The substrate-binding cleft, being the 'business end' of the molecule deserves special mention. Studies with phage display libraries to screen for substrates compatible with DnaK have indicated a preference for Leu, Ile and Lys and a disposition against Phe, Glue and His [63; 64; 65]. The construct studied by Zuiderweg and co-workers had a partially destabilized C-terminal helix, which interacted with the substrate-binding pocket. Leu539 was found to be bound to the β-basket region of the SBD, surrounded by a hydrophobic cluster of residues: Ile472, Val436, Ile438, Ala404, Thr403 and Phe426. This Leucine residue could not be displaced by competing peptides titrated into solution, indicating that the binding was tight and specific, which tallies with the already known preference of the substrate binding cleft for Leu [66]. This

is consistent with the phage display studies mentioned earlier.

1.8.1 STRUCTURE OF THE SUBSTRATE BINDING DOMAIN (SBD): CRYSTALLOGRAPHY STUDIES

Hendrickson and co-workers solved the structure of the substrate binding domain of *E.coli* Hsp70 (DnaK) to a resolution of 2 Å [67]. This study was carried out on a full SBD construct (residues 384-638) in complex with substrate peptide (NRLLLTG). The overall structure was consistent with the NMR derived topological model suggested by Zuiderweg and co-workers (in fact, this crystal structure was published shortly after the topological model was reported and before the NMR structure of Rat Hsp70 SBD was solved). The SBD can be broadly divided into two topological regions. The first, which has a β-basket or β-sandwich architecture, runs from residue 393 to 502. The β-sandwich has two distinct sheets, as described earlier. Residue 502 leads via a seven-residue loop to a five-segment helical domain. The first helix, αA runs from Glu509 to Ala521 and provides a bridge for the rest of the α-helical units, which form a 'lid' of the β-sandwich. The second (and longest) helix, αB runs from Ala532 to Lys556 and is the longest element of the lid, which extends straight out over the β-sandwich. Helix αC runs from Ala559 to Lys577 and loops backward (at a 151 degree angle) from helix αB. A 169 degree turn leads to helix αD which is from Lys581 to Ser595, and runs back 'out' in the direction of αB. Finally, helix αE which runs from Gln596 to His606 lies at 68 degrees w.r.t αB at the very end of a three-helix anti-parallel bundle. This architecture is somewhat similar to transducin.

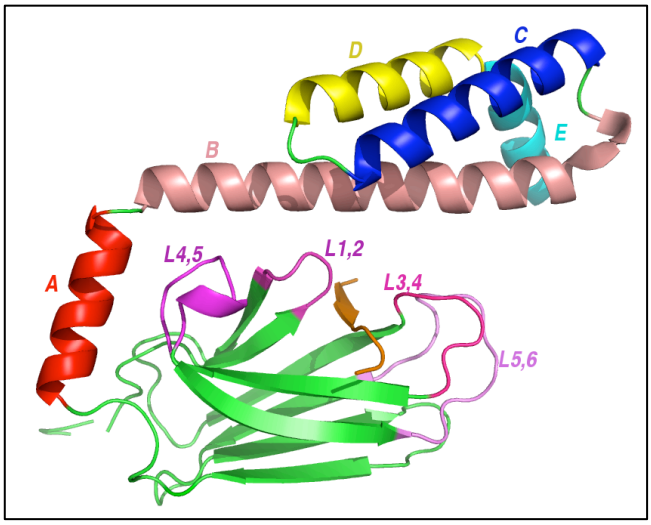

Figure 1.11: Crystal structure of the SBD of E.coli DnaK

Helix αA – red, helix αB – salmon, helix αC – blue, helix αC – yellow, helix αE – cyan, β-sandwich in green, Loops (L4,5, L1,2, L3,4, and L5,6) – magenta shades, substrate peptide (NRLLLTG) – orange [67].

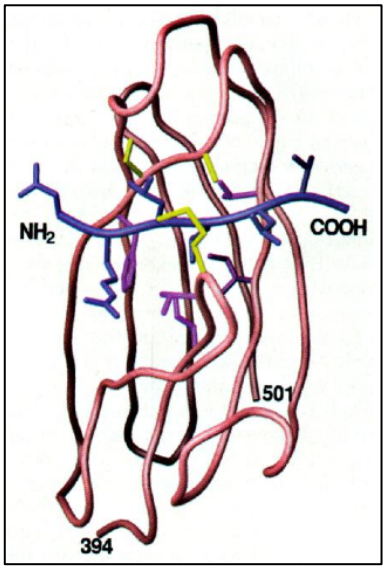

Figure 1.12: Ribbon diagram of the DnaK Substrate Binding Domain showing the Substrate-peptide (NRLLLTG) binding channel {Zhu, 1996 #1523} .

The α-helical domain interacts with the β-sandwich domain at multiple points. Helix αA is most closely integrated with the β-sandwich via hydrophobic contacts with Loop L4,5. Helix αB which overhangs the β-sandwich contacts all four loops (L4,5, L1,2, L3,4, and L5,6) and thus 'protects' the substrate binding pocket.

It is seen from the structure that the substrate peptide (NRLLTG – chosen on the basis of phage display screens) lies in an extended conformation in a channel that slices transversely through the β-sandwich. This channel is formed by loops L1,2, L3,4 and has a cross section of 5 Å X 7 Å.

The authors also speculated about the types of substrate peptides that could be accommodated in the substrate cleft, which they designate as 'hsp70 site 0'. The centre of the pocket, or site 0 is home to a Leu, which is surrounded by a number of hydrophobic residues: Phe426 (β3), Val436 and Ile438 (β4), Ile401 and Thr403 (β1). The extended chain conformation of the peptide might be a clue to the mechanism of how Hsp70 interacts with actual 'client' proteins. The selectivity of site-0 to Leu (and Met/Ile) influences the type of client peptide that can be 'loaded' on the substrate-binding domain. Charge distribution analyses indicate that negatively charged residues are not favoured, large hydrophobic sidechains are preferred as the central residues and positively charged residues preferred at termini of heptamers.

1.8.2 THE QUESTION OF SUBSTRATE BINDING

Zhu and Hendrickson's work identified a very fundamental problem: that of 'substrate loading'. The shape of the substrate cleft and the narrow channel that the heptapeptide is located in is not easily accessible to solution. It is possible, but unlikely, that client peptides will 'thread' the substrate channel in order to bind to the SBD. In the words of Zhu and Hendrickson, the structural evidence points to conformational change on a 'grand scale' to accommodate substrate peptide. They postulate that the region defined by Arg536-Gln538 (part of helix αB) melts away to provide a flexible hinge about which the rest of helix αB (and subsequent helical elements) pivot to allow entry of substrate peptide to the binding cleft.

The NMR structure solved by Zuiderweg and co-workers also argues for large scale conformational changes associated with substrate binding, but with one important

caveat: the location of the hinge point is different. NMR relaxation studies indicate that all five helices comprising the α-helical domain have the same tumbling rate as the β-sandwich domain in solution. This indicates that conformational changes probably involve entire subdomains moving at once. Thus, a hinge point for a lid opening/closing would be between residues 500 and 520. But this naturally gives rise to the question: if lid opening and closing determines substrate binding, then what regulates lid opening and closing? To address this question, we have to look at the operation of this molecule in its entirety.

1.9 ALLOSTERY IN Hsp70s

Thus, through a fairly involved discussion of the structure of the component domains of Hsp70, we have finally arrived at a subject of paramount interest: communication between the NBD and the SBD. To understand the communication between the NBD and the SBD, it is important to settle upon a suitable metric to determine whether the molecule is 'working'. An obvious metric for the NBD is ATP hydrolysis.

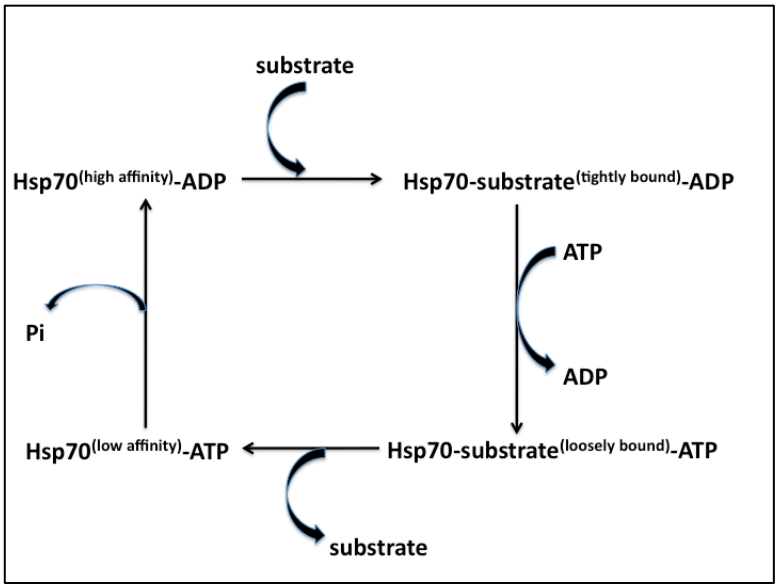

Figure 1.13: ATP binding driven conformation change leads to substrate release in Hsp70. Ref: [56]

One of the first substantive efforts in elucidating the allosteric communication

came from the work of Palleros and Fink in 1993 [56]. Their studies indicated that Mg-ATP *binding*, but not hydrolysis is necessary for substrate peptide release, and further that K^+ is required for this cycle to occur. Their ideas are encapsulated in the simplified cycle (Fig 1.13).

Other studies have also shown that substrate binding/release is coupled to the ATP hydrolysis cycle. Large concentrations of Mg-ATP stimulate substrate release, while high concentrations of substrate peptide stimulate ATP hydrolysis [68]. This is also inferred from the work of Liberek and Georgopolous in 1991, where they show that the basal ATPase rate of a mutant DnaK (DnaK765) that is locked into a substrate-binding-domain open conformation is about 50 times higher than WT DnaK (which is about 0.1 to 1 molecules of ATP hydrolyzed per minute per molecule of DnaK). DnaK765, being locked into an 'open SBD' conformation corresponds to rapid exchange of substrate peptide with solution [69]. Extensive work by McKay and co-workers led to a more exhaustive attempt at understanding the mechanism of this system. Their studies were based on Thin Layer Chromatography (TLC) studies of radio labeled ATP (γ 32P-ATP) hydrolysis and phosphate release [70]. The highlights of their work are that, firstly, the 44kDa NBD fragment by itself hydrolyses ATP about 4 times faster than the full-length Hsc70, and that secondly, inorganic phosphate is released first, and ADP afterward.

So, summarizing what we have discussed so far:

1. The ATP hydrolysis cycle in the NBD is coupled to substrate binding/release by the SBD.
2. Mg-ATP binding induces conformational changes in Hsp70.
3. The release of phosphate occurs before that of ADP.
4. The NBD by itself hydrolyses ATP faster than the full length Hsp70.
5. Addition of specific peptides increases the rate of ATP hydrolysis by the full length Hsp70 by a factor of 2 – to a level approximately equal to the basal ATP hydrolysis rate by the NBD alone [71; 72; 73].

Based on this, we conclude that the allosteric coupling between the NBD and the SBD is two-way. If we think of the NBD as an automobile engine and the SBD as the wheels, then the operation of the engine rotates the wheels, but conversely, free rotation of the wheels will also drive the engine. But the picture we have of the Hsp70 cycle is far from complete. The basal ATP turnover rate is slow (~0.02 to 0.2 /min [48]), in the context of chaperone function. Thus, there are several co-chaperones and effector molecules, which act in concert with Hsp70.

1.10 OTHER PARTICIPANTS IN THE Hsp70 CYCLE: CO-CHAPERONES

Liberek and co-workers demonstrated in 1991 that the ATPase activity of Hsp70 can be stimulated up to 50-fold in the presence of DnaJ and GrpE [74]. They found that the addition of DnaJ (but not GrpE) to Hsp70 stimulated ATP hydrolysis, but not ADP release, and that the addition of GrpE (but not DnaJ) stimulated ADP/ATP release, but not ATP hydrolysis. This led them to suggest that these two 'effector' proteins work at 'opposite' ends of the Hsp70 cycle. The authors also speculated about the possible role of DnaJ and/or GrpE in substrate peptide release. It is apposite, at this point, to discuss these co-chaperones briefly.

1.10.1 A BRIEF LOOK AT DnaJ

DnaJ is a 376-residue constitutively expressed *E.coli* protein, which along with DnaK was initially identified as one of the proteins which assist in DNA replication of lambda-phage [75]. Later, this protein was shown to work in conjunction with DnaK and GrpE to assist in the folding of rhodanase [37]. It should also be noted that DnaJ by itself also has some chaperone activity; in that it bound unfolded rhodanase, thus preventing aggregates from forming. DnaJ has been shown to act in concert with DnaK and ATP to promote binding of RepA to oriP1 during replication of bacteriophage P1 miniplasmid [76; 77]. Furthermore, studies by Liberek and Georgopolous have also found a DnaK - σ32-DnaJ complex, which forms in the presence of ATP. σ32 is a

polypeptide involved in the autoregulation of the heatshock response in *E.coli*, and hence a substrate for the action of DnaK [78]. All of these studies collectively suggest that DnaJ assists in the mediation of the interaction of DnaK with its substrate, or to put it differently, 'loads' client peptides on to DnaK. Thus, it behooves us to discuss the structure of DnaJ briefly.

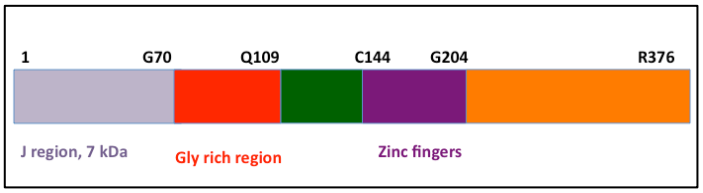

Figure 1.14 E.coli DnaJ sequence schematic

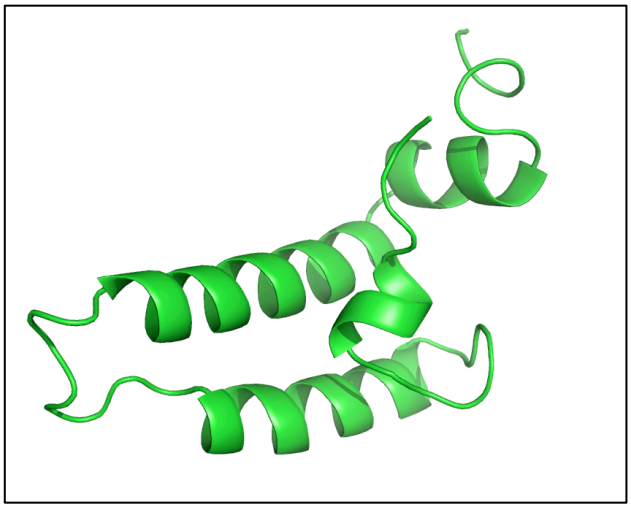

Figure 1.15: NMR solution structure of the J-domain from E.coli DnaJ [79; 80]

There are at least 3 J-protein types: Class I, containing DnaJ, comprises all domains shown in Fig 1.14. In class II, the Zn finger domain is missing. Class III contains J-domains only.

Bacterial DnaJ is a 376-residue protein that has a highly conserved N-terminal region of about 70 residues, called the J-region [81; 82]. Following the J-region, there is a conserved Gly rich stretch, which ends at residue 109. The J-region and the Gly-rich region are sometimes collectively called the J-domain. Downstream of the J-domain, there is a putative zinc finger with four CXXCXGXG repeats starting at C144 and ending at G204 [83].

Work by Wall and Georgopolous showed that the N-terminal fragment (108 residues) of DnaK was sufficient to stimulate ATP hydrolysis by DnaK, as well as for replication of λ-phage [83]. This lent new biological significance ongoing to efforts to solve the structure of the J-Domain. These efforts came to fruition when Pellecchia, Wuthrich and co-workers solved an NMR structure of the J-Domain (2-108 residues) of bacterial DnaJ [79; 80]. Their analysis identified four α-helices in the regions 6-11, 18-31, 41-55, and 61-68. They characterized the backbone as two long antiparallel helices (numbered 3 and 4) that are stabilized by a hydrophobic core of conserved sidechains. The Gly-rich region was found to be disordered in solution.

Stepping away briefly from our discussion of Hsp70s in the context of protein re-folding, we will discuss another function of Hsp70s, namely clathrin uncoating. As briefly touched upon earlier in this chapter, Clathrin is a structural protein, about 1700 residues in length. It forms a triskelion (three legged) structure, and associates in polyhedrals to form vesicle coats. Coalescence of the clathrin-coated vesicle with its target membrane will require re-engineering of the clathrin coat: specifically disassembly of the polyhedral lattice. This is where Hsp70 comes into play. The uncoating of the clathrin polyhedral structure is accomplished by the action of a protein called auxilin [84]. This is a 86 kDa protein of about 910 residues which was first studied by Ahle and Ungewickell [84]. Auxilin possesses a clathrin-binding domain that contains long stretches of Cys and Pro-rich regions, and also a C-terminal J-Domain. The clathrin binding region of auxilin interacts with triskelia both as a part of intact auxilin as well as on its own [85; 86]. Just as the N-terminal J-region of bacterial DnaJ was shown to have biological activity, the auxilin C-terminal J-domain was also shown to disassemble clathrin vesicle coats in association with the 60 kDa construct of Hsp70 (NBD+SBD, but not lid). This activity was not seen when the activity of the auxilin J-domain was assayed with exclusively Hsp70 NBD or Hsp70 SBD [87]. Sousa and co-workers solved the first structure of the auxilin J-domain in 2003 [88].

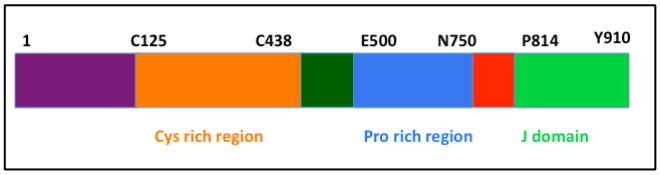

Figure 1.16: Sequence schematic for bovine Auxilin

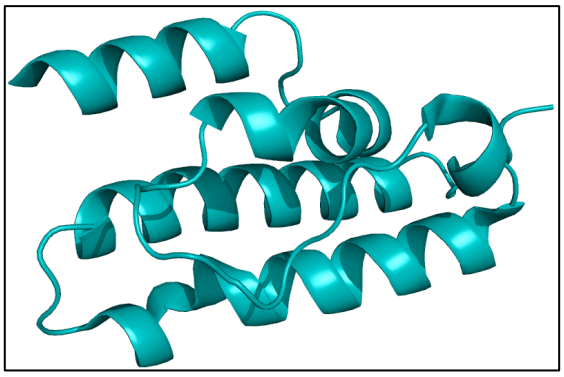

Figure 1.17: Crystal structure of the Auxilin J-Domain [88]

1.10.2 A BRIEF LOOK AT GrpE

GrpE is the other co-chaperone that acts in conjunction with Hsp70. This molecule is an exchange factor, which acts at the opposite end of the Hsp70 functional cycle. The presence of GrpE accelerates ADP exchange in DnaK by a factor of 5000 [89]. GrpE is a 197-residue protein that is usually found as a tightly bound homodimer. The dimer interface is made up of two long paired helices - one contributed by each monomer, leading to a four-helix bundle where each monomer contributes two short helices. GrpE binds to DnaK with a stoichiometry of GrpE: DnaK=2:1 [90].

The interhelical loop is partially disordered in one monomer and completely disordered in the other. Harrison, Kuriyan and co-workers were able to solve a crystal structure of GrpE bound to the DnaK NBD at a resolution of 2.8 Angstroms [90]. This complex was crystallized in the absence of any nucleotide. DnaK NBD (construct 1-383 residues) was found to interact with the GrpE dimer at multiple points – but the principal area of interaction is between the β-sheet surfaces belonging to the proximal GrpE and

subdomains IB and IIB of the DnaK NBD.

Kuriyan and co-workers observed that Phe86 in the proximal GrpE long helix stabilizes Arg183, also in the proximal GrpE (β-strand), which in turn forms an intermolecular hydrogen bond with Glu28 of the DnaK NBD. This interaction is possible due to the asymmetric nature of GrpE dimer, which also accounts for the binding stoichiometry.

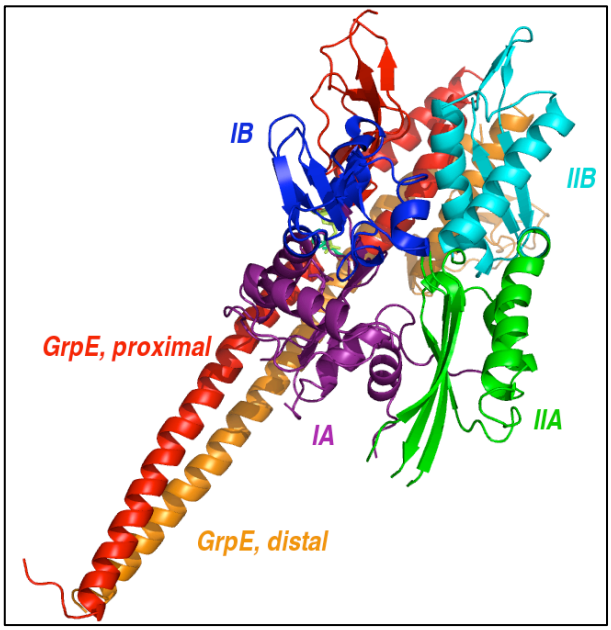

Figure 1.18: Crystal structure of DnaK NBD in complex with GrpE [90]

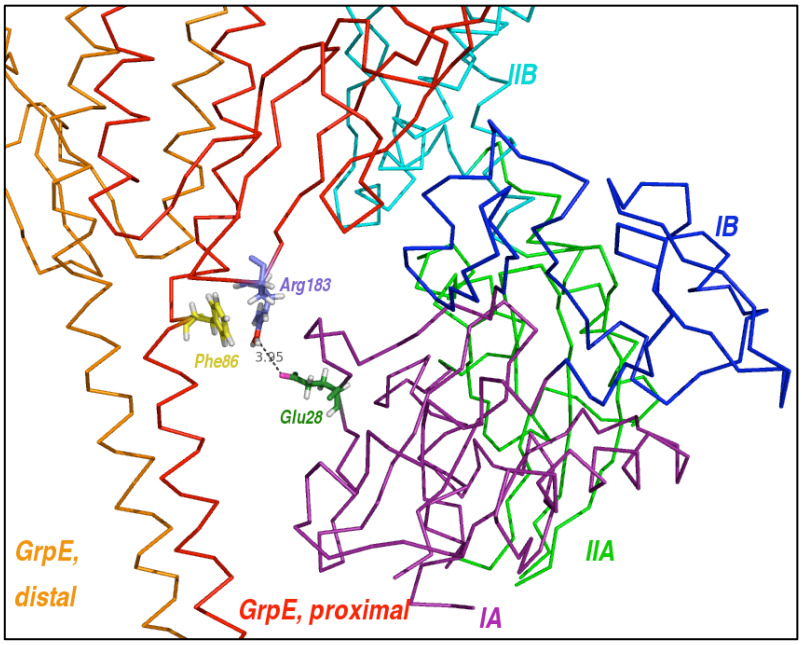

Figure 1.19: Hydrogen bond between Arg183 of the proximal GrpE β-sheet and Glu28 of the DnaK NBD

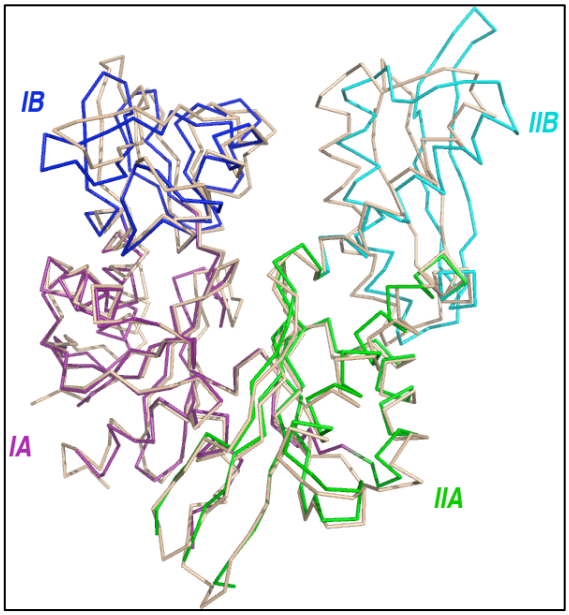

Figure 1.20: Overlay of the DnaK NBD from the crystal structure in complex with GrpE with the bovine Hsc70 NBD (with ADP.Pi) crystal structure (shown here in light salmon colour) {Harrison, 1997 #10372}

Subdomain IIB of the DnaK/GrpE complex (cyan) is tilted outwards by 14 degrees w.r.t the corresponding part of the bovine Hsc70 NBD.

Surface Plasmon Resonance studies carried out by the same group indicated that in the absence of nucleotide, the GrpE/DnaK complex has a KD of 30 nM, which is increased to 600 nM in the presence of ADP [90]. The presence of ATP seems to block the formation of any complex. The authors further postulated that the long helices of the GrpE dimer might be involved in interactions with the DnaK SBD. The principal role of this molecule, however, was thought to be as a nucleotide exchange factor, and the mechanism of nucleotide exchange is thus of prime interest. An overlay of the NBD from this complex with the bovine Hsc70 NBD structure crystallized by McKay and co-workers indicates a rotation of subdomain IIB by up to 14 degrees. This was thought to be due to the influence of GrpE binding, indeed the authors mention that the mechanism of nucleotide exchange is a 'straightforward' opening of the nucleotide-binding cleft of DnaK.

The paper compares and contrasts the action of GrpE in this context to that of *E.coli* elongation factor EfTu, in complex with its corresponding exchange factor EfTs. The authors postulate a change of conformation and a displacement of secondary structure elements required for nucleotide binding in both systems. This is a crucial point which we will discuss in much detail later – whether GrpE in fact forces conformational changes in the DnaK NBD or not. For now, it suffices to mention that unlike in the EfTu/EfTs complex, no residue of GrpE comes closer than 13 Angstroms to the nucleotide-binding cleft of DnaK.

1.11 ASSEMBLING THE Hsp70 CYCLE

Collating the information we have about the Hsp70 operational cycle so far:

1. Hsp70 has two distinct domains, the NBD and the SBD.

2. The NBD hydrolyzes ATP and the SBD exchanges substrate peptide with solution. These two processes are allosterically coupled.

3. There are large conformational differences between the ATP and ADP.Pi [91] bound states.

4. A co-chaperone (called DnaJ in *E.coli*) assists in the loading of client peptide on the SBD.

5. An exchange factor (called GrpE in *E.coli*) is involved in the exchange of ADP for ATP in the NBD – presumably by inducing conformational changes in the NBD.

Armed with this information, we now proceed to draw out an operational cycle for the Hsp70 system (Fig 1.21). There is dynamical evidence from NMR spectroscopy that indicates that in the ATP state of Hsc70, the NBD and the SBD are docked, while in the ADP.Pi state, they are connected by a flexible linker, and tumble in solution as two semi-independent entities [91; 92].

A great deal of effort has gone into elucidating the structure of the different components of this system, either individually, or in complex. Some of these, we have discussed at length, such as the crystal structure of the bovine Hsc70 NBD in the ADP.Pi state [49], solved by McKay and co-workers, the crystal structure of the DnaK SBD, solved by Hendrickson and co-workers [67] as well as the crystal structure of the *E.coli* DnaK NBD, in complex with *E.coli* GrpE [90]. Until very recently, however, no research group had met with success while trying to solve the structure of full length Hsp70 in any nucleotide state. This was first accomplished by Bertelsen and Zuiderweg [92].

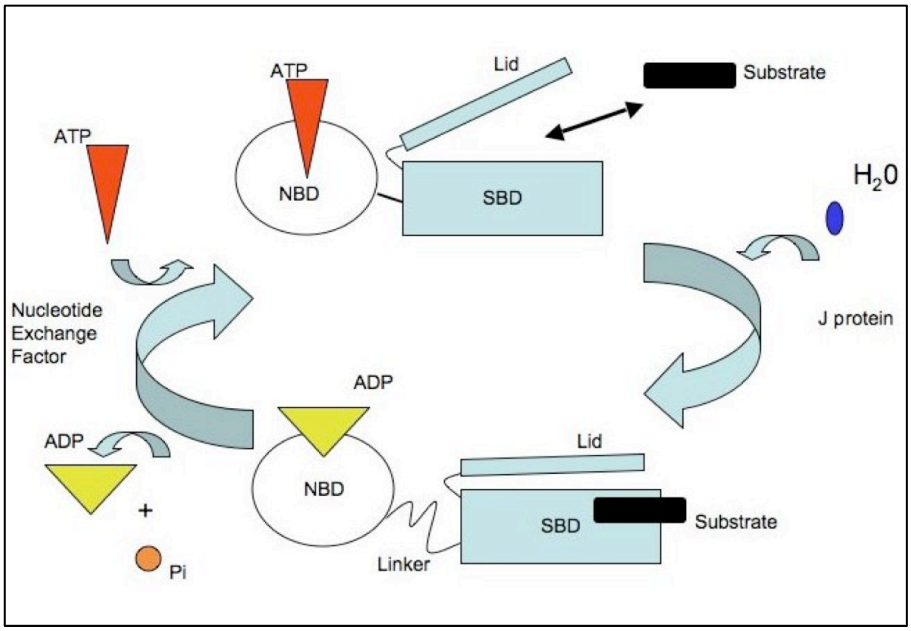

Figure 1.21: Schematic of the Hsp70 operational cycle

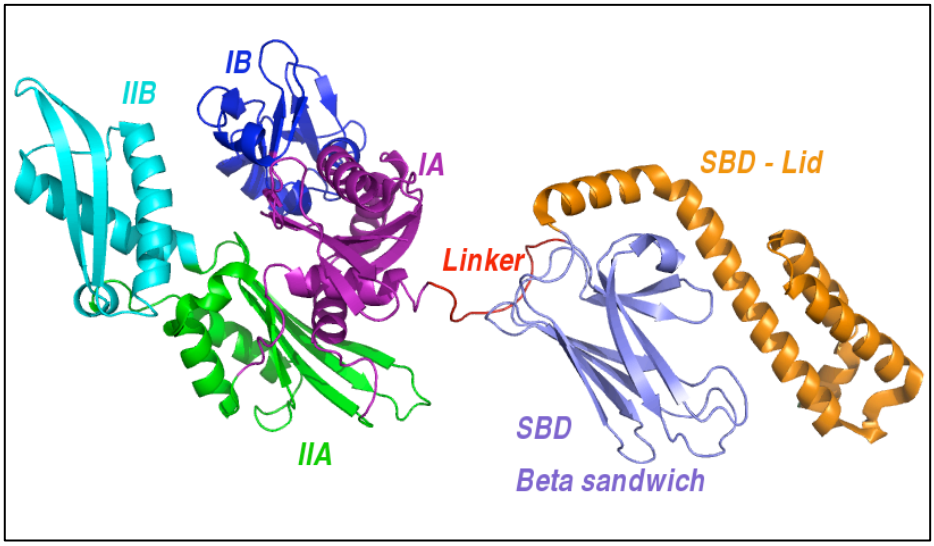

Figure 1.22: NMR solution structure of full length DnaK in the ADP.Pi state [92]

Bertelsen and Zuiderweg used residual dipolar couplings, a novel technique in NMR spectroscopy, to solve the ADP.Pi state of WT DnaK (residues 1-605). This structure indicated that the NBD and the SBD are, as anticipated, connected by a flexible linker, and that the lid of the SBD is clamped down upon substrate peptide (also NRLLLTG).

There have also been efforts to solve the structure of Hsp70 in the ATP state. However, so far, these efforts have not met with success. There is a candidate for the full length ATP state structure – the crystal structure of Hsp110 in the ATP state [93]. The human Hsp110 is an 858-residue (105 kDa) protein that is expressed in many mammalian tissues, especially the brain. Sse1 is the equivalent protein sourced from *S. cerevisiae,* which has 693 residues. Hsp110 has about 35% homology with the NBD and between 10-15% with the SBD of the Hsp70 of the same species. Hendrickson and co-workers were successful in solving a 2.4-Angstrom structure of this protein in the ATP bound state. Remarkably, Hsp110 does not function as a Hsp70 chaperone, but rather as Hsp70 nucleotide exchange factor, which has about 35% homology with the NBD and between 10-15% with the SBD. For this reason, as well as the fact that the linker is not conserved in Hsp110, it is difficult to consider this structure as a suitable model for the ATP state. With that caveat in mind, however, it must be noted that this structure does have the NBD and the SBD docked together.

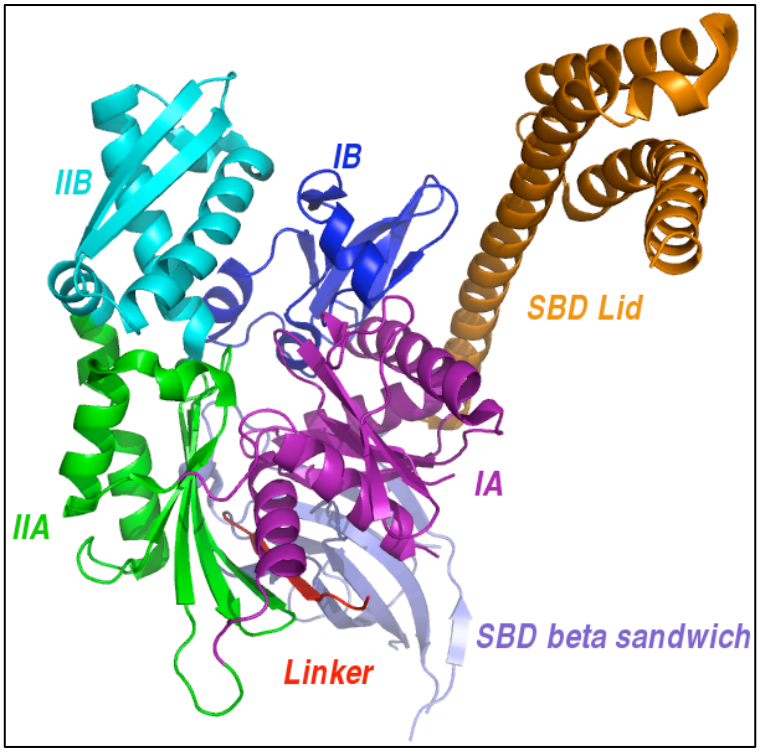

Figure 1.23: Crystal structure of Hsp110 (yeast) solved by Hendrickson and co-workers. This is a candidate model for the ATP state of Hsp70 [93].

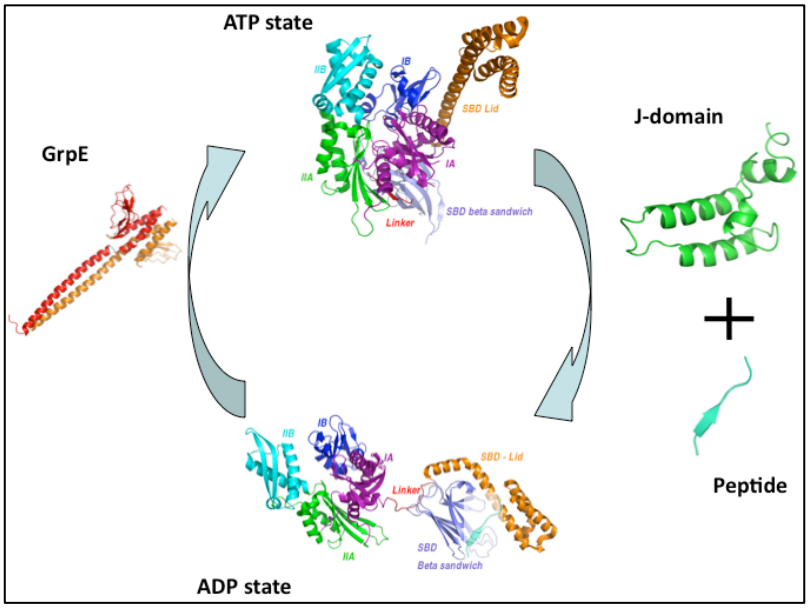

Figure 1.24: Hsp70 operational cycle with the structures investigated so far

With this candidate ATP structure, we are now in a position to draw out the Hsp70 operational cycle with all the structures that have been investigated to far. Apropos this cycle, we are now in a position to ask questions of interest, both in terms of the biochemistry as well as the biophysics of this system.

1.12 THE QUESTIONS ADDRESSED IN THIS DISSERTATION

1.12.1 HOW DOES GrpE (OR ANY NUCLEOTIDE EXCHANGE FACTOR) ACT TO EXCHANGE ADP FOR ATP IN THE NUCLEOTIDE BINDING CLEFT?

The conventional understanding of this process, as discussed by Kuriyan and Harrison [90] is that GrpE binding induces a large-scale conformational change in the nucleotide-binding domain, which leads to the exchange of ADP for ATP. As we will discuss at length in Chapter 2, there is not enough structural evidence to back up this

assertion, and our investigations using solution NMR techniques reveal a hitherto unanticipated mechanism of action for GrpE.

1.12.2 HOW IS THE ALLOSTERIC SIGNAL COMMUNICATED FROM THE NUCLEOTIDE BINDING DOMAIN TO THE SUBSTRATE BINDING DOMAIN?

Investigators have long believed that the linker between the NBD and the SBD plays a crucial role in allosteric signaling [91]. Here we present (Chapter 2 and 3) a hypothesis for how the behaviour of the linker changes when the molecule transitions from ATP to ADP binding. This leads us to propose a mechanism for the conformational changes that drive the Hsp70 cycle.

1.12.3 HOW DOES DnaJ COMMUNICATE WITH DnaK DURING 'LOADING' OF SUBSTRATE PEPTIDE?

This is also a problem that had been attacked by structural biologists. However, the only study [94] in the literature which attempts to answer this question is seriously flawed on two grounds:

1. The 'complex' of the J-domain and Hsp70 in question is not a complex at all, but a disulphide linked structure.
2. Only the NBD of Hsp70 has been studied, whereas biochemical evidence exists which indicates that the DnaJ/J-domain interacts with both the NBD and the SBD [95].

In Chapter 4, we will present results based on solution NMR studies of the interaction of full length DnaK and DnaJ (J-domain) that reveal the nature of the complex structure. This study also allows us to further elaborate our model of the allosteric communication between the NBD and the SBD (as discussed in Chapters 2 and 3).

1.13 ALLOSTERY AND MOLECULAR COMMUNICATION

Having introduced the topic of allostery, it is worthwhile to briefly discuss what this means in living system. Allostery stands for 'other space', which means that action at one end of a biomolecule may induce changes at another end. The first epochal work in this context was the MWC model of haemoglobin action, proposed in 1965. In this seminal work [96], the authors characterize allostery by the statement that '*indirect* interactions between *distinct* specific binding sites are responsible for performance of (the haemoglobin's) regulatory function'. Interestingly, researchers had already postulated the existence of motions in proteins. Before the elucidation of the mechanism of haemoglobin's action, Linderstrom-Lang, amongst others had already discussed 'breathing' motions in molecules [97]. Indeed, close to end of the 1950's, the idea of induced-fit as a possible mechanism of molecular interaction was proposed by Koshland as an alternative to Emil Fischer's venerable model of lock-and-key interactions [98; 99]. Thus, the ground was prepared for the first true discussion of communication over distances within molecular systems. Haemoglobin was an excellent candidate, for the dual reasons that it's structure was under study by X-ray crystallographers and that it is the classic example of a molecular transport mechanism which has to 'know' when to 'load/unload' oxygen and carbon-dioxide as part of its functional cycle. Briefly, the MWC model postulated that the subunits of haemoglobin all exist in the same conformation at any given time, but the overall molecule can switch between a Tense (T) and a Relaxed (R) state, both of which bind different ligands differently [96]. The KNF model, postulated shortly afterward stated that haemoglobin subunits can exist in different conformations side-by-side, thus allowing for a hybrid conformation for the overall molecule. Here [100], the allosteric transition was seen as sequential, rather than concerted, as was in the MWC model. Many other allosteric systems have been investiageted by methods of structural biology, and evidence for both MWC and KNF models, sometimes even mixed (such as in GroEL) has been found But the study of molecular communication-at-a-distance remained a comparison of end-point crystal structures for a long time. Nevertheless, motion is

needed to effectuate the conformational changes. The analysis of crystallographic temperature-factors was used as a tool to to obtain insight into these processes [101; 102]. The evolution of solution NMR methodology to the point where motion-function correlation could be carried out for proteins in solution took a couple of decades.

More recent work has combined solution NMR techniques with computation to explore the role of dynamics in allostery and protein function [103]. The decoupling of molecular motions into small-amplitude motions and large, collective motions helps in identifying motional modes which are associated with conformational changes that are relevant to the functioning of the system under study [104]. To conclude this discussion, dynamics in biomoleculess is increasingly considered as a carrier for entropic component of the free energy changes, which drive allosteric transitions. Elucidating the connection between structure, dynamics, allosteric changes and function is one of the main goals of biophysics [11].

1.14 METHODS

We have used solution NMR spectroscopy as a tool in investigating molecular structure. At this juncture, it is apposite to briefly discuss the basic principles of NMR spectroscopy, especially as applied to proteins. Chapter 3 contains a more detailed exposition of the experimental tools used.

1.15 A BRIEF DISCUSSION OF THE BASIC PRINCIPLES OF NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

Certain isotopes of nuclei such as Hydrogen (^{1}H), Nitrogen (^{15}N), Carbon (^{13}C), Phosphorus (^{31}P), ubiquitous in biomolecules, have magnetic properties [105]. These properties are a result of the combination of charge and spin of all the protons and neutrons, which constitute the nucleus in question. Proton/neutron spin is an intrinsic

quantum mechanical property that cannot be understood in terms of classical electrodynamics [106]. Under the influence of a large external magnetic field (~18 Tesla, provided by a superconducting magnet), these nuclear magnetic moments will statistically align parallel to the field. Nuclei aligned parallel to the external field will have a lower energy than nuclei aligned antiparallel to the field, according to the energy expression:

$$E = -\vec{\mu}.\vec{B}_0 = -\mu B\cos(\theta)$$

Equation 1.1: Energy stored in a magnetic dipole

Here, θ is the angle between the nuclear spin and the external magnetic field. This leads to a split of the energy level (and at a deeper philosophical level, a breaking of space symmetry) for the spin-active nucleus. The nucleus aligned parallel to the external field is referred to as the 'up spin' and the nucleus aligned antiparallel as the 'down spin'. The energy difference (ΔE) between the up and down spins scales linearly with the strength of the external field (B0).

$$\Delta E = -\mu_z B_0 - (-\mu_z B_0) = -2\mu_z B_0$$
$$\Delta E = -2gm_S\hbar B_0 = -g\hbar B_0$$
$$\Delta E = \omega\hbar$$
$$\omega = -gB_0$$

Equation 1.2: Magnetic dipole energy to frequency of on resonance irradiation

This energy difference generates a Boltzmann distribution of nuclear spins with a slightly greater population at the lower energy level (ground state) than at the higher energy level (excited state) [3]. If the system is then exposed to electromagnetic radiation at a frequency ω, corresponding to this energy difference, then certain nuclei will be promoted from the ground state to the excited state.

$$\frac{N_{excited}}{N_{ground}} = e^{-(\Delta E/k_B T)}$$

Equation 1.3: Boltzmann distribution and spin temperature

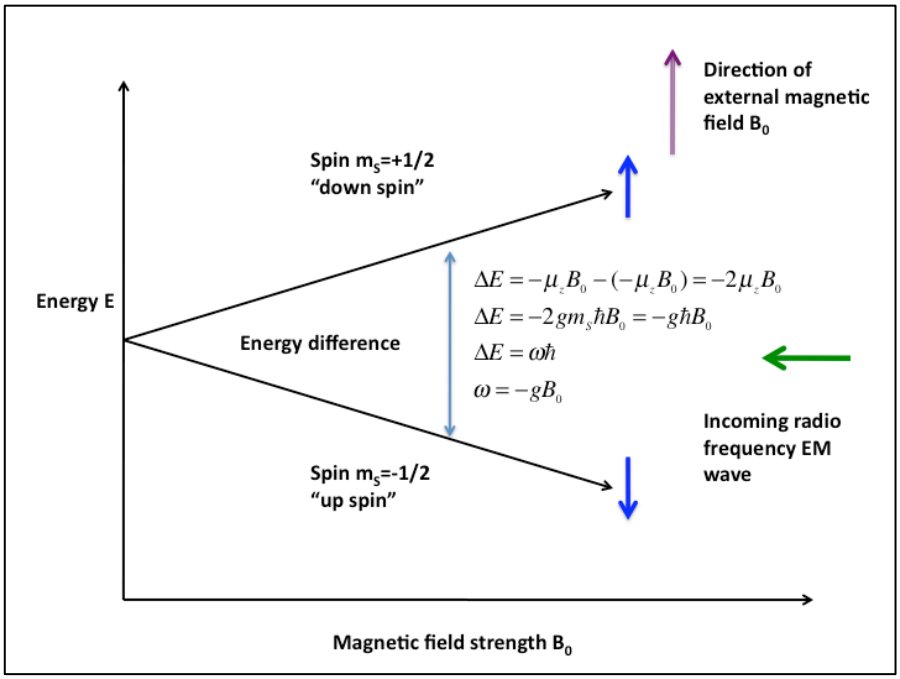

Figure 1.25: NMR fundamentals

These excited state nuclei will soon drop down to the ground state (a process called relaxation), leading to the original equilibrium. During the process of relaxation however, they also emit photons that correspond to the energy difference that these nuclei are transiting. The description of NMR up to this point has corresponded to a physicist's technique for investigating atomic nuclei.

What makes matters immeasurably more complex (and interesting) is the role that electron clouds play [107]. The electron clouds, which surround atomic nuclei are set into motion by the external magnet field, rotating around the nuclei so as to generate tiny currents which in turn create transient electromagnets which oppose the external magnetic field. This can be mathematically worked as a trivial application of Faraday's laws of electromagnetic induction [108]. The fact that these atomic magnets oppose the external field instead of adding on to them is a consequence of the law of energy conservation – the opposite situation would have led to a positive feedback of magnetic field strength, which we do not see in experiments (Lenz's law). But the fact

that these electron cloud generated magnets affect the external magnetic field also means that they affect the frequency of electromagnetic absorption and emission for different nuclei. The final piece in this puzzle is the fact that electron clouds are a reflection of the chemistry of a particular system. Thus, the characteristic frequency of every signal nuclei is, in fact a spectroscopic 'signature' of its chemical properties.

As we look at other factors which may influence the 'resonance frequency' (or signature) of a particular nucleus, we begin to uncover a treasure trove of physical and chemical parameters – direct electron-cloud mediated interaction with other nuclei, through-space interaction with other nuclei, degree of dynamics, solvent accessibility, and many more [105]. Thus, NMR offers a nonpareil tool in structural biophysics [107].

This is an extremely hand-waving description of the fundamentals of NMR spectroscopy, but a more in-depth discussion is out of the scope of this work. Chapter 3 contains a more elaborate analysis of the NMR experiments involved. We will conclude this section with the remark that most other spectroscopic techniques (such as fluorescence) deal with systems where the chemical nature of the species under investigation has created a large scale energy distribution which we proceed to perturb by external means (such as a laser), and gauge the system response. NMR spectroscopy is unique in that exposing the system to an external magnetic field creates the large-scale energy distribution. The chemistry of the system creates fine perturbations on this macroscopic energy difference, which we then proceed to analyze.

1.16 A MORE DETAILED EXPOSITION OF THE NMR METHODS USED IN THIS PROJECT

Here, we present a slightly more elaborate discussion of the NMR methods used in the determination of structural changes in the Hsp70 system. The standard *modus operandi* in any NMR based structural-biology project is the following:

1. Express and purify the protein in solution NMR conditions
2. Record NMR spectra for the protein under suitable conditions of ligand/nucleotide/substrate and assign the spectra
3. Compare spectra recorded under different sample conditions to determine chemical shift changes
4. Extract structural information from techniques like NOEs, RDCs, paramagnetic relaxation, etc.
5. Use structural information to develop a resultant (complex) structure of the protein system
6.Use NMR relaxation measurements to assess the dynamical parameters of the molecule

The first challenge is assigning the protein where a wide range of two and three dimensional NMR experiments are used [109; 110; 111; 112; 113; 114; 115; 116]. This is akin to the process of fingerprinting and can be considered the starting point of any NMR based investigation. Structural information from NOE (Nuclear Overhauser Effect) measurement is in the form of distance constraints [117; 118; 119; 120]. These can be used to refine protein structures [121]. Unfortunately, extracting NOEs for large (~50 kDa) proteins is difficult. Hence, we have chosen to use dipolar couplings and paramagnetic relaxation.

1.16.1 RESIDUAL DIPOLAR COUPLINGS IN STRUCTURAL ANALYSIS

The Hamiltonian for the dipolar coupling between two magnetic nuclei has the mathematical form

$$H_D = \frac{\hbar\gamma_I\gamma_S}{4\pi r_{IS}^3}\left(1-3\cos^2\theta\right)\left(3I_ZS_Z-\bar{I}.\bar{S}\right)$$

Equation 1.4: Dipolar coupling Hamiltonian

Here, γ's refer to gyromagnetic ratios of the nuclei in question, r_{IS} is the distance

between nuclei I and S, θ is the angle between the internuclear vector and the external magnetic field (see Fig. 1.26), vectors I and S refer to magnetization vectors for the two nuclei, and I_z/S_z refer to their Z-components.

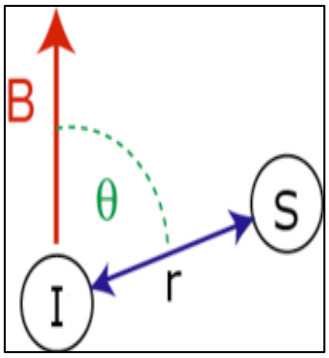

Figure 1.26: Two spin system

Under conditions of weak coupling (where the transverse components of the nuclear spins have different frequencies), this can be reduced to

$$H_D = \frac{\hbar\gamma_I\gamma_S}{4\pi r_{IS}^3}\left(1-3\cos^2\theta\right)\left(2I_ZS_Z\right)$$

Equation 1.5: Dipolar coupling Hamiltonian under the weak coupling approximation

From this equation, it is easy to identify that the eigenvalues of this Hamiltonian will behave like the J-coupling [105]. This can be written as:

$$D_{IS} = \frac{\hbar\gamma_I\gamma_S}{4\pi r_{IS}^3}\left(1-3\cos^2\theta\right)$$

Equation 1.6: The dipolar coupling term

Under normal NMR conditions, a 45 kDa protein tumbles isotropically in solution with a correlation time of about 22 nsec. This is fast enough to average out the orientation dependence of the dipolar coupling (which appears as the second order Legendre polynomial) to zero. That is: $\int P^2(\cos\theta)\sin\theta\, d\theta = 0$, using isotropic distribution of orientations.

However, if isotropic conditions are disturbed, by say the addition of alignment media in solution, the dipolar coupling no longer averages to zero [122; 123]. The

remnant of the dipolar coupling (which does not average to zero) is called the residual dipolar coupling (RDC). Typically it is between 0.01 and .1% of the dipolar coupling. For the NH spin system, that is between 2 and 20 Hz.It can be demonstrated that its contribution to the overall NMR Hamiltonian is mathematically similar to the J-coupling. Thus, the same experiments that can be used to extract J-couplings can be adapted to extract RDCs [124; 125]. As can be seen from the preceding equations, the RDC between two nuclei (typically ^{1}H and ^{15}N) on a peptide backbone is dependant on the orientation of the internuclear vector w.r.t the external field. Thus, if the RDC is measured for a sufficient number of such spin pairs, and it is known how these spin pairs are embedded within a molecular fragment, the 0.01% time-averaged orientation of the molecular fragment itself w.r.t. the magnetic field can thereby be obtained. Streamlining this process has been the subject of intensive work [126; 127; 128; 129; 130]. Today, RDCs are increasingly being used as a powerful tool to probe structural changes in large biomolecules [131; 132; 133].

1.16.2 PARAMAGNETIC RELAXATION IN STRUCTURAL STUDIES

Paramagnetic Relaxation Enhancement (PRE) is the NMR technique that we have used to determine the complex interaction between DnaK and the J-protein system. It is worthwhile to briefly touch upon this subject. More detailed discussions are included in Chapter 4 of this dissertation.

A paramagnetic centre in a protein is a large spin magnet, which will generate additional interaction terms in the spin Hamiltonian of any nearby nuclear spin. As the paramagnetic centre is itself a spin dipole, which switches between the up and down states (w.r.t the external magnetic field) very rapidly, this generates a relaxation mechanism for neighbouring nuclear spins. Magnetization due to these nuclear spins thus decays much faster, which is manifested as a much larger R_2 (transverse

relaxation rate). In practical terms, this is seen as the disappearance of crosspeaks (on an NMR spectrum) corresponding to residues that are close to the paramagnetic centre. This phenomenon is hence called paramagnetic relaxation enhancement. It provides distance restraints at ranges longer (~15-30 Angstrom) than are available from NOEs (~5 Angstrom) [134; 135; 136]. These restraints are then used to accurately determine the structures of protein complexes.

Possible paramagnetic centres include lanthanides or other metals[137; 138] that can be chelated to proteins or nucleic acids by EDTA, or nitroxide labels, such as MTSL. Engineering a suitable paramagnetic tag on the protein under investigation without affecting its biochemical function is not always easy [139]. This is where surface exposed cysteine residues are frequently used to attach MTSL tags by disulphide chemistry. Paramagnetic tags are also used to explore transient intermediates in molecular interactions [140]. This technique continues to find application in different areas of structural biophysics[141; 142] and is one of the most promising avenues to investigate interaction in large molecular systems.

With these techniques (RDCs and PREs), we have investigated structural changes in the Hsp70 system and its interaction with the J-protein system.

1.17 REFERENCES

[1] D.L.N. Albert L. Lehninger, Michael M. Cox, Lehninger Principles of Biochemistry, Fourth Edition, 2005.
[2] C.B. Anfinsen, Principles That Govern Folding of Protein Chains. Science 181 (1973) 223-230.
[3] P.E. Leopold, M. Montal, and J.N. Onuchic, Protein folding funnels: a kinetic approach to the sequence-structure relationship. Proc Natl Acad Sci U S A 89 (1992) 8721-5.

[4] R.K. Pathria, Statistical mechanics, 2nd ed. (1996).
[5] A.L. Fink, Natively unfolded proteins. Curr Opin Struct Biol 15 (2005) 35-41.
[6] M. Karplus, and J.A. McCammon, The dynamics of proteins. Sci Am 254 (1986) 42-51.
[7] M. Karplus, and J.A. McCammon, Dynamics of proteins: elements and function. Annu Rev Biochem 52 (1983) 263-300.
[8] K. Henzler-Wildman, and D. Kern, Dynamic personalities of proteins. Nature 450 (2007) 964-72.
[9] J. Cavanagh, and R.A. Venters, Protein dynamic studies move to a new time slot. Nat Struct Biol 8 (2001) 912-4.
[10] D. Kern, and E.R. Zuiderweg, The role of dynamics in allosteric regulation. Curr Opin Struct Biol 13 (2003) 748-57.
[11] V.I. Abkevich, A.M. Gutin, and E.I. Shakhnovich, Specific nucleus as the transition state for protein folding: evidence from the lattice model. Biochemistry 33 (1994) 10026-36.
[12] H.S. Chan, S. Bromberg, and K.A. Dill, Models of cooperativity in protein folding. Philos Trans R Soc Lond B Biol Sci 348 (1995) 61-70.
[13] K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas, and H.S. Chan, Principles of protein folding--a perspective from simple exact models. Protein Sci 4 (1995) 561-602.
[14] A.R. Fersht, Nucleation mechanisms in protein folding. Curr Opin Struct Biol 7 (1997) 3-9.
[15] M.J. Gething, and J. Sambrook, Protein folding in the Cell. Nature 355 (1992) 33-45.
[16] S.W. Englander, L. Mayne, and M.M. Krishna, Protein folding and misfolding: mechanism and principles. Q Rev Biophys 40 (2007) 287-326.
[17] F.U. Hartl, and M. Hayer-Hartl, Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol 16 (2009) 574-81.
[18] V.I. Abkevich, A.M. Gutin, and E.I. Shakhnovich, Domains in folding of model proteins. Protein Sci 4 (1995) 1167-77.
[19] F.U. Hartl, and M. Hayer-Hartl, Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295 (2002) 1852-8.
[20] C.M. Dobson, Protein folding and misfolding. Nature 426 (2003) 884-90.
[21] B. Bukau, J. Weissman, and A. Horwich, Molecular chaperones and protein quality control. Cell 125 (2006) 443-51.
[22] J. Lu, and C. Deutsch, Folding zones inside the ribosomal exit tunnel. Nat Struct Mol Biol 12 (2005) 1123-9.
[23] C.M. Kaiser, H.C. Chang, V.R. Agashe, S.K. Lakshmipathy, S.A. Etchells, M. Hayer-Hartl, F.U. Hartl, and J.M. Barral, Real-time observation of trigger factor function on translating ribosomes. Nature 444 (2006) 455-60.
[24] A. Yonath, Large facilities and the evolving ribosome, the cellular machine for genetic-code translation. J R Soc Interface 6 Suppl 5 (2009) S575-85.

[25] L. Ferbitz, T. Maier, H. Patzelt, B. Bukau, E. Deuerling, and N. Ban, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. Nature 431 (2004) 590-6.
[26] A.V. Ludlam, B.A. Moore, and Z. Xu, The crystal structure of ribosomal chaperone trigger factor from Vibrio cholerae. Proc Natl Acad Sci U S A 101 (2004) 13436-41.
[27] S. Vorderwulbecke, G. Kramer, F. Merz, T.A. Kurz, T. Rauch, B. Zachmann-Brand, B. Bukau, and E. Deuerling, Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK. FEBS Lett 559 (2004) 181-7.
[28] A. Tissieres, H.K. Mitchell, and U.M. Tracy, Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J Mol Biol 84 (1974) 389-98.
[29] M.J. Schlesinger, Heat shock proteins. J Biol Chem 265 (1990) 12111-4.
[30] T. Hesterkamp, and B. Bukau, Identification of the prolyl isomerase domain of Escherichia coli trigger factor. FEBS Lett 385 (1996) 67-71.
[31] T. Hesterkamp, and B. Bukau, The Escherichia coli trigger factor. FEBS Lett 389 (1996) 32-4.
[32] G. Kramer, D. Boehringer, N. Ban, and B. Bukau, The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat Struct Mol Biol 16 (2009) 589-97.
[33] H. Schroder, T. Langer, F.U. Hartl, and B. Bukau, DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J 12 (1993) 4137-44.
[34] C.P. Georgopoulos, R.W. Hendrix, S.R. Casjens, and A.D. Kaiser, Host participation in bacteriophage lambda head assembly. J Mol Biol 76 (1973) 45-60.
[35] J. Zeilstra-Ryalls, O. Fayet, and C. Georgopoulos, The universally conserved GroE (Hsp60) chaperonins. Annu Rev Microbiol 45 (1991) 301-25.
[36] L. Ditzel, J. Lowe, D. Stock, K.O. Stetter, H. Huber, R. Huber, and S. Steinbacher, Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell 93 (1998) 125-38.
[37] T. Langer, C. Lu, H. Echols, J. Flanagan, M.K. Hayer, and F.U. Hartl, Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356 (1992) 683-689.
[38] J. Frydman, E. Nimmesgern, K. Ohtsuka, and F.U. Hartl, Folding of Nascent Polypeptide-Chains in a High-Molecular-Mass Assembly with Molecular Chaperones. Nature 370 (1994) 111-117.
[39] Z. Xu, A.L. Horwich, and P.B. Sigler, The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature 388 (1997) 741-50.
[40] W.A. Fenton, and A.L. Horwich, GroEL-mediated protein folding. Protein Sci. 6 (1997) 743-760.
[41] A.L. Horwich, K.B. Low, W.A. Fenton, I.N. Hirshfield, and K. Furtak, Folding in vivo of bacterial cytoplasmic proteins: role of GroEL. Cell 74 (1993) 909-917.

[42] P.B. Sigler, Z. Xu, H.S. Rye, S.G. Burston, W.A. Fenton, and A.L. Horwich, Structure and function in GroEL-mediated protein folding. Annu Rev Biochem 67 (1998) 581-608.
[43] Z. Xu, and P.B. Sigler, GroEL/GroES: structure and function of a two-stroke folding machine. J Struct Biol 124 (1998) 129-41.
[44] Y.C. Tang, H.C. Chang, A. Roeben, D. Wischnewski, N. Wischnewski, M.J. Kerner, F.U. Hartl, and M. Hayer-Hartl, Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. Cell 125 (2006) 903-14.
[45] J. Ma, P.B. Sigler, Z. Xu, and M. Karplus, A dynamic model for the allosteric mechanism of GroEL. J Mol Biol 302 (2000) 303-13.
[46] M.J. Kerner, D.J. Naylor, Y. Ishihama, T. Maier, H.C. Chang, A.P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, and F.U. Hartl, Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell 122 (2005) 209-20.
[47] A. Brinker, G. Pfeifer, M.J. Kerner, D.J. Naylor, F.U. Hartl, and M. Hayer-Hartl, Dual function of protein confinement in chaperonin-assisted protein folding. Cell 107 (2001) 223-33.
[48] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-366.
[49] K.M. Flaherty, C. Delucaflaherty, and D.B. McKay, 3-DIMENSIONAL STRUCTURE OF THE ATPASE FRAGMENT OF A 70K HEAT-SHOCK COGNATE PROTEIN. Nature 346 (1990) 623-628.
[50] C. DeLuca-Flaherty, K.M. Flaherty, L.J. McIntosh, B. Bahrami, and D.B. McKay, Crystals of an ATPase fragment of bovine clathrin uncoating ATPase. J Mol Biol 200 (1988) 749-50.
[51] T.G. Chappell, B.B. Konforti, S.L. Schmid, and J.E. Rothman, The ATPase core of a clathrin uncoating protein. J Biol Chem 262 (1987) 746-51.
[52] T.G. Chappell, W.J. Welch, D.M. Schlossman, K.B. Palter, M.J. Schlesinger, and J.E. Rothman, Uncoating ATPase is a member of the 70 kilodalton family of stress proteins. Cell 45 (1986) 3-13.
[53] S.M. Wilbanks, C. DeLuca-Flaherty, and D.B. McKay, Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. I. Kinetic analyses of active site mutants. J Biol Chem 269 (1994) 12893-8.
[54] K.M. Flaherty, S.M. Wilbanks, C. DeLuca-Flaherty, and D.B. McKay, Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. J Biol Chem 269 (1994) 12899-907.
[55] M.C. O'Brien, K.M. Flaherty, and D.B. McKay, Lysine 71 of the chaperone protein Hsc70 Is essential for ATP hydrolysis. J. Biol. Chem. 271 (1996) 15874-15878.
[56] D.R. Palleros, K.L. Reid, L. Shi, W.J. Welch, and A.L. Fink, ATP-induced protein-Hsp70 complex dissociation requires K+ but not ATP hydrolysis. Nature 365 (1993) 664-6.
[57] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-66.

[58] R.C. Morshauser, H. Wang, G.C. Flynn, and E.R. Zuiderweg, The peptide-binding domain of the chaperone protein Hsc70 has an unusual secondary structure topology. Biochemistry 34 (1995) 6261-6.
[59] G.M. Clore, and A.M. Gronenborn, Applications of three- and four-dimensional heteronuclear NMR spectroscopy to protein structure determination. Progr. Nucl. Magn. Reson. Spectrosc. 23 (1991) 43-92.
[60] H. Wang, and E.R. Zuiderweg, HCCH-TOCSY spectroscopy of 13C-labeled proteins in H2O using heteronuclear cross-polarization and pulsed-field gradients. J Biomol NMR 5 (1995) 207-11.
[61] S.W. Van Doren, A.V. Kurochkin, Q.Z. Ye, L.L. Johnson, D.J. Hupe, and E.R.P. Zuiderweg, Assignments for the main chain nuclear magnetic resonances and delineation of the secondary structure of the catalytic domain of human stromelysin-1 as obtained from triple resonance 3D NMR experiments. Biochemistry 32 (1993) 13109-13122.
[62] A. Bax, and S. Grzesiek, Methodological Advances in Protein NMR. Acc. Chem. Res. 26 (1993) 131-138.
[63] S. Rudiger, L. Germeroth, J. Schneider-Mergener, and B. Bukau, Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. Embo J 16 (1997) 1501-7.
[64] A. Gragerov, Z. Li, Z. Xun, W. Burkholder, and M.E. Gottesman, Specificity of DnaK Peptide Binding. J. Mol. Biol. 235 (1994) 848-854.
[65] A. Gragerov, and M.E. Gottesman, Different peptide binding specificities of hsp70 family members. J Mol Biol 241 (1994) 133-5.
[66] R.C. Morshauser, W. Hu, H. Wang, Y. Pang, G.C. Flynn, and E.R. Zuiderweg, High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. J Mol Biol 289 (1999) 1387-403.
[67] X.T. Zhu, X. Zhao, W.F. Burkholder, A. Gragerov, C.M. Ogata, M.E. Gottesman, and W.A. Hendrickson, Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272 (1996) 1606-1614.
[68] S. Sadis, and L.E. Hightower, Unfolded proteins stimulate molecular chaperone Hsc70 ATPase by accelerating ADP/ATP exchange. Biochemistry 31 (1992) 9406-9412.
[69] K. Liberek, D. Skowyra, M. Zylicz, C. Johnson, and C. Georgopoulos, The Escherichia coli DnaK chaperone, the 70-kDa heat shock protein eukaryotic equivalent, changes conformation upon ATP hydrolysis, thus triggering its dissociation from a bound target protein. J Biol Chem 266 (1991) 14491-6.
[70] J.H. Ha, and D.B. McKay, ATPase kinetics of recombinant bovine 70 kDa heat shock cognate protein and its amino-terminal ATPase domain. Biochemistry 33 (1994) 14625-14635.
[71] H. Theyssen, H.P. Schuster, L. Packschies, B. Bukau, and J. Reinstein, The second step of ATP binding to DnaK induces peptide release. J. Mol. Biol. 263 (1996) 657-670.
[72] G.C. Flynn, T.G. Chappell, and J.E. Rothman, Peptide binding and release by proteins implicated as catalysts of protein assembly. Science 245 (1989) 385-90.

[73] B. Gao, L. Greene, and E. Eisenberg, Characterization of nucleotide-free uncoating ATPase and its binding to ATP, ADP, and ATP analogues. Biochemistry 33 (1994) 2048-54.
[74] K. Liberek, J. Marszalek, D. Ang, C. Georgopoulos, and M. Zylicz, Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proc Natl Acad Sci U S A 88 (1991) 2874-8.
[75] M. Zylicz, D. Ang, K. Liberek, and C. Georgopoulos, Initiation of lambda DNA replication with purified host- and bacteriophage-encoded proteins: the role of the dnaK, dnaJ and grpE heat shock proteins. Embo J 8 (1989) 1601-8.
[76] S. Wickner, J. Hoskins, and K. McKenney, Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA. Nature 350 (1991) 165-7.
[77] S. Wickner, J. Hoskins, and K. McKenney, Monomerization of RepA dimers by heat shock proteins activates binding to DNA replication origin. Proc Natl Acad Sci U S A 88 (1991) 7903-7.
[78] K. Liberek, and C. Georgopoulos, Autoregulation of the Escherichia coli heat shock response by the DnaK and DnaJ heat shock proteins. Proc Natl Acad Sci U S A 90 (1993) 11019-23.
[79] T. Szyperski, M. Pellecchia, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure determination of the Escherichia coli DnaJ molecular chaperone: secondary structure and backbone fold of the N-terminal region (residues 2-108) containing the highly conserved J domain. Proc Natl Acad Sci U S A 91 (1994) 11343-7.
[80] M. Pellecchia, T. Szyperski, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure of the J-domain and the Gly/Phe-rich region of the Escherichia coli DnaJ chaperone. J. Mol. Biol. 260 (1996) 236-250.
[81] P.A. Silver, and J.C. Way, Eukaryotic DnaJ homologs and the specificity of Hsp70 activity. Cell 74 (1993) 5-6.
[82] C. Alfano, and R. McMacken, Heat shock protein-mediated disassembly of nucleoprotein structures is required for the initiation of bacteriophage lambda DNA replication. J Biol Chem 264 (1989) 10709-18.
[83] D. Wall, M. Zylicz, and C. Georgopoulos, The NH2-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. J Biol Chem 269 (1994) 5446-51.
[84] S. Ahle, and E. Ungewickell, Auxilin, a newly identified clathrin-associated protein in coated vesicles from bovine brain. J Cell Biol 111 (1990) 19-29.
[85] J. Jiang, A.B. Taylor, K. Prasad, Y. Ishikawa-Brush, P.J. Hart, E.M. Lafer, and R. Sousa, Structure-function analysis of the auxilin J-domain reveals an extended Hsc70 interaction interface. Biochemistry 42 (2003) 5748-53.
[86] S.E. Holstein, H. Ungewickell, and E. Ungewickell, Mechanism of clathrin basket dissociation: separate functions of protein domains of the DnaJ homologue auxilin. J Cell Biol 135 (1996) 925-37.
[87] E. Ungewickell, H. Ungewickell, S.E. Holstein, R. Lindner, K. Prasad, W. Barouch, B. Martin, L.E. Greene, and E. Eisenberg, Role of auxilin in uncoating clathrin-coated vesicles. Nature 378 (1995) 632-5.

[88] E. Ungewickell, H. Ungewickell, and S.E. Holstein, Functional interaction of the auxilin J domain with the nucleotide- and substrate-binding modules of Hsc70. J Biol Chem 272 (1997) 19594-600.
[89] L. Packschies, H. Theyssen, A. Buchberger, B. Bukau, R.S. Goody, and J. Reinstein, GrpE accelerates nucleotide exchange of the molecular chaperone DnaK with an associative displacement mechanism. Biochemistry 36 (1997) 3417-22.
[90] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-5.
[91] J.F. Swain, G. Dinler, R. Sivendran, D.L. Montgomery, M. Stotz, and L.M. Gierasch, Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol Cell 26 (2007) 27-39.
[92] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc Natl Acad Sci U S A 106 (2009) 8471-6.
[93] Q. Liu, and W.A. Hendrickson, Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-20.
[94] J. Jiang, E.G. Maes, A.B. Taylor, L. Wang, A.P. Hinck, E.M. Lafer, and R. Sousa, Structural basis of J cochaperone binding and regulation of Hsp70. Mol Cell 28 (2007) 422-33.
[95] W.C. Suh, W.F. Burkholder, C.Z. Lu, X. Zhao, M.E. Gottesman, and C.A. Gross, Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. Proc Natl Acad Sci U S A 95 (1998) 15223-8.
[96] M.H. Levitt, Spin Dynamics: Basic Principles of NMR Spectroscopy. (2001).
[97] A.L. Ghatak, S, Quantum Mechanics: Theory And Applications. (2004).
[98] W.J.F. John Cavanagh, Arthur G. Palmer III, Nicholas J. Skelton, Protein NMR Spectroscopy: Principles and Practice. (2007).
[99] D.J. Griffiths, Introduction to Electrodynamics (3rd Edition). (1999).

CHAPTER 2 ALLOSTERY IN THE HSP70 CHAPERONES IS TRANSDUCED BY SUBDOMAIN ROTATIONS.

Most of the material in chapter has been published in J Mol Biol. 2009 May 8;388(3):475-90. Epub 2009 Feb 4.

2.1 ABSTRACT

Hsp70s are central to protein folding, refolding, and trafficking in organisms ranging from *Archaea* to *Homo sapiens*, both at normal and at stressed cellular conditions. Hsp70s are comprised of a nucleotide-binding domain (NBD), and a substrate-binding domain (SBD). The nucleotide binding site in the NBD and the substrate-binding site in the SBD are allosterically linked: ADP binding promotes substrate binding, while ATP binding promotes substrate release. Hsp70s have been linked to inhibition of apoptosis, i.e., cancer, and diseases associated with protein misfolding such as Alzheimer's, Parkinson's, and Huntington's.

It has long been a goal to characterize the nature of the allosteric coupling in these proteins. However, earlier studies of the isolated NBD could not show any difference in overall conformation between the ATP and ADP state. Hence the question: how is the state of the nucleotide communicated between NBD and SBD?

Here we report a solution NMR study of the 44 kDa NBD of the Hsp70 from *Thermus thermophilus,* in the ADP and AMP-PNP states. Using the solution NMR methods of residual dipolar coupling analysis, we determine that significant rotations occur for the different subdomains of the NBD upon exchange of nucleotide. These rotations modulate the access to the nucleotide-binding cleft, in absence of a nucleotide exchange factor. Moreover, the rotations cause a change in accessibility of a hydrophobic surface cleft remote from the nucleotide-binding site, which previously has been identified as essential to the allosteric communication between NBD and

SBD. We propose that it is this change in the NBD surface cleft that constitutes the allosteric signal that can be recognized by the SBD.

2.2 INTRODUCTION

Hsp70s (heat shock 70 kDa) chaperone proteins are central to protein folding, refolding, and trafficking in organisms ranging from *Archaea* to *Homo sapiens*, both at normal and at stressed cellular conditions [48]. The Hsp70's recognize misfolded substrate proteins by their exposed hydrophobic amino acids. In a process that requires ATP and two co-chaperones [18; 143], the Hsp70's unfold the substrate proteins by cycles of binding and release. The unfolded proteins then refold on their own accord. Much of the mechanism by which this takes places remains unknown. Recently, Hsp70s have been linked to cancer [144; 145; 146; 147] and diseases associated with protein misfolding such as Alzheimer's [148], Parkinson's [149] , and Huntington's [150; 151]. It has been suggested that modulation of Hsp70 activity with small compounds may form an avenue to treat these diseases [152].

Hsp70s are comprised of two main domains: a 44 kDa N-terminal nucleotide-binding domain (NBD), and a 25 kDa substrate-binding domain (SBD) that harbors the substrate-binding site. The Hsp70's are allosteric molecules: with ADP bound to the NBD, substrate binds tightly to the SBD (KD=1.5x10-7 M); ATP binding to the NBD promotes substrate release from the SBD (KD=4.6x10-5 M) [153; 154].

While several high-resolution crystal and solution structures are available for the NBD [54; 155; 156] and SBD [58; 66; 67; 157; 158; 159] domains of the Hsp70's of several species, no high-resolution structure for a complete wild-type Hsp70 has been reported. Hence, the allosteric mechanism is not yet known. However, several recent studies have shown that a conserved, hydrophobic and flexible linker (residues 388VQDLLLLDVTP (Hsc70 *H. sapiens*), 380VRDVVLLDVTP (DnaK *T. thermophilus*)) loosely tethers the NBD and SBD in the ADP-substrate-bound state [91; 160; 161]. In

addition, it is anticipated that NBD and SBD are docked together in the ATP state [91; 162], but this still needs to be demonstrated for the Hsp70s with an actual structure.

2.2.1 CURRENT MODEL FOR THE ALLOSTERIC MECHANISM OF Hsp70s

Currently, the following overall model for the allosteric mechanism seems credible: with substrate bound in the SBD and ADP.Pi in the NBD, the SBD collides repeatedly with the NBD, but does not dock [161]. Upon ATP binding to the NBD, a change takes place in the NBD such that the collisions lead to stable docking of SBD and NBD. The docking event propagates through the SBD and leads to release of substrate. A key feature of this model is that the NBD conformational change upon ADP --> ATP change takes place while the SBD is loosely tethered.

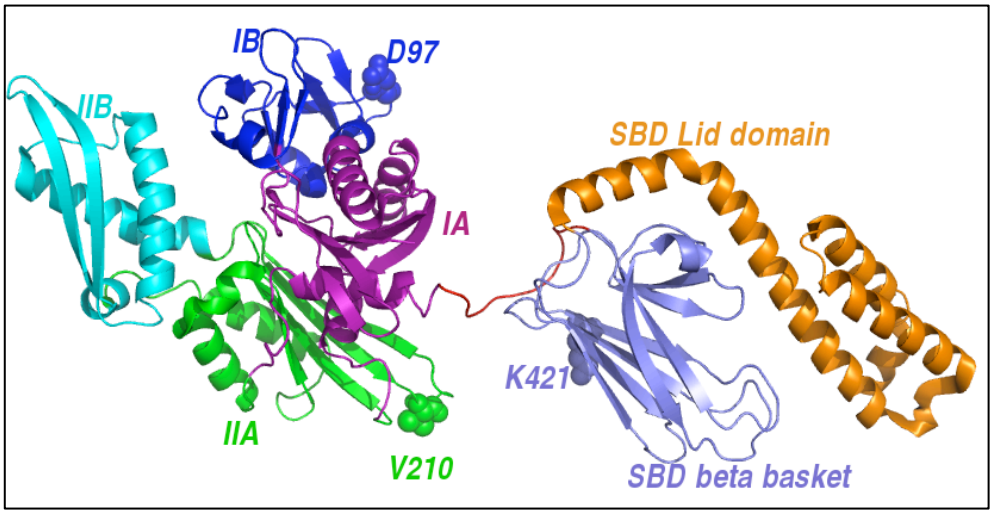

Figure 2.1: Collision model of the ADP.Pi state of DnaK based on RDC solution NMR spectroscopy

Hence, it is expected that the isolated NBD should also show such changes between different nucleotide states.

2.2.2 CRYSTAL STRUCTURES OF THE NUCLEOTIDE BINDING DOMAIN DO NOT SHOW ANY CHANGE AS Hsp70 TURNS OVER NUCLEOTIDE

As has been just discussed, we expect to see some motion as the NBR turns over nucleotide. This should show up as conformational change in different nucleotide

states of the NBD. However, this notion is seriously challenged by X-ray diffraction studies of the isolated NBD. They show no obvious difference in overall conformation of the NBD of bovine (identical to human) Hsc70 between the ATP and ADP state [55; 155] (see Figure 2.2) in the crystalline state. Indeed, the various efforts that have gone into solving crystal structures of the NBD have yielded remarkably similar structures. In addition, small-angle X-ray scattering methods could not detect a difference in the radius of gyration for the ATP and ADP state of the isolated NBD of Hsc70 in solution either [163], apparently supporting the crystallographic studies.

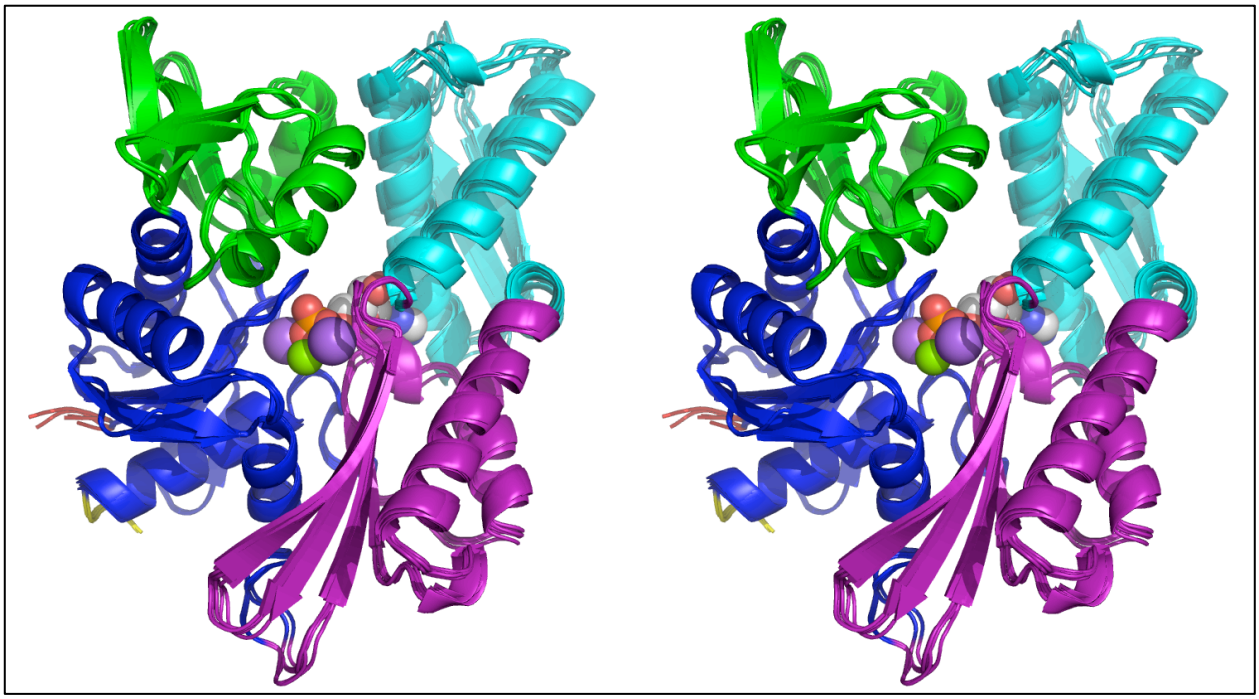

Figure 2.2: Stereograph of crystal structures of different nucleotide states of the Hsp70 NBD

Stereograph of the superposition of five X-ray crystallography structures for bovine Hsc70 NBD (4-380). The proteins were crystallized in the following states: wt-APO (2QW9.pdb, chain A), wt-ADP.PO4 (3HSC.pdb and 2QWL.pdb, chain A), wt-ADP.VO4 (2QWM.pdb, chain A) and K71M-ATP (1KAX.pdb). The proteins were overlaid on the secondary structure elements of sub domain IA (Hsc70-count residues 1-39,116-188, 361-381; blue) to accentuate potential changes in sub domains IB (residues 40-115; green), IIA (residues 189-228, 307-360; purple) and IIB (residues 229-306; cyan). The N-terminus is in red; the C-terminus of this domain is shown in yellow. The nucleotide, PO43-, Mg2+ and

two Na+ present in 3HSC.pdb are shown in space fill. The figure was prepared with MacPyMOL.

An investigation of the different crystal structures of Hsp70 reveals that when the NBD is crystallized in isolation, the structures are remarkably similar regardless of nucleotide state. We have looked at several distinct "families" [55; 164; 165; 166], which incorporate different nucleotide states, different mutational forms, one structure (1YUW [167]) which contains part of the SBD β sandwich domain, and a family of structures of the NBD disulphide linked to bovine Auxilin J-domain [94].

Family	PDB	MOLECULE	GROUP	NUCLEOTIDE STATE	MUTATIONS	LENGTH	SPECIES	RELEASE DATE	RESOL (ANGSTR)
1	3HSC	Hsc70 NBD	McKay	ADP.Pi	WT	386	Bovine	03/24/95	1.93
	2QW9	Hsc70 NBD + part of linker	Sousa	*apo*	WT	394	Bovine	08/10/07	1.85
	2QWL	Hsc70 NBD + part of linker	Sousa	ADP	WT	394	Bovine	08/11/07	1.75
2	2QWM	Hsc70 NBD + part of linker	Sousa	$ADP.VO_4$	WT	396	Bovine	08/12/07	1.86
3	1KAX	Hsc70 NBD	McKay	ADP	K71M	381	Bovine	04/15/96	1.7
4	1YUW	Hsc70 NBD + linker + beta-sandwich region of SBD	Sousa	*apo*	E213A, D214A	554	Bovine	02/14/05	2.6
5	1ATR	Hsc70 NBD	McKay	ADP	WT	386	Bovine	08/09/93	2.34
	2QWN	Hsc70 NBD disulphide linked to Bovine Auxilin J-domain	Sousa	ADP.Pi	R171C	394: NBD, 94: auxilin	Bovine	08/10/07	2.4
	2QWQ	Hsc70 NBD disulphide linked to Bovine Auxilin J-domain	Sousa	AMPPNP hydrolyzed	R171C	396: NBD, 92: auxilin	Bovine	08/10/07	2.21
6	2QWR	Hsc70 NBD disulphide linked to Bovine Auxilin J-domain	Sousa	AMPPNP intact	R171C	397: NBD, 92: auxilin	Bovine	08/10/07	2.21
7	1HPM	Hsc70 NBD	McKay	ADP.Pi	WT	386	Bovine	03/24/95	1.7

Figure 2.3: Different crystal structures used to evaluate changes in the NBD

Different X-ray crystallography based structures of bovine Hsc70 NBD constructs in various nucleotide states, different mutational forms, etc [55; 94; 164; 166; 167; 168]

These structures were aligned based on backbone atoms from NBD residues using Pymol. The results are shown in the matrix in Fig 2.4. Three results are immediately obvious:

1. Overall backbone alignment RMSD across families is remarkably small, indicating that the changes in the NBD conformation are minimal.

2. Within a particular family (say NBD + part of the linker = 394 residues,

represented by 2QW9, 2QWL and 2QWM), the backbone alignment RMSD is always less than 0.34 Angstroms, indicating that for a particular construct, the nucleotide state does not matter at all: the NBD is always the same.

However, X-ray studies of NBD's in complex with nucleotide exchange factors show that the NBD is capable of major conformational changes.

3. The only structure where the NBD is different from other structures is 1YUW: this is 554 residue double mutated structure solved by Sousa and co-workers [167]. This is indicative of the fact that changes in the NBD are coupled to changes in the SBD; this two-way allostery has been introduced in the Chapter 1. Thus, in order to see changes in the molecule, it is preferable to work with the full length Hsp70, or failing that, use conditions where possible conformation changes are not 'frozen out', by, say crystal packing forces.

	3HSC	2QW9	2QWL	2QWM	1KAX	*1YUW*	1ATR	2QWN	2QWQ	2QWR	1HPM
3HSC	0.000	0.593	0.453	0.420	0.369	***0.949***	0.207	0.358	0.372	0.374	0.331
2QW9	0.593	0.000	0.272	0.331	0.493	***0.828***	0.579	0.662	0.667	0.664	0.580
2QWL	0.453	0.272	0.000	0.121	0.368	***0.951***	0.462	0.450	0.495	0.492	0.405
2QWM	0.420	0.331	0.121	0.000	0.345	***0.962***	0.427	0.412	0.463	0.459	0.350
1KAX	0.369	0.493	0.368	0.345	0.000	***0.888***	0.393	0.399	0.424	0.425	0.215
1YUW	***0.949***	***0.828***	***0.951***	***0.962***	***0.888***	***0.000***	***1.009***	***1.020***	***1.029***	***1.026***	***0.982***
1ATR	0.207	0.579	0.462	0.427	0.393	***1.009***	0.000	0.404	0.436	0.436	0.359
2QWN	0.358	0.662	0.450	0.412	0.399	***1.020***	0.404	0.000	0.245	0.243	0.287
2QWQ	0.372	0.667	0.495	0.463	0.424	***1.029***	0.436	0.245	0.000	0.049	0.342
2QWR	0.374	0.664	0.492	0.459	0.425	***1.026***	0.436	0.243	0.049	0.000	0.347
1HPM	0.331	0.580	0.405	0.350	0.215	***0.982***	0.359	0.287	0.342	0.347	0.000

Figure 2.4: Backbone alignment RMSD (Angstrom) of NBD residues from different crystal structures

The results from crystallography seem to indicate that the NBD does not undergo any conformational change as it turns over nucleotide. However, analysis of the NBD solution NMR conditions lead to a different conclusion.

2.2.3 CHANGES IN Hsp70 NUCLEOTIDE BINDING DOMAIN CHEMICAL SHIFTS SEEN VIA SOLUTION NMR EXPERIMENTS

We report here that subtle but extensive chemical shift changes occur in the NMR spectra of the isolated NBD of Hsc70 between the ADP.Pi and ATP state in neutral aqueous solution (see Fig 2.5). Clearly, change in nucleotide causes changes

in shifts that extend far from the nucleotide site. This, indeed, suggests that the isolated NBD changes its conformation upon nucleotide change, even in the absence of the SBD.

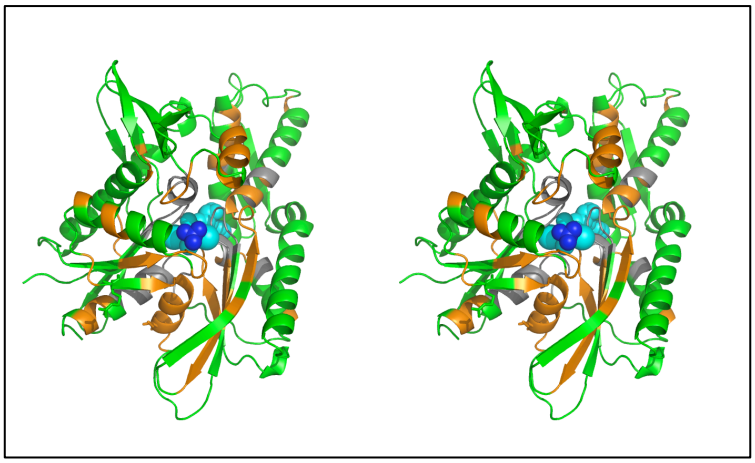

Figure 2.5: Chemical shift differences between different nucleotide states of Bovine Hsc70 NBD in solution

^{15}N-^{1}H chemical shift differences between the ATP and ADP.Pi conformation of Hsc-70-NBD. Orange: significant shift, green: no shift, grey: not known. Residues R171 (R167 in DnaK E. coli) and I181 (L177 in DnaK E. coli) are rendered as sticks.

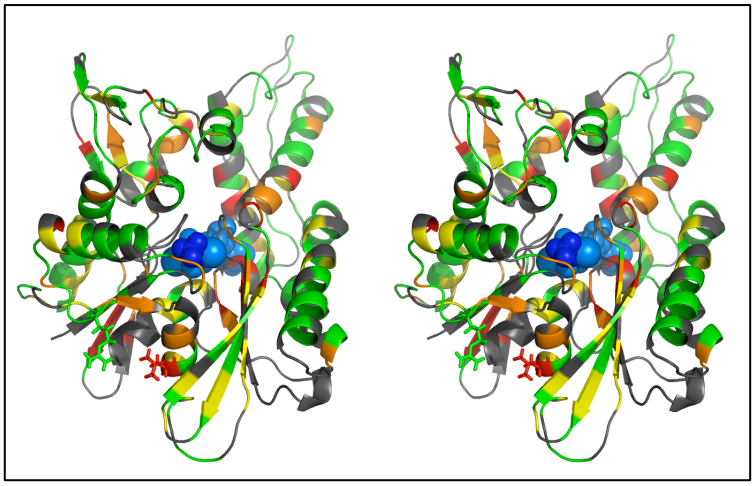

Figure 2.6: Chemical shift differences between different nucleotide states of Thermus thermophilus DnaK NBD in solution

^{15}N-^{1}H chemical shift differences ($\Delta\delta$) between the AMPPNP and ADP.Pi conformation of TTh-NBD. Red: $\Delta\delta > 2\sigma\Delta\delta$; Orange: $\sigma\,\Delta\delta < \Delta\delta < 2\,\sigma\,\Delta\delta$; yellow:

0.5 σ Δδ < Δδ < σ Δδ; green: Δδd < 0.5 σ Δδ ; grey: not known, where Δδ = ((ΔδH^2 + ΔδN^2))$^{1/2}$. ADP is in light blue, PO_4^{3-} in dark blue. Residues R164 (R167 in DnaK E. coli) and L174 (L177 in DnaK E. coli) are rendered as sticks.

2.3 MOTIVATING THIS PROBLEM

Summarizing the situation as discussed so far: the evidence from crystallography indicates that there are no changes in the Nucleotide Binding Domain when it is studied in isolation. There may be changes when it is observed in the context of a larger construct of Hsp70, which contains part of the SBD. In addition, X-ray studies of NBD's in complex with nucleotide exchange factors show that the NBD is capable of major conformational changes. However, solution NMR indicates that there are changes in chemical shift between different nucleotide states of the bovine Hsc70 NBD. Now, the chemical shift is in itself only a moderately useful parameter: it is best described as a fingerprint of the protein. Changes in the chemical shift can be due to any number of reasons:

1. Ligand binding.
2. Change in the basic chemistry of the species studied (such as a change in oxidation state).
3. Structural changes: movement of the residue, either by itself or in concert with a larger fragment of the protein, such as an alpha helix.
4. Dynamics.

Based on the biochemical evidence that the Hsp70 is a highly allosteric system [57] [91], it seems apposite to concentrate on # 3 listed above, with the caveat that we might expect changes to occur close to the active site (in this case – the nucleotide binding cleft), or further away. The question is: what is the extent of any such change? In order to answer this question, we changed the focus of study to the NBD of the Hsp70 homologue DnaK of *Thermus thermophilus* [169] (TTh-NBD), because the collection of NMR data at elevated temperature (50 ^{0}C) allows better quality spectra than for the mesophilic Hsc70. TTh-NBD is highly related to human Hsc70 (49%

identity, 72% homology). As is shown in Fig 2.6, quite similar chemical shift changes occur in the NMR spectrum of TTh-NBD in the ADP and AMP-PNP states, where the latter is a non-hydrolysable ATP analogue.

2.4 A BRIEF DESCRIPTION OF THIS PROJECT

In this work, we use residual dipolar coupling analysis [170] [171], and discover that subdomain IIB in this species (see Fig. 2.1/2) rotates up by as much as 20 degrees between the AMP-PNP and ADP state. Such conformational changes have been observed before, but were thought to be caused by the binding of a nucleotide exchange factor; such as BAG [156] or GrpE [90], rather than by nucleotide change itself. We also discover that subdomains IA and IIA rotate significantly with respect to each other, opening (AMPPNP) and closing (ADP) a hydrophobic surface cleft between these subdomains. Significantly, this cleft area has been hypothesized to be of importance to the communication between NBD and SBD [91; 160]. The current work demonstrates that a change in this cleft occurs upon nucleotide exchange. It is likely that this is the allosteric change on the surface of the NBD that is recognized by the SBD and/or the hydrophobic linker between NBD and SBD.

2.5 RESULTS

The NBDs of the Hsp70s are comprised of four subdomains [155] called IA, IB, IIA and IIB, as defined in Figure 2.1/2, and earlier in Chapter 1. The nucleotide, ATP or ADP.PO_4, together with several cations, is located deep in the central cleft and interacts with all four subdomains. AMPPNP binds at the same location {Jiang, 2007 #7436}. Several crystal and some NMR structures are available for several domains of Hsp70's of several species (called DnaK in bacteria, Hsc70 for the bovine protein and SS1 in fungi) molecules, but no de-novo structure is available for any domain of the DnaK of *Thermus Thermophilus*.

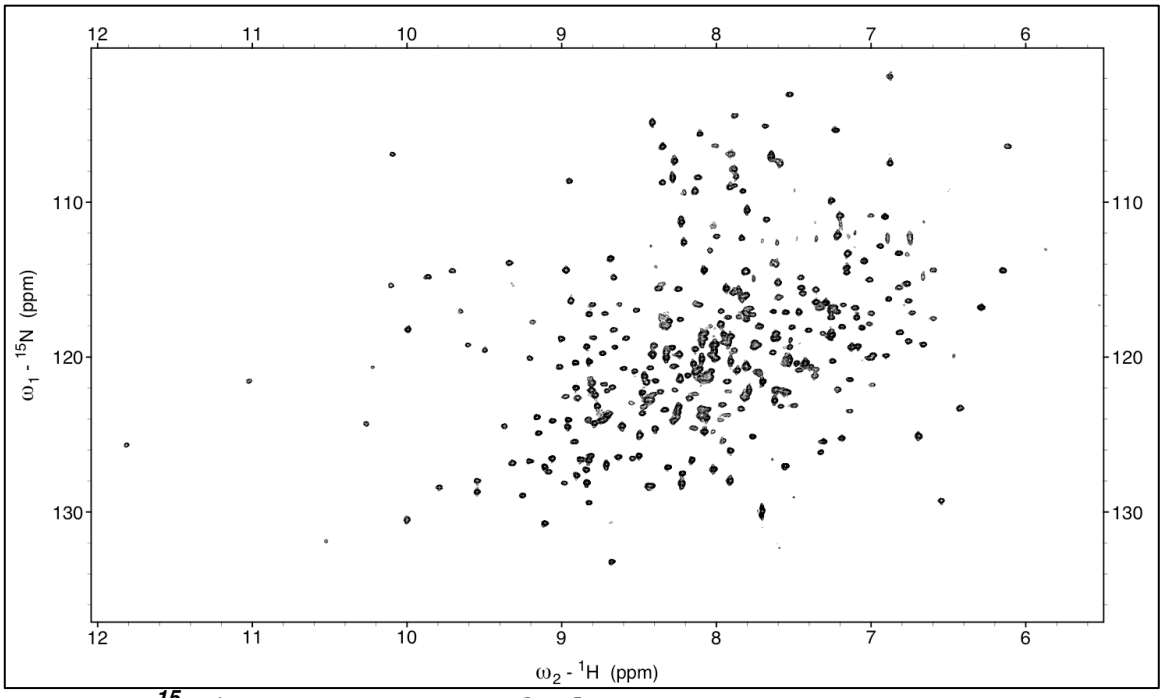

Figure 2.7: ^{15}N-^{1}H TROSY spectrum of TTh-NBD

^{15}N-HSQC-TROSY of TTh-NBD, 800 MHz, 50 ^{0}C. Sample: 150 uM triple labeled TTh-NBD, pH 7.2 in 20mg/ml phage.

We used the construct [172] of TTh-NBD containing residues 1-381, which gives rise to excellent ^{15}N-^{1}H HSQC-TROSY NMR spectra at 50 ^{0}C (see Fig 2.7).

2.6 ASSIGNING TTh-NBD

The assignments of the backbone resonances for this construct were based on an assignment of this protein in the ADP.AlFx bound state, which in turn were based on the published assignment by Revington and Zuiderweg {Revington, 2004 #2622}. The AlF_x assignment was used as a template, and three dimensional assignment experiments (HNCAs) were carried out on all three nucleotide states of TTh-NBD. These HNCAs were used along with the existing assignment template to generate complete assignments of TTh-NBD in all the available nucleotide states. The NMR spectra of TTh-NBD in the ADP.PO_4 and AMPPNP state show many chemical shift differences (Figure 2.6/8). The amide ^{1}H and ^{15}N NMR assignments were confirmed for each state by analysis of 3D HNCA-TROSY spectra. The shift changes are subtle but

significant, as is shown in Figure 2.6/8. Mapping the shift differences on a TTh-NBD homology model shows that the shifts extend far beyond the location of the nucleotide (see Figure 2.6/8). The shifts disclose subtle but widespread changes in structure and/or dynamics between the states.

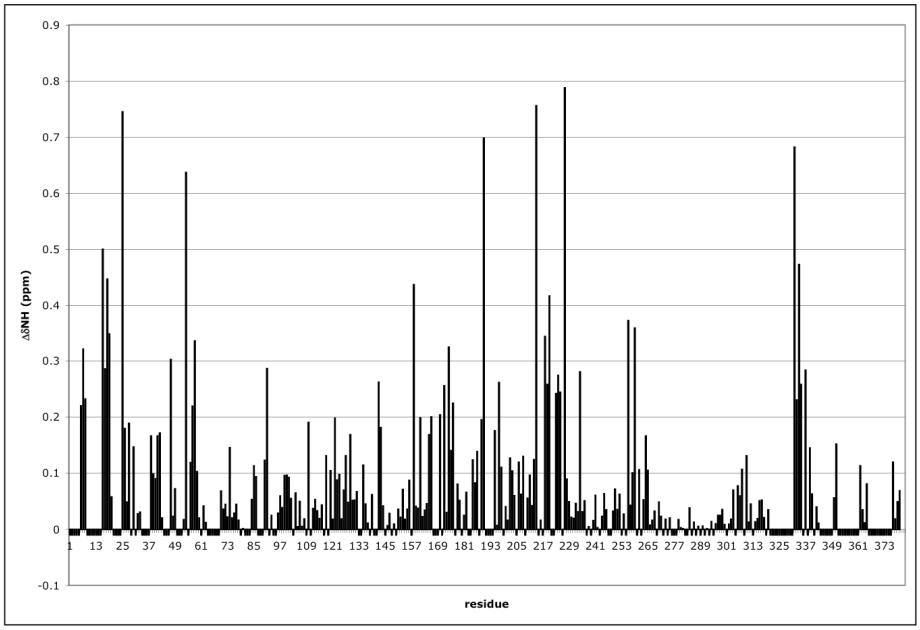

Figure 2.8: Chemical shift differences ΔδNH between the ADP.Pi and the AMPPNP states of Tth-NBD (in phage)

Chemical shift differences $(\Delta\delta NH=(\Delta\delta^{1}H^{2} + \Delta\delta^{15}N^{2})^{1/2})$

2.7 USING RESIDUAL DIPOLAR COUPLINGS TO PROBE STRUCTURE

At 44 kDa, TTh-NBD is too large for a high-resolution structure determination by NMR. However, the reorganization of subdomains can be probed, with excellent precision, using residual dipolar coupling NMR (RDC) analysis in solution [171]. Briefly, the method is as follows. The target protein is aligned in neutral aqueous buffer to which Pf1 phage is added [173]. The rod-like phages align in the magnetic field, providing an anisotropic environment for the protein occupying the buffer between the phage molecules. The target protein aligns, dynamically, about 0.2% of the time, allowing the magnetic dipolar coupling of the amide nitrogen and hydrogen nuclei to be measured by solution NMR techniques [174] [175]. We use modified HSQC-TROSY

experiments that shift the resonances in the ^{15}N dimension proportionally to ^{1}JNH + ^{1}DNH (described in detail in Chapter 3). The comparison of those spectra with unshifted TROSY spectra reveals the residual dipolar couplings [174] (see Fig 2.9).

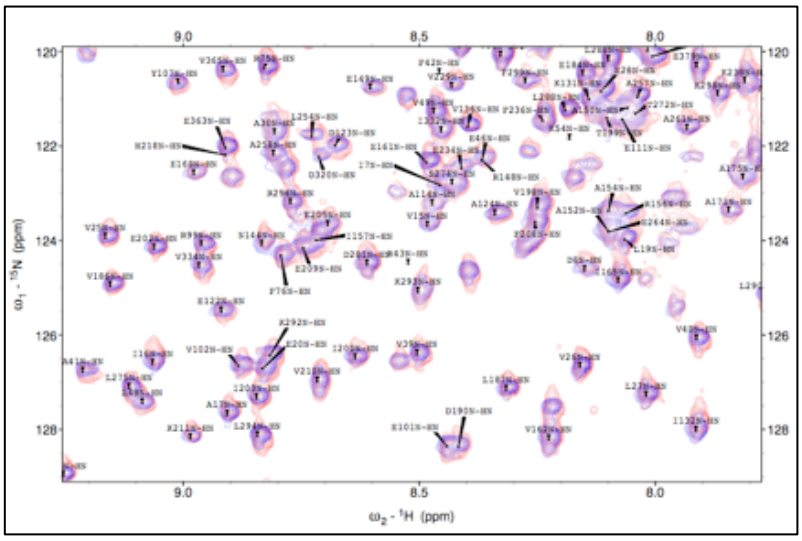

Figure 2.9: RDC-TROSY spectrum of TTh-NBD used to extract RDCs

^{15}N-HSQC-TROSY with kappa-shift element [176] of TTh-NBD, 800 MHz, 50 ^{0}C. Sample: 150 uM triple labeled TTh-NBD, pH 7.2 in 20mg/ml phage. Region of an overlay of the kappa = 0.00 and kappa=0.75 spectra.

This information is, in turn, used to obtain the time-averaged orientation of the protein's sub-domains with respect to the magnetic field, making use of the available high-resolution (X-ray) structures and homology models of the sub-domains. Subsequently, this information is used to reconstruct the relative orientation of the sub-domains. Sub-domain orientation is characterized by three axes corresponding to a rectangular parallelepiped with the longest axis defined as Szz, the shortest as Sxx with Syy in between.

On average, 60 reproducible NH RDCs per subdomain were obtained, with an uncertainty of ~4 Hz. Only RDCs corresponding to residues that are homologous between the NBD's of DnaK of *T. thermophilus*, DnaK of *E. coli*, and Hsc70, Hsp70 and mt-Hsp70 of *H. sapiens* were used. This reduced the number of RDCs available for analysis to about 50 per subdomain (see Table 1).

Table 1									
Input RDCa	NRDC[b]	α	β	γ	Szz	Syy	Sxx	RMSD[c]	Q[d]
		(o)	(o)	(o)				Hz	
AMPPNP state									
IA commone	45	161.8	87.6	63.1	-8.8E-04	8.6E-04	2.1E-05	8.1	0.65
IA allf	58	161.2	86.9	61.8	-8.8E-04	8.1E-04	6.6E-05	9.8	0.73
IA selfval averageg	34	160.2	87.4	61.9	-9.2E-04	8.6E-04	5.7E-05	9.1	0.68
IA selfval RMSDg	4	6.2	4.5	5.2	6.3E-05	1.0E-04	7.4E-05	1.6	0.10
IB common	26	136.4	76.4	70.0	-1.1E-03	7.5E-04	3.1E-04	10.5	0.67
IB all	37	143.3	81.5	66.8	-8.5E-04	7.0E-04	1.5E-04	11.4	0.76
IB selfval average	23	145.3	81.4	66.2	-9.0E-04	7.4E-04	1.6E-04	11.1	0.74
IB selfval RMSD	3	26.3	5.3	4.5	6.4E-05	9.3E-05	8.6E-05	0.9	0.06
IIA common	24	163.0	90.6	61.1	-8.5E-04	5.8E-04	2.6E-04	7.2	0.60
IIA all	45	145.6	87.5	59.7	-9.1E-04	7.2E-04	1.9E-04	7.3	0.56
IIA selfval average	27	144.6	88.0	58.7	-9.3E-04	7.4E-04	1.9E-04	7.0	0.53
IIA selfval RMSD	4	6.9	5.9	4.6	5.1E-05	6.4E-05	8.9E-05	0.6	0.06
IIB common	34	159.8	90.3	56.9	-1.1E-03	1.0E-03	4.0E-05	7.6	0.54
IIB all	45	165.1	92.2	66.1	-7.9E-04	8.0E-04	-1.7E-05	9.2	0.71
IIB selfval average	28	166.7	92.9	69.8	-8.1E-04	8.1E-04	1.0E-06	9.0	0.71
IIB selfval RMSD	4	7.0	3.7	10.0	6.1E-05	6.6E-05	3.1E-05	1.2	0.10
ADP.Pi state									
IA common	45	150.4	85.5	67.4	-1.1E-03	8.7E-04	1.9E-04	8.0	0.59
IA all	66	155.3	86.7	68.3	-8.5E-04	6.8E-04	1.7E-04	8.8	0.70
IA selfval average	41	154.4	87.1	68.2	-8.8E-04	6.9E-04	1.9E-04	8.5	0.67
IA selfval RMSD	3	12.3	5.3	4.8	5.7E-05	9.5E-05	7.6E-05	0.6	0.05
IB common	26	115.6	82.1	70.3	-9.5E-04	7.3E-04	2.2E-04	8.4	0.58
IB all	38	107.2	85.2	71.4	-9.4E-04	6.0E-04	3.5E-04	9.0	0.66
IB selfval average	24	165.4	86.7	71.8	-9.7E-04	6.7E-04	3.1E-04	8.5	0.62
IB selfval RMSD	4	65.9	6.8	4.2	6.6E-05	8.0E-05	7.0E-05	1.3	0.08
IIA common	24	180.5	78.2	66.2	-9.3E-04	6.1E-04	3.2E-04	5.5	0.43
IIA all	43	175.8	83.4	72.1	-1.0E-03	7.4E-04	2.7E-04	6.7	0.47
IIA selfval average	26	177.1	83.3	71.1	-1.0E-03	7.6E-04	2.4E-04	6.2	0.45
IIA selfval RMSD	4	8.7	3.6	4.8	5.4E-05	6.1E-05	6.0E-05	0.6	0.05
IIB common	34	161.5	87.5	50.1	-1.1E-03	1.1E-03	9.0E-05	7.9	0.51
IIB all	50	159.1	80.1	48.4	-8.7E-04	8.7E-04	-1.5E-06	8.3	0.60
IIB selfval average	30	156.2	79.9	50.1	-8.6E-04	8.5E-04	5.7E-06	7.9	0.58
IIB selfval RMSD	4	7.6	5.2	6.3	5.6E-05	7.7E-05	6.4E-05	1.0	0.06

Table 2.1: Solution statistics for RDC analysis of TTh-NBD subdomains

The calculations were carried out using an in-house written grid search and optimization program, which optimized α, β and γ (Euler rotational angles in the z-y-z convention describing the orientation of the subdomain in the aligned protein with respect to the coordinates of the reference structure based on DnaK E. coli and Hsc70 H. sapiens), the overall alignment (Szz) and the rhombicity ((Syy-Sxx)/Szz).

a) The input RDC set used

b) Number of RDCs in that set

c) RMSD of the best fit between experimental and calculated RDC

d) $$Q = \frac{RMSD}{\sqrt{\frac{\sum_{i=1}^{N_{RDC}} \left(RDC(i)_{\exp} \right)^2}{N_{RDC}}}}$$

e) Consists of experimentally reproducible RDCs, for NH groups of residues that are homologous between the following Hsp70's: DnaK of T.Thermophilus, DnaK of E. coli, Hsc70, Hsp70, mt-Msp70 and Bip of H. sapiens. This set was further reduced by taking only those residues for which RDC data was available in both the ADP and AMPPNP state.

f) Consists of experimentally reproducible RDCs, for NH groups of residues that are homologous between the following Hsp70's: DnaK of T.Thermophilus, DnaK of E. coli, Hsc70, Hsp70, mt-Msp70 and Bip of H. sapiens.

g) 40% of the RDCs of (f) were omitted at random for an ensemble of 30 calculations per subdomain. The average and RMSD of the fitting parameters for these ensembles are given.

We used a homology model for a reference structure of the subdomains (see Materials and Methods).

2.8 VISUALISING THE DIFFERENCES IN SUBDOMAIN ORIENTATIONS BETWEEN DIFFERENT NUCLEOTIDE STATES OF TTh-NBD: SANSON FLAMSTEED SINUSOIDAL PROJECTIONS.

As is shown in Figure 2.10/11, our data is of sufficient quality, so as to define with

precision, orientations of the different subdomains for the two nucleotide states. In these so-called Sanson-Flamsteed (sinusoidal projection) globe plots, the main orientational axis Szz for all subdomains is located approximately at 40^{o} West, 5^{o} South. This average location by itself is irrelevant, as it just reports the difference between the physical Szz axis orientation of the aligned protein and the arbitrary PDB-file Z-axis direction of the model structure used. However, the differences between the globe locations of the Szz axes of the different subdomains are highly relevant: they indicate that not all subdomains are oriented as in the model structure. Moreover, the relative orientations of the subdomains change between the ADP and AMPPNP state. The spread in the Szz directions of the subdomains represents the uncertainty in the raw data as determined by the Monte-Carlo routine in the REDCAT program used, using RDCs corresponding to homologous positions without any further editing. According to the REDCAT program, for the AMPPNP state the subdomains are, within experimental error, oriented similarly to those in the model structure, which is based on bovine Hsc70 in the ADP state. Surprisingly, this is not the case for the ADP state, where the orientation of subdomain IIB has moved away by approximately 20 degrees. In addition, the relative orientations of the other subdomains change as well.

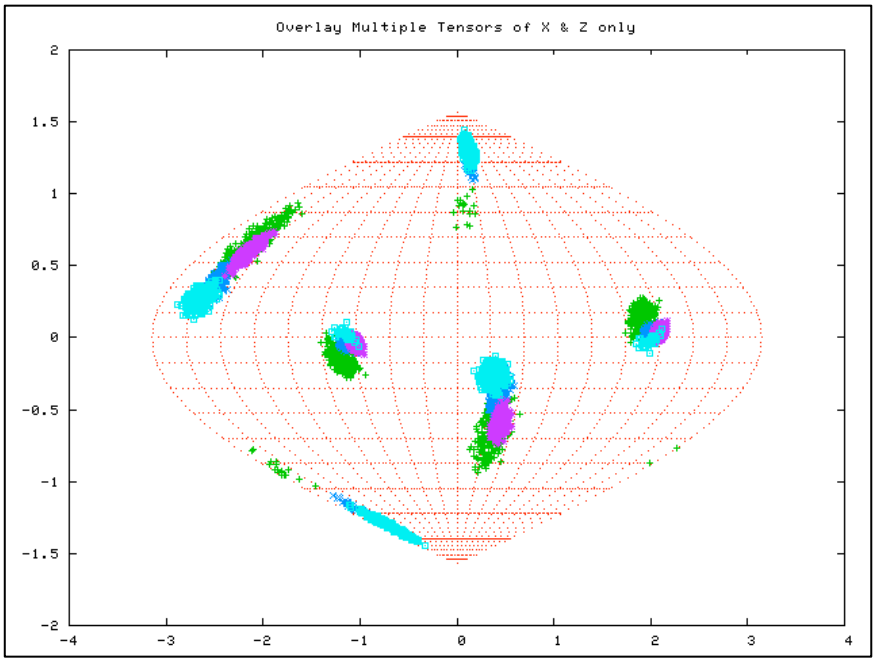

Figure 2.10: Sanson Flamsteed projection of subdomain orientation for TTh-NBD in the AMPPNP state

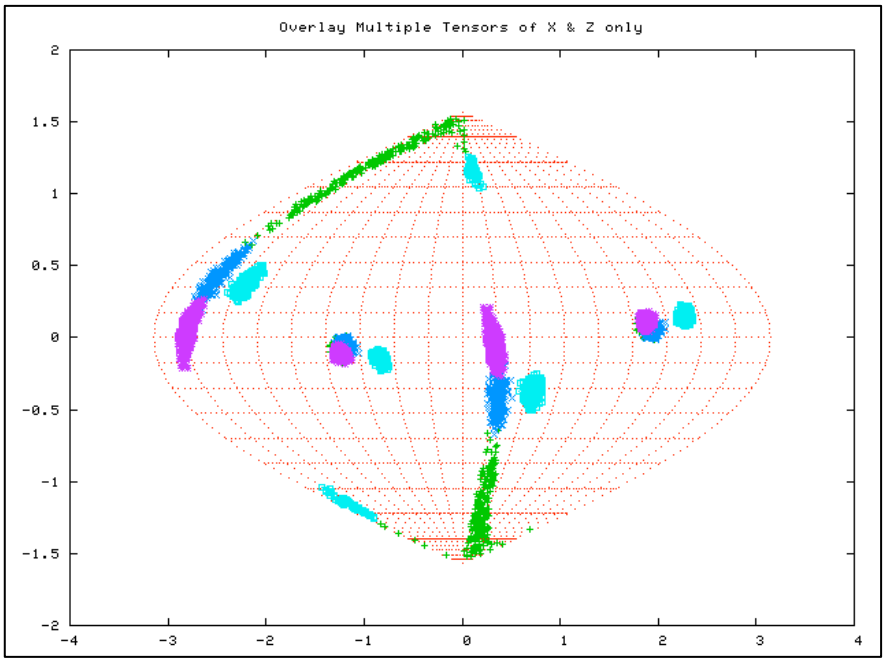

Figure 2.11: Sanson Flamsteed projection for subdomain orientation of TTh-ND in the ADP.Pi state

Globe graphs (Sanson-Flamsteed plots) showing the orientations and experimental uncertainties of the Szz principal alignment axes (around 60^{o}

West and 5^{o} South) of the different subdomains of NBD of DnaK-Tth as derived from the NMR RDC measurements using REDCAT[177]. RDC input: conserved residues and experimentally reproducible Fig 2.10(top) AMP-PNP state; Fig 2.11(bottom) ADP.Pi state. Color code: IA (1-37,109-180, 357-377; blue); IB (residues 38-108; green); IIA (residues 181-219, 307-356; purple) and IIB (residues 220-304; cyan). The reference structure for the NBD of DnaK-Tth was modeled on the NBD of DnaK E. coli (1DKG.pdb) and bovine Hsc70.ADP.Pi (3HSC.pdb). Scale: Horizontal: 20^{o} per gridline, vertical, 10^{o} per gridline. (The Szz principal alignment axes also appear at the other side of the globe (120^{o} East, and 5^{o} North). The Sxx axes appear as smears around 150^{o} West, 45^{o} North and 30^{o} East, 45^{o} South.

2.9 FROM RDCs TO STRUCTURES

As a next step in the analysis, the orientational changes were modeled as changes in the 3D structure. As explained in detail in the materials and methods, the models for the ADP and AMPPNP states were obtained by superposing the rotated subdomains on the model structure for minimal positional RMSD using translation of coordinates only.

A comparison of the best-fit structures based on all RDCs (as defined in the legend of Table 1) is shown in Figure 2.12. These best-fit models are referred to as eigenstructures; in light of the fact that the solution to an RDC based orientation analysis is essentially an eigenvalue decomposition problem. Major shifts are seen for the relative orientations of all domains between the two independent nucleotide states.

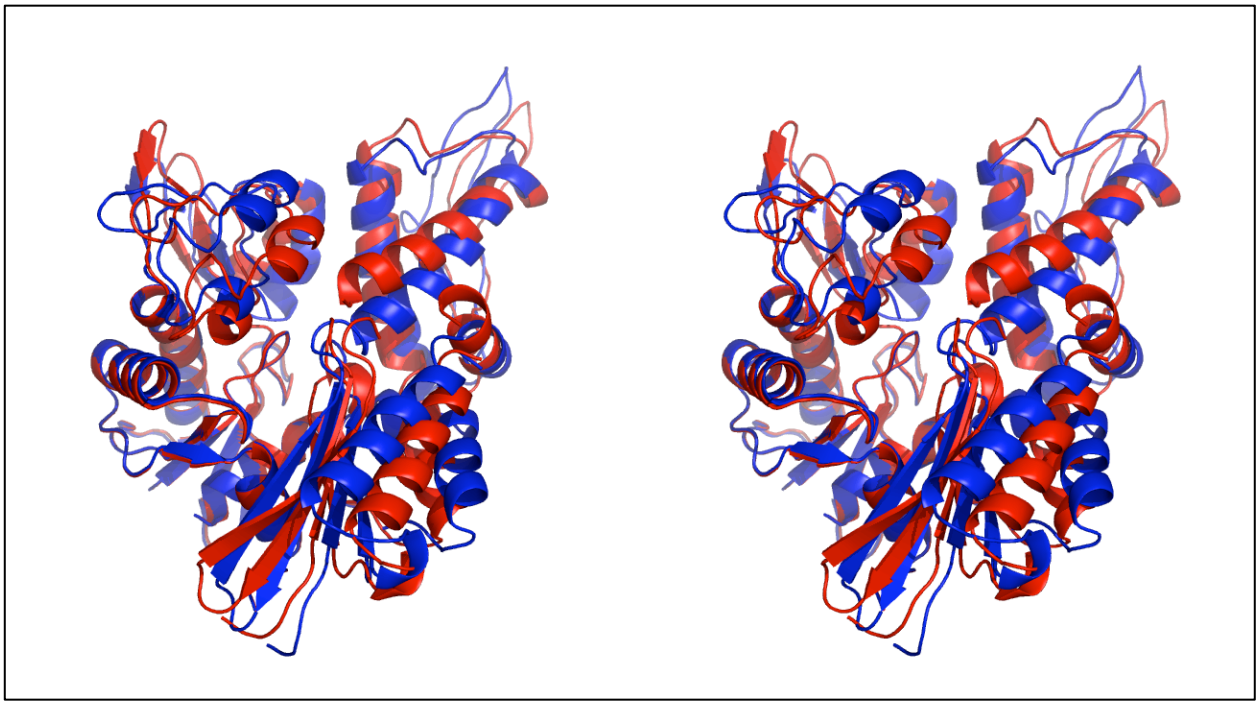

Figure 2.12: RDC derived eigenstructures of TTh-NBD in the ADP.Pi and AMPPNP bound states

Model of the ADP state of the NBD of DnaK T.Th (RED) superposed on domain IA of the model for the AMPPNP state (BLUE) calculated using all RDCs

2.10 ARE THE STRUCTURAL DIFFERENCES IN THE TWO NUCLEOTIDE STATES STATISTICALLY SIGNIFICANT?

In order to assess the statistical significance of the differences seen in Figure 2.12, we performed a "Jack-Knife" self-validation procedure. As shown in the Materials and Methods section, it was determined that the calculation of model structures obtained from RDC calculations in which 40% of the RDC data was deleted at random, yields an ensemble that represents a reliable measure of the precision of the structure determination.

One of the ensembles is presented in Figures 2.13, in which subdomains IA are superimposed. The figures clearly show that the average position of subdomain IIB in

the ADP state (relative to the orientation of IA) lies outside of the ensemble of possible subdomain orientations in the AMPPNP state.

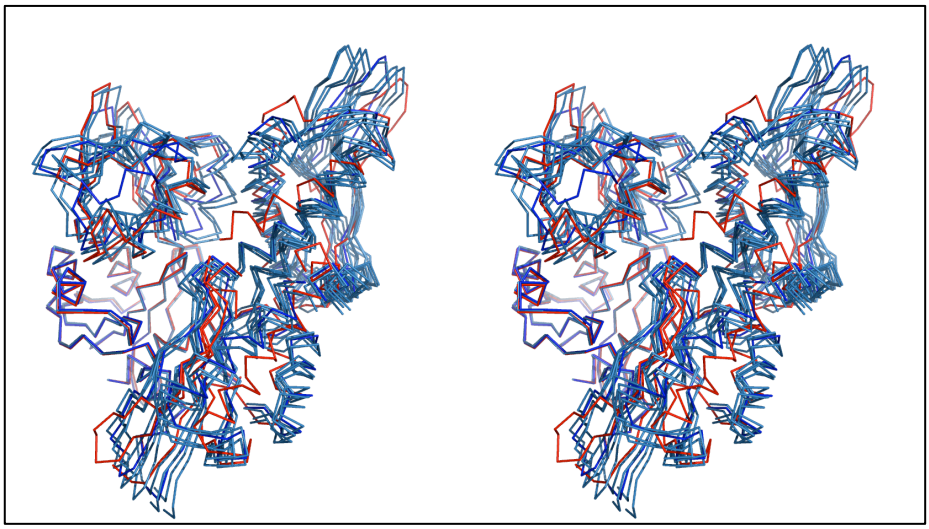

Figure 2.13: Comparing the average ADP.Pi structure of TTh-NBD with the ensemble of AMPPNP structures

The average of the Tth-NBD self-validation ensemble for AMPPNP (blue) and Tth-NBD ADP (red) compared to the self-validation ensemble of Tth-NBD ADP (salmon)

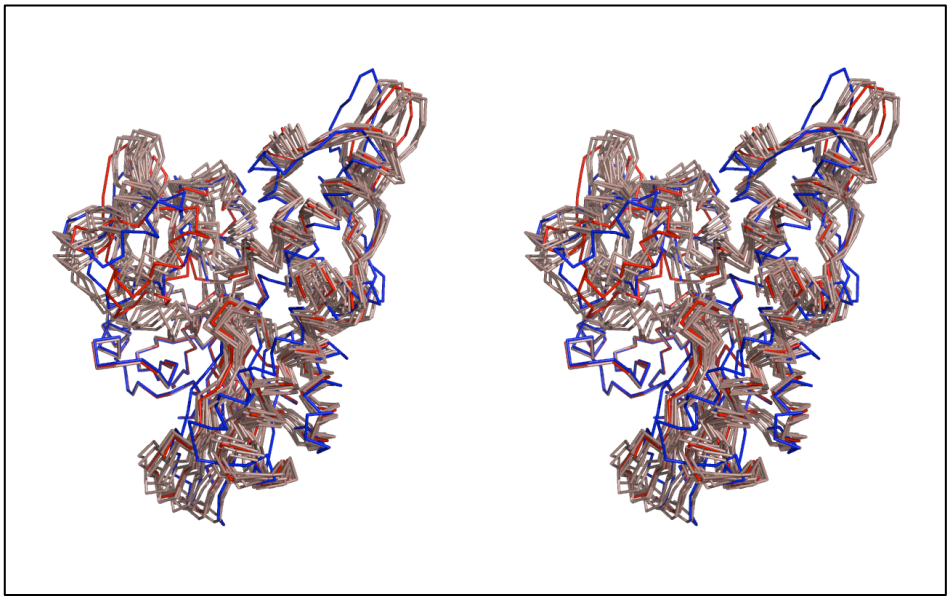

Figure 2.14: Comparing the average of the AMPPNP structure of TTh-NBD with the ensemble of ADP.Pi structures

The average of the Tth-NBD self-validation ensemble for AMPPNP (blue) and Tth-NBD ADP (red) compared to the self-validation ensemble of Tth-NBD ADP (salmon). The conformations were computed using the in-house written grid search program.

Figure 2.14 shows the reverse is also true. Hence, the differences in orientation between these states as seen in Figure 2.12 are statistically highly significant. The average position of subdomain IB in the ADP state lies well inside the ensemble of possible subdomain IB positions in the AMPPNP state. Hence, the differences in orientation between these states as seen in Figure 2.12 are statistically NOT significant. The average position of subdomain IIA in the ADP state lies at the boundary of the ensemble of possible subdomain positions in the AMPPNP state. And vice-versa. According to T-statistics, there is just a 3 % probability that the observed difference in orientations of subdomains IIA in the different nucleotide states is by chance (see Materials and Methods). Hence, the statistical significance of the differences for subdomains IIA and IIB as seen in Figure 2.12 is high.

Summarizing, our data shows that the subdomains in the AMP-PNP state are, within experimental error, similarly oriented as in the model structure, which is based on an Hsc70 NBD crystal structure in the ADP state (3HSC.pdb, ref [155]). In contrast, in the ADP state, one observes statistically significant deviations, by as much as 20^{o}, for the Szz axes of subdomains IIB. In addition we observe statistically significant rotation of subdomain IIA with respect to subdomain IA between the different states, as disclosed in Figure 2.12.

2.11 DISCUSSION

The RDC data reports on the orientations and changes in orientations of the NBD subdomains. The data gives no information whether the subdomains shift in a lateral way with respect to each other or not. Therefore, it is important to emphasize that the structures shown in the figures are just convenient models; they show the orientational

changes, but changes in translation have been minimized by superposition. The models have not been further minimized, and contain areas in which atoms clash (see also Materials and Methods).

Figure 2.12 shows that domain IIB rotates about 20 degrees clockwise going from AMP-PNP to ADP.Pi state. As was shown in Figure 2.13 and 2.14, this change is significant within the accuracy of the experimental data. This movement is the likely explanation for the chemical shift changes for residues in domains IB and IIB that line the nucleotide binding cleft (see Figure 2.6). We suggest that the origin of this rotation is as follows:

From the available crystal structures, it is seen that the nucleotide contacts all subunits of the NBD. It therefore is reasonable to assume that in the ATP and AMP-PNP state, a rigid nucleotide bridges the left and right halves of the protein. For ADP.Pi, however, the γ-phosphate bond is hydrolyzed, which breaks the molecule and thus the bridge between left and right, allowing a rotation of the subunits with respect to eachother. It appears that the left and right halves can move relatively independently in the ADP.Pi state. It is possible (and perhaps likely, see below) that the ADP.Pi structure can open dynamically even further than displayed in the models.

Available crystal structures of the isolated NBD correspond closely to each other regardless their nucleotide state. This is illustrated in Figure 2.2, which is an overlay of five crystal structures of the NBD of bovine Hsc70 in APO, ADP (twice), ADP-V2O5 state and the mutant K71M complexed with ATP from two different laboratories. Small Angle X-ray Scattering (SAXS) measurements [178] showed no significant difference in radii of gyration for the ADP and ATP states of the human Hsc70 NBD fragment either (see Table 2 below). This appears to confirm that the conformations of these states are also identical in solution. However, there is also no difference in the computed radii of gyration for the AMPPNP and ADP structures of TTh-NBD as derived from the RDC data, despite the rather large (visual) differences (Table 2).

Table 2: Computed and experimental radii of gyration (Å)			
	ADP	AMPPNP	ATP
Computed values for DnaK-TTh-NBD	21.55	21.63	
Experimental values for Hsc70-NBD Ref [178], Table 1	22.8 ± 0.2		22.6± 0.1
Experimental values for Hsc70-NBD Ref [178], Table 2	23.1± 0.1		

Table 2.2 Computed and experimental radii of gyration of different NBD constructs from different species [163]

As it turns out, the X-ray structures of the isolated NBD all correspond much closer to the AMPPNP state of TTh-NBD in solution then to the solution ADP state (Figure 2.15). The fit between TTh-NBD on the ADP state with HSC70 in ADP state is inferior (Figure 2.16), mostly because of the clockwise rotation of IIB. The rotation of subdomain IIB in the ADP state in solution is very reminiscent of the rotation of this domain upon binding of a nucleotide exchange factor (NEF).

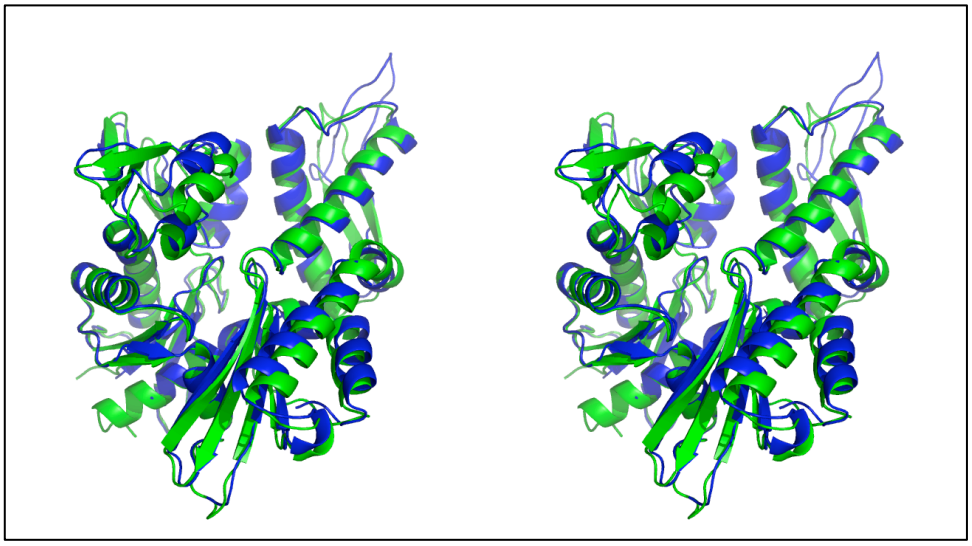

Figure 2.15: Overlay of the TTh-NBD AMPPNP (blue) eigenstructure with the crystal structure of bovine Hsc70 NBD in ADP.Pi state (3HSC.pdb, red).

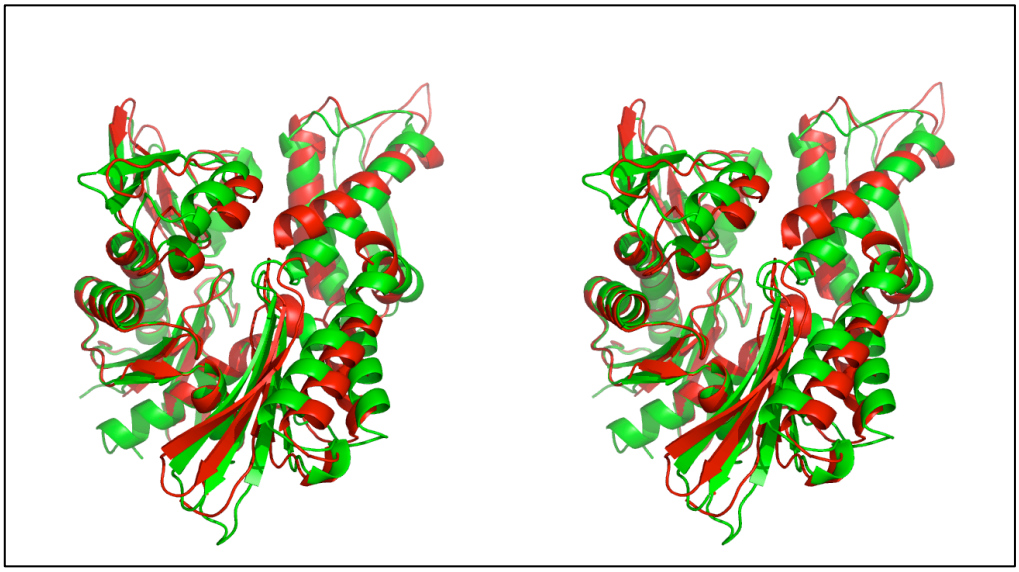

Figure 2.16: Overlay of the TTh-NBD ADP.Pi eigenstructure (red) with the crystal structure of bovine Hsc70 NBD in complex with ADP.Pi (3HSC.pdb, red).

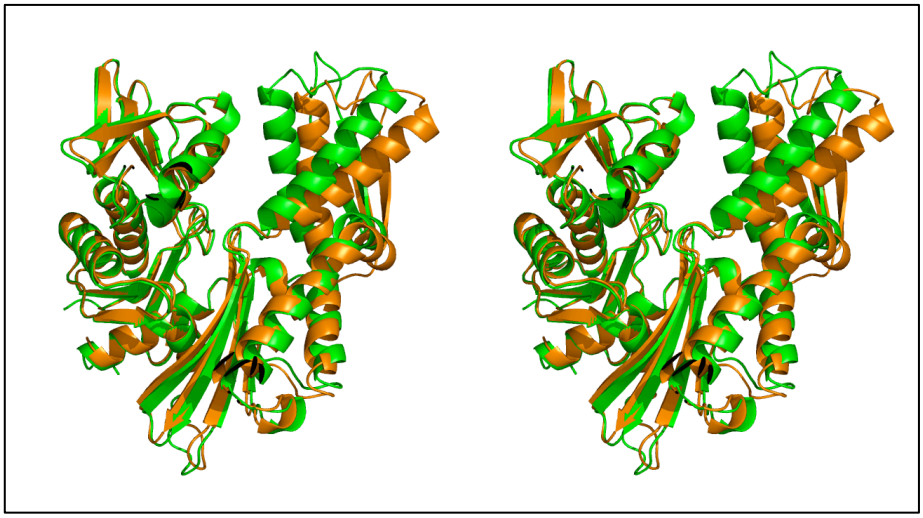

Figure 2.17: Overlay of the bovine Hsc70 NBD in the ADP.Pi state (3HSC.pdb, green) and in complex with BAG-1 (1HX1.pdb, brown)

Figures 2.15 – 2.17: Comparison of structures. The figures were made by superposing the CA-positions of the secondary structure elements of domains IA. Fig. 2.17: BAG-1 is not shown.

Figure 2.17 shows a superposition the crystal structures of bovine Hsc70-NBD

with and without its NEF BAG-1 [156]. Similar rotations were observed for *E.coli* DnaK NBD upon binding to its NEF GrpE [179] (not shown) and in the very recent co-crystal structures of yeast Hsp70 NBD with Hsp110 as exchange factor [180] [181] (not shown). Hence, the rotated state that exists in solution in for the ADP state for DnaK TTh-NDB can also be observed in the crystal, provided it is stabilized by a NEF.

2.11.1 A NEW PARADIGM FOR NUCLEOTIDE EXCHANGE IN THE NBD

The results just discussed - that the subdomain rotation that we observe to occur in different nucleotide states is also seen in the context of exchange factors - is of profound importance. This means that earlier notions that it is the interaction between Hsp70 and NEF that forces an induced fit which actively promotes nucleotide exchange [90], should likely be modified. It seems more appropriate to assume that the NEF selectively captures and stabilizes the ADP state. The nucleotide exchange process in the Hsp70 chaperones (abundant ATP replacing ADP) in this model is then catalyzed by other processes, which may include the interaction of GrpE's tail with the SBD, as has been suggested before [179]. Selective interaction of the NEF with the ADP conformation also makes sense from a functional point of view: it avoids unnecessary interference with the ATP state and provides a driving force for NEF release when the nucleotide has been exchanged from ADP to ATP. The process of selective capture which we here suggest to be operant for the chaperone NEF interaction, has recently been recognized as an common interaction paradigm for several other allosteric proteins [11] and nucleic acids [182].

However, it has been found that GrpE interacts more tightly with the *apo* than the ADP state (REF). Hence, it is likely that the NEF, while recognizing the ADP state, may further open the molecule and bring it in an even more fully opened state such as seen is the apo form (see chapter 3). This mechanism is a combination of selective capture followed by induced fit.

The chemical shift changes in Figure 2.6 suggest that major changes take place in domains IA and IIA when changing from the AMP-PNP to the ADP.Pi state. The globe plots in Figures 2.10/11 and the models based on this data in Figure 2.12 suggest that these changes are due to a clockwise rotation of domain IIA with respect to domain IA when going from AMPPNP to ADP state. Figures 2.13/14 prove that such a rotation is significant with a 97 % probability.

This is the first time that such a conformational change has been demonstrated for nucleotide change in a single Hsp70 species. But, as is illustrated in Figure 2.18, similar differences are present between the published structures of different species in different nucleotide states. For instance, a crystal structure for a yeast Hsp110 dimer, locked in the ATP state, was recently solved [162]. Hsp110 has been recognized as a protein with a strong Hsp70 homology, but it functions as a NEF rather than as a chaperone. It was suggested that the structure of this related protein might provide insight in the structure of an Hsp70 protein in the ATP state. If one associates the structure of Hsc70 NBD in complex with its NEF as a true representative of the ADP state, and Hsp110 as a true representative of the ADP state, the ADP - ATP change becomes quite apparent.

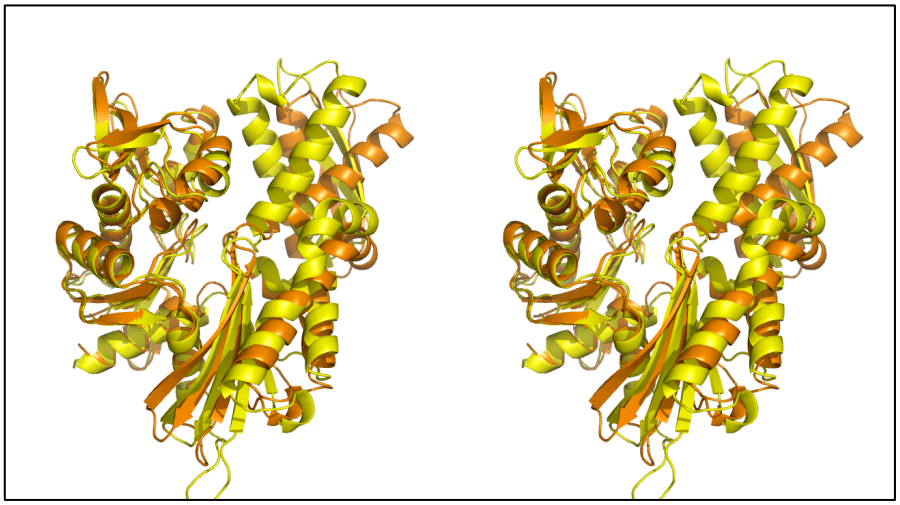

Figure 2.18: Overlay of Hsc70 NBD in complex with BAG-1(brown) with Hsp110 NBD, ATP state (yellow)

Comparison of structures. The figures were made by superposing the CA-positions of the secondary structure elements of domains IA. Hsc70-bovine complexed with BAG (1HX1.pdb; orange; BAG not shown) and Hsp110. (1QXL.pdb; NBD of one monomer shown, in brown).

Figure 2.18 shows a dramatic clockwise rotation of domain IIA of Hsc70 (complexed with BAG) as compared to Hsp110 structure. Remarkably, these intermolecular ADP-ATP changes are very similar to the intramolecular changes seen for Tth-NBD as shown in Figure 2.112. That conformational changes are possible in the Hsc70 NBD, has been suggested by this research group before [183]. However, the experimental data at the time was of insufficient quality to disclose what the changes entailed.

2.12 SUMMARISING THE RESULTS

At this point of the discussion it is worthwhile to take stock. Our studies have shown that significant conformational changes occur in the nucleotide-binding domain of DnaK of *Thermus thermophilus* upon nucleotide change. This is confirmed by and does explain the observed chemical shift changes for this domain in its different states. It is of importance that similar chemical shift changes were observed between the ADP and ATP state of the Hsc70 NBD; confirming that the similar conformational changes take place for the different species. Analysis of the RDC data and subsequent modeling, discloses significant and substantial clockwise rotations for domains IIA and IIB when changing from AMPPNP state to ADP state (relative to IA). Because of these findings, we can be now also confident that the differences seen between structures of different species in different nucleotide states are most likely due to changes in nucleotide, and not to changes in species. Hence, those differences may now be interpreted in terms of conformational changes and the mechanism of allostery. The remainder of this discussion is a first attempt.

2.13 INTERPRETING STRUCTURAL DIFFERENCES BETWEEN NUCLEOTIDE STATES IN TERMS OF ALLOSTERY

The relative rotations of subdomains IA and IIA are of great interest, since it affects a hydrophobic surface cleft between them (see Figure 2.19/20), which likely is involved in linker binding, and substrate domain binding [91; 160]. In Hsp110, this surface cleft is open and is occupied by the linker between the NBD and SBD. The cleft is closed in the ADP state and the linker cannot be docked. It has been shown before for DnaK of *E.coli* that the linker moves freely in the ADP state and allows the SBD to move relatively freely [91] [161], while the linker becomes immobilized in the ATP state [91]. Several biochemical studies also showed that the linker is exposed in the ADP state, and not in the ATP state[160] [184] [185]. It has been speculated [91] that changes in the surface cleft between IA and IIA are responsible for the docking in the ATP state, but it is only in our current work that this is substantiated with an actual observation of change in that area for a single species.

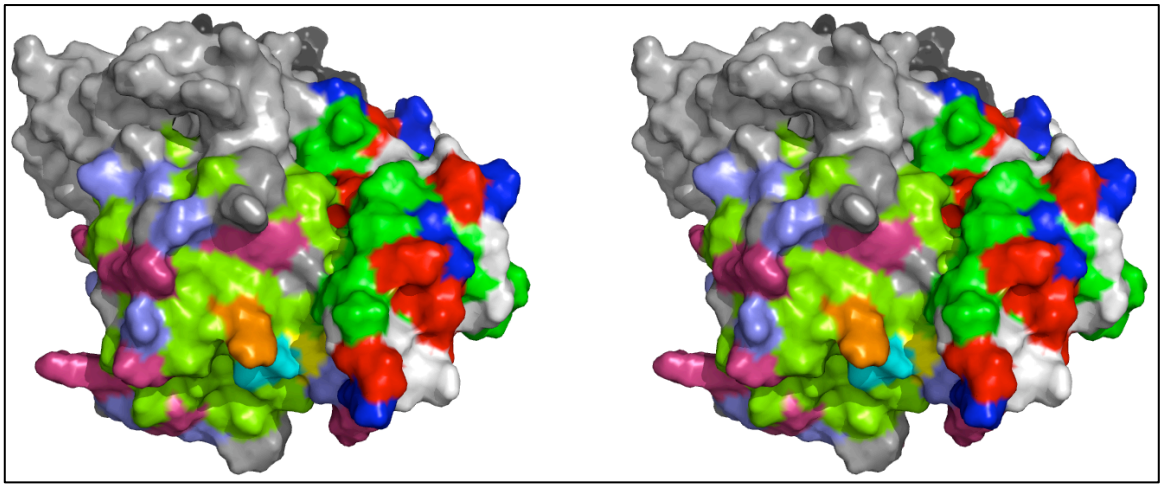

Figure 2.19: The IA/IIA interface of TTh-NBD in the ADP.Pi state

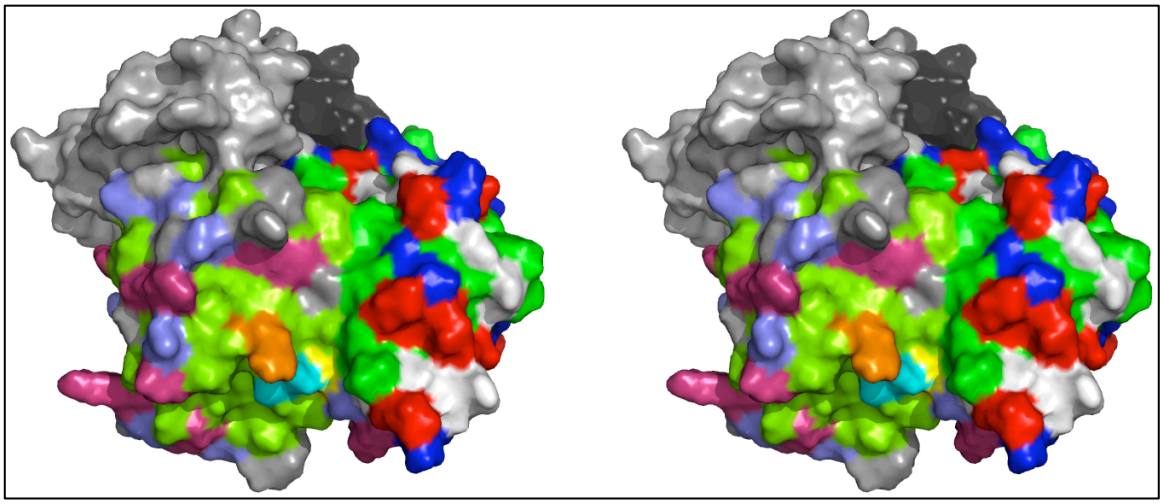

Figure 2.20: The IA/IIA interface of TTh-NBD in the AMPPNP state

View of the IA / IIA interface of DnaK-TTh. This view is a 90-degree rotation around the horizontal axis as compared to the other figures. It shows the "bottom" as compared to these other figures. Colour coding: In domain IA (left) hydrophobics, light-green; positive, light-blue; negative, purple; polar, grey

In domain IIA (right) hydrophobics, green; positive, blue; negative, red; polar, white. The C-terminus (residue 372 in this construct) is rendered in cyan. Domain IB (at left top) is in grey; domain IIB (at right top) is in black.

On the basis of the current work, literature reports on a confoundingly large number of Hsp70 species, states, complexes and artificial truncations, studied by different experimental techniques, can now be merged into the following general allosteric model: in the ATP-substrate-free state, NBD and SBD are docked. The NBD is relatively rigid[184] [186]. ATP hydrolysis leads to a loosening of the junction between the "left" (subdomains IA and IB) and "right" (subdomains IIA and IIB) of the NBD. Overall, the NBD has become more flexible [184; 186]. In this process, subdomain IIA rotates clockwise and closes the IA/IIA surface cleft. As a consequence, the linker is expelled and the SBD dissociates from the NBD. We suggest that the ATP hydrolysis energy is spent on this step to compensate for the lost SBD-NBD binding energy. In the SBD, the dissociation from the NBD is transduced through the beta sheet [158], and affects the substrate-binding cleft, which rigidifies. Meanwhile, subdomain IIB rotates clockwise as well, and is predisposed to bind to the NEF, which

further stabilizes the open, ADP state and promotes nucleotide exchange. The re-binding of ATP causes NBD and SBD to dock, and expels the substrate. We suggest that the regained SBD-NBD docking energy compensates for the lost SBD-substrate binding energy.

2.13.1 INTERACTION OF THE LINKER WITH THE IA/IIA INTERFACE CLEFT

While the scenario discussed is likely to be correct in the overall sense, many of the details are still missing. For instance, we don't really know how the NBD/SBD linker binds in the ATP state for the true Hsp70's. The Hsp110 model [162] is useful, but there is only marginal sequence homology: true Hsp70's have a fully hydrophobic linker such as 388VQDLLLLDVTP (Hsc70 *H. sapiens*), ^{380}VRDVVLLDVTP (DnaK *T. thermophilus*), ^{387}TQDILLLDVAP (*Candida albicans* SSA-1) while yeast Hsp110 has the sequence ^{389}VRPFKFEDIH. Moreover, Hsp110 is locked in the ATP state, and the divergent linker may actually be the cause of its inability to change conformation.

Being fully aware of the limits on the precision of current NMR structure calculations and models, we tentatively show in Figure 2.19/20 a comparison of the IA/IIA cleft region for the ADP and AMPPNP state of DnaK *T. thermophilus*. The surface residues visible on this side of the protein are completely conserved in over 300 sequences of *Archaea*, bacteria and animals. The hydrophobic nature of the cleft is evident. Access to the cleft becomes hindered in the ADP state. This is especially obvious for Leu 174, rendered in cyan, which forms the bottom of the cleft. In the TROSY spectrum of NBD-TTh, L174 shows a large shift upon nucleotide exchange. Its position is also indicated in Figure 2.6. The homologous residue, I181 in Hsc70, also shifts (Fig 2.6). NMR chemical shifts were also observed for this residue (L177) in DnaK of *E. coli* upon nucleotide exchange [91].

Mutagenesis studies [160] strongly suggest that the universally conserved R167 of DnaK *E. coli* interacts with the universally conserved D393 of the linker sequence in

the ATP state. The NMR shift data in Figure 2.6 and the surface view in Figure 2.19/20 shows that this residue is unaffected by nucleotide exchange (R164 and R171 in DnaK TTh. and Hsc70 *B. taurus*, respectively). This is not surprising since the constructs studied here are truncated before the linker sequence. However, the finding suggests that this Arginine residue is "passive" and just serves to steer the sense of linker insertion. This sense corresponds to the insertion of the linker in Hsp110 [162].

2.14 MATERIALS AND METHODS

2.14.1 SAMPLES

DnaK-Tth 1–381 (DnaK-Tth NBD), cloned into Pet22-b (Novagen) with an N-terminal His-tag, was expressed in *E. coli* BL21 cells at 37^{o} C. Expression was induced at O.D. 600 = 0.5 in a triple labeled M9 medium containing 98% D20, ^{13}C-glucose and ^{15}N ammonium chloride. Cells were harvested by spinning down at 15,000 g, were resuspended and subsequently lysed using a microfluidizer. The protein was purified in two steps using Ni-NTA agarose with an imidazole gradient and FastFlo Q- ion exchange at pH 7.2 with a linear KCl gradient [186]. The purified protein was extensively dialyzed against NMR buffer (see below) and concentrated using Centricon micro filters.

2.14.2 NMR EXPERIMENTS

NMR samples contained 0.15-0.2 mM protein in 50 mM HEPES, pH 7.4, 10 mM KCl, 5 mM $MgCl_2$, and 5 mM sodium phosphate. ADP or AMP-PNP concentration was 10 mM. Experiments were performed at 50^{o} C on an 800 MHz Varian Inova spectrometer, using a triple resonance cold-probe. Backbone resonance assignments were obtained from a single 3D HNCA-TROSY experiment for each nucleotide state, using a previously obtained peak list for DnaK-Tth in the ADP-AlF_x state as a template [172]. 281 and 310 assignments were obtained for the ADP and AMP-PNP form, respectively.

For RDC measurements, Pf1 bacteriophage in the aforementioned buffer was added to the NMR samples to a concentration of 20 mg/ml for partial alignment [123]. The $^{2}H_2O$ quadrupolar splitting was 8 Hz at these circumstances. RDCs were extracted from a series of 2D TROSY experiments with a [$\kappa t_1/2$-180(N,H)-κ $t_1/2$] sequence[176] at the beginning of the ^{15}N chemical shift labeling period , with κ=0, 0.75 and 1.5. This method is better suited for larger proteins than the IP/AP analysis [187] or the TROSY/HSQC comparison [188], since this method retains the full TROSY narrowing in the ^{1}H dimension and partial TROSY narrowing in the ^{15}N-dimension. The 2D NMR spectra were recorded at 50 0C. 10, 20 and 40 hours of data acquisition was used for the κ=0, 0.75 and 1.5 experiments, respectively, for both the ADP and AMPPNP state. Further details of the κ experiment will be revealed in Chapter 3.

2.14.3 DATA ANALYSIS: EXTRACTING RDCs

The experiments yield TROSY spectra, in which the ^{15}N coordinate of each cross peak is shifted by κ*(^{1}JNH+^{1}DNH)/2 (see Figure 2.9). Using a uniform value of ^{1}JNH 90.5 Hz, the ^{1}DNH was extracted from the differences between the κ=0 and κ=0.75 and between the κ=0 and κ=1.5 spectra. The expected differences in real ^{1}JNH ($\pm$ 2 Hz) for the different residues is significantly less than the experimental precision in the data ($\pm$ 4 Hz in the ^{15}N dimension) for this large protein.

2.14.4 DATA ANALYSIS: CULLING THE RDCs

The available RDC data for each state was trimmed down in two successive steps. First, the RDCs obtained from the differences between the κ=0 and κ=0.75 and between the κ=0 and κ=1.5 spectra were compared. Those residues, for which the scaled differences were larger than the scaled RMS difference between these values, were rejected. The scaled differences were defined as difference=abs(2*(value1 -

value2)/(value1+value2)). The remaining RDCs were labeled as "experimentally sound". Second, only the RDCs were corresponding to NH groups of residues that are homologous between the following Hsp70s: DnaK of *T.Thermophilus*, DnaK of *E.coli*, Hsc70, Hsp70, mt-Hsp70 and Bip of *H. sapiens*.

The remaining RDC data, termed experimentally and homology sound, was used as input to orient subdomain IA, IB, IIA and IIB of the DnaK-Tth homology model. It was assumed that each subdomain was a rigid unit by itself. While most of the data shown in this paper is based on this relatively unedited RDC input data, several calculations were performed with a subset of RDC data in which only those residues for which RDC data was available in both the ADP and AMPPNP state were retained. The results of these calculations, shown in Figure 2.21, shows that the differences in conformations is essentially the same as for the larger RDC dataset, establishing that the differences are not due to the selection of the RDCs.

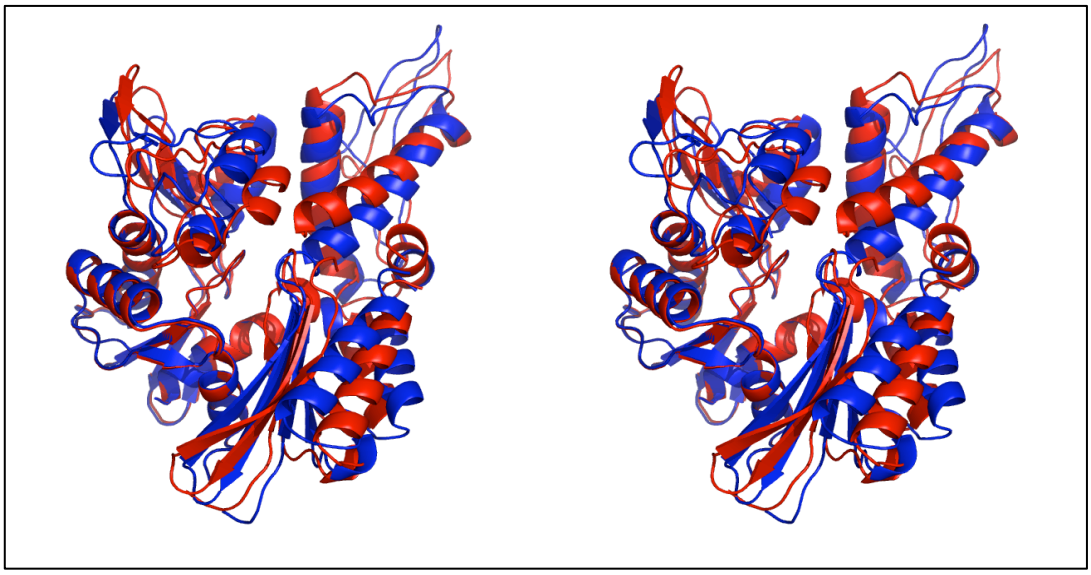

Figure 2.21: RDC derived eigenstructures of Tth-NBD in ADP.Pi (red) and AMPPNP (blue) states, calculated with a 'culled' set of RDCs

2.14.5 DATA ANALYSIS: RDCs TO STRUCTURE

We used in-house written programs and REDCAT [177] (A Residual Dipolar

Coupling Analysis Tool) to transform the RDCs to the orientational data. REDCAT's solution algorithm relies on singular value decomposition and Monte Carlo error estimation to generate an ensemble of 1000 structures compatible with the input structures and the set of RDCs provided, based on an experimental error range of 4 Hz. The in-house written Fortran program is based on a grid and minimization search of all possible Euler rotations, overall alignment and rhombicity, to find the best fit to the experimental data. The results are identical to the REDCAT solutions. The structural data shown in this report were all derived from the results of Fortran program, since it gave easier access to the distribution of Euler angles.

2.14.6 DATA ANALYSIS: HOMOLOGY MODEL

The orientations of the subdomains were computed from the structures of the subdomains. Since no coordinates are available for Tth-NBD, we used a homology reference model. It was constructed in two steps. First, we threaded the sequence of Tth-NBD on the coordinates of a crystal structure of the NBD of DnaK-*E.coli*, the protein most homologous to DnaK *T. thermophilus* for which coordinates are available [90] (1DKG.pdb, with a resolution of 2.8 Å). However, this structure was obtained in the presence of a nucleotide exchange factor, which likely rearranged the subdomain orientations. Several crystal structures [155] without co-chaperones are available for the bovine Hsc70 NBD, in which the subdomain orientations, especially for IIB, are different from that seen in 1DKG.pdb. In order to obtain the required reference model for NBD-TTh, we adjusted the orientations of the subdomains in the model to correspond with those in 3HSC.pdb, a structure of Hsc70-NBD nominally in the ADP state at a resolution of 1.90 Å. The adjustment was based on a superposition of the Ca atoms of the secondary structure elements in the subdomains. The quality of the superposition of the Tth-NBD reference model with 3HSC.pdb can be seen in Figure 2.22.

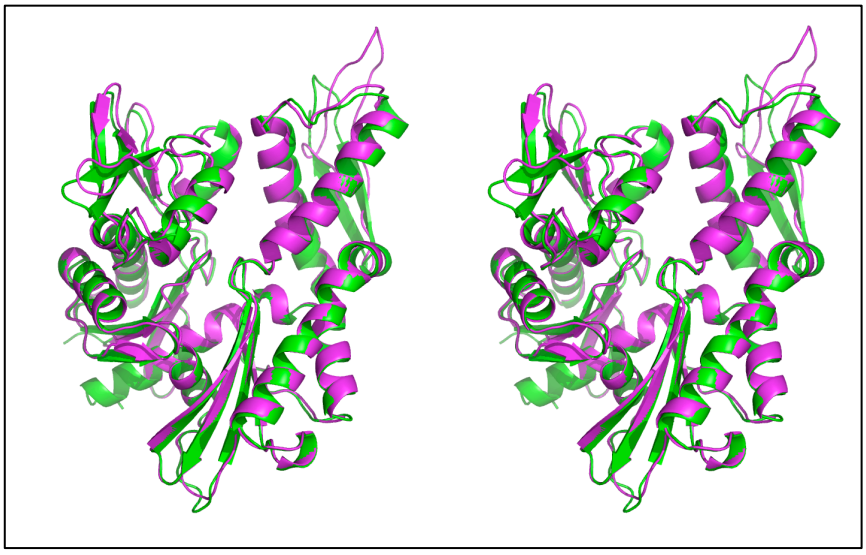

Figure 2.22: Reference model for Tth-NBD (purple), overlaid on bovine Hsc70 NBD (3HSC.pdb, shown in green)

2.14.7 SELF VALIDATION

An alternative estimation of the significance of the differences observed in Figures 2.12/13/14 can be provided by a "Jack-Knife" self-validation procedure. In self-validation, a certain fraction of the data is omitted at random, and the remaining data is then fitted to the relevant mathematical model. This process is carried out many times, and a distribution in fitting parameters is obtained. In order to assess the relationship of the self-validation data retention percentage, the distribution in the fitting parameters and the real error, we carried out the following model calculation. To a 3rd order polynomial function, random noise was added to an amplitude that allowed the data to be fitted back to a 3rd polynomial with a R^2 of 0.65 (see Figure 2.23). This was repeated ten times to generate ten independent rather noisy "data sets". The distribution in the ten sets of fitted parameters was taken as a measure of the achievable precision due to the "real" noise in the "data" (see Table 2.3). Subsequently, one of the "data sets" was analyzed using self-validation. When 60% of this single "data set" is kept at random and fitted back to a 3rd polynomial, and doing this ten times, one ends up with a similar distribution in the ten sets of fitted parameters

as for the "real" noise (see Table 2.3). Hence, we view self-validation with retention of 60% of the data as a realistic measure of the influence of measurement error on the precision of the fitted parameters.

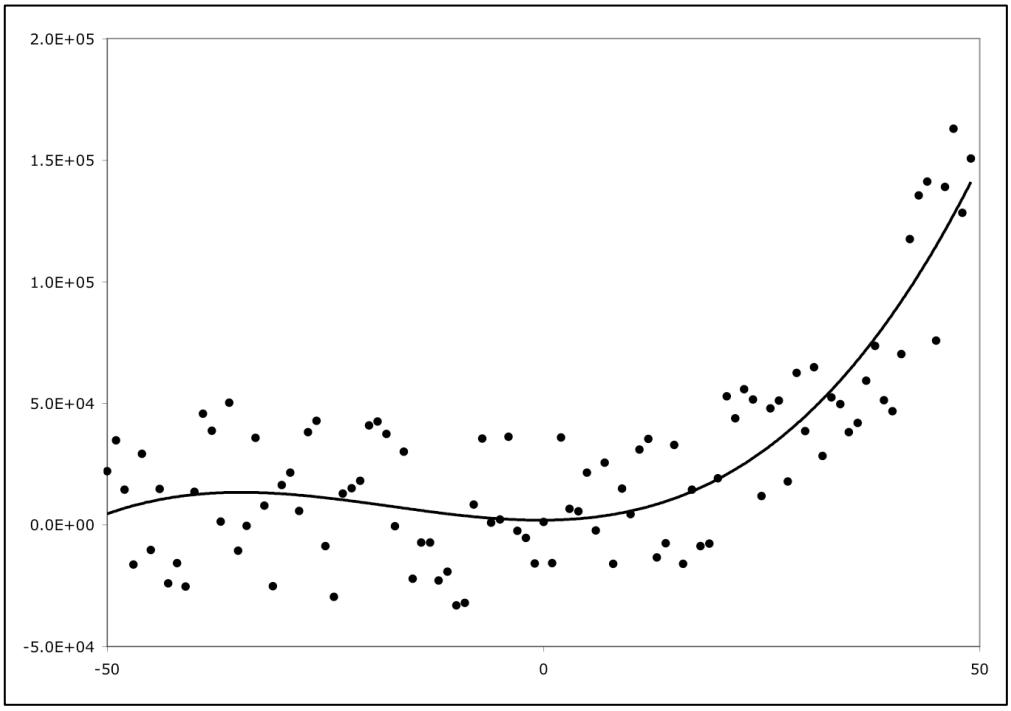

Figure 2.23: Testing the self-validation procedure

Data used to test the self-validation procedure illustrated in Table, below. The shown data points were generated as $y=0.5x^3+30x^2+100x+1000$ + noise. The noise was in the range ±40,000. The fitted line is $y=0.57x^3+29.8x^2+22x+1954$ with $R^2=0.70$.

Table S1.
Fit of polynomial data with random noise

Fit parameters	A	B	C	D	R^2
"data"1	0.60	34.55	-130.45	-4246.40	0.72
"data" 2	0.31	34.21	368.60	215.56	0.70
"data" 3	0.55	30.78	115.35	303.12	0.70
"data" 4	0.42	26.09	156.20	3773.10	0.63
"data" 5	0.57	29.76	21.77	1953.60	0.70
"data" 6	0.38	28.96	383.31	2848.80	0.69
"data" 7	0.79	28.89	-340.08	6941.70	0.69
"data" 8	0.56	28.87	24.17	-1637.30	0.70
"data" 9	0.49	31.97	18.58	5654.50	0.69
"data" 10	0.74	28.07	-339.59	-1692.50	0.70

average	0.54	30.22	27.79	1411.42	0.69
sigma	0.14	2.55	237.10	3318.49	0.02
Self-evaluation retaining, at random, 60% of "data" 5					
Fit parameters	A	B	C	D	R^2
try1	0.48	30.06	224.76	-755.49	0.68
try2	0.56	31.16	83.89	-1268.70	0.75
try3	0.31	29.26	573.67	3558.60	0.75
try4	0.70	26.71	-308.10	-702.05	0.66
try5	0.56	31.57	-85.56	-1595.90	0.70
try6	0.56	34.62	108.45	344.39	0.75
try7	0.71	29.06	-253.57	75.42	0.72
try8	0.66	29.78	-133.57	-1679.80	0.73
try9	0.60	30.73	-16.28	-3353.30	0.70
try10	0.55	24.27	118.81	5827.00	0.72
average	0.57	29.72	31.25	45.02	0.72
sigma	0.11	2.79	255.48	2708.18	0.03

Table 2.3: Tabulation of data used to verify the self validation procedure

The polynomial $y=0.5x^3+30x^2+100x+1000$ + noise ("data" 1-10) was fitted with the function $y=Ax^3+Bx^2+Cx+D$ using the routines available in Microsoft Excel. The noise was in the range ±40,000 Bottom: "data" 5 was fitted 10 times, retaining, at random, 60% of the entries.

Figures 2.24/25 show the results of 100 self-validations per domain using 60% of the RDC data, selected at random, computed with the REDCAT program. Even when using only 60% of the data, one obtains statistically significant differences in sub domain orientations as compared to the model structure. For each of the datasets used (with 60%) of total available RDCs, the best-fit solution was found and plotted. This procedure samples a fairly large input space. Consider for 40 RDCs (typically available in subdomain IA), 60% of this input set is 24 RDCs. As the arrangement of RDCs does not matter (this is a solution of a set of simultaneous linear equations), the number of ways in which this can be chosen is:

$$W_{60} = \frac{40!}{(40-24)!(24!)} = \frac{40!}{16!.24!} = 62852101650$$

Equation 2.1: Jacknifing choices

Thus, the input space contains billions of searchable combinations, and hence the jackknifing procedure is not biased to any particular solution. This is an excellent measure of how robust our input dataset happens to be.

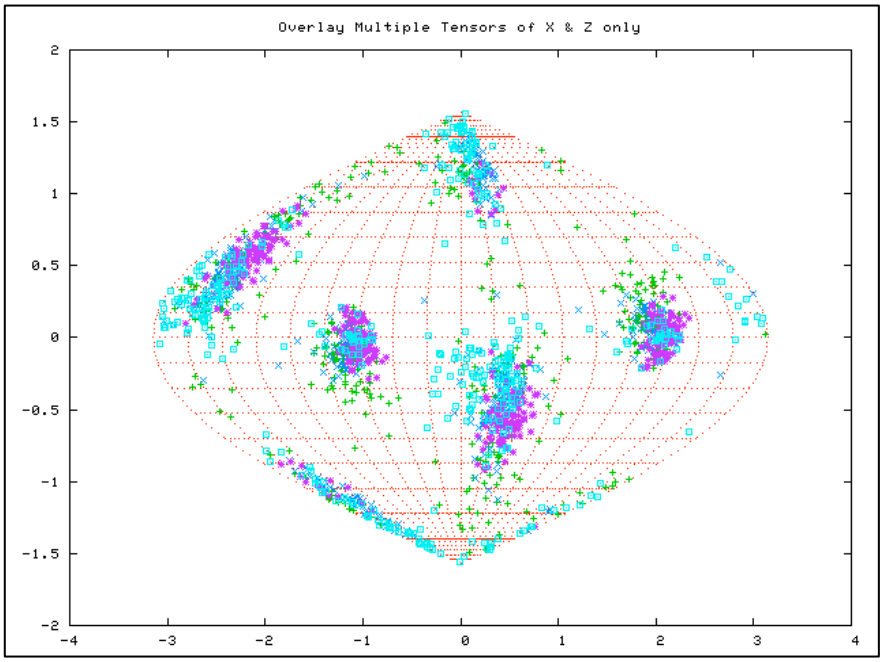

Figure 2.24: Sanson Flamsteed sinusoidal projection of subdomain orientation for Tth-NBD in the AMPPNP bound state with 60% data retained.

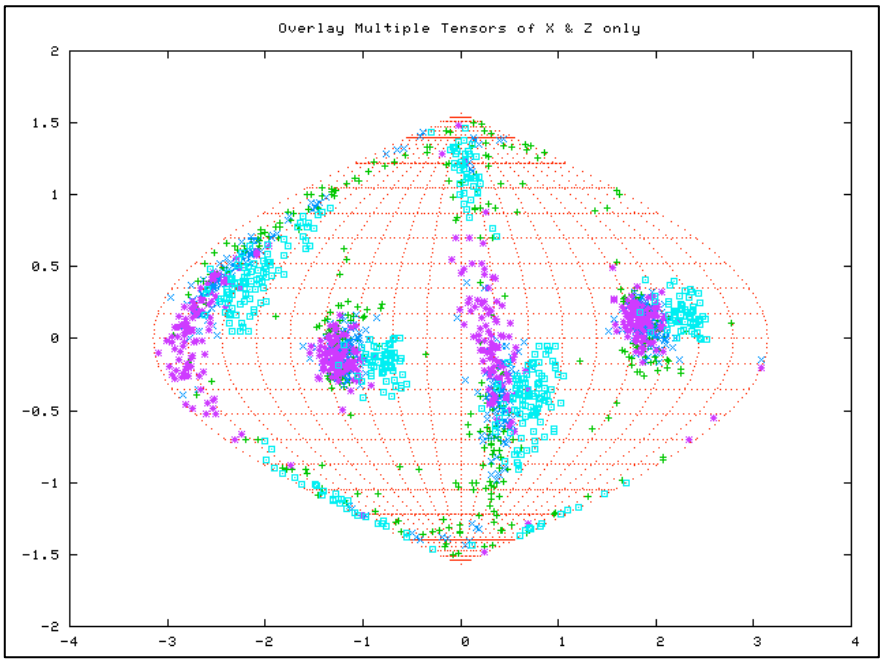

Figure 2.25: Sanson Flamsteed sinusoidal projection of subdomain orientation for Tth-NBD in the ADP.Pi bound state with 60% data retained.

Globe graphs (Sanson-Flamsteed plots) showing the distribution of the 60% (retained) self-validation ensembles of the principal alignment axes of the different subdomains of NBD of DnaK-Tth as derived from the NMR RDC measurements, using REDCAT [177]. Color code: sub domain IA (DnaK-Tth-count residues 1-37,109-180, 357-377; blue); sub domain IB (residues 38-108; green); sub domain IIA (residues 181-219, 307-356; purple) and sub domain IIB (residues 220-304; cyan). Scale: Horizontal: 20 ^{o} per gridline, vertical, 10 ^{o} per gridline. (The Szz principal alignment axes also appear at the other side of the globe (120 ^{o}East, and 5 ^{o}North). The Sxx axes appear as smears around 150 ^{o} West, 45 ^{o} North and 30 ^{o}East, 45 ^{o}South. The polar presences are low abundant alternative 90 ^{o} rotated solutions)

2.15 MOLECULAR MODELS

We obtained the molecular models for each of the states in five steps:

1. Rotate each subdomain into its Principle Axes System, (PAS) based on the average Euler angles obtained from the 60% retained self validation procedure (obtaining IA-PAS, IIA-PAS, IB-PAS, IIA-PAS)
2. Rotate the reference structure into the PAS of domain IA (TThREFPAS-IA)
3. Optimize the superposition of the Ca's in the 2nd structure of each PAS subdomain onto the corresponding subdomain of TThREFPAS-IA, using translation only
4. Reassemble the domains into a single structure
5. Check that this process has not altered the relative subdomain orientations by carrying out a REDCAT computation on the obtained models. Zero degree Euler angles should be and were obtained.

The obtained model structures were made available in pdb format in the supplementary materials of the results communicated to JMB.

The model of Tth-NBD in the AMPPNP state was used as a template for all super positions of structures shown. Corresponding Cα atom positions in secondary structure

elements of subdomain IA were used for all overlays.

2.15.1 ENSEMBLE OF RDC STRUCTURES

Using the in-house written Fortran programs, we calculated self-validations at the 60% level (retained) for an ensemble of 30 structures for the ADP state and for the AMPPNP state (see Table 2.1). The RMSDs of the three Euler angles were computed from these ensembles as shown in Table 2.1. Next, for each subdomain we calculated a molecular structure, as described above, for each of the 8 different sets of Euler angles: $<\alpha>\pm RMSD\alpha$, $<\beta>\pm RMSD\beta$ and $<\gamma>\pm RMSD\gamma$ as shown in Table 2.1.

2.15.2 STATISTICS OF THE ENSEMBLE

The 30 structures in the ensembles are represented by 8 structures describing the boundaries of the ensemble, like the 8 vertices of a cube. These boundaries represent the average RMSD multiplied by $\sqrt{3}$. While the ensemble contains 30 structures, it should actually be seen as ensemble of just 6 independent structures, since it takes 5 RDCs to determine an orientation (the 4-fold degeneracy in orientation is not an issue here). Under the assumption that this distribution of 6 structures can be described as a Gaussian around an average, the T-statistic can be used.

In the present case, the average orientation of subdomain IIA in the ADP state lies at the edge of a similar ensemble of that domain in the AMPPNP state, and vice versa (Figures 2.13/14). The T-statistic for similar ensembles, is defined as (http://en.wikipedia.org/wiki/Student%27s_t-test)

$$t_{1,2} = \frac{\langle X_1 \rangle - \langle X_2 \rangle}{\sigma_{X_1 X_2} \sqrt{\frac{2}{N}}}$$

Equation 2.2: T-statistic

Where $\langle X_1 \rangle$ and $\langle X_2 \rangle$ are the averages of the ensembles, where N is the number

of elements in each ensemble (the same), and where

$$\sigma_{X_1 X_2} = \sqrt{\frac{\sigma^2_{x_1} + \sigma^2_{x_2}}{2}}$$

Equation 2.3: Squared standard deviation

,

with $\sigma^2_{x_1} (\cong \sigma^2_{x_2})$ as the squared standard deviation of the distributions.

For the present case, with $\langle X_1 \rangle - \langle X_2 \rangle = \sigma_{X_1 X_2} . \sqrt{3}$ and N=6 one finds t1,2=3, which means that there is just a 3% probability that the observed difference in orientations for subdomains IIA between the two states is by chance (http://changingminds.org/explanations/research/analysis/t-test_table.htm). This provides a firm statistical underpinning to the structural differences observed in this study.

2.16 SUMMARY

We have detected structural differences in Tth-NBD between the ADP.Pi and the AMPPNP bound states. Specifically, the nucleotide binding cleft is more open in the ADP.Pi state, as a result of subdomain IIB rotating by ~20 degrees. This cleft opening has not been hitherto seen in X-ray studies of the isolated NBD. However, it has been seen in an NBD-exchange factor complex. Our result indicates that exchange factors are not required to force open the cleft, enabling nucleotide exchange. Thus, the role of nucleotide exchange factors must be rethought from an induced fit to a conformation capture model. This answers the initial question, posed in the introduction: what is the exact role of the exchange factor and how does it work? To wit: the internal nucleotide state of the NBD is communicated to the outside environment by a change in global conformation. This is recognized by nucleotide exchange factors (such as GrpE) which then stabilize the open cleft state of the NBD, and facilitate nucleotide exchange. However, NEFs do not actively force this very conformation change by themselves binding to the NBD. The NBD has already changed its global conformation when an

exchange factor binds. These ideas are encapsulated in Figs 2.26/26.

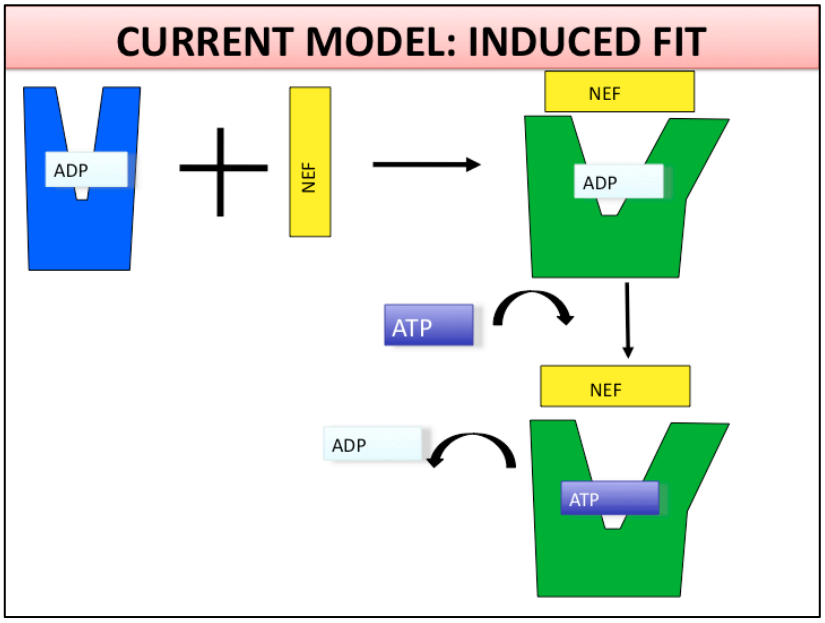

Figure 2.26: Conventional model of the exchange factor mechanism of Hsp70: Induced Fit

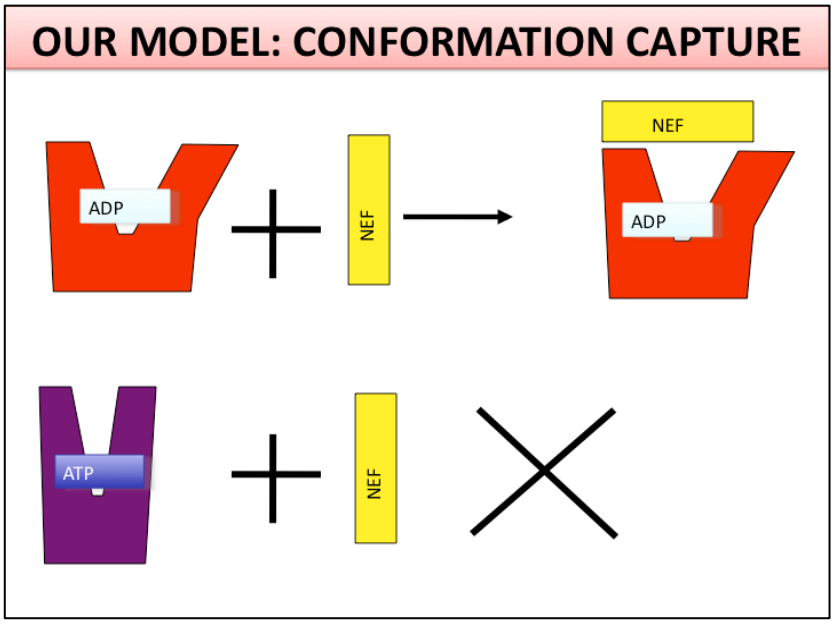

Figure 2.27: Our model of the exchange factor mechanism: Conformation Capture

The changes in the nucleotide binding cleft occur in tandem with changes in the bottom of the NBD: the IA/IIA interface cleft is more open in the AMPPNP bound state

as compared to the ADP.Pi state. Thus, the linker has space enough to be docked in the IA/IIA cleft in the AMPPNP (ATP mimic) state and is expelled in the ADP.Pi state. This provides a mechanistic explanation for why the NBD and SBD are loosely tethered and rotate semi-independently in the ADP.Pi state [92], while there is evidence for them to be docked in the ATP state [93]. Furthermore, it is immediately apparent that now a composite model can be tentatively drawn up for the two-way allostery between the NBD and the SBD. The changes in the nucleotide-binding cleft are meshed with changes at the linker interface. The ATPase activity of the NBD is boosted in the presence of substrate peptide in the substrate-binding cleft, which leads to dissociation of the SBD from the NBD docking site, promoting linker release and allowing IA/IIA cleft closure. This can be explained in terms of the syncromeshed conformational changes at the nucleotide binding cleft and the linker interface; this is illustrated in the figure below. Obviously, there are lacunae in this model, which need to be investigated in depth. One obvious target of investigation is how J-proteins come into the picture. This is the subject matter of Chapter 4.

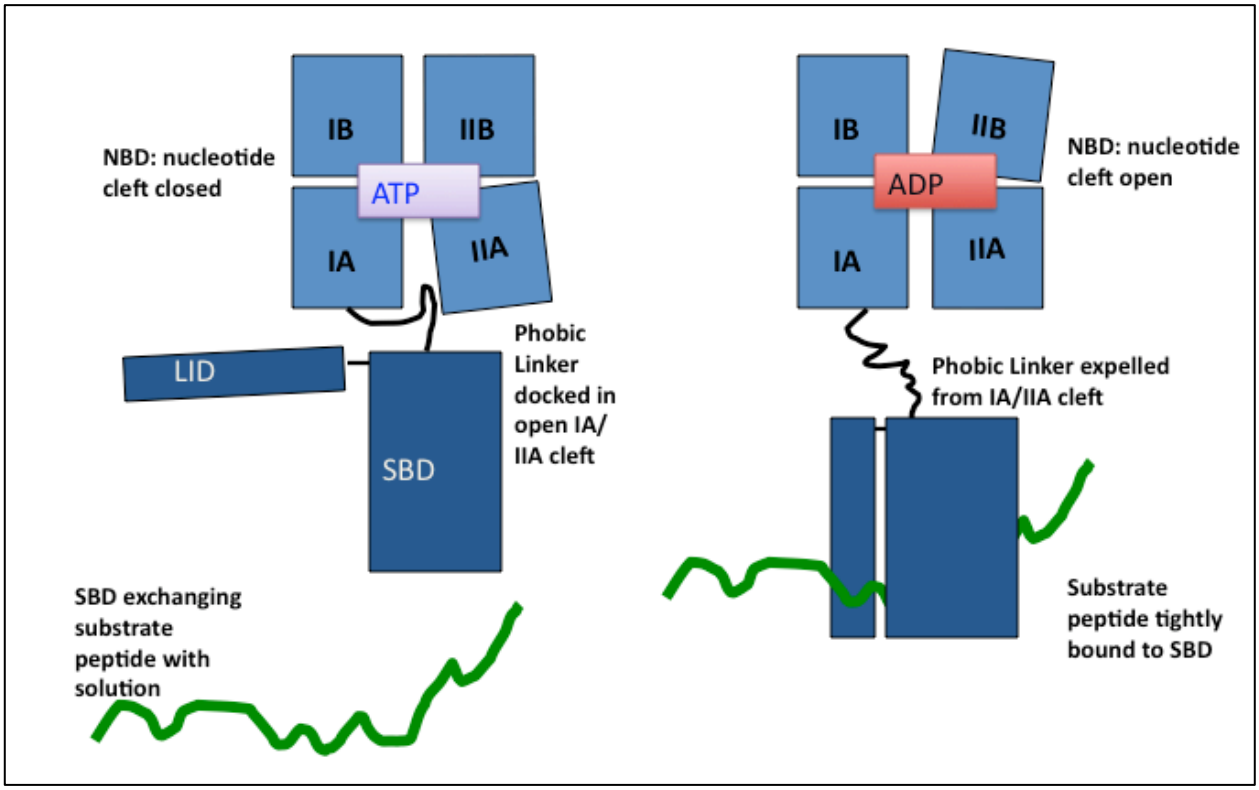

Figure 2.28: Proposed allosteric mechanism of Hsp70s

2.17 REFERENCES

[1] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-366.
[2] H. Schroder, T. Langer, F.U. Hartl, and B. Bukau, Dnak, Dnaj and Grpe Form a Cellular Chaperone Machinery Capable of Repairing Heat-Induced Protein Damage. Embo J. 12 (1993) 4137-4144.
[3] F.U. Hartl, and M. Hayer-Hartl, Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol 16 (2009) 574-81.
[4] C. Garrido, M. Brunet, C. Didelot, Y. Zermati, E. Schmitt, and G. Kroemer, Heat shock proteins 27 and 70: anti-apoptotic proteins with tumorigenic properties. Cell Cycle 5 (2006) 2592-601.
[5] J. Nylandsted, K. Brand, and M. Jaattela, Heat shock protein 70 is required for the survival of cancer cells. Ann N Y Acad Sci 926 (2000) 122-5.
[6] R. Wadhwa, S.C. Kaul, and Y. Mitsui, Cellular mortality to immortalization: mortalin. Cell Struct Funct 19 (1994) 1-10.
[7] R. Wadhwa, S. Takano, K. Kaur, C.C. Deocaris, O.M. Pereira-Smith, R.R. Reddel, and S.C. Kaul, Upregulation of mortalin/mthsp70/Grp75 contributes to human carcinogenesis. Int J Cancer 118 (2006) 2973-80.
[8] J.E. Gestwicki, G.R. Crabtree, and I.A. Graef, Harnessing chaperones to generate small-molecule inhibitors of amyloid beta aggregation. Science 306 (2004) 865-9.
[9] K.K. Chung, and T.M. Dawson, Parkin and Hsp70 sacked by BAG5. Neuron 44 (2004) 899-901.
[10] S. Krobitsch, and S. Lindquist, Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. Proc Natl Acad Sci U S A 97 (2000) 1589-94.
[11] T.V. Novoselova, B.A. Margulis, S.S. Novoselov, A.M. Sapozhnikov, J. van der Spuy, M.E. Cheetham, and I.V. Guzhova, Treatment with extracellular HSP70/HSC70 protein can reduce polyglutamine toxicity and aggregation. J. Neurochemistry 94 (2005) 597-606.
[12] J.L. Brodsky, and G. Chiosis, Hsp70 molecular chaperones: emerging roles in human disease and identification of small molecule modulators. Curr Top Med Chem 6 (2006) 1215-25.
[13] M.P. Mayer, S. Rudiger, and B. Bukau, Molecular basis for interactions of the DnaK chaperone with substrates. Biol Chem 381 (2000) 877-85.
[14] J.S. McCarty, A. Buchberger, J. Reinstein, and B. Bukau, The role of ATP in the functional cycle of the DnaK chaperone system. J. Mol. Biol. 249 (1995) 126-137.
[15] K.M. Flaherty, C. Deluca-Flaherty, and D.B. McKay, 3-Dimensional Structure of the ATPase Fragment of a 70k Heat-Shock Cognate Protein. Nature 346 (1990) 623-628.
[16] K.M. Flaherty, S.M. Wilbanks, C. DeLuca-Flaherty, and D.B. McKay, Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II.

Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. J Biol Chem 269 (1994) 12899-907.
[17] H. Sondermann, C. Scheufler, C. Schneider, J. Hohfeld, F.U. Hartl, and I. Moarefi, Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. Science 291 (2001) 1553-7.
[18] X.T. Zhu, X. Zhao, W.F. Burkholder, A. Gragerov, C.M. Ogata, M.E. Gottesman, and W.A. Hendrickson, Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272 (1996) 1606-1614.
[19] R.C. Morshauser, H. Wang, G.C. Flynn, and E.R. Zuiderweg, The peptide-binding domain of the chaperone protein Hsc70 has an unusual secondary structure topology. Biochemistry 34 (1995) 6261-6.
[20] H. Wang, Y. Pang, A.V. Kurochkin, W. Hu, G.C. Flynn, and E.R.P. Zuiderweg, The solution structure of the 21 kDa chaperone protein DnaK substrate binding domain: a preview of chaperone - protein interaction. Biochemistry 37 (1998) 7929-7940.
[21] R.C. Morshauser, W. Hu, H. Wang, Y. Pang, G.C. Flynn, and E.R. Zuiderweg, High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. J Mol Biol 289 (1999) 1387-403.
[22] M. Pellecchia, D.L. Montgomery, S.Y. Stevens, C.W. Vander Kooi, H.P. Feng, L.M. Gierasch, and E.R. Zuiderweg, Structural insights into substrate binding by the molecular chaperone DnaK. Nat Struct Biol 7 (2000) 298-303.
[23] S.Y. Stevens, S. Cai, M. Pellecchia, and E.R. Zuiderweg, The solution structure of the bacterial HSP70 chaperone protein domain DnaK(393-507) in complex with the peptide NRLLLTG. Protein Sci 12 (2003) 2588-96.
[24] M. Vogel, M.P. Mayer, and B. Bukau, Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. J Biol Chem 281 (2006) 38705-11.
[25] J.F. Swain, G. Dinler, R. Sivendran, D.L. Montgomery, M. Stotz, and L.M. Gierasch, Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol Cell 26 (2007) 27-39.
[26] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E.coli Hsp70 (DnaK) chaperone
complexed with ADP and substrate.
. Proc. Natl. Acad. Sci. U. S. A. (2009) in the press.
[27] Q. Liu, and W.A. Hendrickson, Insights into hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-120.
[28] M.C. O'Brien, K.M. Flaherty, and D.B. McKay, Lysine 71 of the chaperone protein Hsc70 Is essential for ATP hydrolysis. J. Biol. Chem. 271 (1996) 15874-15878.
[29] S.M. Wilbanks, L.L. Chen, H. Tsuruta, K.O. Hodgson, and D.B. Mckay, Solution Small-Angle X-Ray-Scattering Study of the Molecular Chaperone Hsc70 and Its Subfragments. Biochemistry 34 (1995) 12095-12106.
[30] M.C. O'Brien, and D.B. McKay, Threonine 204 of the chaperone protein Hsc70 influences the structure of the active site, but is not essential for ATP hydrolysis. J. Biol. Chem. 268 (1993) 24323-24329.
[31] D.B. McKay, Structure and mechanism of 70-kDa heat-shock-related proteins. Adv. Protein Chem. 44 (1993) 67-98.

[32] S.M. Wilbanks, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. II. Potassium binds specifically in the ATPase active site. J Biol Chem 270 (1995) 2251-7.
[33] J. Jiang, K. Prasad, E.M. Lafer, and R. Sousa, Structural basis of interdomain communication in the Hsc70 chaperone. Mol. Cell 20 (2005) 513-524.
[34] J. Jiang, E.G. Maes, A.B. Taylor, L. Wang, A.P. Hinck, E.M. Lafer, and R. Sousa, Structural basis of J cochaperone binding and regulation of Hsp70. Mol Cell 28 (2007) 422-33.
[35] M.C. O'Brien, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. I. Potassium is required for optimal ATPase activity. J. Biol. Chem. 270 (1995) 2247-2250.
[36] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-66.
[37] D. Klostermeier, R. Seidel, and J. Reinstein, Functional properties of the molecular chaperone DnaK from Thermus thermophilus. J Mol Biol 279 (1998) 841-53.
[38] 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. Proc. Natl. Acad. Sci. USA 92 (1995) 9279-9283.
[39] M.W. Fischer, J.A. Losonczi, J.L. Weaver, and J.H. Prestegard, Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. Biochemistry 38 (1999) 9013-9022.
[40] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-5.
[41] M. Revington, and E.R. Zuiderweg, TROSY-driven NMR backbone assignments of the 381-residue nucleotide-binding domain of the Thermus Thermophilus DnaK molecular chaperone. J Biomol NMR 30 (2004) 113-4.
[42] N. Tjandra, J.G. Omichinski, A.M. Gronenborn, G.M. Clore, and A. Bax, Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat. Struct. Biol. 4 (1997) 732-738.
[43] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, and J.H. Prestegard, Structural and dynamic analysis of residual dipolar coupling data for proteins. J. Am. Chem. Soc. 123 (2001) 1416-1424.
[44] D. Yang, J.R. Tolman, N.T. Goto, and L.E. Kay, An HNCO-based Pulse Scheme for the Measurement of 13C[agr]-1H[agr] One-bond Dipolar couplings in 15N, 13C Labeled Proteins. J. Biomol. NMR 12 (1998) 325-332.
[45] D. Yang, R.A. Venters, G.A. Mueller, W.Y. Choy, and L.E. Kay, TROSY-based HNCO pulse sequences for the measurement of 1HN-15N, 15N-13CO, 1HN-13CO, 13CO-13C[agr] and 1HN-13C[agr] dipolar couplings in 15N, 13C, 2H-labeled proteins. J. Biomol. NMR 14 (1999) 333-343.
[46] H. Valafar, and J.H. Prestegard, REDCAT: a residual dipolar coupling analysis tool. J Magn Reson 167 (2004) 228-41.

[47] S.M. Wilbanks, L. Chen, H. Tsuruta, K.O. Hodgson, and D.B. McKay, Solution small-angle X-ray scattering study of the molecular chaperone Hsc70 and its subfragments. Biochemistry 34 (1995) 12095-12106.
[48] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-435.
[49] Z. Polier S Fau - Dragovic, F.U. Dragovic Z Fau - Hartl, A. Hartl Fu Fau - Bracher, and A. Bracher, Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. Cell 133 (2008).
[50] J. Schuermann Jp Fau - Jiang, J. Jiang J Fau - Cuellar, O. Cuellar J Fau - Llorca, L. Llorca O Fau - Wang, L.E. Wang L Fau - Gimenez, S. Gimenez Le Fau - Jin, A.B. Jin S Fau - Taylor, B. Taylor Ab Fau - Demeler, K.A. Demeler B Fau - Morano, P.J. Morano Ka Fau - Hart, J.M. Hart Pj Fau - Valpuesta, E.M. Valpuesta Jm Fau - Lafer, R. Lafer Em Fau - Sousa, and R. Sousa, Structure of the Hsp110:Hsc70 nucleotide exchange machine. Molecular Cell 31 (2008).
[51] D. Kern, and E.R. Zuiderweg, The role of dynamics in allosteric regulation. Curr Opin Struct Biol 13 (2003) 748-57.
[52] H.M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, and D.J. Patel, Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. J Mol Biol 315 (2002) 95-102.
[53] Y. Zhang, and E.R. Zuiderweg, The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. Proc Natl Acad Sci U S A 101 (2004) 10272-7.
[54] W. Rist, C. Graf, B. Bukau, and M.P. Mayer, Amide hydrogen exchange reveals conformational changes in hsp70 chaperones important for allosteric regulation. J Biol Chem 281 (2006) 16493-501.
[55] A. Buchberger, H. Theyssen, H. Schroder, J.S. McCarty, G. Virgallita, P. Milkereit, J. Reinstein, and B. Bukau, Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication. J. Biol. Chem. 270 (1995) 16903-16910.
[56] M. Revington, Y. Zhang, G.N. Yip, A.V. Kurochkin, and E.R. Zuiderweg, NMR investigations of allosteric processes in a two-domain Thermus thermophilus Hsp70 molecular chaperone. J Mol Biol 349 (2005) 163-83.
[57] M.R. Hansen, L. Mueller, and A. Pardi, Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol. 5 (1998) 1065-1074.
[58] F. Ottiger M Fau - Delaglio, A. Delaglio F Fau - Bax, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra.
[59] J.A. Lukin, G. Kontaxis, V. Simplaceanu, Y. Yuan, A. Bax, and C. Ho, Quaternary structure of hemoglobin in solution. Proc Natl Acad Sci U S A 100 (2003) 517-20.
[60] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc Natl Acad Sci U S A 106 (2009) 8471-6.

[61] Q. Liu, and W.A. Hendrickson, Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-20.

CHAPTER 3 MEASUREMENT AND INTERPRETATION OF ^{15}N-^{1}H RESIDUAL DIPOLAR COUPLINGS IN LARGER PROTEINS

A large fraction of this material in this chapter has been accepted for publication in the Journal of Magnetic Resonance by A. Bhattacharya, M. Revington and E.R.P. Zuiderweg .

3.1 ABSTRACT

A decade ago, Dr. L.E. Kay and co-workers described an ingenious HNCO-based triple resonance experiment from which several protein backbone RDCs can be measured simultaneously [176]. They implemented a J-scaling technique in the ^{15}N dimension of the 3D experiment to obtain the NH RDCs. We have used this idea to carry out J-scaling in a 2D ^{15}N-^{1}H TROSY experiment and have found it to be an excellent method to obtain NH RDCs for larger proteins up to 70 kDa, far superior to commonly used HSQC in-phase/anti-phase and HSQC/TROSY comparisons. Here, this method, dubbed "RDC-TROSY" is discussed in detail and simulations are used to assess the limits of its utility. Prominent in the latter analysis is the evaluation of the effect of amide proton flips on the "RDC-TROSY" linewidths. The details of the technical and computational implementations of these methods for the determination of domain-orientations in 45-60 KDa Hsp70 chaperone protein constructs are described.

3.2 INTRODUCTION

The basic NMR Hamiltonian for a two-spin system can be written down as follows:

$$\hat{H} = \omega_H I_Z^H + \omega_N I_Z^N + 2\pi \hat{I}^H \, \bar{J}_{HN} \, \hat{I}^N - \frac{\mu_0}{4\pi} \frac{\gamma_H \gamma_N \hbar}{r_{HN}^3} \left(\left(3\hat{I}^H \, \hat{r}_{HN} \right) \left(3\hat{I}^N \, \hat{r}_{HN} \right) - \hat{I}^H \, \hat{I}^N \right)$$

Equation 3.1: The NMR Hamiltonian for a two spin system

In this equation, ω is the chemical shift, J_{HN} is the scalar coupling, γ is the gyromagnetic ratio, r_{HN} is the distance between the two nuclei and I is nuclear magnetization vector.

Here, the two spins in question are a ^{1}H and a ^{15}N nucleus, together constituting the ubiquitous NH group of the protein backbone. NMR spectroscopy is, quite simply the manipulation of this Hamiltonian, as described in Equation 3.1 to measure different parameters, in isolation and in conjunction. The first and most basic such parameter to be measured is the chemical shift, which is described in the first two terms of Equation 3.1, and is the characteristic frequency of the nucleus in question [105].

$$\hat{H} = \omega_H I_Z^H + \omega_N I_Z^N$$

Equation 3.2: The chemical shift Hamiltonian

The interaction terms can be broken up into the J-coupling, which is mediated by electron clouds between nuclei and the dipolar coupling. Since the J-coupling (also called the scalar coupling) is mediated by intervening electrons, it is an indicator of the chemical environment of the nucleus under study. Under conditions of 'weak coupling' [105] and the secular approximation, this term can be written in terms of the Z-magnetization of the interacting nuclei. This simplifies the math considerably.

$$\hat{H}_{J-coupling} = 2\pi J_{HN} I_Z^H I_Z^N$$

Equation 3.3: The scalar coupling Hamiltonian

The dipole-dipole interaction acts directly through space, and is thus a measure of the spatial environment of a nucleus under study: it identifies its neighbours. Thus, this interaction is indicative of the structure of the protein. Again, under the secular approximation, the dipolar coupling can be written as:

$$\hat{H}_{dipolarcoupling} = \frac{\mu_0}{4\pi}\frac{\gamma_H \gamma_N \hbar}{r_{HN}^3}\left(3\cos^2\theta - 1\right) I_Z^H I_Z^N = 2\pi D_{NH} I_Z^H I_Z^N$$

Equation 3.4: Dipolar Coupling Hamiltonian

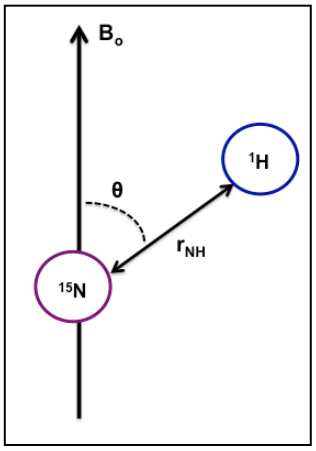

Figure 3.1: Two spin system

Here, θ is the angle (Fig. 3.1) between the inter-nuclear vector and the external magnetic field (Bo). Interestingly, it can be seen that experiments which measure the coupling terms can now be used to measure both the scalar and the dipolar coupling simultaneously. This is of great import.

$$\hat{H}_{Coupling} = 2\pi(J_{HN} + D_{NH})I_Z^H I_Z^N$$

Equation 3.5: Two spin reduced coupling Hamiltonian

The dipolar coupling term D_{NH} depends on $(3\cos^2\theta\text{-}1)$, as shown. Due to the protein tumbling very rapidly in solution (typical rotational correlation time $\tau_C \sim 22$ nsec), this term is averaged out to zero.

$$\int_{\theta=0}^{\theta=\pi} P_2(\cos\theta)\sin\theta d\theta = \frac{1}{2}\int_{\theta=0}^{\theta=\pi}\left(3\cos^2\theta - 1\right)\sin\theta d\theta = 0$$

Equation 3.6: Spatial averaging of dipolar coupling to zero

However, under conditions where the isotropy of the buffer/medium is disturbed, this dipolar coupling is not averaged out to zero, and we are able to measure a 'residual' dipolar coupling.

3.3 ABCs OF PROTEIN STRUCTURE DETERMINATION

The modus operandi in protein structure determination is well established. It consists of the following steps:

1. Assign (or fingerprint) the protein [113; 189; 190]

2. Find distance constraints (usually NOEs [118; 191; 192; 193] – also a dipole-dipole interaction term).

3. Use distance constraints to find a high-resolution structure of the protein (perhaps by using NIH-XPLOR [121; 194])

4. Extract dynamical information

From a very simplified point of view, fingerprinting a protein can be boiled down to finding cross peaks in 2D/3D spectra for each residue (Fig 3.2). The characteristic 2D NMR spectrum of a protein is usually a ^{15}N-^{1}H spectrum recorded so as to generate a single peak for each NH group. Using special pulse sequences, it is possible to record spectra where each NH crosspeak is split into a quartet, based upon the coupling constants. The single characteristic cross peak that is seen in an HSQC experiment [195] is split once in the ^{15}N dimension and once again in the ^{1}H dimension, both times by (J+D) (in Hz) , which is the total coupling. For small proteins, all four members of this quartet are of the same intensity and linewidth. For large proteins, relaxation effects start playing a more significant role and the four peaks in the quartet have different profiles.

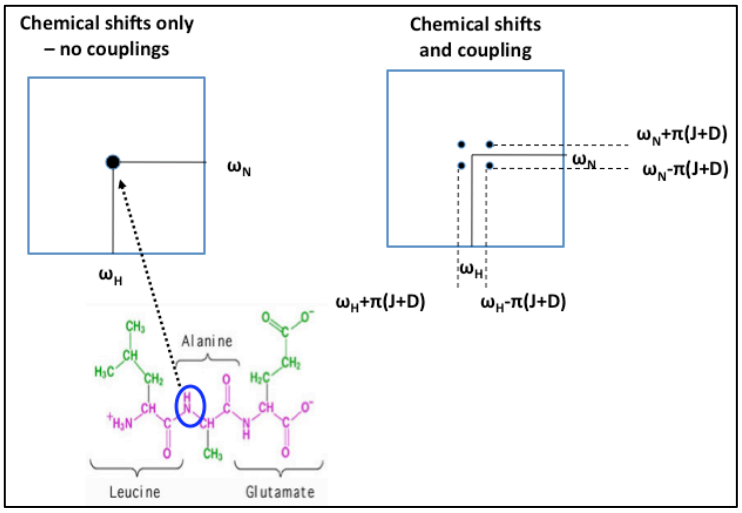

Figure 3.2: Fingerprinting a protein: each NH group corresponds to a crosspeak

3.4 STRUCTURE DETERMINATION WHEN NOEs ARE NOT AVAILABLE

NOEs are difficult to extract for large proteins (~45 kDa). At such molecular weights it becomes more convenient to use other techniques to refine structures based

on, say, existing X-ray derived structures or homology models rather than try to solve a protein structure de novo. The use of residual dipolar couplings (RDC) in biomolecular NMR can enhance the precision of local structure determination [196], and perhaps more importantly, allows unique access to meso-scale structural features of biomolecules in solution, such as relative domain orientations [170; 171] and dynamics[197] [182]. In principle, RDCs exist for every pair of nuclear dipoles in the aligned macromolecule, and several methods to measure these have appeared in the literature. In proteins, NH residual dipolar couplings are the most easily accessible. For very small proteins, they can be directly measured using a ^{15}N-^{1}H HSQC spectrum without ^{1}H decoupling in the ^{15}N dimension [198] (see Fig 3.3). The difference frequency between these NHa and NHb components is equal to ^{1}JNH + RDCNH (Hz) . The experiment in question is modified from the basic HSQC (Heteronuclear Single Quantum Spectroscopy) experiment, a schematic of which is shown here.

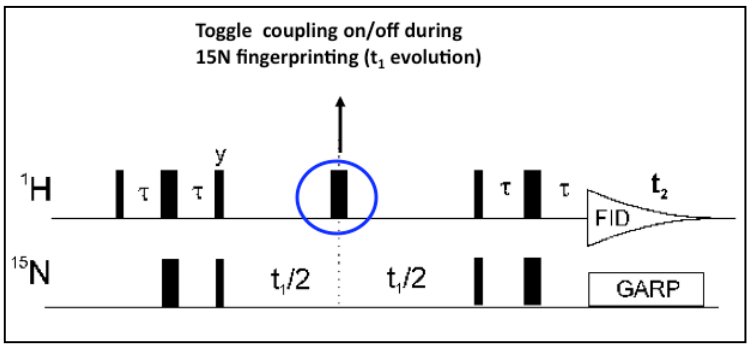

Figure 3.3: HSQC pulse sequence

3.4.1 EXTRACTING COUPLINGS VIA AN IP/AP EXPERIMENT

This pulse sequence (Fig 3.3) allows us to toggle the coupling between ^{15}N and ^{1}H on and off (with appropriate phase corrections, not shown here). Thus, we can either resolve each NH group as a single Lorentzian peak, or as a Lorentzian split by a frequency difference proportional to the coupling (see Fig 3.3).

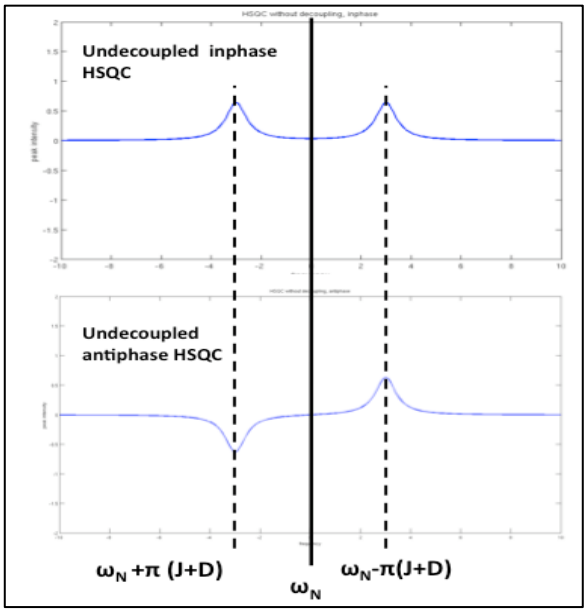

Figure 3.4: One-dimensional projection of cross peaks in a decoupled and coupled HSQC experiment

Hence, by comparing the spectra recorded with and without decoupling (see Fig 3.3/4), we can easily measure J_{NH} ($+D_{NH}$). However, this technique does have its limitations: the spectra become rapidly crowded with increasing protein size and number of peaks.

Resolution can be improved by separating the two doublet components by the in-phase/anti-phase (IP/AP) technique [199]. Adding and subtracting in-phase and anti-phase experiments results in spectra where we observe either α or β peaks. In Fig 3.5, we see that adding the IP and AP peaks results in the α-peak, and the difference results in the β-peak. Overlays will result in peaks for corresponding residues being separated by the amount of coupling. These methods start to fail for proteins of increasing size, which are investigated using large static magnetic fields (B_0). Because of ^{15}N CSA / ^{15}N-^{1}H dipole-dipole cross-correlated relaxation, the NHb doublet line becomes much broader than the NHa line (see Figure 3.6).

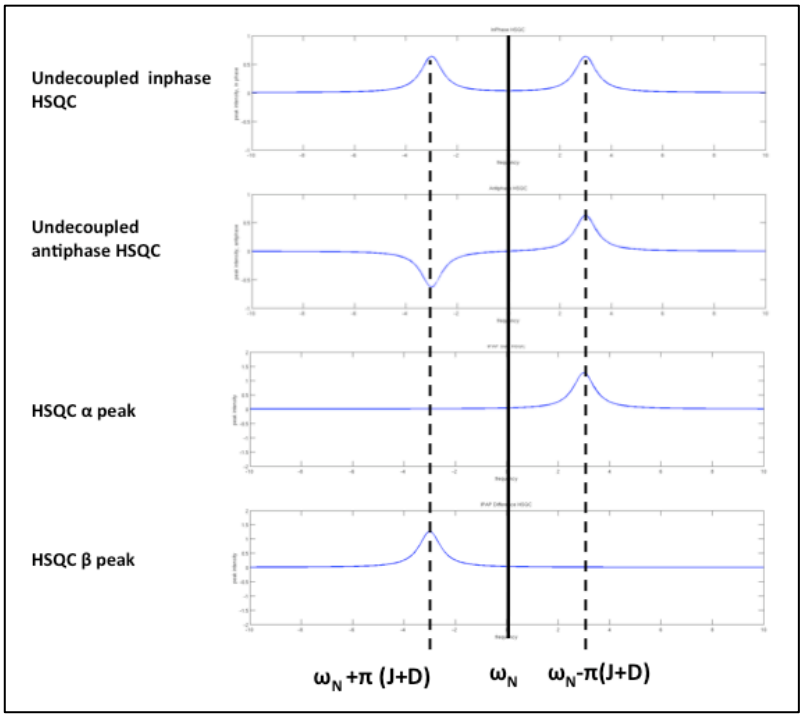

Figure 3.5: IP/AP sum and difference spectra leading to alpha and beta peaks

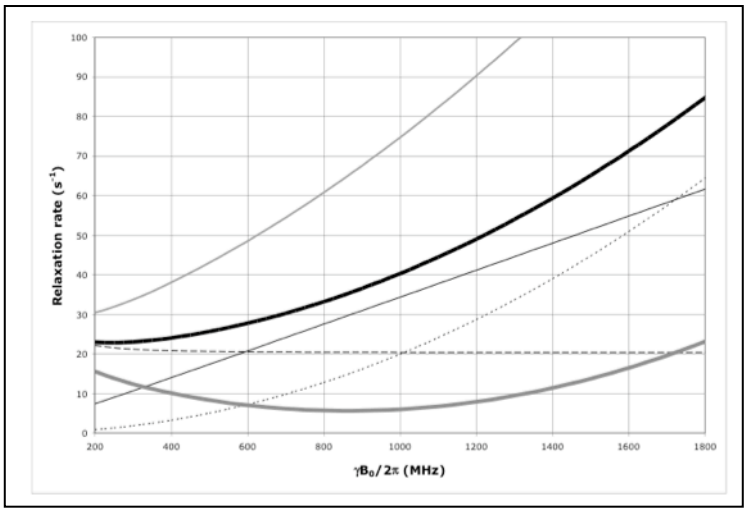

Figure 3.6: Field dependence of the ^{15}N-TROSY effect

The field dependence of the protein ^{15}N TROSY effect for τ_c =22 ns calculated using an rhombic ^{15}N CSA tensor with σ_{11}=230 σ_{22}=87 and σ_{33}=60 ppm, with the direction of σ_{11} 17 degress off the NH-bond of 1.04 Å using Eqs. (3.7-3.10). Thick solid line: total ^{15}N R_2 as encountered in a HSQC or HMQC; Thick dashed line: ^{15}N R_2 due to ^{15}N-^{1}H dipolar relaxation; Thin dashed line: ^{15}N R_2 due to ^{15}N CSA relaxation; Thin grey line: the ^{15}N R_2 ^{15}N-^{1}H dipolar / ^{15}N CSA cross correlated

relaxation rate; Thick grey line: ^{15}N R_2 as encountered in a TROSY (i.e. total ^{15}N R_2 minus ^{15}N R_2 ^{15}N-^{1}H dipolar / ^{15}N CSA cross correlation). Thin grey line: ^{15}N R_2 of the anti-TROSY line (i.e. total ^{15}N R_2 plus ^{15}N R_2 ^{15}N-^{1}H dipolar / ^{15}N CSA cross correlated relaxation rate).

3.4.2 WHEN IP/AP HSQC FAILS

The result of the NHβ peak being much broader than the NHα peak is that finding the coupling between them (JNH+DNH) is very difficult (see Fig 3.7).

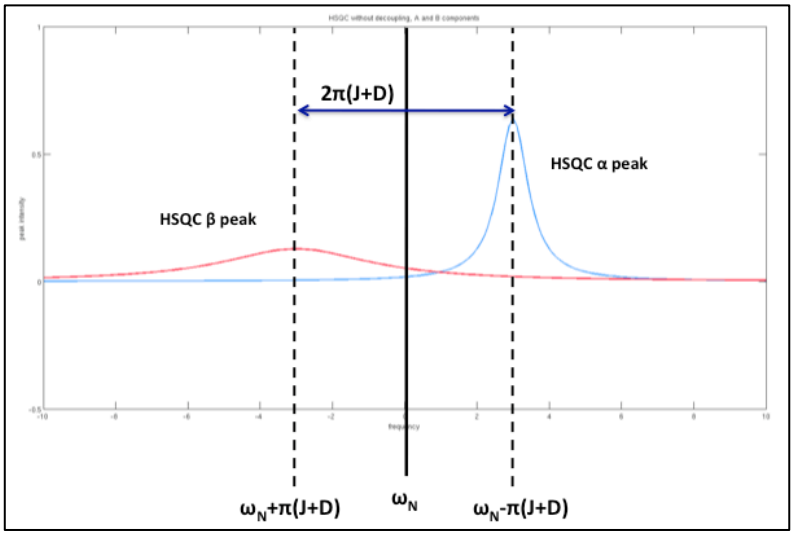

Figure 3.7: Undecoupled HSQC on a large protein

The situation is even direr than has been just discussed. Apart from the different peak heights of the HNα and HNβ peaks, a decoupled HSQC (which contains magnetization from both peak components) is itself rendered dramatically ineffective as molecular size increases. An increase in molecular size leads to an increase in the protein's rotational correlation time, which, in turn affects its transverse relaxation rate and hence linewidth [107]. Before long, large protein spectroscopy becomes impossible as all crosspeaks are broadened out into oblivion.

But fortuitously, the very effect that leads to the failure of HSQC (and IP/AP techniques) can be used to resurrect some of the magnetization that is lost to relaxation. The result of the twin processes of ^{15}N Chemical Shift Anisotropy (CSA) and

^{15}N-^{1}H dipole-dipole coupling result in the quartet of peaks in the undecoupled HSQC having different linewidths and intensities. The two processes act individually and together to broaden out two of the peaks in one dimension, and one peak in both dimensions. This is illustrated in Fig. 3.9. But the last peak is one where these two effects oppose each other to lead to a smaller linewidth.

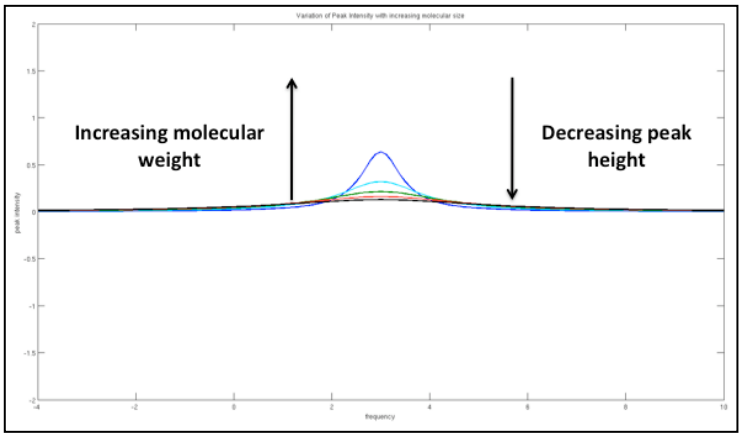

Figure 3.8: Peak height drops as molecular weight increases

Thus, ^{15}N CSA / ^{15}N-^{1}H dipole-dipole cross correlated relaxation can be made to work 'against' each other, thus negating some of relaxation effects experienced by large proteins [194; 195]. Indeed, the effect is exploited in TROSY spectroscopy, which selects for the narrow component [200], shown as the circled black peak in the right spectrum in Fig 3.9. Simultaneously, this also means that it becomes very difficult to obtain the frequency of the "anti-TROSY" component for larger systems, needed for the RDC determination, whether one selects for this peak directly, or by an IP/AP difference method. This is the circled red peak in the right spectrum of Fig 3.9. Figure 3.10 illustrates this difficulty for a 45 kDa protein (τ_c = 22 ns) shown in an actual spectrum.

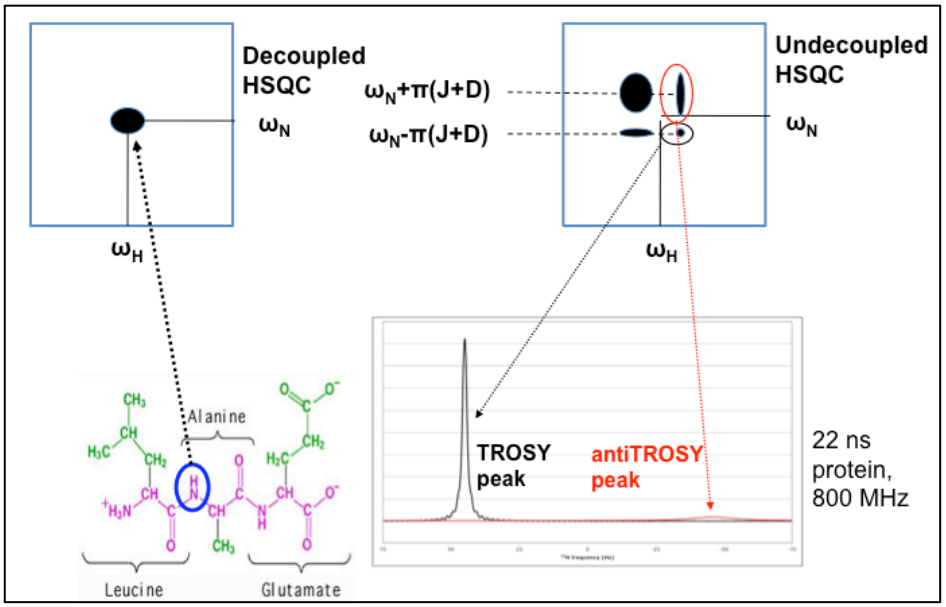

Figure 3.9: Quartet of peaks in an undecoupled HSQC seen for a large protein

One approach is to obtain the NHa - NHb difference frequency from measuring the ^{15}N frequency difference between a TROSY spectrum and a decoupled HSQC spectrum [188]. The difference is equal to (^{1}JNH + ^{1}RDCNH)/2. The improvement is substantial (see Figure 3.11), but still leaves much to be desired.

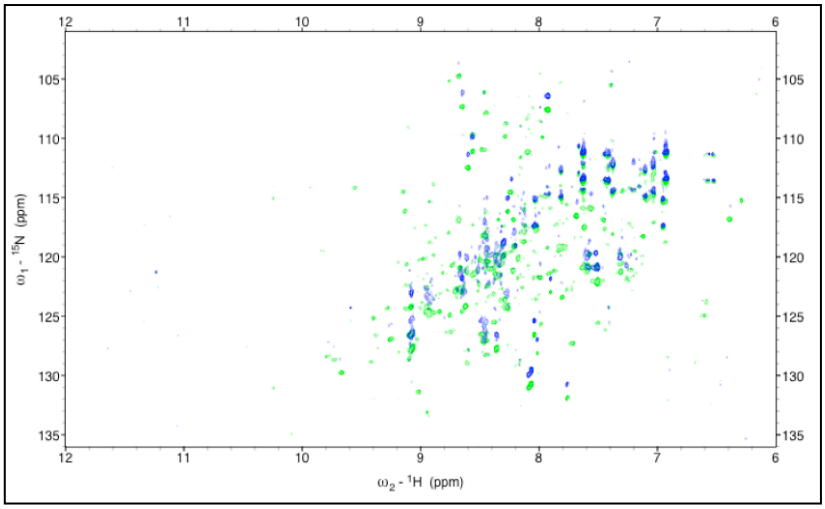

Figure 3.10: IPAP spectra run on Tth-NBD

In-phase anti-phase difference spectra of DnaKTth-NBD at 15 ^{0}C (τ_c =22ns) Green: NH a peaks; Blue: NH b peaks.

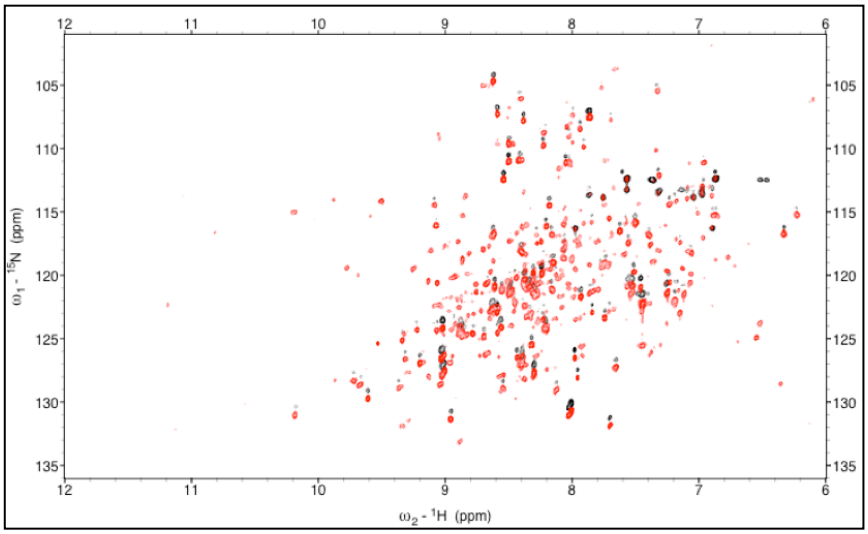

Figure 3.11: Comparison of decoupled HSQC and TROSY spectra for Tth-NBD

Red: TROSY spectrum of DnaK Tth-NBD at 15 ^{0}C (τ_c =22ns). Black: HSQC spectrum of DnaK-TTh NBD (shifted 45 Hz in the ^{1}H dimension)

3.4.3 USING TROSY EFFECTS TO MEASURE COUPLINGS – AN OVERVIEW

A decade ago, Kay and co-workers described an ingenious HNCO-based triple resonance experiment from which several protein backbone RDCs can be measured at once [176]. They implemented a J-scaling technique [201] in the ^{15}N dimension of the 3D experiment to obtain the NH RDC. We have used this idea to carry out J-scaling in a 2D TROSY experiment and have found it to be an excellent method to obtain NH RDCs for larger proteins as described in several publications of our group [92; 183; 202; 203] . We call the experiment "RDC-TROSY". Here, we describe this method in detail and assess the limits of its utility by simulations. Prominent in the latter analysis is the assessment of the effect of amide proton flips [204] on the RDC-TROSY linewidths.

3.5 A BRIEF DESCRIPTION OF THIS PROJECT

We describe the technical and computational implementation of our methods by determining domain-orientations in an Hsp70 chaperone protein. The 70 kDa Hsp70 chaperone protein family plays fundamental roles in protein refolding and regulation

using ATP-driven cycles of protein binding and release [205]. Hsp70 proteins consist of a 45 kDa N terminal Nucleotide Binding (ATPase) Domain (NBD) and a 25kDa C-terminal Substrate Binding Domain (SBD). The NBD is divided into four subdomains: IA, IB, IIA and IIB[49].

We have studied four forms of the TTDnaK Nucleotide Binding Domain denoted as DnaK Tth-NBD: 1) The *apo* form, with no nucleotide bound; 2) The ADP.PO_4 bound form and 3) The AMPPNP bound form (AMPPNP being a slowly hydrolyzed ATP analogue, we expect this to be similar to the ATP bound state). An ADP form of a construct that also contained the substrate-binding domain (DnaK Tth-NBD-SBD) was also studied.

We observe significant, and different, NBD sub-domain rearrangements in these four states. In particular, subdomain IIB can change its orientation by as much as 20 degrees. Interestingly, such orientational changes were not observed in earlier X-ray structures of these proteins in different nucleotide states[55; 155; 164; 166; 168; 206].

Measurement of the RDCs for these orientational changes for a protein of this size required the use of the RDC-TROSY, which should be a useful general approach for other larger proteins as well.

3.6 THEORY

3.6.1 COMPUTATIONS OF TROSY LINE WIDTHS

The ^{15}N linewidths of the TROSY, anti-TROSY and HSQC linewidths were computed for a protein with 22 ns correlation time, using the standard equations for the ^{15}N-^{1}H dipolar relaxation

$$R_2^{DD} = d^2\left\{4J(0) + 3J\left(\omega_N\right) + 6J\left(\omega_H\right) + J\left(\omega_H - \omega_H\right) + 6J\left(\omega_H + \omega_H\right)\right\}$$

Equation 3.7: Transverse relaxation due to dipole-dipole coupling

for ^{15}N CSA relaxation due to a rhombic CSA tensor

$$R_2^{CSA} = c^2 \left\{ \left[\sigma_{11} - \sigma_{33} \right]^2 + \left[\sigma_{22} - \sigma_{33} \right]^2 - \left[\sigma_{11} - \sigma_{33} \right] \left[\sigma_{22} - \sigma_{33} \right] \right\} \left\{ 4J(0) + 3J\left(\omega_N \right) \right\}$$

Equation 3.8: Transverse relaxation due to ^{15}N CSA

and the CSA / dipole-dipole cross correlated relaxation rate for a rhombic CSA tensor

$$R_2^{CC} = 2cd \left\{ \left[\sigma_{11} - \sigma_{33} \right] \frac{3\cos^2 \theta_{11} - 1}{2} + \left[\sigma_{22} - \sigma_{33} \right] \frac{3\cos^2 \theta_{22} - 1}{2} \right\} \left\{ 4J(0) + 3J\left(\omega_I \right) \right\}$$

Equation 3.9: CSA/dipole-dipole cross-correlated relaxation rate

In the above equations the following definitions were used:

$$d^2 = \frac{1}{8} \left(\frac{\mu_0}{4\pi} \frac{\gamma_H \gamma_N \hbar}{r_{NH}^3} \right)^2$$

$$c^2 = \frac{1}{18} \left(\omega_N \right)^2$$

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

Equation 3.10: Definitions

and θ indicates the angle between the ^{15}N CSA principal axes and the NH vector. Other symbols have their usual meanings [207].

As is shown in Figure 3.6, the ^{15}N TROSY R_2 rate for a protein with τ_c=22 ns (the N^+Ha transition) at 800 MHz is 5.7 s^{-1} ($\Delta\nu 1/2$=1.8 Hz), for the anti-TROSY line (the N^+Hb transition) the R_2 is 60.8 s^{-1} (19.4 Hz). The HSQC line has an R_2 of 33.25 s^{-1} ($\Delta\nu 1/2$=10.6 Hz). Similar equations were used for the calculation of the TROSY effect on the amide proton shown in Figure 3.6. The ^1H-TROSY R_2 rate for the same protein (the H^+Nb transition) at 800 MHz is 9 s^{-1}, the anti-TROSY R_2 rate is 47 s^{-1} (the H^+Na transition) and the HSQC R_2 rate is 28 s^{-1}.

3.6.2 PULSE SEQUENCE OVERVIEW

Residual dipolar couplings (RDCs) were extracted from a 2 dimensional

experiment that is a ^{15}N-^{1}H TROSY with a [$\kappa t_1/2$-180(N,H)-κ $t_1/2$] sequence at the beginning of the ^{15}N chemical shift labeling period (Fig.3.12). This sequence (hereafter referred to as 'RDC-TROSY') mixes the TROSY and anti TROSY components.

With reference to the HN spin system energy level diagram shown in Fig. 3.12, TROSY involves selecting the slowly relaxing $N^{\pm}_{H\alpha}$ transition and transferring it to the slowly relaxing $H^{\pm}_{N\beta}$ transition (Fig.3.12/13).

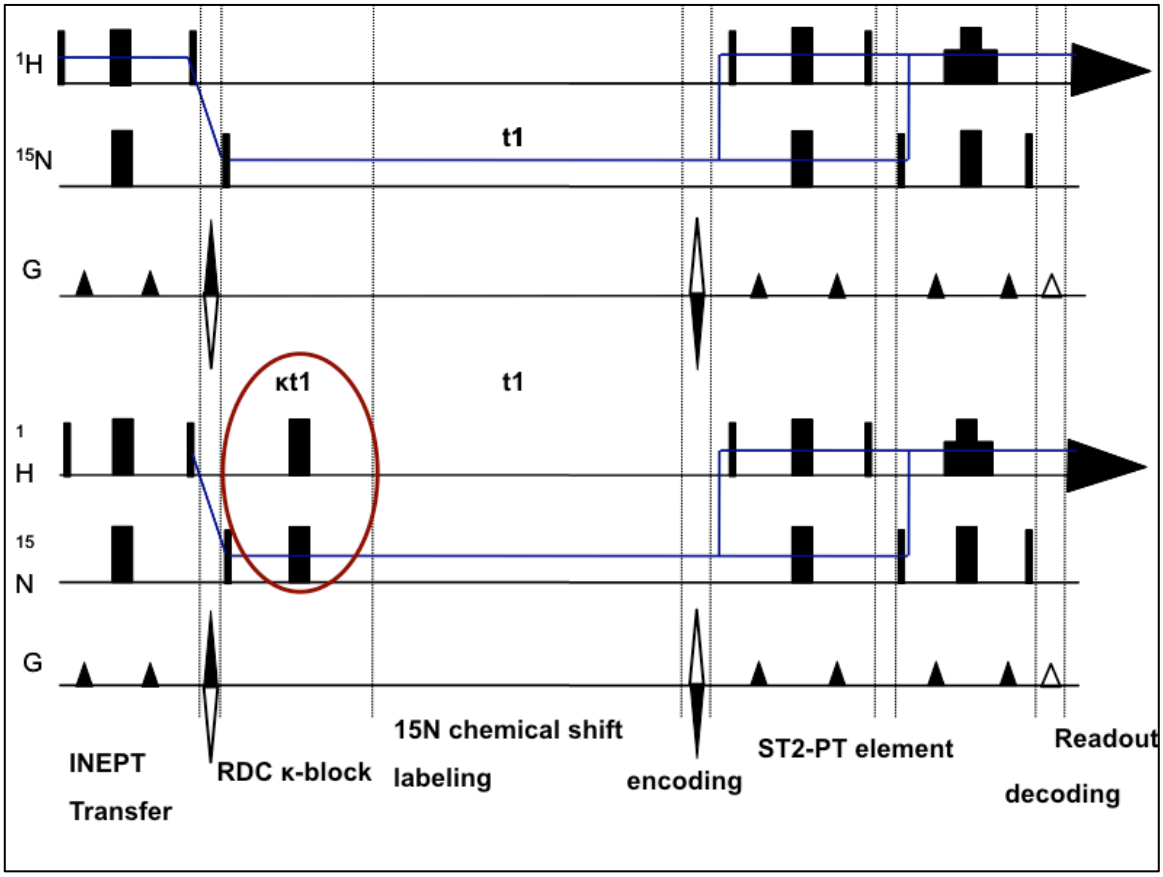

Figure 3.12: the RDC-TROSY (kappa block) pulse sequence

Top: TROSY pulse sequence implementation following [208]. Bottom: RDC-TROSY pulse sequence to extract residual dipolar couplings

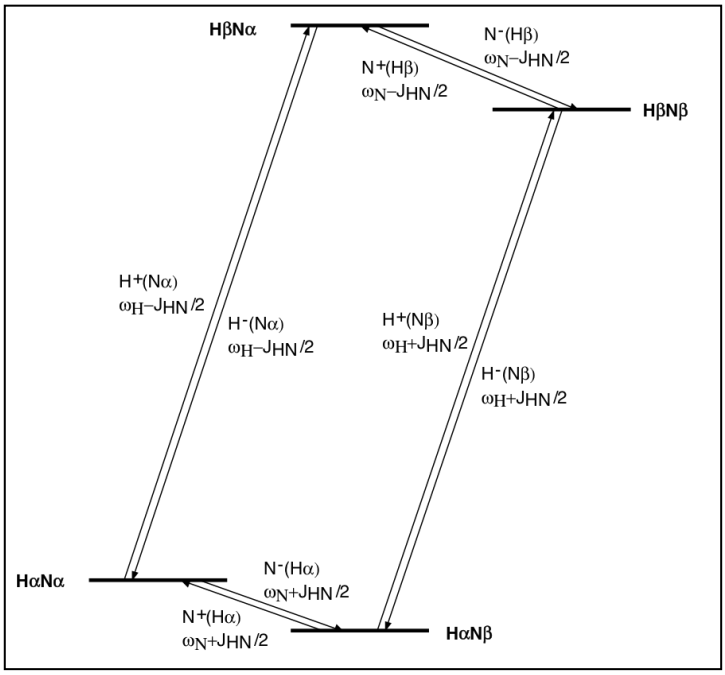

Figure 3.13: Energy level diagram for an NH two-spin system

3.6.3 MATHEMATICAL ANALYSIS OF THE RDC-TROSY (κ-BLOCK) PULSE SEQUENCE USING THE STATES/TRANSITIONS FORMALISM

The κ-block incorporates a $\kappa t_1/2$-πHN-$\kappa t_1/2$ period before the actual t_1 period. Using the common the Cartesian operators for the initial INEPT block, one obtains:

$$H_Z \xrightarrow{pulse\ H\ 90(x)} -H_Y$$
$$-H_Y \xrightarrow{scalar} 2H_X N_Z$$
$$2H_X N_Z \xrightarrow{pulse\ H\ 90(-y)} 2H_Z N_Z$$
$$2H_Z N_Z \xrightarrow{pulse\ N\ 90(x)} -2H_Z N_Y$$

Equation 3.11: The initial INEPT transfer

Switching to transition operators (also known as ladder or shift operators), we use the relation

$$N_{Y=}\frac{-i}{2}\left(N^{+}-N^{-}\right)$$

to write the output of the INEPT block as

$$-2H_Z N_Y = -2H_Z\left(\frac{-i}{2}\left(N^+ - N^-\right)\right) = iH_Z\left(N^+ - N^-\right)$$

Equation 3.12: Rewriting the output of the initial INEPT block

Now, recognizing that the longitudinal magnetization term HZ can be written as: $H_Z = H_\alpha - H_\beta$, the Cartesian term $-2H_Z N_Y$ can now be expanded to

$$-2H_Z N_Y = i\left(H_\alpha - H_\beta\right)\left(N^+ - N^-\right)$$

Equation 3.13: Expanding from Cartesian to Ladder operators

which can be rewritten in terms of ^{15}N coherences as

$$-2H_Z N_Y = i(N^+_{H\alpha} - N^-_{H\alpha} - N^+_{H\beta} + N^-_{H\beta})$$

Equation 3.14: Rewriting the INEPT block output in terms of ^{15}N coherences

Therefore, realizing that the 'i' multiplier is merely a globally adjustable phase factor, the terms contributing to the density operator (σ) after the initial INEPT block, terminated by π/2 pulses on H and N are:

$$\sigma(0) = N^+_{H\alpha} - N^-_{H\alpha} - N^+_{H\beta} + N^-_{H\beta}$$

Equation 3.15: Density matrix elements at the end of the first INEPT block

Note that time (t) has been set to zero at the end of the INEPT. This is for mathematical convenience, as well as the fact that the INEPT applies the same relaxation decay to all coherence terms. Now comes the κ-block, or the RDC-measurement block. This has the effect of interconverting from the TROSY to the antiTROSY state (see Fig 3.13/14)

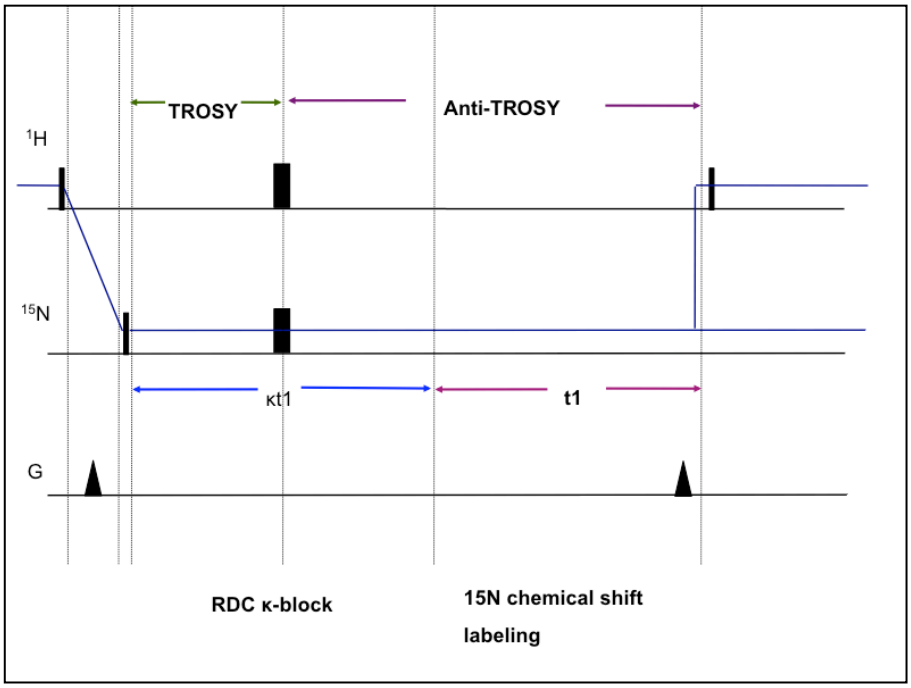

Figure 3.14: Detailed view of the RDC-kappa block

To describe the effect of the κ-block, we will use the single-transition-operator formalism, for which the following transformation rules hold:

$$N^{+}_{H\alpha} \xrightarrow{shift} N^{+}_{H\alpha}\exp(-i\Omega_N t)$$
$$N^{+}_{H\alpha} \xrightarrow{scalar} N^{+}_{H\alpha}\exp(-i\pi J_{HN} t)$$
$$N^{+}_{H\alpha} \xrightarrow{pulse\ N180(x)} N^{-}_{H\alpha}$$
$$N^{+}_{H\alpha} \xrightarrow{pulse\ N180(y)} -N^{-}_{H\alpha}$$
$$N^{+}_{H\alpha} \xrightarrow{pulse\ H180} N^{+}_{H\beta}$$

$$N^{-}_{H\alpha} \xrightarrow{shift} N^{-}_{H\alpha}\exp(+i\Omega_N t)$$
$$N^{-}_{H\alpha} \xrightarrow{scalar} N^{-}_{H\alpha}\exp(i\pi J_{HN} t)$$
$$N^{-}_{H\alpha} \xrightarrow{pulse\ N180(x)} N^{+}_{H\alpha}$$
$$N^{-}_{H\alpha} \xrightarrow{pulse\ N180(y)} -N^{+}_{H\alpha}$$
$$N^{-}_{H\alpha} \xrightarrow{pulse\ H180} N^{-}_{H\beta}$$

$$N^{+}_{H\beta} \xrightarrow{shift} N^{+}_{H\beta}\exp(-i\Omega_N t)$$
$$N^{+}_{H\beta} \xrightarrow{scalar} N^{+}_{H\beta}\exp(i\pi J_{HN} t)$$
$$N^{+}_{H\beta} \xrightarrow{pulse\ N180(x)} N^{-}_{H\beta}$$
$$N^{+}_{H\beta} \xrightarrow{pulse\ N180(y)} -N^{-}_{H\beta}$$
$$N^{+}_{H\beta} \xrightarrow{pulse\ H180} N^{+}_{H\alpha}$$

$$N^{-}_{H\beta} \xrightarrow{shift} N^{-}_{H\beta}\exp(+i\Omega_N t)$$
$$N^{-}_{H\beta} \xrightarrow{scalar} N^{-}_{H\beta}\exp(-i\pi J_{HN} t)$$
$$N^{-}_{H\beta} \xrightarrow{pulse\ N180(x)} N^{+}_{H\beta}$$
$$N^{-}_{H\beta} \xrightarrow{pulse\ N180(y)} -N^{+}_{H\beta}$$
$$N^{-}_{H\beta} \xrightarrow{pulse\ H180} N^{-}_{H\alpha}$$

Equation 3.16: Transition rules for single transition operators

After evolution for a period $\kappa t_1/2$, the different ^{15}N coherences contributing to the density operator are:

$$\sigma\left(\frac{\kappa t_1}{2}\right)=$$

$$N_{H\alpha}^{+}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\alpha}^{-}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\beta}^{+}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$+N_{H\beta}^{-}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

Equation 3.17: Density matrix elements halfway through the RDC-kappa block

Here, $R_{2N}^{H\alpha}$ and $R_{2N}^{H\beta}$ are the relaxation rates of coherences $N_{H\alpha}^{\pm}$ and $N_{H\beta}^{\pm}$; J_{NH} is the scalar coupling and D_{NH} is the residual dipolar coupling between spins N and H. Ω_N is the chemical shift of ^{15}N.

Application of the π_H pulse results in

$$\sigma\left(\frac{\kappa t_1}{2},\pi_H\right)=$$

$$N_{H\beta}^{+}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\beta}^{-}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\alpha}^{+}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$+N_{H\alpha}^{-}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

Equation 3.18: Interconversion of density matrix elements due to the 180(H) pulse

Application of the π_N pulse along the x-axis results in

$$\sigma\left(\frac{\kappa t_1}{2},\pi_H,\pi_N^X\right)=$$

$$N_{H\beta}^{-}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\beta}^{+}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$-N_{H\alpha}^{-}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

$$+N_{H\alpha}^{+}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}$$

Equation 3.19: Interconversion of density matrix elements due to the 180(N) pulse along x

These coherences now evolve for a time $\kappa t_1/2+t_1$, leading to:

$$\sigma\left(\frac{\kappa t_1}{2},\pi_H,\pi_N^X,\left(1+\frac{\kappa}{2}\right)t_1\right)=$$

$$N_{H\beta}^{-}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}\times e^{\left(+i\pi\Omega_N(1+\frac{\kappa}{2})t_1\right)}e^{\left(-i\pi(J+D)_{NH}(1+\frac{\kappa}{2})t_1\right)}e^{\left(-R_{2N}^{H\beta}(1+\frac{\kappa}{2})t_1\right)}$$

$$-N_{H\beta}^{+}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}\times e^{\left(-i\pi\Omega_N(1+\frac{\kappa}{2})t_1\right)}e^{\left(+i\pi(J+D)_{NH}(1+\frac{\kappa}{2})t_1\right)}e^{\left(-R_{2N}^{H\beta}(1+\frac{\kappa}{2})t_1\right)}$$

$$-N_{H\alpha}^{-}e^{\left(-i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}\times e^{\left(+i\pi\Omega_N(1+\frac{\kappa}{2})t_1\right)}e^{\left(+i\pi(J+D)_{NH}(1+\frac{\kappa}{2})t_1\right)}e^{\left(-R_{2N}^{H\alpha}(1+\frac{\kappa}{2})t_1\right)}$$

$$+N_{H\alpha}^{+}e^{\left(+i\pi\Omega_N\frac{\kappa t_1}{2}\right)}e^{\left(-i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)\right)}\times e^{\left(-i\pi\Omega_N(1+\frac{\kappa}{2})t_1\right)}e^{\left(-i\pi(J+D)_{NH}(1+\frac{\kappa}{2})t_1\right)}e^{\left(-R_{2N}^{H\alpha}(1+\frac{\kappa}{2})t_1\right)}$$

Equation 3.20: Density matrix elements at the end of the ^{15}N chemical shift-labeling period (t_1)

Combining terms, one obtains

$$\sigma\left(\frac{\kappa t_1}{2},\pi_H,\pi_N^X,\left(1+\frac{\kappa}{2}\right)t_1\right)=$$

$$N_{H\beta}^{-}e^{\left(+i\pi\Omega_N t_1\right)}e^{\left(-i\pi(J+D)_{NH}(1+\kappa)t_1\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)-R_{2N}^{H\beta}(1+\frac{\kappa}{2})t_1\right)}$$

$$-N_{H\beta}^{+}e^{\left(+i\pi(J+D)_{NH}\frac{\kappa t_1}{2}\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)\right)}e^{\left(-R_{2N}^{H\alpha}\left(\frac{\kappa t_1}{2}\right)-R_{2N}^{H\beta}(1+\frac{\kappa}{2})t_1\right)}$$

$$-N_{H\alpha}^{-}e^{\left(+i\pi\Omega_N t_1\right)}e^{\left(+i\pi(J+D)_{NH}(1+\kappa)t_1\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)-R_{2N}^{H\alpha}(1+\frac{\kappa}{2})t_1\right)}$$

$$+N_{H\alpha}^{+}e^{\left(-i\pi\Omega_N t_1\right)}e^{\left(-i\pi(J+D)_{NH}(1+\kappa)t_1\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)-R_{2N}^{H\alpha}(1+\frac{\kappa}{2})t_1\right)}$$

Equation 3.21: Density matrix elements after simplification

The TROSY element now extracts the $N_{H\alpha}^{\pm}$ coherences and transfers those to $H_{N\beta}^{\pm}$ coherences in the detection period. The time-domain signal in t_1 is:

$$S^{\pm}(t_1)\propto e^{\left(\mp i\pi\Omega_N t_1\right)}e^{\left(\mp i\pi(J+D)_{NH}(1+\kappa)t_1\right)}e^{\left(-R_{2N}^{H\beta}\left(\frac{\kappa t_1}{2}\right)-R_{2N}^{H\alpha}(1+\frac{\kappa}{2})t_1\right)}$$

Equation 3.22: Density matrix element, which is extracted by the ST2-PT element and read out

For κ=0, a standard TROSY spectrum is recovered:

$$S^{\pm}(t_1)\propto e^{\left(\mp i\pi\Omega_N t_1\right)}e^{\left(\mp i\pi(J+D)_{NH}t_1\right)}e^{\left(-R_{2N}^{H\alpha}t_1\right)}$$

Equation 3.23: Reduction to a standard TROSY signal

For κ=2, one obtains

$$S^{\pm}(t_1)\propto e^{\left(\mp i\pi\Omega_N t_1\right)}e^{\left(\mp i\pi(J+D)_{NH}3t_1\right)}e^{\left(-R_{2N}^{H\beta}t_1-R_{2N}^{H\alpha}2t_1\right)}$$

Equation 3.24: Kappa shifted TROSY signal

A frequency shift of $2\pi(J+D)_{NH}$ is observed between the κ=0 and the κ=2 spectra. Varying κ allows this shift to be chosen at will. Thus, by collecting spectra for different values of κ and eliminating the scalar coupling contribution to

the shift, the residual dipolar coupling (D_{NH}) can be recovered. It is also obvious that undesirable relaxation terms are incorporated as κ increases. Thus a tradeoff between measurable RDCs and resolvable peaks exists. This will be discussed in detail in the Results and Discussion section.

3.6.4 HOW THE κ-BLOCK RDC-TROSY EXPERIMENT SHIFTS CROSSPEAKS

From Equations 3.22 and 3.23, it is seen that at κ=0, the standard TROSY spectrum is recovered. But at κ=2, while the peak is shifted by an amount equal to $2\pi(J+D)$, it is shifted away from the centre in the same direction as the TROSY peak. This means that the crosspeak actually moves further away from the centre peak(location shown in Fig 3.15 as the blue line) and in the opposite direction as the antiTROSY peak(location shown in Fig 3.15 as the purple line). Along with this shift, which is the basis of the RDC measurement, the peak also picks up the undesirable broad shape of the anti-TROSY peak. This is illustrated below.

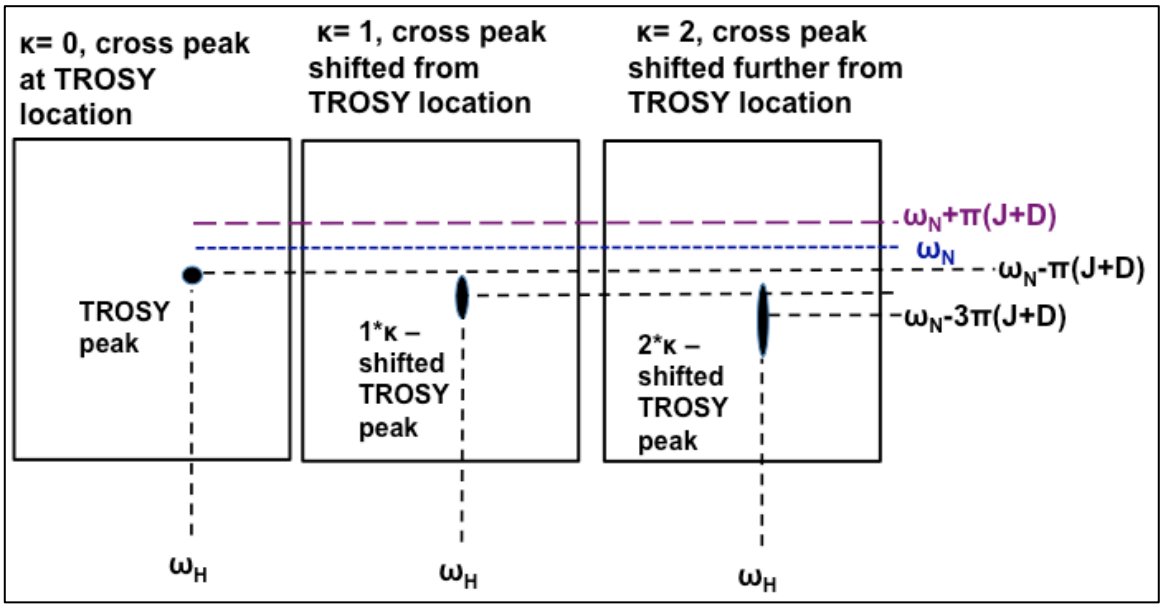

Figure 3.15: Shifting of crosspeaks in a RDC-TROSY experiment

3.6.5 EFFECT OF AMIDE PROTON SPIN FLIPS

A key tenet of the description of the κ-block in the RDC-TROSY experiment above is the assumption that the NHa and NHb coherences do not mix, except by the effects of ^{1}H r.f. pulses. However, amide protons also flip spontaneously between states by relaxation mechanisms, dominated by zero-quantum flip-flops (NOEs) due to dipolar interactions with other protons in the vicinity. These rates can become quite fast [204] (up to 30 s^{-1} for larger and protonated proteins) and cause life-time broadening on the TROSY and anti-TROSY lines. The mechanism is referred to as "scalar relaxation of the second kind" [209], and is identical to chemical exchange broadening in slow-exchange regime [210]. The ^{15}N evolution of the two doublet components in presence of proton spin flips of rate $k_{\alpha\beta}\left(=k_{\beta\alpha}\right)$ is described by the following Bloch-McConnell equations (here expressed in Cartesian single spin operators for computer coding ease, and omitting chemical shifts)

$$\frac{d}{dt}\begin{bmatrix} N_{H\alpha}^{X} \\ N_{H\alpha}^{Y} \\ N_{H\beta}^{X} \\ N_{H\beta}^{Y} \end{bmatrix} = -\begin{bmatrix} R_{2N}^{H\alpha}+k_{\alpha\beta} & -\pi J_{HN} & -k_{\beta\alpha} & 0 \\ \pi J_{HN} & R_{2N}^{H\alpha}+k_{\alpha\beta} & 0 & -k_{\beta\alpha} \\ -k_{AB} & 0 & R_{2N}^{H\beta}+k_{\beta\alpha} & -\pi J_{HN} \\ 0 & -k_{\alpha\beta} & \pi J_{HN} & R_{2N}^{H\beta}+k_{\beta\alpha} \end{bmatrix}\begin{bmatrix} N_{H\alpha}^{X} \\ N_{H\alpha}^{Y} \\ N_{H\beta}^{X} \\ N_{H\beta}^{Y} \end{bmatrix}$$

Equation 3.25: Evolution of ^{15}N doublet components due to amide proton spin flips

where $N_{H\alpha}^{X,Y}$ represent the TROSY components and $N_{H\beta}^{X,Y}$ the anti-TROSY components, with the associated auto relaxation rates $R_{2N}^{H\alpha}$ and $R_{2N}^{H\beta}$.

We calculate the evaluation of these coherence components from these equations using numerical integration, for all ^{15}N dwell times in the RDC-TROSY experiments for several k values. The effect of the ^{15}N, 180 degree (x-phase) pulse between the two κ $t_1/2$ periods is computed as

$$\begin{bmatrix} N_{H\alpha}^{X} \\ N_{H\alpha}^{Y} \\ N_{H\beta}^{X} \\ N_{H\beta}^{Y} \end{bmatrix}_{\text{Post-180-15N}} = \begin{bmatrix} N_{H\alpha}^{X} \\ -N_{H\alpha}^{Y} \\ N_{H\beta}^{X} \\ -N_{H\beta}^{Y} \end{bmatrix}_{\text{Pre-180-15N}}$$

Equation 3.26: Evaulating the effect of the (^{15}N) 180 degree pulse halfway through the RDC-kappa block

while the effect of the ^{1}H 180 degree pulse is obtained by setting

$$R_2^A = R_2^B, \quad R_2^B = R_2^A, \quad J_{AB} = -J_{AB}$$

Equation 3.27: Evaluating the effect of the (^{1}H) 180 degree pulse halfway through the RDC-kappa block

The time domain signals $N_{H\alpha}^{X}(t_1), N_{H\alpha}^{Y}(t_1)$ and $N_{H\beta}^{X}(t_1), N_{H\beta}^{Y}(t_1)$ were transformed to the frequency domain using a complex Fourier transform.

3.7 MATERIALS AND METHODS

3.7.1 PROTOCOLS FOR DnaK Tth-NBD .

The plasmid for wild type DnaK Tth-NBD (residues 1- 382) was expressed in *E.coli* BL21 cells at 37 °C in M9 media with 99%$^{2}H_20$, 98% $^{15}NH_4Cl$, and 99% ^{13}C,^{1}H glucose. All isotopes were purchased from Cambridge Isotope Laboratory (Andover, MA). The His-tagged protein was purified as described in Chapter 2 and also in the references [172; 211]. Guanidine-HCl and heating was required to denature DnaK Tth-NBD to allow re-protonation of all amide groups in the perdeuterated protein and complete removal of the bound nucleotide. Samples for NMR experiments were made at a protein concentration of 0.34 mM in a buffer containing 50mM HEPES, 10mM KCl, 5mM $MgCl_2$, and 5mM sodium phosphate. Nucleotide concentration (ADP or AMPPNP) was 10 mM.

DnaK Tth-NBD behaves well in a solution of 10 mg/ml Pf1 phage at 50o C, allowing the measurement of ^{15}N-^{1}H RDCs using RDC-TROSY. The RDC-TROSY experiments were recorded at 50o C on an 800 MHz Varian Inova spectrometer, using

a triple resonance cold-probe, for κ=0, κ = ¾ and κ =3/2. 10, 20 and 40 hours of data acquisition was used for the κ =0, 0.75 and 1.5 experiments, respectively.

3.7.2 PROTOCOLS FOR DnaK Tth-NBD-SBD .

The truncated 1-501 form of DnaK Tth was constructed by PCR from a plasmid containing the full length DnaK Tth gene supplied by Dr. A. Joachimiak at Argonne National Laboratory. The resulting plasmid, pMR23, had the 1-500 sequence of wild type DnaK Tth with the L501E mutation and a 6-residue histidine tail. Introduction of the ΔT422, A423E mutations was accomplished using mutagenic primers and the Quickchange mutagenesis Kit (Stratagene). *E.coli* strain BL21 (DE3) cells containing the pMR23 plasmid were grown at 37 °C in M9 media with $^{15}NH_4Cl$ or with 99%$^{2}H20$, 98% $^{15}NH_4Cl$, and 99% ^{13}C,^{1}H glucose. All isotopes were purchased from Cambridge Isotope Laboratory (Andover, MA). The His-tagged protein was purified as described previously. Guanidine-HCl and heating was required to denature DnaK Tth-NBD-SBD to allow re-protonation of all amide groups in the perdeuterated protein and complete removal of the bound nucleotide. TROSY spectra of the protein confirmed that it was properly re-folded from these conditions. NMR spectra were collected at 55 °C using an 800 MHz Varian Inova spectrometer, using a triple-resonance Z-gradient probe. Samples were approximately 400 mM DnaK Tth-NBD-SBD in 50 mM HEPES pH 7.4, 10 mM KCl, 5 mM $MgCl_2$, 5 mM ADP and 5mM Na_2PO_4. At these conditions, the protein was monomeric as derived from the rotational correlation time of approximately ^{15} ns determined from 1D ^{15}N NMR R_1 and R_2 relaxation data.

DnaK Tth-NBD-SBD ADP.Pi /*apo* behaves well in a solution of 20 mg/ml Pf1 phage at 50o C, allowing the measurement of ^{15}N-^{1}H RDCs using RDC-TROSY. RDC-TROSY experiments were recorded at 50^{o} Celsius on an 800 MHz Varian Inova spectrometer, using a conventional triple resonance probe, for κ =0, κ = 1 and κ =2.

3.7.3 BACKBONE RESONANCE ASSIGNMENTS

A peak list (Revington and Zuiderweg, unpublished) prepared for DnaKTth-NBD in the ADP-AlF$_x$ state, based on earlier assignment work [172] was used as a template for assignment of the peptide backbone of DnaKTth-NBD in the different nucleotide states. The raw data was processed and transformed using NmrPipe [212]. The assignment 'walk' through the peptide backbone was done using Sparky [213]. A total of 223 assignments for the APO form, 281 for the ADP form and 310 for the AMPPNP form of the protein were obtained.

The backbone resonance assignments of the 501-residue DnaKTth-NBD-SBD in the presence of ADP.Pi and in the absence of peptide substrate at pH 7.4 and 55 ^{o}C were extended from earlier work on the 381-residue DnaKTth NBD. A total of 415 of the 475 non-proline residues in the native sequence (87%) could be unambiguously assigned.

3.8 RESULTS AND DISCUSSION

3.8.1 SPECTRA FROM RDC-TROSY κ-BLOCK EXPERIMENTS

As is shown in Figure 3.16/17, the RDC-TROSY experiments yield good sensitivity for the extraction of RDCs. Residual CSAs [214] are identical for all values of κ, since the ^{15}N 180 degree pulse in the κ-block refocuses them. Hence residual CSAs do not appear in the differences between the k experiments. The detailed view (seen in Fig 3.17) shows that relative shifts between different κ-block experiment spectra can be positive, negative, or zero. This is consistent with the possible RDC amplitudes and indicates that there is no detectable systematic error in measurement.

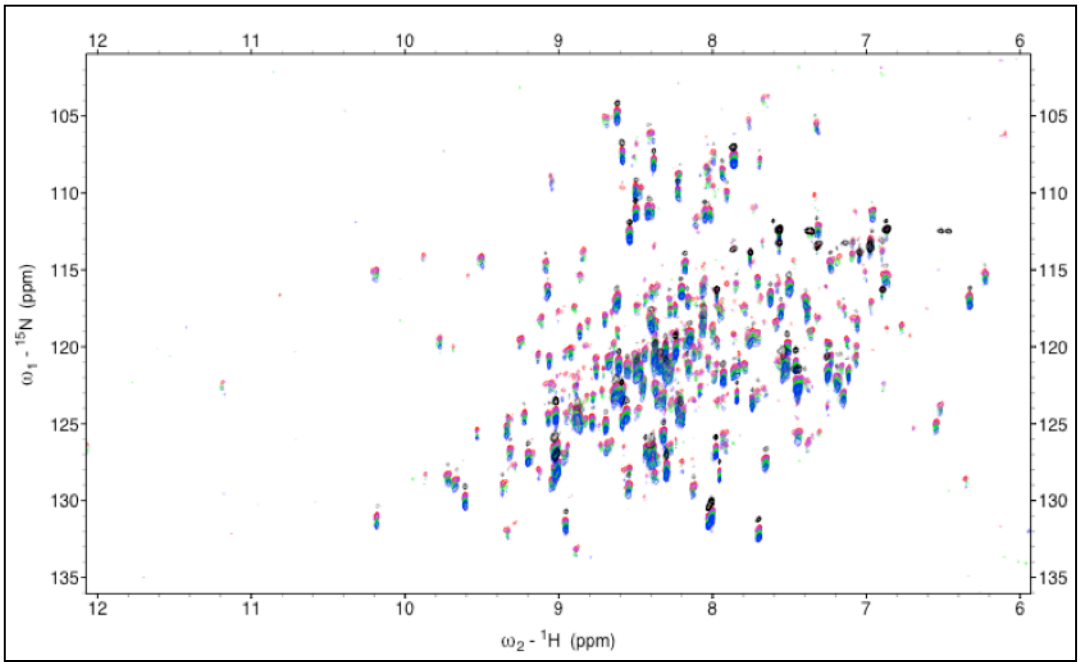

Figure 3.16: RDC-Kappa block spectra for Tth-NBD

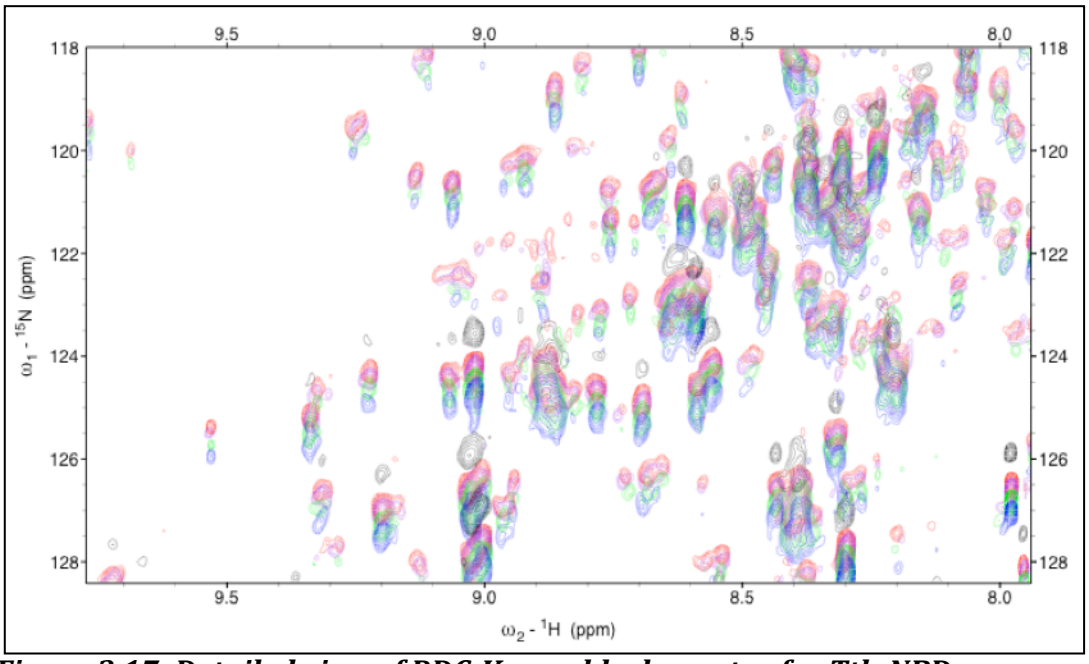

Figure 3.17: Detailed view of RDC-Kappa block spectra for Tth-NBD

Detail of RDC-TROSYs of DnaKTth-NBD. Red: k=0 , Purple: k=0.2, green: k=0.6 Blue: k=1. In black is the HSQC spectrum.

3.8.2 SIMULATIONS: OPTIMUM VALUES OF κ

The effective ^{15}N relaxation rate of the detected TROSY line is given as

$$R_2^{EFF} = \frac{\kappa}{2} R_{2N}^{H\beta} + \frac{2+\kappa}{2} R_{2N}^{H\alpha}$$

Equation 3.28: Effective ^{15}N transverse relaxation rate

Where $R_{2N}^{H\alpha}$ is the ^{15}N TROSY (NHa) linewidth, and $R_{2N}^{H\beta}$ is the ^{15}N Anti-TROSY (NHb) linewidth. R_2^{EFF} varies between $R_{2N}^{H\alpha}$ and $R_{2N}^{H\beta} + 2R_{2N}^{H\alpha}$ in the range $0 \le \kappa \le 2$.

Table 1 gives the computed R_2^{eff} values (based on Eqs (3.6 – 3.9)) for different k values and for HSQC, for different rotational correlation times, at 600 and 800 MHz ^{1}H. It shows that the ^{15}N linewidth for the κ=0.8 experiment (shown in bold) is smaller than that for the HSQC experiment (last row) for all correlation times at all fields.

Table 1. ^{15}N Ha effective R_2 rates					
Kappa	τ_c=5 ns 600 MHz	τ_c =5 ns 800 MHz	τ_c =10 ns 800 MHz	τ_c =20 ns 800 MHz	τ_c =30 ns 800 MHz
0.00	1.88	1.44	2.66	5.20	7.77
0.20	3.34	3.09	5.75	11.25	16.82
0.40	4.80	4.75	8.83	17.30	25.86
0.60	6.25	6.40	11.92	23.35	34.91
0.80	**7.71**	**8.06**	**15.00**	**29.40**	**43.96**
1.00	9.17	9.71	18.09	35.45	53.01
1.20	10.63	11.36	21.18	41.50	62.05
1.40	12.09	13.02	24.26	47.55	71.10
1.60	13.54	14.67	27.35	53.60	80.15
1.80	15.00	16.33	30.43	59.65	89.19
2.00	16.46	17.98	33.52	65.70	98.24
Anti-Trosy	12.70	15.10	28.20	55.30	82.70
HSQC	7.29	8.27	15.43	30.25	45.24

Table 3.1: ^{15}N H-alpha effective linewidths showing the limitations of RDC-TROSY vs HSQC experiments

Relaxation rates calculated for the "^{15}N TROSY" multiplet components in the RDC-TROSY experiment using Equations 3.6-3.9.

Table 2 shows that the relative ^{15}N intensity for the κ=1.2 experiment (bold) is larger than that for the HSQC (last row) experiment for all correlation times using an 800 MHz system.

Table 2. ^{15}N Ha line relative peak height					
Kappa	τ_c =5 ns 600 MHz	τ_c =5 ns 800 MHz	τ_c =10 ns 800 MHz	τ_c =20 ns 800 MHz	τ_c =30 ns 800 MHz
0.00	100.00	100.00	100.00	100.00	100.00
0.20	56.32	46.54	46.29	46.22	46.20
0.40	39.20	30.33	30.12	30.06	30.04
0.60	30.06	22.49	22.32	22.27	22.26
0.80	24.38	17.87	17.73	17.69	17.68
1.00	20.50	14.83	14.70	14.67	14.66
1.20	**17.69**	**12.67**	**12.56**	**12.53**	**12.52**
1.40	15.56	11.06	10.96	10.94	10.93
1.60	13.88	9.81	9.73	9.70	9.69
1.80	12.53	8.82	8.74	8.72	8.71
2.00	11.42	8.01	7.94	7.91	7.91
Anti-Trosy	5.49	2.42	1.95	1.80	1.78
HSQC	27.92	14.08	11.80	11.07	10.96

Table 3.2: ^{15}N H-alpha effective intensities (peak heights) for RDC-TROSY vs. HSQC experiments

Relative peak heights, setting the κ=0 line to 100 a.u. for each correlation time. The Peak intensity of the Anti-TROSY line is reduced by the ratio ^{1}H ANTI-TROSY / ^{1}H TROSY intensity; the peak intensity of the HSQC line is reduced by the ratio 2* ^{1}H HSQC / ^{1}H TROSY intensity (see Figure 3.6).

The change in linewidths/intensity is further affected by the ^{1}H TROSY effect as well. Here the ^{1}H-^{15}N dipole-dipole coupling interacts with the ^{1}H CSA relaxation to produce a TROSY effect that results in maximal narrowing of the ^{1}H linewidth at around 1 GHz external magnetic field (see Fig. 3.18)

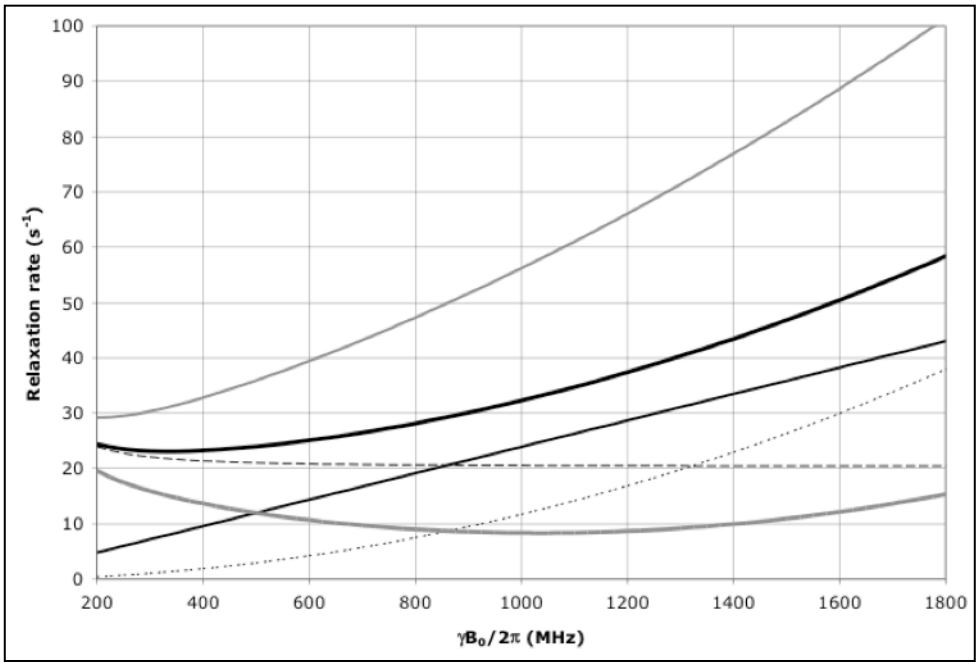

Figure 3.18: ^{1}H TROSY effect simulated for a protein tumbling with a correlation time of ~ 22 nsec

The field dependence of the protein ^{1}H TROSY effect for τ_c=22 ns calculated using a rhombic ^{1}H CSA tensor with σ_{11}=3, σ_{22}=8 and σ_{33}=17 ppm, with the direction of σ_{11} aligned with the NH-bond of 1.04 Å, using Eqs. (3.6-3.9). Thick solid line: total ^{1}H R_2 as encountered in a HSQC or HMQC; Thick dashed line: ^{1}H R_2 due to ^{15}N-^{1}H dipolar relaxation; Thin dashed line: ^{1}H R_2 due to ^{1}H CSA relaxation; Thin grey line: the ^{1}H R_2 ^{15}N-^{1}H dipolar / ^{1}H CSA cross correlated relaxation rate; Thick grey line: ^{1}H R_2 as encountered in a TROSY (i.e. total ^{1}H R_2 minus ^{1}H R_2 ^{15}N-^{1}H dipolar / ^{1}H CSA cross correlation). Thin grey line: ^{1}H R_2 of the anti-TROSY line (i.e. total ^{1}H R_2 plus ^{1}H R_2 ^{15}N-^{1}H dipolar / ^{1}H CSA cross correlation).

Explicit simulation of the RDC-TROSY experiments yields spectra as shown in Figure (3.19). They convey the same message as shown in Tables 1 and 2: RDC-TROSY experimentation is superior to the TROSY- HSQC comparison up to k values of 1.0.

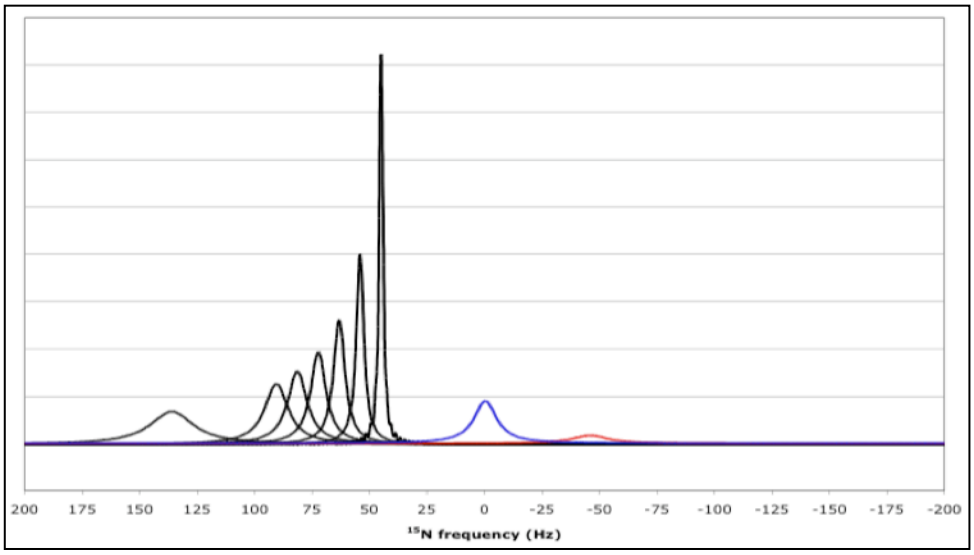

Figure 3.19: Simulation of RDC-TROSY vs. HSQC peakshapes

^{15}N resonance positions, intensities and linewidths for the NH doublet components in various experiments, calculated for protein with τ_c =22 ns, at 800 MHz ^{1}H, no amide proton flips. In black, RDC-TROSY peaks, with k values of 0, 0.2, 0.4, 0.6, 0.8, 1.0 and 2.0 from right to left. In red, the anti-TROSY peak, scaled down by a factor of 5 due to additional broadening in the ^{1}H dimension (see Fig 3.18) in blue, the HSQC peak with intrinsic double intensity, but scaled down by a factor of 3 due to additional broadening in the ^{1}H dimension (see Fig 3.16).

3.8.3 THE EFFECT OF AMIDE PROTON FLIPS

However, the ^{15}N linewidth and peak intensity of cross peaks in the RDC-TROSY experimentation is affected by amide proton flip-flop rates, which occur predominantly because of zero-quantum flip-flops (large-molecule regime NOEs) with other protons in the vicinity [204]. These rates can become quite fast (up to 30 s^{-1} for larger and especially protonated proteins) and cause life-time broadening on the TROSY and anti-TROSY lines (see Theory section). The effect can easily double the TROSY line width, and hence, reduce the peak detection sensitivity by equal amounts. Perdeuteration of the protein does reduce the effect substantially, but the lifetime broadening can never be completely overcome for amide protons that are close to each other such as in alpha helices.

Figure 3.20/21 shows that for amide proton flip-flop rates less than 10 s^{-1} the RDC-TROSY experiment remains superior to the TROSY-HSQC comparison.

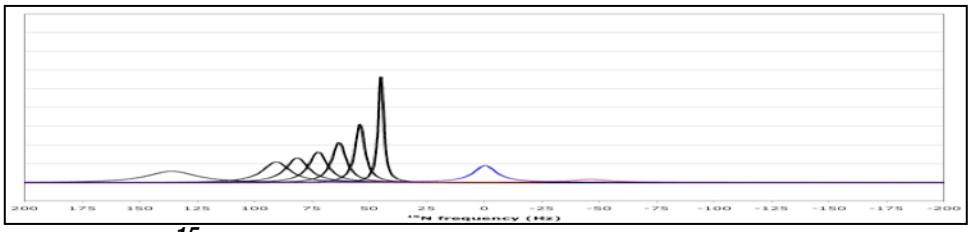

Figure 3.20: ^{15}N lineshapes for an amide proton flip rate of 3/sec

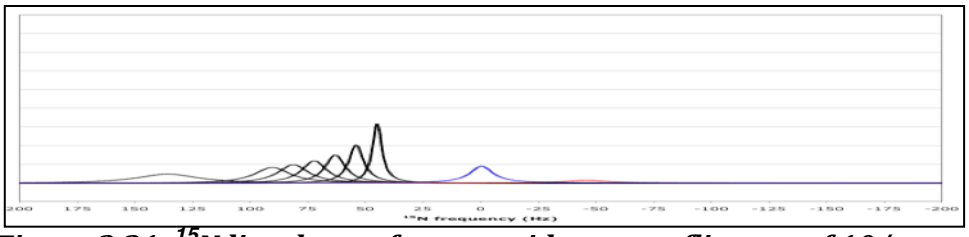

Figure 3.21: ^{15}N lineshapes for an amide proton flip rate of 10/sec

For rates between 30 and 100 s^{-1}, the TROSY-HSQC comparison becomes equal or superior.

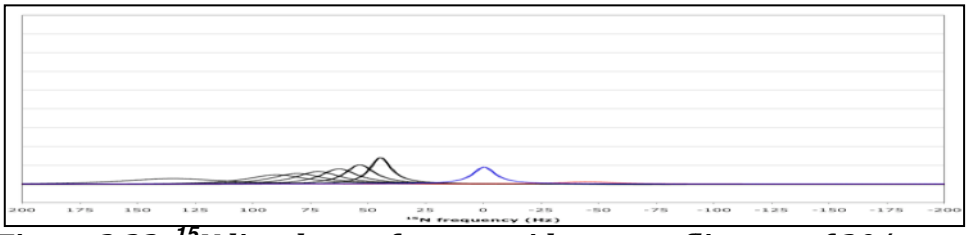

Figure 3.22: ^{15}N lineshapes for an amide proton flip rate of 30/sec

For amide proton flip-flop rates faster than 100 s^{-1}, the NHa-NHb doublet "auto-decouples" and collapses to the HSQC position. In this case ^{1}H-^{15}N RDCs cannot be measured anymore from the ^{15}N dimension.

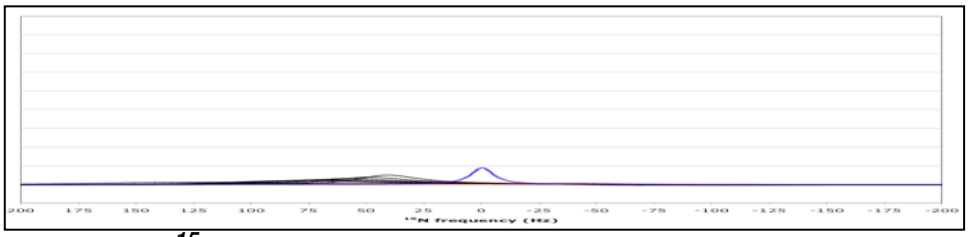

Figure 3.23: ^{15}N lineshapes for an amide proton flip rate of 100/sec

^{15}N resonance positions, intensities and linewidths for the NH doublet components in various RDC-TROSY experiments, calculated for protein with

τ_c=22 ns, at 800 MHz ^{1}H. Color coding and scaling is the same as in Fig 3.17. From 3.18 to 3.21: $k_{ab}=k_{ba}$=3 s^{-1} ; 10 s^{-1}; 30 s^{-1}; 100 s^{-1}.

In the kinetic regime just before the collapse occurs, intermediate exchange causes small shifts of the lines from their true positions [210] . Our simulations bear this out: whereas the slope of the line of the ^{15}N resonance position vs. κ is the expected 45.5 Hz / unit of κ for κ_{ab}=0, 3 and 10 s^{-1}, small changes (45.08 Hz / unit of κ for κ_{ab}=30 s^{-1}) and substantial changes (40.5 Hz / unit of κ for κ_{ab}=100 s^{-1}) occur (see Fig 3.24). But, the shifts are not dramatic enough to be a true cause for concern in any practical case for which RDCs can actually be measured.

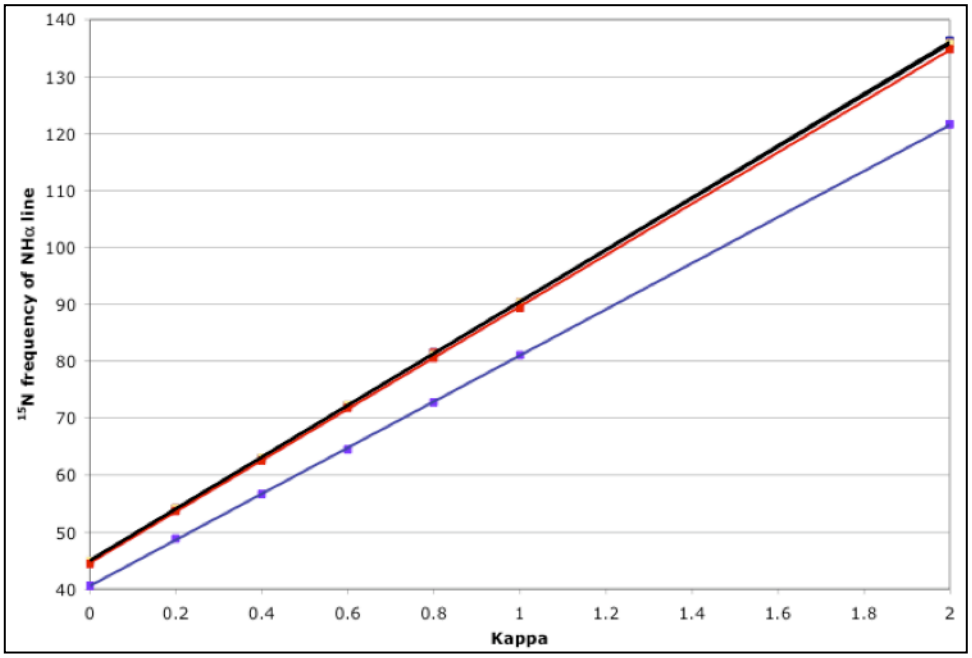

Figure 3.24: Shift of the TROSY line with increasing kappa in an RDC-TROSY experiment

^{15}N peak frequency for the NHa (TROSY) line, calculated for a protein with τ_c=22 ns, at 800 MHz ^{1}H for the RDC-TROSY experiments. The lines in the figure are least squares fits, with slopes of 45.63, 45.39, 45.37, (all black) 45.08 (red) and 40.5 (blue) Hz / unit of k for k_{ab}=0, 3, 10, 30 and 100 s^{-1}, respectively.

3.8.4 SIMULATIONS SUMMARY

Summarizing, the simulations bear out that the RDC--TROSY experimentation for

the extraction of RDCs is to be preferred over TROSY – HSQC comparisons for proteins with τ_c > 10 ns, provided that the amide proton flip-flop rates are not faster than 30 s^{-1}. Our simulations also point to the importance of protein perdeuteration in all TROSY experimentation: not only does it improve the amide ^1H line width, but it also improves the ^{15}N line width by suppressing much of the proton flip-flop processes.

3.9 APPLICATION OF RDC-TROSY TO SUBDOMAIN ORIENTATION OF THE NBD OF DnaK *Thermus thermophilus*

The method described in preceding sections was applied to the nucleotide-binding domain of *Thermus thermophilus* DnaK. Two nucleotide states of this protein were already discussed in Chapter 2. Here, we report results on another two nucleotide states and expand the analysis.

3.9.1 EXTRACTING RDCS FOR DnaK Tth-NBD AND DnaK Tth-NBD-SBD

Comparison of 3 datasets collected for three κ values showed excellent stability and reproducibility of spectra. Peak positions were reproducible within 4 Hz. Files corresponding to each κ value were co-added for higher S/Ns. Sequence-specific variations in $^1J_{NH}$ are expected to be less than 1 Hz as was determined for smaller proteins. Since this variation is smaller than the accuracy of the peak position determinations in the spectra for DnaK Tth-NBD, the RDC-TROSY spectra were superposed after shifting them by $\kappa/2 * {}^1J_{NH}$, where we assumed a uniform $^1J_{NH}$ of 90.5 Hz. What remains is the RDC. Sparky functions [213] were used to volume integrate and peak center each peak. Peak-lists were generated with the chemical shift in Hz and compiled in a spreadsheet. Differences in ^{15}N shift between each κ-value and κ=0 yield the RDC multiplied by a suitable scaling factor (0.375 or 0.75).

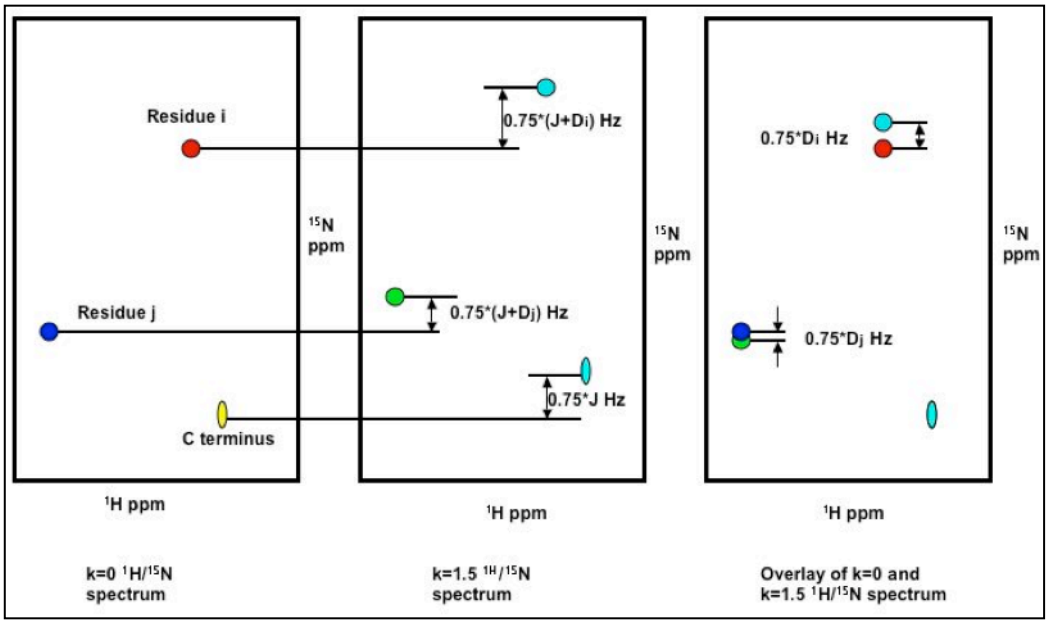

Figure 3.25: Shifting spectra to eliminate J-coupling and extract RDCs

Shifting the spectra by $\kappa/2*J_{NH}$ Hz from different RDC-TROSY spectra to eliminate the scalar-coupling component (J_{NH}) of the ^{15}N shift. What remains is the RDC (D_{NH}), scaled by a factor.

3.9.2 PROCEDURE USED TO DETERMINE A SUBSET OF EXPERIMENTAL RDCS OF MAXIMAL MUTUAL CONSISTENCY

1. Based on the way that the constants are defined in the RDC equation, the difference of ^{15}N shift between the κ=0.75 and the κ=0 block yields 0.375*RDC. 2. The difference between the κ=1.5 block and the κ=0 block yields 0.75*RDC.

3. The quantity $\left|\frac{0.75 * RDC}{2} - 0.375 * RDC\right| * \frac{8}{3}$ yields the properly scaled 'modulo differential error' in RDC. The square of this quantity is defined as the 'square differential error'.

4. The square differential error was plotted for each residue. Residues for which the square differential error deviates more than 2*RMSD from the mean value were rejected.

5. This filter was applied twice, to yield a dataset designated 'statistically sound'.

Shown in Fig.3.26 to 3.28 are three plots of 0.375*RDC vs. 0.75*RDC, determined as discussed above. These plots correspond to the dataset before filtering, and after the first and second filtering process. The improvement in the R^2 value from 0.712 to 0.9393 indicates mutually consistent RDCs. The ideal case would correspond to a straight line through the origin with a slope of 0.5.

We used a homology model of DnaK Tth-NBD for our computations. To avoid RDC violations due to uncertainties in the model, we only used RDCs for those residues that correspond to maximally conserved residues across different members of the Hsp70 family were retained from the statistically sound dataset. The resulting dataset was designated 'experimentally sound'.

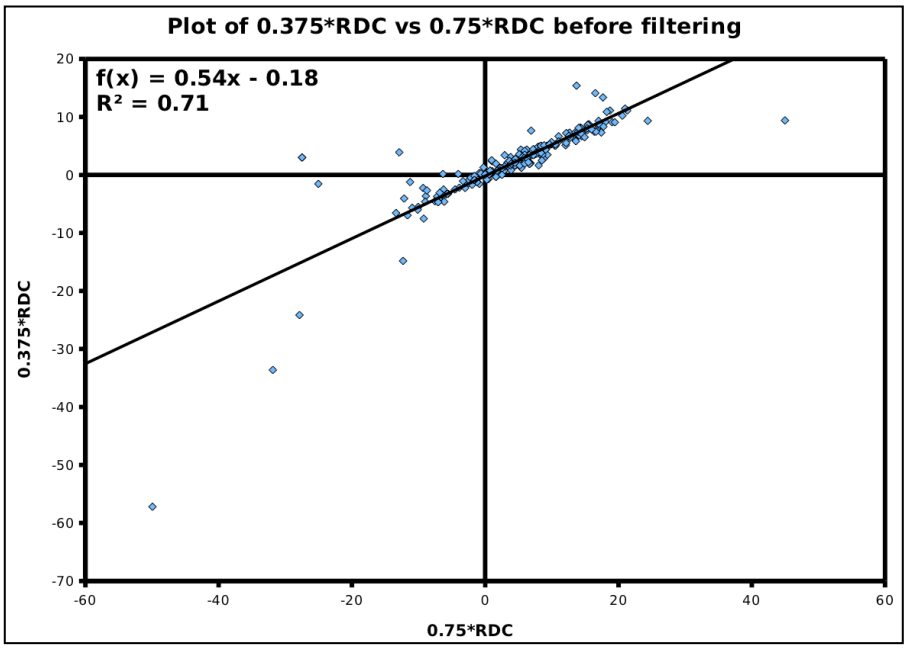

Figure 3.26: Consistency check plot of experimental RDCs before the first filter

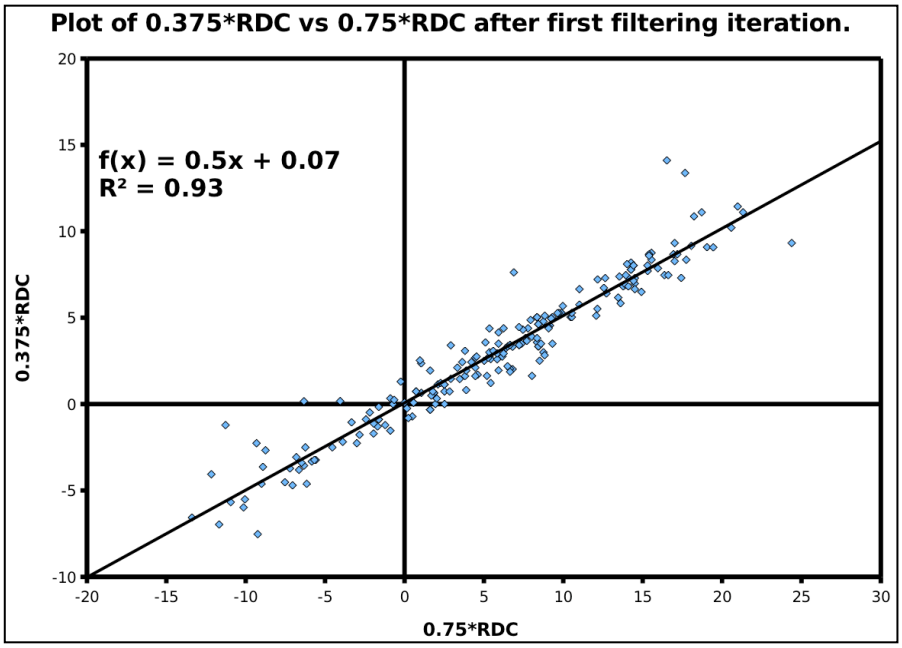

Figure 3.27: Consistency check plot of experimental RDCs after the first filter

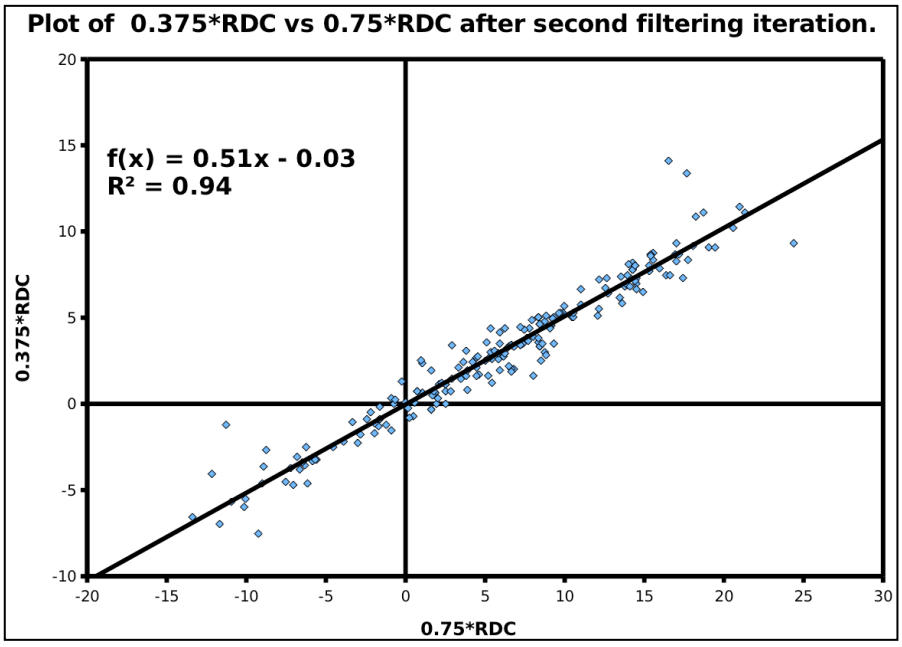

Figure 3.28: Consistency check plot of experimental RDCs after the second filter

This led to a subset of RDCs, enumerated in Table 3 per sub-domain, for each form of the protein.

3.9.3 STATISTICS OF THE COMPUTATION

We used in-house written Fortran 90 programs and REDCAT [177] (A Residual Dipolar Coupling Analysis Tool) to transform the RDCs to orientational data for the NBD subdomains. RDC data can be transformed to a set of simultaneous linear equations using spherical harmonics. These simultaneous linear equations are then numerically solved. REDCAT's solution algorithm relies on singular value decomposition and Monte Carlo error estimation to generate an ensemble of 1000 structures compatible with the input structures and the set of RDCs provided [215], in our case based on an experimental error range of 4 Hz. Along with the 1000 nontrivial solutions,10 'null' set solutions are also generated. The in-house written Fortran program is based on a grid and minimization search of all possible Euler rotations, overall alignment and rhombicity, to find the best fit to the experimental data. The results are identical to the REDCAT solutions for high quality data. However, for noisy data, the Fortran program proved to be the more robust approach to find solutions. The structural data shown in this report were all derived from the results of Fortran program, since it gave easier access to the distribution of Euler angles. The Fortran program is available upon request. The Sanson-Flamsteed plots were derived from REDCAT.

For our purposes, each sub-domain of the DnaKTth-NBD is viewed as an internally 'rigid' fragment. The rationale is that rigid-body reorientations of the individual sub-domains have been observed in crystal structures of complexes of chaperones and co-chaperones such as BAG-1[156], GrpE[90] and HspBP[216]. Furthermore, we observed significant exchange broadening on the interfaces between the sub-domains, indicating true boundaries in solution as well. Hence it was possible to use the residual dipolar couplings to provide orientation information by identifying each sub-domain of DnaKTth-NBD as a rigid fragment.

Identifying each sub-domain of DnaK Tth-NBD as a rigid fragment, residual

dipolar couplings provide orientation information. This information once extracted was applied to the homology model of DnaKTth-NBD to detect conformational changes. The homology model was constructed in two steps. First, we threaded the sequence of DnaKTth-NBD on the coordinates of a crystal structure of the NBD of DnaK- *E. coli*, the protein most homologous to DnaK *T. thermophilus* for which coordinates are available[90] (1DKG.pdb, with a resolution of 2.8 Å). However, this structure was obtained in the presence of a nucleotide exchange factor, which likely rearranged the subdomain orientations. Several crystal structures [155] without co-chaperones are available for the bovine Hsc70 NBD, in which the subdomain orientations, especially for IIB, are different from that seen in 1DKG.pdb. In order to obtain the required reference model for DnaKTth-NBD, we adjusted the orientations of the subdomains in the model to correspond with those in 3HSC.pdb, a structure of Hsc70-NBD nominally in the ADP state at a resolution of 1.90 Å. The adjustment was based on a superposition of the Ca atoms of the secondary structure elements in the subdomains.

3.9.4 SUBDOMAIN ORIENTATIONS

Table 3 shows the statistics of the sub-domain orientation determination using the grid-search Fortran-90 programs with a Jack knife cut off of 60% of the experimental data.

Table 3: Statistics of RDC calculations of DnaK-TTh-NDB

		N[a]	α[b]	β[b]	γ[b]	RMSD	Q[c]	Szz	Syy	Sxx
			Deg.	Deg.	Deg.	Å				
DnaKTTh-NBD AMPPNP IA	Alld	45	161.83	87.59	63.07	8.11	0.65	-8.61E-04	8.83E-04	-2.14E-05
	<SV>[e]	34	160.23	87.41	61.89	9.10	0.68	-9.21E-04	8.64E-04	5.68E-05
	σ[f]	4	6.18	4.48	5.20	1.63	0.10	6.31E-05	1.02E-04	7.41E-05
DnaKTTh-NBD AMPPNP IB	All	26	136.39	76.41	69.98	10.52	0.67	-1.07E-03	7.54E-04	3.13E-04
	<SV>	23	145.3	81.4	66.1	11.06	0.74	-8.97E-	7.38E-	1.59E-

			4	1	7			04	04	04
	σ	3	26.26	5.27	4.51	0.91	0.06	6.43E-05	9.32E-05	8.65E-05
DnaKTTh-NBD AMPPNP IIA	All	34	163.02	90.63	61.09	90.63	118.91	7.30E+01	1.51E+02	1.63E+02
	<SV>	27	144.59	88.00	58.68	6.97	0.53	-9.27E-04	7.37E-04	1.90E-04
	σ	4	6.89	5.91	4.57	0.62	0.06	5.13E-05	6.43E-05	8.87E-05
DnaKTTh-NBD AMPPNP IIB	All	34	159.84	90.29	56.92	90.29	123.08	6.98E+01	1.47E+02	1.60E+02
	<SV>	28	166.71	92.91	69.78	9.05	0.71	-8.08E-04	8.07E-04	1.04E-06
	σ	4	6.99	3.68	9.96	1.20	0.10	6.12E-05	6.56E-05	3.15E-05
DnaKTTh-NBD ADP IA	All	45	157.95	84.39	67.53	7.94	0.60	-1.02E-03	7.65E-04	2.58E-04
	<SV>	40	163.96	85.36	67.06	7.92	0.64	-8.42E-04	6.50E-04	1.92E-04
	σ	5	13.94	6.62	5.75	0.87	0.06	6.12E-05	7.26E-05	7.98E-05
DnaKTTh-NBD ADP IB	All	26	117.43	81.50	68.30	8.46	0.57	-1.03E-03	7.68E-04	2.65E-04
	<SV>	22	154.49	87.39	69.54	9.34	0.65	-9.72E-04	6.76E-04	2.97E-04
	σ	3	60.48	5.19	4.05	1.10	0.09	6.48E-05	7.00E-05	5.49E-05
DnaKTTh-NBD ADP IIA	All	24	183.93	75.88	62.49	5.77	0.43	-9.78E-04	6.48E-04	3.30E-04
	<SV>	27	180.00	81.18	68.08	6.97	0.48	-9.94E-04	7.64E-04	2.30E-04
	σ	4	11.63	3.39	5.92	0.78	0.06	4.18E-05	7.17E-05	6.13E-05
DnaKTTh-NBD	All	34	165.1	89.5	49.8	8.04	0.53	-1.06E-	9.86E-	7.54E-

ADP IIB			3	6	0			03	04	05
	<SV>	30	156.91	83.62	48.41	8.11	0.61	-8.48E-04	7.54E-04	9.37E-05
	σ	4	10.35	3.54	4.09	1.13	0.07	5.72E-05	7.79E-05	4.49E-05
DnaKTTh-NBD APO IA	All	66	160.58	81.15	65.76	4.53	0.46	-7.96E-04	6.45E-04	1.51E-04
	<SV>	41	160.52	81.02	66.12	4.31	0.44	-8.06E-04	6.44E-04	1.62E-04
	σ	5	5.09	1.85	3.15	0.36	0.05	5.22E-05	7.22E-05	4.42E-05
DnaKTTh-NBD APO IB	All	32	189.53	76.49	66.36	8.46	0.66	-1.10E-03	6.17E-04	4.81E-04
	<SV>	19	90.00	76.14	66.61	7.90	0.61	-1.12E-03	6.85E-04	4.39E-04
	σ	3	69.98	4.31	3.00	1.04	0.08	9.78E-05	7.13E-05	1.01E-04
DnaKTTh-NBD APO IIA	All	25	136.16	83.95	67.26	5.38	0.79	-5.17E-04	3.60E-04	1.57E-04
	<SV>	15	129.44	84.03	69.49	4.99	0.72	-5.49E-04	4.03E-04	1.46E-04
	σ	3	36.32	22.91	20.36	0.61	0.08	9.11E-05	7.78E-05	4.76E-05
DnaKTTh-NBD APO IIB	All	30	144.58	89.91	51.28	5.42	0.51	-8.83E-04	6.93E-04	1.90E-04
	<SV>	19	143.31	89.38	51.58	4.91	0.47	-8.89E-04	7.34E-04	1.56E-04
	σ	3	8.92	3.18	4.37	0.60	0.07	7.23E-05	9.21E-05	7.24E-05
DnaKTTh_NBD_SBD ADP IA	All	25	97.01	60.93	75.94	7.42	0.46	-1.54E-03	9.82E-04	5.54E-04
	<SV>	16	96.21	60.03	75.90	6.80	0.44	-1.56E-03	1.03E-03	5.25E-04
	σ	2	15.98	4.21	5.44	1.00	0.08	1.20E-	1.45E-	8.41E-

								04	04	05
DnaKTTh_NBD_SBD ADP IB	All	24	135.87	68.47	74.47	7.54	0.48	-1.89E-03	1.25E-03	6.37E-04
	<SV>	15	124.35	68.97	72.87	7.00	0.46	-1.82E-03	1.21E-03	6.16E-04
	σ	2	43.09	2.72	5.04	1.04	0.09	1.32E-04	1.90E-04	1.28E-04
DnaKTTh_NBD_SBD ADP IIA	All	31	97.27	74.12	75.60	7.08	0.38	-1.54E-03	1.04E-03	5.04E-04
	<SV>	20	96.51	75.09	73.69	6.54	0.37	-1.52E-03	1.04E-03	4.79E-04
	σ	3	15.78	2.54	2.97	1.31	0.08	1.02E-04	1.16E-04	1.04E-04
DnaKTTh_NBD_SBD ADP IIB	All	27	94.99	78.53	66.32	7.15	0.38	-1.83E-03	1.37E-03	4.59E-04
	<SV>	16	92.49	79.57	65.57	6.67	0.35	-1.87E-03	1.38E-03	4.91E-04
	σ	2	9.36	3.74	2.25	0.96	0.05	8.54E-05	9.74E-05	1.21E-04

Table 3.3: Statistics of RDC computing for TTh-NBD

Calculations for the subdomains of DnaK-TTh-NDB in various states, using a grid-search algorithm.

(a.) Number of RDCs used

(b.) Euler angles in the z-y-z convention

(c.) Quality factor, defined - $Q = \dfrac{RMSD}{\sqrt{\dfrac{\sum_{i=1}^{N_{obs}} \left(RDC(i)_{exp} \right)^2}{N}}}$

(d.) All RDCs used, best fit result

(e.) Average of self validation at the 60% level

(f.) Standard deviation of the self validation

If all subdomains were oriented identically as in the homology structure, which was based on the crystal structure of Hsc70-NBD, one would expect that the Euler angles needed to rotate these subdomains into their PAS would be same for all (within experimental accuracy). Table 3 shows that this is only (approximately) the case for DnaK Tth-NBD in the AMPPNP state. Figure 3.29 displays this same information in a Sanson Flamsteed projection, which gives the orientations of the Szz and Sxx orientational principal axes of the subdomains with respect to the model structure Cartesian frame (which is the frame of the PDB file of Hsc70). Here, the Szz axis for all subdomains of DnaKTth-NBD in the AMPPNP state all pierce the unit sphere around (15° N, 110 ° E) and (15 ° S, 70 ° W). The Sxx axes of the subdomains are relatively well-defined and pierce the unit sphere around (30 ° S, 20 ° E) and (30 ° N, 160 ° W). But they are not at identical positions.

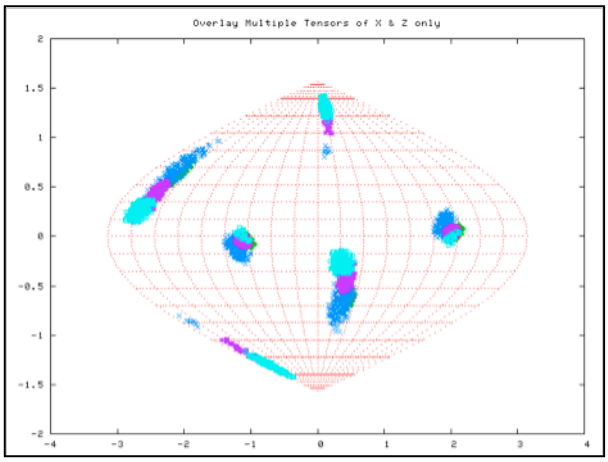

Figure 3.29: Sanson Flamsteed sinusoidal projection for the AMPPNP bound state of Tth-NBD

Fig.3.30 shows that the Szz axis of sub-domain IIB of DnaK Tth-NBD in the ADP state is shifted by nearly 20 degrees from the axes of the other subdomains. The dark blue smears indicate that the Sxx orientation for domain IB cannot be obtained with any certainty.

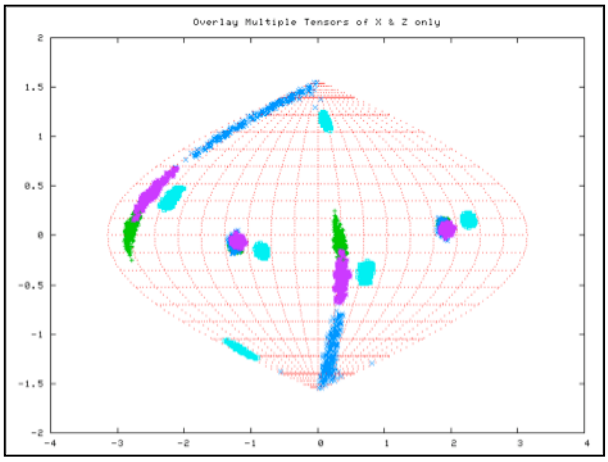

Figure 3.30: Sanson Flamsteed sinusoidal projection for the ADP.Pi bound state of Tth-NBD

For the *apo* state, the Szz axis of domain IIB is also seen to deviate from that of the other three subdomains, but in a different sense than in the ADP state. Here the Sxx-axes cannot be determined with any certainty for subdomains IB and IIA (see Fig. 3.31).

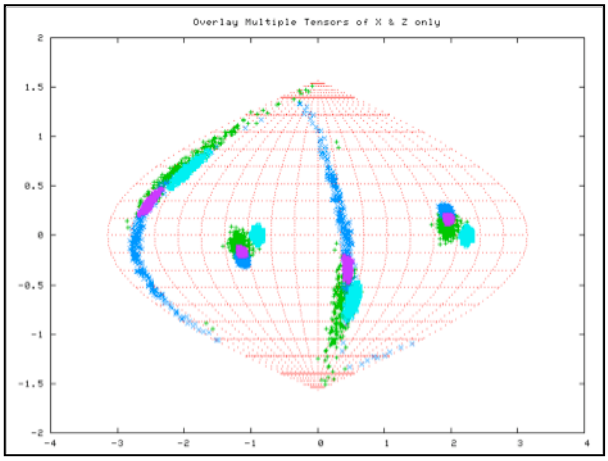

Figure 3.31: Sanson Flamsteed sinusoidal projection for the apo state of Tth-NBD

For DnaK Tth-NBD-SBD, the data is better (Fig 3.32), and Szz and Sxx axes can be defined for all sub domains. They all deviate significantly from the orientations in the model structure.

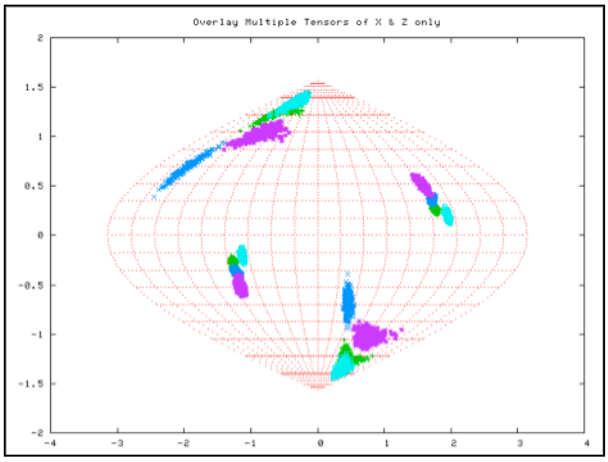

Figure 3.32: Sanson Flamsteed sinusoidal projection for the ADP.Pi bound state of Tth-NBD-SBD

3.10 STATISTICAL VALIDATION

In order to evaluate the statistical significance of the domain reorientations, we performed a 'self validation analysis' on our data. This consists of rejecting 40% of available RDCs at random generating truncated datasets. From an n element dataset an un-ordered combination of k elements can be created by C(k,n) ways. For example, for sub-domain IIB in the ADP state, we can create 1.26 X 10^{14} datasets by rejecting half of the data at random. In practice, 100 datasets are generated for each self-validation run. The best fit solutions for each of these 100 datasets is plotted as a point on the Sanson Flamsteed projection. The choice of a 40% rejection rate for this self-validation (also called Jack-knife) was discussed at length in Chapter 2 [203].

From the self-validation study, our result appears to be quite robust (Fig.3.33 to 3.36): the clustering of all sub-domain alignments in the AMPPNP of DnaKTth-NBD and in DnaKTth-NBD-SBD is still visible at 60% data cutoff. The deviation of subdomain IIB in the ADP and APO states from the other subdomains also holds up to the Jack-knife test. However, no well-defined orientations were found for the Szz axes of subdomain IB in the ADP or APO form. The definition of the Sxx axes is problematic for all subdomains in all states, except for subdomain IIB.

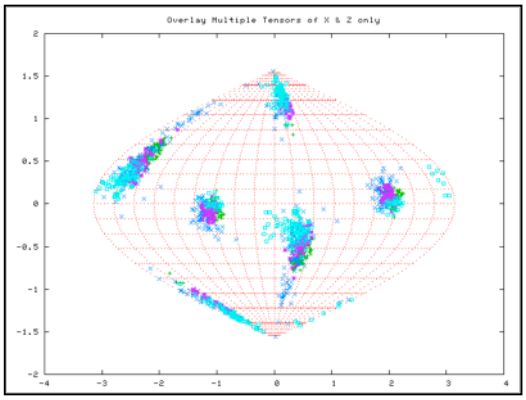

Figure 3.33: Sanson Flamsteed projection for 40% RDCs rejected solutions for Tth-NBD in the AMPPNP bound state

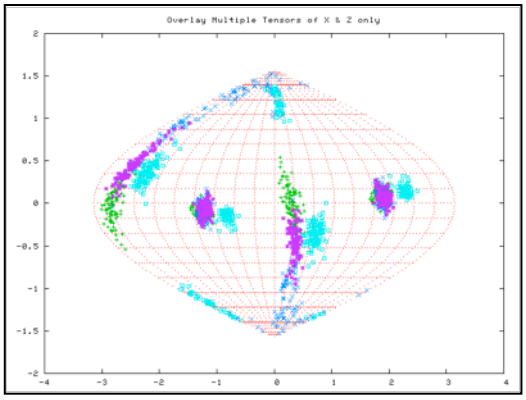

Figure 3.34: Sanson Flamsteed sinusoidal projection for 40% RDCs rejected solution for Tth-NBD in the ADP.Pi bound state

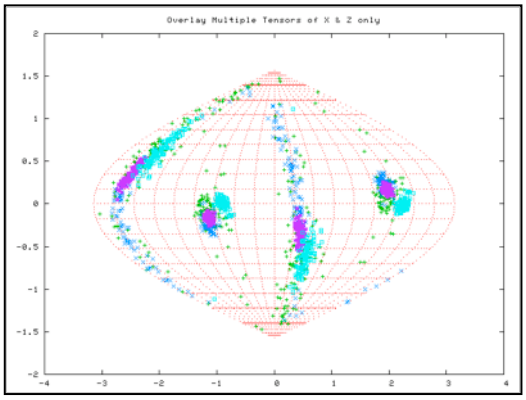

Figure 3.35: Sanson Flamsteed sinusoidal projection of 40% RDCs rejected solution for Tth-NBD in the apo form

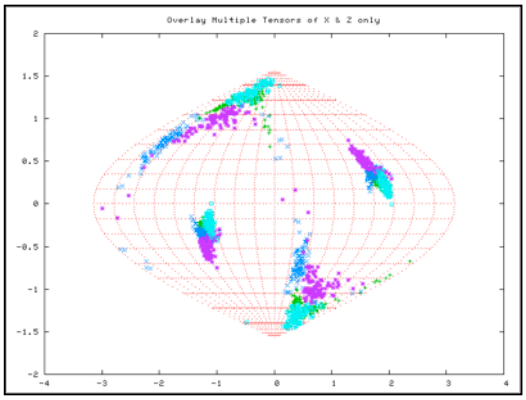

Figure 3.36: Sanson Flamsteed sinusoidal projection of 40% RDCs rejected solution for Tth-NBD-SBD in the ADP.Pi bound state

3.11 STRUCTURAL INSIGHTS FROM RDC ANALYSIS

From this data, we observe that only the AMPPNP bound state of the DnaKTth-NBD closely compares to the reference structure. This is surprising, since the reference structure is a homology model, which is based on the crystal structure of Hsc70 in the ADP state. In contrast, the orientations of domain IIB of DnaKTth-NBD in the ADP-bound state in solution deviates significantly from the crystal structure. These baffling results have been discussed extensively in our earlier work [203], shown in Chapter 2, which concluded that the Hsc70 ADP crystal structure is a closed-cleft structure, while in solution only the ATP-like state is in this closed form, with more degrees of freedom for the subdomain orientations in the ADP state. These findings have also been discussed in the context of the allosteric mechanism of the Hsp70 chaperones [203].

Here, we find that subdomain IIB is also oriented “away” in the APO state of the protein, but not quite the same as in the ADP state. This suggests an “ADP” like behavior for this state, which corresponds to the findings in biochemical studies [154]. However, the apo state has no functional significance, since the nucleotide binding cleft will almost always be occupied by ATP in vivo due to the high cellular concentration of that nucleotide.

The orientation of subdomain IIB in the DnaKTth-NBD-SBD construct in the ADP state, with no substrate bound, is only slightly offset from the other subdomains, and is

intermediate between the open and closed form. This corresponds quite well with what one would expect[91; 92; 184] for this “mixed state”. Peptide binding is known to drive the NBD of the Hsp70’s to the ADP state, but lack of peptide stabilizes the ATP state. Hence, and ADP bound state without peptide is expected to be “intermediate”.

3.12 3D STRUCTURAL INFORMATION – SOLUTION STRUCTURES OF Tth-NBD AND Tth-NBD-SBD

The primary RDC data only yields orientational information. The magnitudes of the alignment tensor also give information about the overall shape of the system under study. This can be exploited [217] when the data is of very high precision and / or when the system deviates very much from spherical shape. Neither is the case for the current data and molecules. Hence, in the current case we have only orientational information.

In an attempt to give a 3D impression of the conformational changes between the different states, we use the following procedure. Given the reorientation information (i.e., averaged Euler angles corresponding to Principal Axes frames or ‘PAS’) for each sub-domain, we rotate each sub-domain, for each nucleotide state, into its principal axis frame, which we call an ‘RDC Eigen structure’. A model for relative translational positions of these sub-domain Eigen structures for each nucleotide state in a full molecule was obtained as follows. First, the entire reference structure was rotated into PAS of the RDCs of subdomain IA for a particular nucleotide state. Subsequently, the ‘RDC Eigen structure’ of each sub-domain was superposed on the corresponding subdomains of the (rotated) reference structure using translational shifts only, using the Cα atom coordinates of the secondary structure elements in those subdomains. Thus, in such a model structure, the RDC-determined orientations of the subdomains are retained, and their translational positions are roughly equal to the reference structure, which was based on experimental X-ray diffraction data. We have chosen to not further refine these models even though significant amount of atom clashes occur at the interfaces of the sub domains. Potentially one can do so, by running a Molecular

Dynamics relaxation computation, in which the relative subdomain orientations are retained (e.g. by restraining an orientational parameter based on the Cα positions of the secondary structure elements).

The overall model structures for the different nucleotide states were oriented in the PAS of the RDCs of their subdomains IA. These orientations should be roughly the same, since the same alignment medium (10 mg/ mL phage) was used, but are not necessarily exactly the same, if conformational changes occur which affect the overall shape of the protein. However, for the AMPPNP, ADP and APO states of DnaKTth-NBD, the overall alignment is very similar, as a comparison of Fig 3.29 to 3.31 shows. This indicates that no large conformational changes, beyond the subdomain orientation changes, take place. This justifies our modeling protocol described above. It is to be expected that the overall orientation of the NBD in the DnaKTth-NBD-SBD construct is different (see Fig 3.32): the overall hydrodynamic shape of the molecule is quite different from that of the isolated NBD because of the presence of the 10 kDa SBD.

In order to compare the different molecular molecules, we have to remove the overall orientational differences. This can be accomplished by a simple superposition of the entire molecules. In order to emphasize the differences in cleft openings, we have chosen to overlay the model structures on the Ca positions of subdomain IA only. The results of these exercises are shown in Fig 3.37 to 3.40. In Figure 3.37/38, one easily identifies that the orientation of domain IIB, at the right top, is very similar in the reference structure (green) and the AMPPNP state (blue) of DnaKTth-NBD. In the ADP (yellow) and APO (red) structure we find that IIB is rotated “outward” in a rather similar fashion, representing an open ATP binding cleft. The orientation of subdomain IIA (bottom right) of the AMPPNP state is most similar to the reference structure, while the ADP and APO forms deviate considerably, and differently from each other (also see Figure 3.35-36).

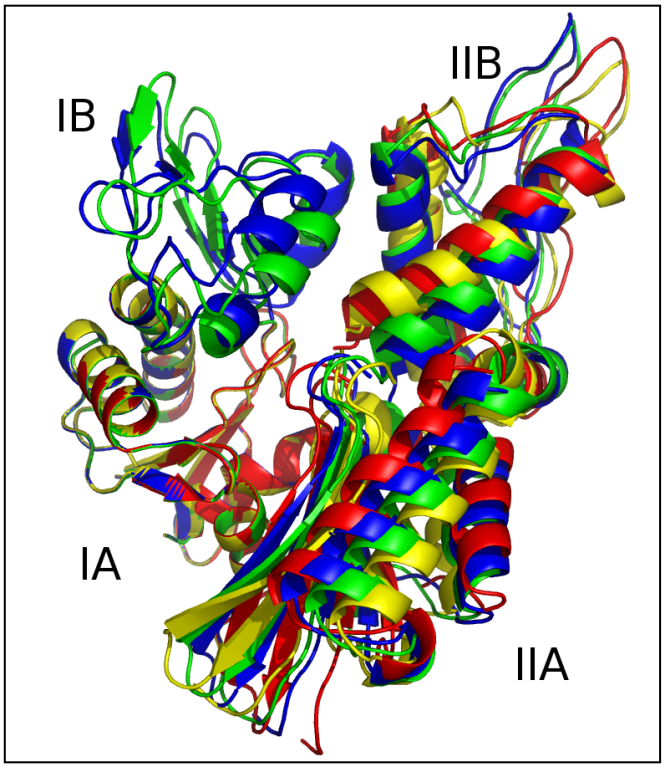

Figure 3.37: Overlay of Tth-NBD RDC eigenstructures (standard view)

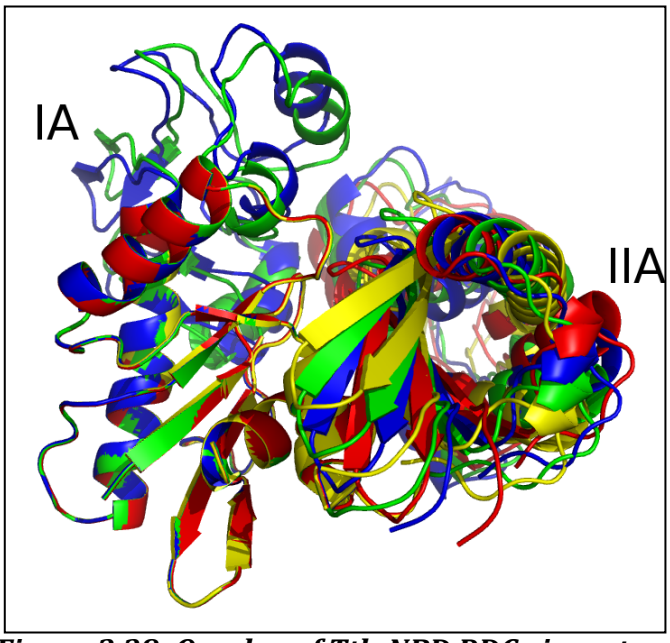

Figure 3.38: Overlay of Tth-NBD RDC eigenstructures (IA/IIA interface view)

Overlays of RDC eigenstructures using secondary structure of sub-domain IA as a reference. DnaK Tth-NBD AMPPMP – blue DnaK Tth-NBD ADP – yellow DnaK Tth-NBD apo – red. The reference structure based on the Hsc70 NBD crystal

structure is in green. The IB domains for the ADP and APO state are not shown. Fig 3.35, "standard" view, with IA at bottom left, IB at top left, IIA at bottom right and IIB at top right. Fig 3.36, "bottom" view, with IA at the left and IIA at the right.

Figure 3.39/40, below, shows that the orientation of subdomain IIB of DnaK DnaKTth-NBD-SBD in the ADP – peptide-free state (cyan) is very similar to that of the ADP state (yellow). However, the orientation of subdomain IIA of DnaK DnaKTth-NBD-SBD is more similar to that of the AMPPNP state. This confirms the "mixed conformation" as expected for this state (see also above).

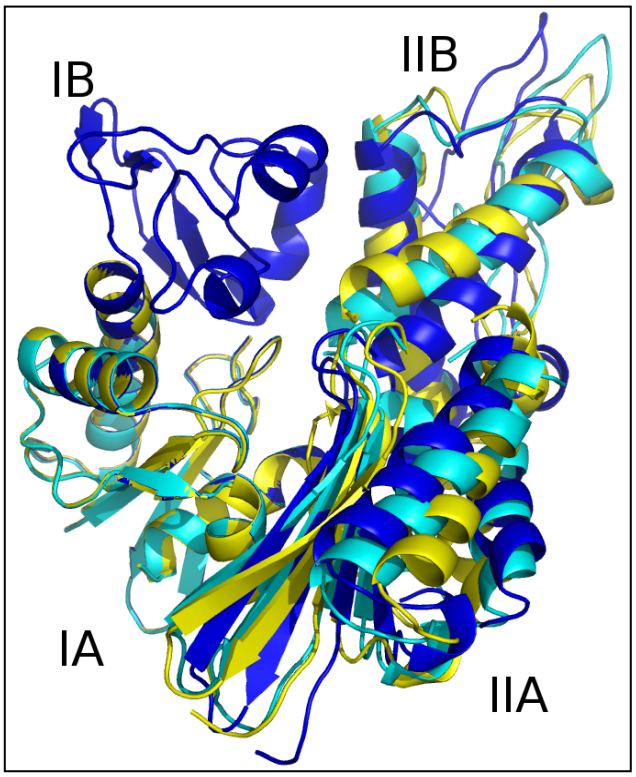

Figure 3.39: Overlay of RDC eigenstructures of nucleotide bound forms of Tth-NBD and Tth-NBD-SBD: standard view

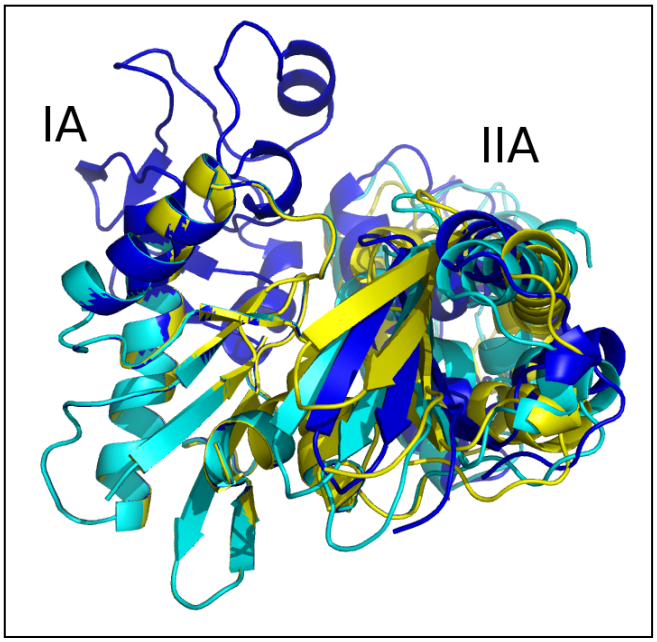

Figure 3.40: Overlay of RDC eigenstructures of nucleotide bound forms of Tth-NBD and Tth-NBD-SBD: bottom view of the IA/IIA interface

Overlays of RDC eigenstructures using secondary structure of sub-domain IA as a reference. DnaKTth-NBD AMPPMP – blue DnaKTth-NBD ADP – yellow DnaKTth-NBD-SBD ADP – cyan The IB domains for the ADP and APO state and DnaKTth-NBD-SBD are not shown. Fig 3.37, "standard" view, with IA at bottom left, IB at top left, IIA at bottom right and IIB at top right. Fig 3.38, "bottom" view, with IA at the left and IIA at the right.

3.13 CONCLUSIONS

^{15}N-^{1}H RDC data can be obtained for relatively large proteins using the RDC-TROSY methodology. The simulations show that κ values should not be extended above 1.0. Simulations also show that this RDC measurement paradigm is superior to other methods (IP-AP and TROSY – HSQC comparisons) for almost all practical molecular weight classes on any practical NMR spectrometer. Our simulations show that amide proton flip-flops, caused by NOEs with nearby protons, quickly deteriorate the quality of the RDC-TROSY, and of TROSY in general. Hence perdeuteration of the protein, which suppresses such NOEs, will improve the quality of the RDC-TROSY, and TROSY in general, not only by improving the direct linewidths, but also by removal

of “lifetime broadening” of the TROSY lines in the ^{15}N dimension.

Despite the improvements afforded by the RDC-TROSY experimentation and perdeuteration, RDC data for large proteins are still noisy. Hence interpretation of this data requires very conservative jack-knife evaluations at the 40% rejection rate to establish statistical significance. We find, for noisy RDC data, that a computer algorithm based on a simple grid-search is more robust than an SVD-based analytical fit. Finally, we show that the combined approaches described herein can yield functionally important conformational information on larger proteins in the size range of 45 – 60 kDa (with rotational correlation times around 20 ns), using samples with ~ 300 uM protein concentration, and experimental times of less than a week, using a 800 MHz NMR instrument equipped with a cold (cryo) probe. In addition to studies described herein, this lab has also obtained usable RDC data with this method for ~200 uM wt-DnaK *E. coli*, a 70 kDa protein, measured at 27 ^{0}C[92].

3.14 REFERENCES

[1] D. Yang, R.A. Venters, G.A. Mueller, W.Y. Choy, and L.E. Kay, TROSY-based HNCO pulse sequences for the measurement of 1HN-^{15}N, ^{15}N-13CO, 1HN-13CO, 13CO-13C[agr] and 1HN-13C[agr] dipolar couplings in ^{15}N, 13C, 2H-labeled proteins. J. Biomol. NMR 14 (1999) 333-343.

[2] M.H. Levitt, Spin Dynamics: Basic Principles of NMR Spectroscopy. (2001).

[3] M. Ikura, L.E. Kay, and A. Bax, A novel approach for sequential assignment of 1H, 13C, and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry 29 (1990) 4659-67.

[4] A. Bax, and M. Ikura, An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly ^{15}N/13C enriched proteins. J Biomol NMR 1 (1991) 99-104.

[5] M. Ikura, L.E. Kay, M. Krinks, and A. Bax, Triple-resonance multidimensional NMR study of calmodulin complexed with the binding domain of skeletal muscle myosin light-chain kinase: indication of a conformational change in the central helix. Biochemistry 30 (1991) 5498-504.

[6] A. Kumar, R.R. Ernst, and K. Wuthrich, A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem Biophys Res Commun 95 (1980) 1-6.

[7] K. Wuthrich, Protein structure determination in solution by nuclear magnetic resonance spectroscopy. Science 243 (1989) 45-50.
[8] K. Wuthrich, Protein structure determination in solution by NMR spectroscopy. J Biol Chem 265 (1990) 22059-62.
[9] K. Wuthrich, Six years of protein structure determination by NMR spectroscopy: what have we learned? Ciba Found Symp 161 (1991) 136-45; discussion 145-9.
[10] C.D. Schwieters, and G.M. Clore, The VMD-XPLOR visualization package for NMR structure refinement. J Magn Reson 149 (2001) 239-44.
[11] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, and G.M. Clore, The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160 (2003) 65-73.
[12] P. Mandal, and A. Majumdar, A comprehensive discussion of HSQC and HMQC pulse sequences. Concepts in Magnetic Resonance Part A 20 (2004) 1-23.
[13] N. Tjandra, and A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278 (1997) 1111-1114.
[14] 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. Proc. Natl. Acad. Sci. USA 92 (1995) 9279-9283.
[15] M.W. Fischer, J.A. Losonczi, J.L. Weaver, and J.H. Prestegard, Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. Biochemistry 38 (1999) 9013-9022.
[16] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard, NMR evidence for slow collective motions in cyanometmyoglobin. Nat. Struct. Biol. 4 (1997) 292-297.
[17] H.M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, and D.J. Patel, Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. J Mol Biol 315 (2002) 95-102.
[18] N. Tjandra, Establishing a degree of order: obtaining high-resolution NMR structures from molecular alignment. Structure 7 (1999) R205-11.
[19] M. Ottiger, F. Delaglio, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J. Magn. Reson. 131 (1998) 373-378.
[20] W.J.F. John Cavanagh, Arthur G. Palmer III, Nicholas J. Skelton, Protein NMR Spectroscopy: Principles and Practice. (2007).
[21] K. Pervushin, R. Riek, G. Wider, and K. Wuthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94 (1997) 12366-71.
[22] J.A. Lukin, G. Kontaxis, V. Simplaceanu, Y. Yuan, A. Bax, and C. Ho, Quaternary structure of hemoglobin in solution. Proc Natl Acad Sci U S A 100 (2003) 517-20.
[23] A. Sheth, M. Ravikumar, and R. Hosur, Application of J scaling in two-dimensional spin-echo-correlated spectroscopy to observation of small coupling correlations. Journal or Magnetic Resonance 74 (1969) 352-355.

[24] Y. Zhang, and E.R. Zuiderweg, The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. Proc Natl Acad Sci U S A 101 (2004) 10272-7.
[25] M. Revington, Y. Zhang, G.N. Yip, A.V. Kurochkin, and E.R. Zuiderweg, NMR investigations of allosteric processes in a two-domain Thermus thermophilus Hsp70 molecular chaperone. J Mol Biol 349 (2005) 163-83.
[26] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc Natl Acad Sci U S A 106 (2009) 8471-6.
[27] A. Bhattacharya, A.V. Kurochkin, G.N. Yip, Y. Zhang, E.B. Bertelsen, and E.R. Zuiderweg, Allostery in Hsp70 chaperones is transduced by subdomain rotations. J Mol Biol 388 (2009) 475-90.
[28] D.S. Weaver, and E.R. Zuiderweg, Protein proton-proton dynamics from amide proton spin flip rates. J Biomol NMR 45 (2009) 99-119.
[29] M.P. Mayer, and B. Bukau, Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci 62 (2005) 670-84.
[30] K.M. Flaherty, C. Delucaflaherty, and D.B. McKay, 3-DIMENSIONAL STRUCTURE OF THE ATPASE FRAGMENT OF A 70K HEAT-SHOCK COGNATE PROTEIN. Nature 346 (1990) 623-628.
[31] K.M. Flaherty, C. Deluca-Flaherty, and D.B. McKay, 3-Dimensional Structure of the ATPase Fragment of a 70k Heat-Shock Cognate Protein. Nature 346 (1990) 623-628.
[32] M.C. O'Brien, K.M. Flaherty, and D.B. McKay, Lysine 71 of the chaperone protein Hsc70 Is essential for ATP hydrolysis. J. Biol. Chem. 271 (1996) 15874-15878.
[33] M.C. O'Brien, and D.B. McKay, Threonine 204 of the chaperone protein Hsc70 influences the structure of the active site, but is not essential for ATP hydrolysis. J. Biol. Chem. 268 (1993) 24323-24329.
[34] M.C. O'Brien, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. I. Potassium is required for optimal ATPase activity. J. Biol. Chem. 270 (1995) 2247-2250.
[35] M.C. Sousa, and D.B. McKay, The hydroxyl of threonine 13 of the bovine 70-kDa heat shock cognate protein is essential for transducing the ATP-induced conformational change. Biochemistry 37 (1998) 15392-9.
[36] S.M. Wilbanks, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. II. Potassium binds specifically in the ATPase active site. J Biol Chem 270 (1995) 2251-7.
[37] M.W.F. Fischer, A. Majumdar, and E.R.P. Zuiderweg, Protein NMR relaxation: theory, applications and outlook. Prog. NMR Spectrosc. 33 (1998) 207-272.
[38] T. Schulte-Herbruggen, and O.W. Sorensen, Clean TROSY: compensation for relaxation-induced artifacts. J Magn Reson 144 (2000) 123-8.
[39] A. Abragam, The Principles of Nuclear Magnetism, Clarenden Press, Oxford, 1961.
[40] A. Carrington, and A. MacLachlan, Introduction to Magnetic Resonance with Applications to Chemistry and Chemical Physics, Harper & Row, New York, 1967.

[41] M. Revington, T.M. Holder, and E.R. Zuiderweg, NMR study of nucleotide-induced changes in the nucleotide binding domain of Thermus thermophilus Hsp70 chaperone DnaK: implications for the allosteric mechanism. J Biol Chem 279 (2004) 33958-67.
[42] M. Revington, and E.R. Zuiderweg, TROSY-driven NMR backbone assignments of the 381-residue nucleotide-binding domain of the Thermus Thermophilus DnaK molecular chaperone. J Biomol NMR 30 (2004) 113-4.
[43] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6 (1995) 277-293.
[44] T.D. Goddard, and D.G. Kneller, SPARKY 3. University of California, San Francisco (2000).
[45] A.L. Hansen, and H.M. Al-Hashimi, Insight into the CSA tensors of nucleobase carbons in RNA polynucleotides from solution measurements of residual CSA: towards new long-range orientational constraints. J Magn Reson 179 (2006) 299-307.
[46] H. Valafar, and J.H. Prestegard, REDCAT: a residual dipolar coupling analysis tool. J Magn Reson 167 (2004) 228-41.
[47] J.A. Losonczi, M. Andrec, M.W. Fischer, and J.H. Prestegard, Order matrix analysis of residual dipolar couplings using singular value decomposition. J. Magn. Reson. 138 (1999) 334-342.
[48] H. Sondermann, C. Scheufler, C. Schneider, J. Hohfeld, F.U. Hartl, and I. Moarefi, Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. Science 291 (2001) 1553-7.
[49] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-5.
[50] C.A. McLellan, D.A. Raynes, and V. Guerriero, HspBP1, an Hsp70 cochaperone, has two structural domains and is capable of altering the conformation of the Hsp70 ATPase domain. J Biol Chem 278 (2003) 19017-22.
[51] J.S. McCarty, A. Buchberger, J. Reinstein, and B. Bukau, The role of ATP in the functional cycle of the DnaK chaperone system. J. Mol. Biol. 249 (1995) 126-137.
[52] W. Rist, C. Graf, B. Bukau, and M.P. Mayer, Amide hydrogen exchange reveals conformational changes in hsp70 chaperones important for allosteric regulation. J Biol Chem 281 (2006) 16493-501.
[53] J.F. Swain, G. Dinler, R. Sivendran, D.L. Montgomery, M. Stotz, and L.M. Gierasch, Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol Cell 26 (2007) 27-39.
[54] M. Zweckstetter, G. Hummer, and A. Bax, Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. Biophys J 86 (2004) 3444-60.

CHAPTER 4 UNDERSTANDING THE Hsp70 CYCLE - INTERACTION OF DnaK AND DnaJ

To be published by A. Bhattacharya, A. Ahmad, R. McDonald, E. B. Bertelsen and E.R.P. Zuiderweg.

4.1 ABSTRACT

As has been discussed in the introduction (Chapter 1), partner molecules called J-proteins are involved with the loading of substrate peptides on Hsp70. The nature of the interaction of J-proteins has been a subject of great interest. Recent work by crystallographers has resulted in an X-ray interaction model of the bovine J-domain and the nucleotide binding domain of Hsc70. However, this structure is a disulphide linked covalent adduct, and hence compromised in that it likely does not portray a genuine complex interaction between the two proteins.

In this study, we have used the technique of paramagnetic relaxation enhancement NMR to study the interaction between the J-domain and full-length DnaK. Our results lead to a tentative model for the complex structure. We assessed the functional validity of this complex interaction by assaying the efficacy of DnaK in ATP hydrolysis imn the presence of the DnaJ constructs used in the NMR studies. Our results indicate that the crystal structure is [94] may not be a valid model for the interaction. The model we propose shows that the J-domain interacts with both the NBD and the SBD, and thus provides further detail to our understanding of the allosteric machinery of the Hsp70 functional cycle.

4.2 INTRODUCTION

4.2.1 RECAPITULATING THE Hsp70 FUNCTIONAL CYCLE

The Hsp70 functional cycle[48] has been discussed at some length in Chapter 1 (Introduction). Hsp70s function along with several partner molecules [32]. One family of such partner molecules is the nucleotide exchange factor family (GrpE, BAG, Hsp110, HspBP [90; 156; 162; 218]). Chapters 2 and 3 contain a detailed analysis of how nucleotide exchange factors interact with the Hsp70 nucleotide-binding domain, wherein we propose a conformation capture mechanism, which is a modification from the currently accepted model of induced fit allostery [203]. However, the other family of co-chaperones that participate in the Hsp70 cycle has not been scrutinized in detail. This is the family of J-proteins, which are associated with loading substrate peptide on to the Hsp70 molecular machines for processing. Let us discuss the role of J-proteins before we go on to talk about structural studies carried out in this work.

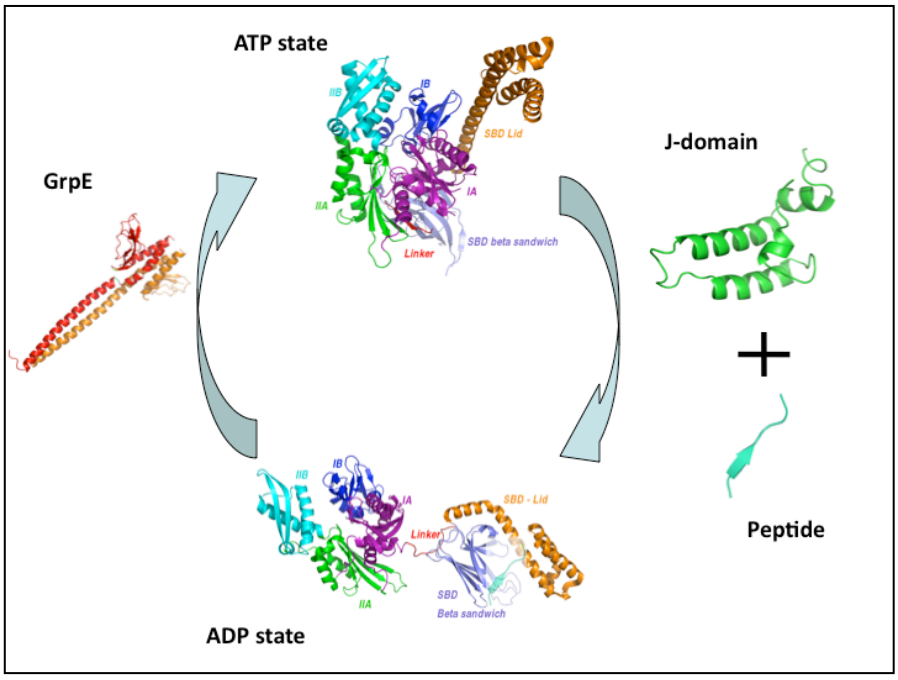

Figure 4.1: Hsp70 functional cycle

4.2.2 THE BIOCHEMICAL ROLE OF J-PROTEINS

J-proteins are not one unified family of biomolecules the way Hsp70s are. This is actually a very diverse group of proteins, which finds application in a myriad of

biological functions.

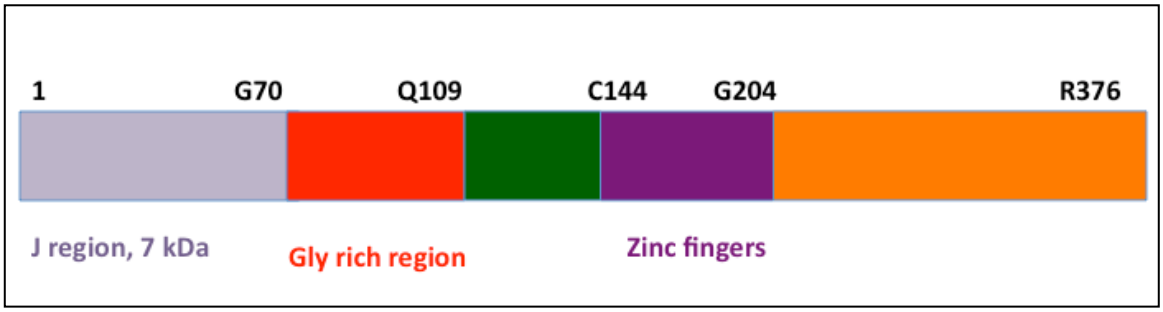

Figure 4.2: Sequence schematic for J-proteins

The main identified role of *E.coli* DnaJ is to stimulate the ATPase activity of the nucleotide-binding domain. In the light of research carried out by various structural biologists [219; 220; 221], and indeed work presented in Chapters 2 and 3 of this dissertation, this should not be viewed as merely enhancing the efficacy of nucleotide turnover. It would be only reasonable to suspect that this would also affect the allosteric mechanism of protein re-engineering carried out by Hsp70s.

Returning to the diversity of the J-protein family, there are at least three classes/families within J-proteins. We have shown a schematic of the sequence of J-proteins in Chapter 1, which is reproduced here for convenience.

4.2.3 DIFFERENT FAMILIES OF J-PROTEINS

Apropos this schematic, there are three regions of interest: the J-region (sometimes called the J-domain), the Gly (and Phe) rich region, and finally the Zinc finger(s) region (rich in Cys). The different families of J-proteins are broadly segregated as below [221]:

DnaJ family	J-region	Gly rich region	Zinc finger(s)
1	Yes	Yes	Yes
2	Yes	Yes	No
3	Yes	No	No

Table 4.1: Different families of J-proteins

Several members of the different J-protein family are listed below. As can be seen, these proteins are present in species ranging from bacteria to mammals.

Accession #	Uniprot Entry	Protein name	Organism	Length
Q03363	DNJH1_ALLPO	DnaJ protein homolog 1	Allium porrum (Leek)	397
P08622	DNAJ_ECOLI	Chaperone protein dnaJ	Escherichia coli (strain K12)	376
O34242	DNAJ_VIBC	Chaperone protein dnaJ	Vibrio cholerae	381
P25491	MAS5_YEAST	Mitochondrial protein import protein MAS5	Saccharomyces cerevisiae (Baker's yeast)	409
Q96EY1	DNJA3_HUMAN	DnaJ homolog subfamily A member 3, mitochondrial	Homo sapiens (Human)	480
Q27237	TID_DROME	Protein tumorous imaginal discs, mitochondrial	Drosophila melanogaster (Fruit fly)	520
P25294	SIS1_YEAST	Protein SIS1	Saccharomyces cerevisiae (Baker's yeast)	352
P25686	DNJB2_HUMAN	DnaJ homolog subfamily B member 2	Homo sapiens (Human)	324
Q9UGP8	SEC63_HUMAN	Translocation protein SEC63 homolog	Homo sapiens (Human)	760
P32527	ZUO1_YEAST	Zuotin	Saccharomyces cerevisiae (Baker's yeast)	433
Q9H3Z4	DNJC5_HUMAN	DnaJ homolog subfamily C member 5 or Cysteine string protein	Homo sapiens (Human)	198
P25685	DNJB1_HUMAN	DnaJ homolog subfamily B member 1	Homo sapiens (Human)	340
O75953	DNJB5_HUMAN	DnaJ homolog subfamily B member 5 or Heat shock protein cognate 40	Homo sapiens (Human)	348
P0A9X9	CSPA_ECOLI	Cold shock protein cspA	Escherichia coli (strain K12)	70

Figure 4.3: Representative list of J-proteins

The J-proteins are somewhat well conserved, although not as well as Hsp70s themselves. A cross-indexed table of J-protein BLAST alignment is shown below:

	Q03363	P08622	O34242	P25491	Q96EY1	Q27237	P25294	P25686	Q9UGP8	P32527	Q9H3Z4	P25685	O75953	P0A9X9
Q03363	**100, 100**	35,53	34,53	43,60	29,46	30,48	36,53	45,61	43,68	38,58	42,59	32,45	34,46	31,42
P08622	35,53	**100, 100**	71,82	35,50	35,55	32,51	56,74	45,60	40,61	37,53	51,73	27,47	59,80	41,66
O34242	34,53	71,82	**100, 100**	36,52	33,52	32,49	52,73	39,52	35,56	37,63	53,76	26,43	40,58	41,66
P25491	43,60	35,50	36,52	**100, 100**	30,47	27,40	64,78	50,70	37,48	43,56	55,69	31,45	29,44	27,50
Q96EY1	29,46	35,55	33,52	30,47	**100, 100**	50,66	46,68	46,61	30,50	27,47	47,69	29,44	30,43	24,51
Q27237	30,48	32,51	32,49	27,40	50,66	**100, 100**	49,67	46,62	31,52	34,49	46,64	30,46	57,74	38,72
P25294	36,53	56,74	52,73	64,78	46,68	49,67	**100, 100**	34,46	32,48	38,51	53,66	35,54	36,53	66,66
P25686	45,61	45,60	39,52	50,70	46,61	46,62	34,46	**100, 100**	31,50	35,55	54,77	52,69	47,61	50,75
Q9UGP8	43,68	40,61	35,56	37,48	30,50	31,52	32,48	31,50	**100, 100**	32,49	44,54	42,60	42,63	17,40
P32527	38,58	37,53	37,63	43,56	27,47	34,49	38,51	35,55	32,49	**100, 100**	42,52	34,57	37,53	
Q9H3Z4	42,59	51,73	53,76	55,69	47,69	46,64	53,66	54,77	44,54	42,52	**100, 100**	50,73	54,70	
P25685	32,45	27,47	26,43	31,45	29,44	30,46	35,54	52,69	42,60	34,57	50,73	**100, 100**	59,73	
O75953	34,46	59,80	40,58	29,44	30,43	57,74	36,53	47,61	42,63	37,53	54,70	59,73	**100, 100**	
P0A9X9	31,42	41,66	41,66	27,50	24,51	38,72	66,66	50,75	17,40					**100, 100**

Figure 4.4: Sequence alignment comparison of different J-proteins

It can be seen from this table that J-proteins share between 30-71 % identity and 40-82 % similarity. Interestingly enough, several J-proteins (especially belonging to families 1 & 2) also share substantial identity with cspA, an *E.coli* coldshock protein. Work by Liberek and Georgopoulos identified DnaJ as one of the proteins which play a role in DNA replication of λ-phage [75]. Subsequent work by McKenney and co-workers

indicated that DnaK and DnaJ mediate an alteration in the P1 initiator protein that makes it more receptive for oriP1 DNA binding during P1 plasmid replication. They suggested that DnaJ operates as a co-factor that enables ATP-hydrolysis dependant activation of RepA by DnaK, and went on to anticipate such relationships between Hsp70s and co-factors in Eukarya [76].

The human genome codes for 32 J proteins, many of which are expressed when under stressed conditions {Craig, 2006 #6593}.

4.2.4 BIOPHYSICAL ASSAYS OF J-FUNCTION

Gross and coworkers used Surface Plasmon Resonance (SPR) to determine the affinity of DnaK for DnaJ. They found that the full length DnaK is required to interact with DnaJ. These results are discussed shortly.

In vitro, the function of J-proteins can be measured as the enhancement of ATP hydrolysis of DnaK, by as much as a factor of 10, depending on conditions. This is the basis for much of the biochemical justification of the disulphide linked covalent adduct between Hsc70 NBD and the Auxilin J-Domain [94].

Early NMR studies by Greene and Landry[222] indicated that DnaJ (2-75) binds tightly to DnaK, with an apparent K_D of 10 μM and to the isolated DnaK NBD with an apparent K_D of 5μM.

4.2.5 STRUCTURAL STUDIES ON J-PROTEINS

The N-terminal J-domain fragment of J-proteins, or DnaJs (these two names will be used interchangeably hereon) is highly conserved. Cyr, Langer and Douglas[221] interpreted this as various J-proteins contain a highly conserved J-domain to interact seamlessly with Hsp70s. The C-terminal part of the J-proteins could have then undergone divergent evolution to participate in different processes ranging from plasmid replication to protein folding, transport and so on [223; 224; 225]. Wuthrich and co-workers were the first to meet with success in solving a J-domain protein structure

[79; 80]. It can be seen from this structure (see Fig 4.5) that the J-domain has 4 characteristic helices. Helix 1(red) runs from residue 6-10, helix 2(green) from 18-32, helix 3(blue) from 41-57 and helix 4(purple) from 61-68. Helices 2 and 3 form an antiparallel coiled-coil pair (Fig 4.5).

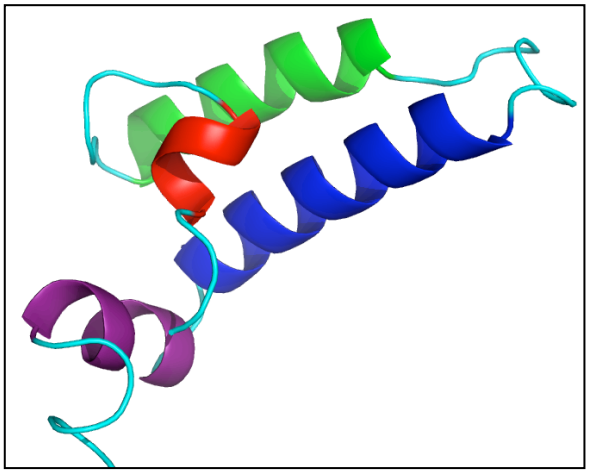

Figure 4.5: J-domain NMR structure solved by Wuthrich and co-workers

Helix 1(red) runs from residue 6-10, helix 2(green) from 18-32, helix 3(blue) from 41-57 and helix 4(purple) from 61-68.

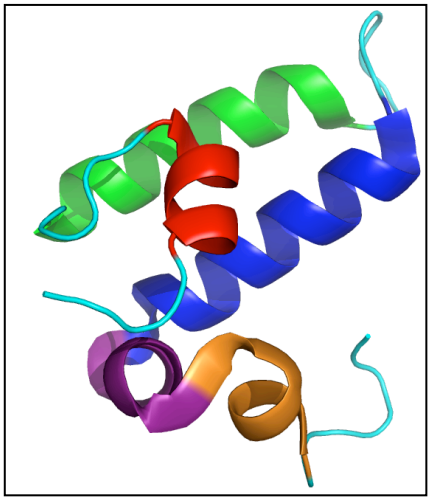

Figure 4.6: Crystal structure of S.cerevisaie DnaJ homolog J-domain

Helix 1(red) runs from residue 6-10, helix 2(green) from 18-32, helix 3(blue) from 41-57, helix 4(purple) from 58-68, and helix 5(orange) from 68-77.

The authors were however, unable to pin down the orientation of helix 4 w.r.t the

rest of the molecule. This highly α-dominated topology leads to a hydrophobic core to which all four helices contribute sidechains. The degree of sequence conservation in this region is also correlated to the structural conservation of the core.

More recently, Joachimiak and co-workers solved a crystal structure of J-domain from *S.cerevisiae* to a resolution of 1.25 Angstroms. This structure also has the same properties of the solution *E.coli* J-domain structure with a few important caveats: helix 4(purple) begins earlier than in the *E.coli* protein – at residue 58, and segues into a kinked helix 5(orange) at residue 68. This structure (2O37.pdb) is shown below in Fig 4.6. Wright & Dyson's group then solved a solution structure of the Cys rich region of *E.coli* DnaJ in 2000 [226]. This structure (Fig 4.7) has an overall V-shape with an extended β-hairpin topology. This domain of DnaJ contains four CXXCXGXG motifs, which constitute two zinc fingers.

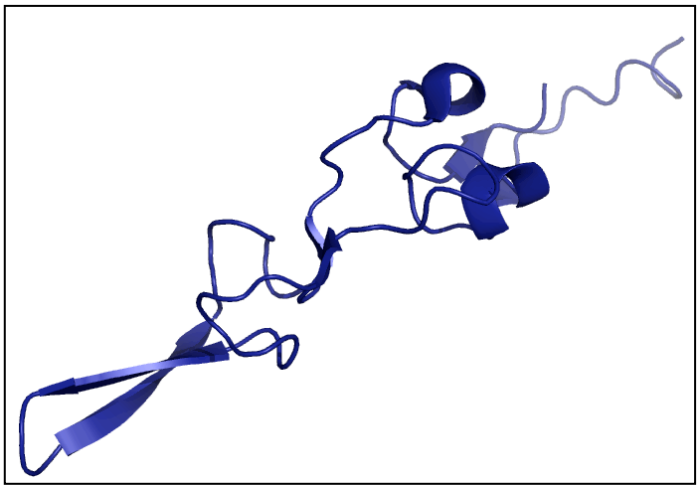

Figure 4.7: Solution structure of the Cys rich region from DnaJ

4.2.6 BIOLOGICAL ROLES OF J-PROTEINS

Georgopoulos and co-workers demonstrated that DnaK, DnaJ and GrpE function together as a complex machine in the replication of λ-DNA. Specifically, these proteins together disassemble the preprimosomal complex, consisting of λO-λB-DnaB-helicase at the origin of λ replication, hence liberating DnaB helicase and allowing it to unwind λ-DNA, thus leading to DNA replication [75; 82; 83]. They also determined that the following mutations would completely kill DnaJ activity (in terms of visible λ-growth at

30 Celsius) -

1. Deletion of a 34 residue N-terminal fragment
2. Asp insertion at residue 35

However, they also noticed that λ-growth was supported at 30 C (but not at 42 C) by the 108 residue N-terminal fragment of DnaJ. This is the J-domain with the Gly-rich region. Interestingly enough, this 108 residue construct of DnaJ (referred to as DnaJ12 by Georgopoulos) did not have the ability to dimerize, which is a property of the full length DnaJ [227]. Furthermore, the interaction with the σ32 peptide, also noted for the full length DnaJ [228] – is lost in the 108 residue J construct. Although DnaJ12 itself is stable at higher temperatures, it does not support λ or bacterial growth at 42 C. Based on this evidence, Geogopoulos and co-workers assigned putative dimerizing/substrate binding roles to the C-terminal fragment of DnaJ and Hsp70 interaction roles to the J-domain. Langer, *et.al*. found evidence that (full length) DnaJ/DnaK and GrpE together assist in the refolding of rhodanase, and that DnaJ in itself has some limited chaperone activity, in that it can bind to unfolded rhodanase and prevent its aggregation [229].

4.2.7 STUDYING THE J-DOMAIN ALONE VS. STUDYING FULL LENGTH DnaJ

Although it is not mentioned explicitly in these early studies, but evidence for a possible model of action is already beginning to become apparent: that DnaJ has two different, but interconnected biophysical roles: the first is as a co-factor to ATP hydrolysis by Hsp70, and the second is as a co-chaperone (there is weak evidence to classify this protein as a 'stand-alone' chaperone {Hendrick, 1993 #10820}) which binds to unfolded polypeptides and herds them to the Hsp70 system for processing. These two roles are likely not mutually independent. We already know from our studies[183; 186] of the Hsp70 system, as well as work by other groups[91; 160; 205; 230] that nucleotide turnover by the NBD is allosterically crosslinked to substrate loading/unloading from the SBD. Thus, a complete picture of the interaction of DnaJ and the Hsp70 system can ideally be gained from a study of the complete DnaJ

molecule and the complete DnaK molecule. Ideally, the Hsp70 functional cycle should contain all its constituent proteins *in vitro* (DnaK/DnaJ/GrpE). However, studying multiprotein complexes of this molecular weight at medium to high resolution with the help of available NMR is challenging enough. Naturally, this begs the question: have any advances been made in crystallizing complexes of DnaK/DnaJ. Before we address this question, it is pertinent to talk about yet another manifestation of the J-domain.

4.2.8 AUXILIN AND THE C-TERMINAL J-DOMAIN

As briefly discussed in Chapter 1, clathrin is a structural protein that is has a triskelion shape. It has three heavy chains and three light chains (about 1700 residues long). Interaction of clathrin triskelia leads to the formation of polyhedral lattices (spherical) that coat vesicles. Endocytosis or any other process, which involves coalescence of such vesicles with their target membranes, involves reengineering of the polyhedral lattices.

Hsp70s participate in these reengineering processes in association with another 86 kDa protein called Auxilin. This protein was first isolated by Ahle and Ungewickel [84]. They identified that auxilin binds to clathrin triskelia and promotes the formation of 50-100 nm cages. Subsequent studies by the Ungewickel group indicated that auxilin binds to clathrin lattices (during lattice disassembly leading up to membrane fusion processes) in association with Hsp70. This process is connected to ATP turnover. Disassembling the clathrin lattice was earlier supposed to depend on the clathrin light chain [231]. This study showed that clathrin disassembly depends on the proximal domain of the heavy chain of clathrin, Hsp70 (with ATP) and auxilin. The binding of Hsp70s to clathrin baskets depends crucially on a C-terminal fragment of auxilin. This C-terminal fragment of auxilin, runs from residue 846 to 910 and bears substantial homology with different J-domains (49 % to DnaJ, 62% to yeast SIS1). A truncated version of auxilin, terminating at Pro814 interacted with clathrin baskets, but failed to promote association of Hsp70s to these baskets and hence disassembly.

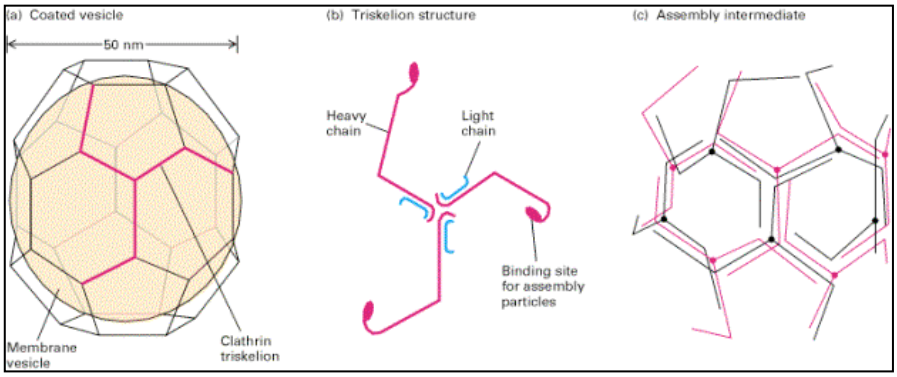

Figure 4.8: Clathrin triskelia and lattice assembly

Ref:

http://tainano.com/Molecular%20Biology%20Glossary.files/image050.gif

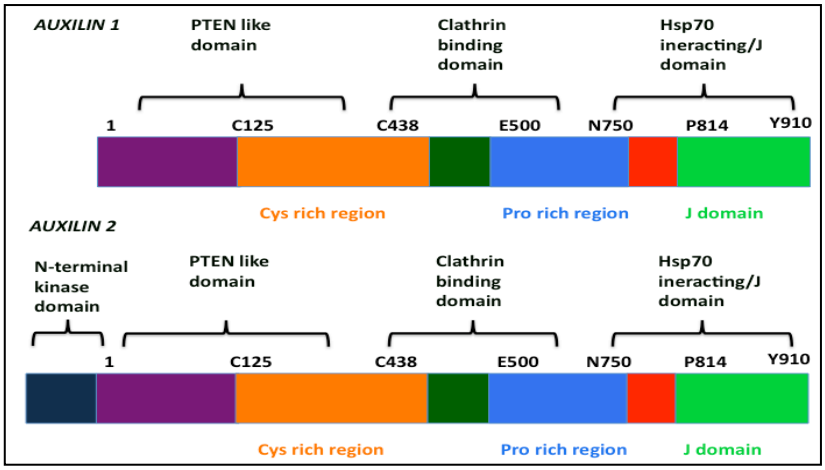

Figure 4.9: Auxilin 1 and Auxilin 2 sequence schematic

Residue count is shown for Auxilin 1

Delving further into the biochemistry of clathrin uncoating and auxilins, it is known that there are two distinct auxilins seen in *Mammalia*: auxilin 1 (general) and auxilin 2(brain specific [232]). Auxilin 1 has been discussed in the introduction as well as in this chapter. It contains an N-terminal PTEN like domain [233], a clathrin binding domain, and a C-terminal J-domain. In addition to this, Auxilin 2 also contains an N-terminal kinase domain [234; 235]. Subsequent cryo-electron microscopy studies by

Kirchhausen and co-workers has resulted in a 12 Angstrom low resolution structure [236] that shows Auxilin binding near the vertices of a clathrin lattice. Auxilin contacts multiple heavy chains and recruits Hsp70 to a critical neighbourhood of contacts. This destabilizes the clathrin lattice and leads to its disassembly. This study showed that the 39 kDa C-terminal fragment of auxlin 1 (residues 547-910) is sufficient to uncoat clathrin cages *in vitro*. This fragment comprises the clathrin binding domain and the J-domain of auxilin. The resultant structure is shown in Fig 4.10, where the different clathrin heavy chains are coloured in shades of red, and the auxilin J-domains (114 residues) in shades of blue. Again, this work drives us to the same conclusion, as do the studies by Ungewickell and co-workers (mentioned earlier): that the J-domain interacts with Hsp70 and preceding N-terminal fragments drive clathrin processing interactions.

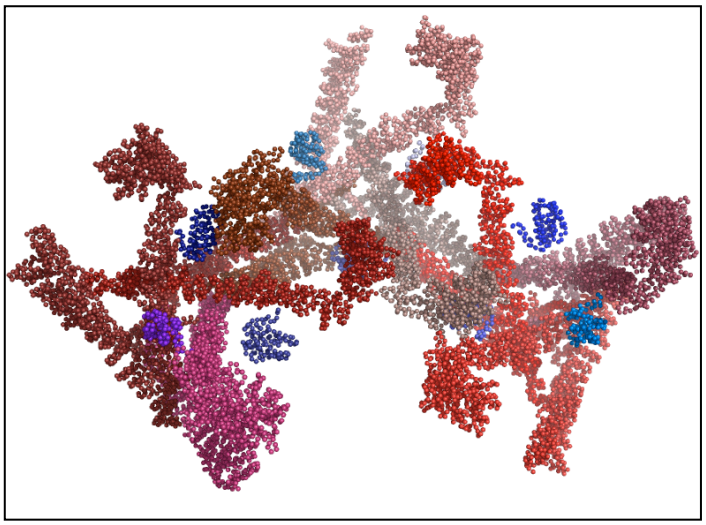

Figure 4.10: Cryo-electron microscopy derived low resolution structure of Clathrin heavy chain lattice with the Auxilin J-domain[236]

4.2.9 SEARCHING FOR CORRELATIONS BASED ON J-DOMAIN POSITION: TOPOLOGY-FUNCTION RELATIONSHIPS

These results [86; 236] are of profound importance: because the J-domain has been identified in a different protein family, with the topology-function character reversed. So far, while discussing J-proteins we have found that a highly conserved N-

terminal J-domain interacts with Hsp70, while a less conserved C-terminal fragment (indeed of varying length) has evolved to fulfill multiple roles. Here, auxilin features a C-terminal J-domain, which interacts with Hsp70, while the N-terminal fragment interacts with the 'true' substrate of the reaction (clathrin). In both cases, however, the J-domain always interacts with Hsp70. Thus, we are motivated to study the interaction of Hsp70 with J-domains, as this will shed light on both the nature of the complex interaction as well as (hopefully) better explain the nature of allostery in the Hsp70 functional cycle. Thus, it is pertinent to discuss efforts made by crystallographers in this direction.

4.2.10 PREVIOUS EFFORTS TO STUDY THE STRUCTURE OF Hsp70-J-DOMAIN COMPLEXES: THE HPD LOOP

Before we discuss the efforts made by structural biologists to explore the nature of Hsp70-J-Domain complexes, it is worthwhile to look at the biochemical evidence, which motivated such studies. Georgopoulos and co-workers did much of the original work in this direction. They first showed that modification of His33 to Gln results in a DnaJ mutant (called DnaJ259) that was unable to stimulate ATP hydrolysis by DnaK. A modified ELISA approach was used to test the interaction of this protein (and compare it with DnaJ Δ77-107, which is the J-domain, but not the Gly/Phe rich region) and DnaK [237]. His33 was identified as a residue essential to the interaction of the J-domain with DnaK. A similar residue was identified in the Simian virus 40-antigen oncoprotein (His42) which was found essential for λ-growth [238]. Taking this information is in the context that His33 is part of a highly conserved triad [83; 237] (His33, Pro34, Asp35) located in the loop between helices two and three of the J-domain – we are on our way to identifying details of a potential DnaJ-DnaK complex.

4.2.11 COMPLEMENTARY MUTATIONS IN DnaJ AND DnaK

Paralleling the studies by Georgopoulos' group, Suh, Gross and co-workers investigated the interaction of DnaJ and DnaK using Surface Plasmon Resonance (SPR). They found, firstly, that the HPD loop is indeed essential for DnaK interaction. Altering Asp35 to Asn abolished binding of full length DnaK to the J-domain. However,

they found that this lost binding could be recovered by counter-mutating Arg167 on DnaK to His or Ala. Their results are summarized in the table shown below, where dissociation constants (KD) are in nM:

DnaK mutant ⇒ DnaJ mutant ⇓	WT	R167H	R167A
WT	70	331	463
D35N	Negligible signal	167	163

Table 4.2: SPR results on interaction of various DnaK and DnaJ mutants

Two residues were identified on DnaK, close to R167: N170 and T173, which, if mutated to Ala; lead to severely, reduced DnaJ binding. Furthermore, several mutations were identified on the DnaK substrate-binding domain, which affect DnaJ binding. G400D and G539D reduced the affinity and D526N reduced the rate of both peptide and DnaJ binding to DnaK. Thus Gross and co-workers postulated the propagation of conformational changes throughout DnaK, upon the binding of DnaJ [95]. This was bolstered by their subsequent work, where the performed SPR interaction experiments with DnaJ and different constructs of DnaK. Here, they were able to show that ATP hydrolysis is required for the formation of a DnaK-DnaJ complex (complex formation is hindered in the presence of slowly hydrolyzed ATP analogues, such as ATP γS or AMPPNP). Their results are tabulated below:

DnaK construct ⇒	WT	1-628 (full length minus floppy tail)	1-538 (NBD+linker+SBD, but no lid)	1-403 (NBD + linker + part of SBD)	386-638 (SBD+lid)
Binds to DnaJ ⇒	Yes	Yes	Yes	No	No

Table 4.3: SPR results on binding of DnaJ to various fragments of DnaK

As the table (4.3) shows, DnaJ binding requires the presence of both the NBD and the SBD, but not the C-terminal floppy tail. The role of the 10 kDa α-helical tail near the C-terminal also appears to be relatively unimportant. In this work, Gross and co-authors also identified a conserved QKRAA motif (residues 61 to 65) on DnaJ, which appears to be important to DnaK binding, but not as crucial as the HPD loop [239].

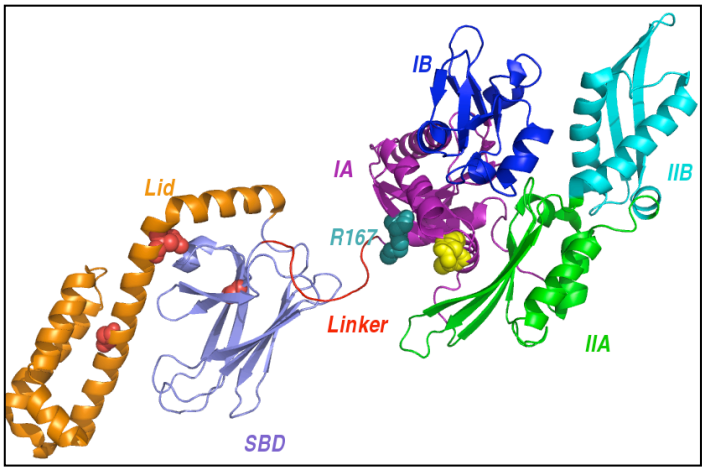

Figure 4.11: Residues on DnaK, which affect binding to DnaJ

Subdomain IA – purple, IB – blue, IIA – green, IIB – cyan, linker – red, SBD – slate, lid – orange, R167 – light blue spheres, N170 & T173 – yellow spheres, G400, G539, D526 – red spheres.

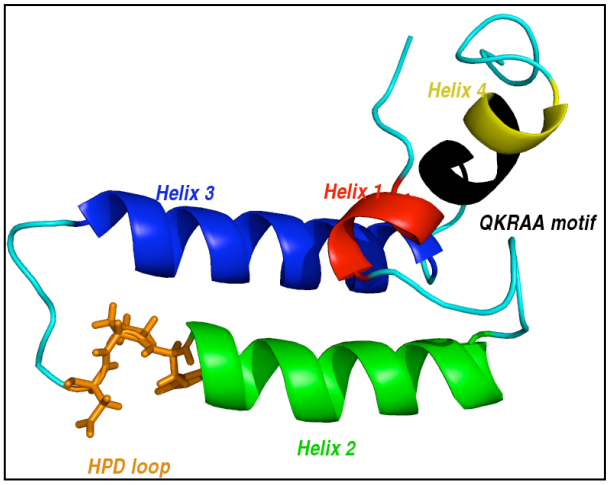

Figure 4.12: Residues on the J-domain critical for DnaK binding

Helix 1(red) runs from residue 6-10, helix 2(green) from 18-32, helix 3(blue) from 41-57 and helix 4(yellow) from 61-68. The HPD loop is shown in orange sticks and the QKRAA motif is shown in black.

4.3 CRYSTAL STRUCTURE OF A DISULPHIDE LINKED ADDUCT BETWEEN BOVINE Hsc70 NBD AND THE AUXILIN J-DOMAIN

Sousa and co-workers attempted to co-crystallize and solve structures of various NBD fragments of bovine Hsc70 and the Auxilin J-domain. While crystals were formed, no complex structures were seen in diffraction patterns. They mutated R171 on Hsc70 NBD (which corresponds to R167 in *E.coli* DnaK) and D876 on bovine auxilin J-domain

(which corresponds to D35 in *E.coli* J domain) to cysteines. Air oxidation between these two cys-mutant proteins led to the formation of 'disulphide linked complexes', which yielded workable diffraction patterns. This structure was published recently (Fig 4.13) [94].

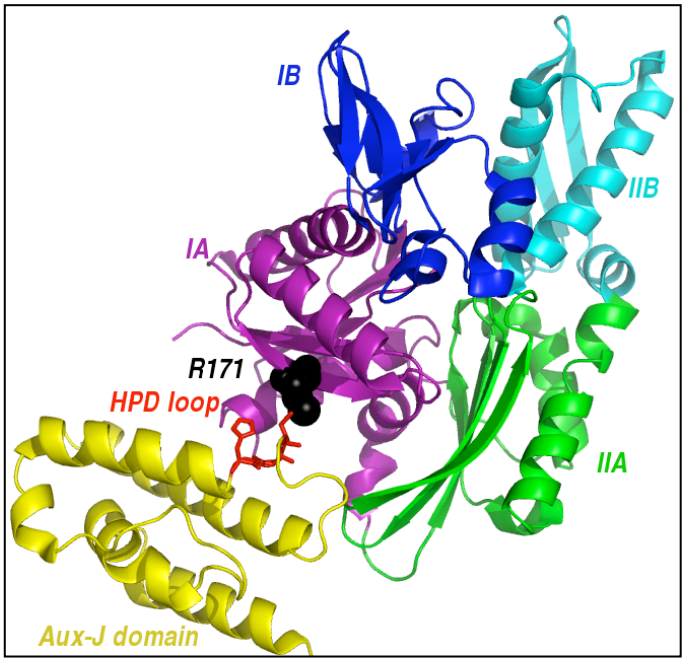

Figure 4.13: Disulphide linked 'complex' of Hsc70 NBD with the Auxilin J – domain

Subdomain IA – purple, IB – blue, IIA – green, IIB – cyan, Auxilin J-domain – yellow, HPD loop – red sticks, R171 (NBD) – black spheres

4.3.1 CRITICISM OF THE DISULPHIDE LINKED Hsc70 NBD – AUXILIN J DOMAIN STRUCTURE

Sousa and co-workers tested the efficiency of ATP hydrolysis with different constructs of the Hsc70 NBD with Auxilin J-domain either added in solution or disulphide bound [94]. Their results are summarized below (ATPase activity in relative units of /min):

Hsc70 construct (1 μM)	WT	1-554(excluding lid)	1-386(NBD)	1-394(NBD+linker)
Cys Mutant	No	No	No	No
ATPase activity	< 0.025	~0.0625	< 0.025	~0.38
ATPase activity with	~0.0625			

Peptide				
ATPase activity with J-domain (10 μM)	~0.13	***~0.375***	< 0.025	1.875

Table 4.4: ATPase rates of various bovine Hsc70 fragments and stimulation by the Auxilin J-domain

It is immediately apparent that the authors were able to see a five-fold enhancement upon addition of J-domain to Hsc70, but only when the linker was present, or with a larger construct of Hsc70 (including the SBD). This indicates that the NBD alone does not experience significant ATP hydrolysis enhancement due to the J-domain. Indeed, the isolated NBD does not even have any significant ATP turnover as measured by their assay. Interestingly enough, the NBD + linker construct has a much higher basal ATP hydrolysis rate than the NBD + SBD construct. This phenomenon is not adequately explained in this work. The peptide used in this study is ALLLSAPRRGAGKKC, as compared to NRLLLTG, the peptide used as a inhibitor of the SBD in the work of Hendrickson and Zhu [67], Stevens et al [159] and Bertelsen et al [92]. The information provided in this analysis is diverse, but also somewhat fragmentary. The natural conclusion would be that a complex of full length Hsc70 and the J-domain is the ideal subject of study; but crystallizing such a complex has already proven to be very difficult. The authors have also presented ATPase data pertinent to the disulphide-linked structures that they have solved. This is shown below (ATPase activity in relative units of /min):

Hsc70 construct (1 μM)	1-394(NBD+linker)	1-386(NBD)
Cys Mutant	R171C	R171C
ATPase activity without J-domain	~0.19	< 0.03
ATPase activity with J-domain (10 μM)	~0.19	
ATPase activity with J-domain (500 μM)	~0.65	< 0.03
ATPase activity with J-domain (D876C) cross linked to Hsc70 R171C	~1.87	< 0.03

Table 4.5: ATPase activity of different Hsc70 fragments with R171C mutation and stimulation by the Auxilin J-domain disulphide linkage

Summarizing the points on which we feel that this study is compromised:

1. There is sufficient evidence to indicate that the J-domain interacts with both the NBD and the SBD [87; 95; 239]. Hence, studying a complex of the NBD or NBD + linker with the J-domain will be of some limited utility, but does not present the complete picture.

2. The structure presented as a complex is a disulphide linked adduct of the Hsc70 NBD and the Auxilin J-domain. This is not a true complex at all.

3. It can be seen from the crystal structure that the J-domain cross-linked to a particular Hsc70 shares a substantial interaction surface with the adjacent Hsc70 NBD. In Fig 4.14, observe how the J-domain of complex 1 interacts with the Hsc70 NBD of complex 2. Thus, the orientation of the J-domain might be marred by crystal packing influences.

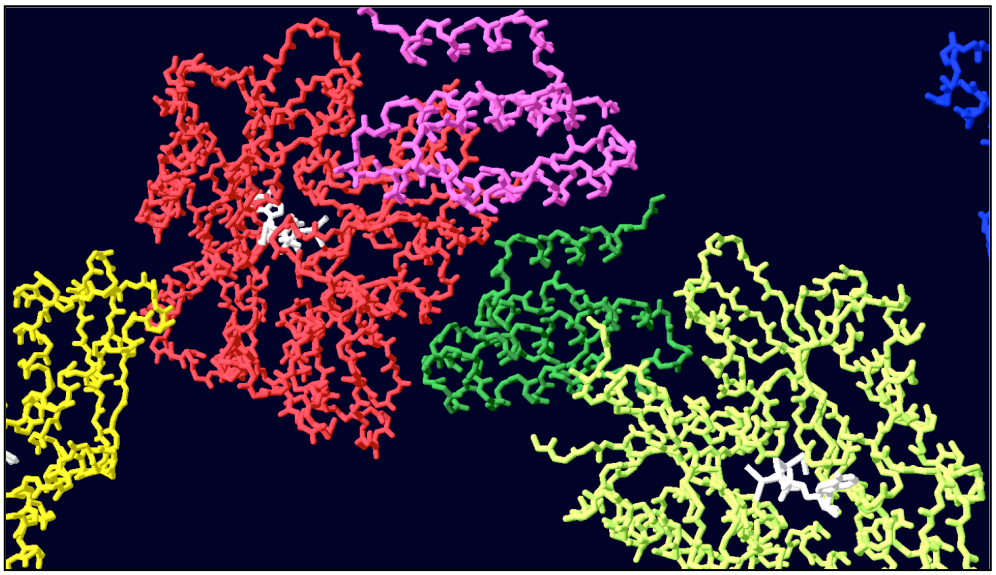

Figure 4.14: Disulphide linked structure of bovine Hsc70 NBD and the Auxilin J-domain showing interaction interfaces

Hsc70-NBD1 – light green, J-domain 1 – dark green, Hsc70-NBD1 – red, J-domain 1 – pink

4.4 A BRIEF DESCRIPTION OF THIS PROJECT

In this project, we have studied the interaction between *E.coli* DnaK with *E.coli* J-domain (1-108 residues). The DnaK construct used is 1-605 residues: this is the

complete wild-type molecule, excluding the last 33 residues which exists as a floppy tail and severely affects the hydrodynamic properties of this molecule, rendering it intractable to NMR analysis. The DnaJ construct used is the complete J-domain, residues 1-78, and a 20-residue extension comprising part of the Gly-Phe region, and has a His-tag at the C-terminus. We have used 2 dimensional NMR fingerprinting experiments (TROSY [240]) to study the effect of this J-domain on the NMR spectra of DnaK in the ADP.PO_4 /peptide state. Paramagnetic labels were attached to the J-domains by disulphide chemistry. NMR fingerprinting experiments were run on complexes of DnaK and J-domain in solution. The paramagnetic tag on DnaJ (otherwise NMR silent) induced changes in the characteristic spectra of DnaK. These changes were identified and analyzed in the context of the DnaK structure (solved by this group earlier [92]). This led to an interaction map of DnaK and the J-domain, which translates to a model for their complex interaction.

4.5 THEORY: PARAMAGNETIC LABELING AS A TOOL IN STRUCTURAL BIOLOGY

Unpaired electrons characterize paramagnetic tags. The electron magnetic moment is approximately 1850 times that of the proton. Thus, a paramagnetic centre in a protein or nucleic acid is a large magnet in the vicinity of many smaller nuclear magnetic moments. The fluctuating magnetic field created by electron spin T_1 processes and due to the molecular motion of the paramagnetic tag causes local field variations at neighbouring nuclear magnetic moments – which in turn manifests itself in more rapidly relaxing nuclear magnetic moments near the paramagnetic tag. This effect is known as Paramagnetic Relaxation Enhancement (PRE). The enhanced rate of relaxation will lead to a broader linewidth for NMR crosspeaks corresponding to residues, which have their intrinsic nuclear magnetic moments (i.e. $^{15}N^{1}H$ groups) close to the paramagnetic tag. This broader linewidth (or transverse relaxation rate) can be written as:

$$R_2^{PARA} = \frac{1}{15} S(S+1) \frac{\gamma^2 g^2 \beta^2}{r^6} \left(4\tau_C + \frac{3\tau_C}{1+\omega^2 \tau_C^2} \right)$$

Equation 4.1: Paramagnetic Relaxation Enhancement of linewidth

In this equation, S is the electron spin, γ is the proton gyromagnetic ratio, g is the electron g factor, β is the Bohr magneton, r is the distance between the electron and the nuclear spin, ω is the proton resonance frequency and τ_C is the correlation time of the dynamics of the vector between the electron and nuclear spins [134]. The correlation time τ_C can be broken down as follows:

$$\frac{1}{\tau_C} = \frac{1}{\tau_R} + \frac{1}{\tau_S} + \frac{1}{\tau_m}$$

Equation 4.2: Electron-proton correlation time

Here, τ_R is the rotational correlation time of the electron – nuclear vector. It is composed of two terms: the correlation time associated with the local motion of the spin label, and the overall rotational correlation time of the protein / protein-complex to which it is attached (typically > 20 ns, for this protein size ~ 65 kDa), τ_S is the electronic relaxation time, typically ~100 ns [134], and τ_m is the lifetime of the complex between of the paramagnetically labeled DnaJ and DnaK , which, at seconds, is much longer than either τ_R or τ_S and hence does not contribute to τ_C . These numbers are typical of nitroxide radicals. Thus, the actual correlation time is mostly dominated by protein tumbling, and leads to substantial increase in peak width. Hence, crosspeaks corresponding to residues close to the paramagnetic tag will be broadened. But we do not expect the peaks to show chemical shift changes: these occurs only by direct hyperfine coupling or by dipolar coupling when the electronic spin quantization axis is partially determined by the molecular frame (anisotropic g-tensor), neither of which is applicable for the systems under study.

There is another class of paramagnetic tags, namely transition metals (Co^{+2}, Ni^{+2}, Fe^{+3} and most lanthanides) which have anisotropic g-tensors and which have much

smaller τ_S (~10^{-13} sec). These species result in chemical shift changes for proximal crosspeak residues, but not necessarily linebroadening. The chemical shifts are proportional to 1/r3, as opposed to linebroadening, which is proportional to $1/r^6$.

The spin label that we have chosen is MTSL (S-(2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate), which is a nitroxide label.

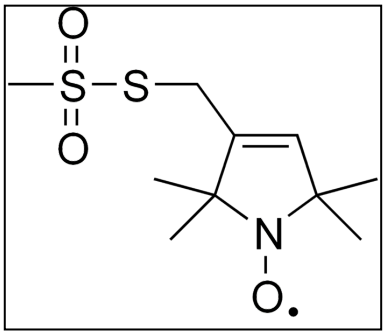

Figure 4.15: MTSL chemical structure

This label attaches to Cysteine residues by disulphide chemistry, and can be displaced by high concentrations of Dithiothreitol (DTT). It can be reduced by the addition of ascorbic acid, which turns MTSL into a diamagnetic species (as it no longer has a single unpaired electron). Hence MTSL, which has been treated with ascorbic acid no longer leads to line broadening in NMR spectra.

4.6 MATERIALS AND METHODS

4.6.1 PREPARATION OF DnaK

We first attempted to express protein using the plasmid for DnaK1-605 in *E.coli* (BL21 DE3 cells). Transformation was done according to the Strategene competent cells protocol. The transformed cell culture was grown in M9 minimal media (with $^{15}NH_4Cl$) at 37 C with shaking at 250 rpm for aeration, and induced at OD_{600} = 0.6 with 0.7 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were pelletted by spinning down at 35000 rpm, and then lysed in a microfluidizer. Protease inhibitor cocktail tablets (Biocompare) were added to prevent enzymatic degradation.

Purification was in three stages:

1. Passage through Fast Flo QTM (GE Healthcare Lifesciences). This is an ion exchange column (quaternary ammonium ion): the protein was eluted by a KCl gradient (max salt concentration: 0.5 M).

2. Dialysis into a low salt TRIS/KCl/MgCl2 based buffer at pH 7.4.

3. Passage through ATP agarose column. Protein was eluted rapidly by a step-function of high ADP (20 mM).

The problem with this protocol was that the yields were somewhat low (< 4mg of DnaK from 1 l of M9 media), and the protein was highly susceptible to lysis, especially at the floppy linker. Hence, we decided to try an alternate strategy for purification.

Site directed mutagenesis (Stratagene Quikchange^{TM}) was used to clone a six-residue C-terminal Histidine tag into the pET22b vector (Novagen^{TM}) for DnaK 1-605. Armed with this His-tag, we now used the following protocol for purification after cell lysis and pelletting:

1. Passage through a Ni-NTA column. Elution by a step function of high Imidazole (0.3 M).

2. Dialysis into a low salt TRIS/KCl/MgCl2 /PMSF/EDTA based buffer at pH 7.4.

3. Passage through a Sephacryl S-100 (GE Healthcare Lifesciences TM) size exclusion column. Protein was eluted slowly by a TRIS/KCl/MgCl2 /PMSF/EDTA based buffer at pH 7.4 which also contained high ADP (20 mM).

This protocol yielded DnaK of high purity, suitable for NMR.

4.6.2 PREPARATION OF DnaJ

Dr. Eric Bertelsen (former postdoc in the Zuiderweg lab) had cloned DnaJ J-domain (1-108 residues) into pET22b vector with an N-terminal six residue Histidine tag. A library of cysteine mutants of DnaJ 1-108 His-tag were then made using site directed mutagenesis (Stratagene QuikchangeTM). These Cys-mutants were then transformed into *E.coli* (BL21 DE3 cells). Transformed cell cultures were grown at 37 C

in LB media [241] with shaking @ 250 rpm and induced with IPTG at OD600 = 0.65. Cells were harvested and spun down at 35000 rpm. Cell pellets were lysed in a microfluidiser and loaded on a Ni-NTA column. Purification was as described above for DnaK, except that size exclusion chromatography was not required. Thus, DnaJ, as prepared is NMR silent, but is suitable for tagging with a spin label via disulphide chemistry. This will then affect interface residues on DnaK (which is NMR active) in a complex interaction, and lead to substantial line broadening.

4.6.3 RESIDUES CHOSEN FOR CYSTEINE MUTATION

We chose residues that gave us a broad spatial distribution of points over DnaJ, while not choosing residues belonging to the conserved HPD loop[220; 237; 242] or the QKRAA motif [95; 239]. We were also careful to pick residues that belong to elements of defined secondary structure. The points picked are:

Residue	location
R19	Helix 2
M30	Helix 2
K31	Helix 3
K41	Helix 3

Table 4.6: Residues on DnaJ J-domain chosen for Cys mutations

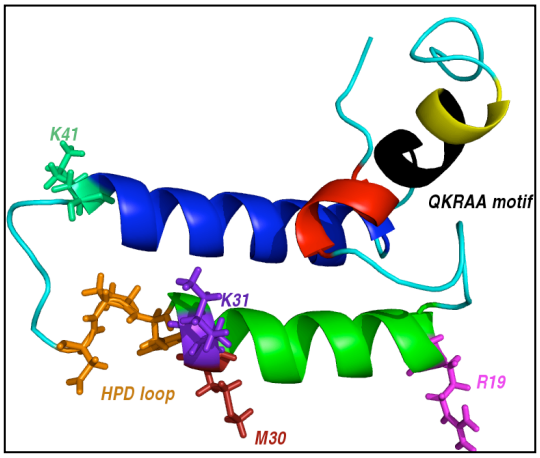

Figure 4.16: DnaJ J-domain showing residues chosen for Cysteine mutations

4.6.4 MTSL-TAGGING OF DnaJ

This was performed in stages, based on the fact the DnaJ has a His-tag. The protocol is described below:

1. Add (2 mM) DTT to 1ml of DnaJ (Cys mutant, concentration of protein – 400 μM).

2. Incubate at room temperature for 30 minutes with rocking.

3. Add 1 ml of homogenized Ni-NTA resin.

4. Incubate at room temperature for 10 minutes with rocking.

5. Spin down at 4000 g for 5 minutes.

6. Extract supernatant (1ml), set aside.

7. Wash with 1 ml of TRIS buffer, spin down, and extract supernatant. Repeat twice.

7. Add MTSL (2 mM), incubate at room temperature for 30 minutes with rocking.

8. Wash with 1 ml of TRIS buffer, spin down, and extract supernatant. Repeat twice.

9. Add TRIS buffer with 0.3 M Imidazole; incubate at room temperature for 10 minutes with rocking.

10. Extract supernatant. This is DnaJ-labeled with MTSL.

11. Dialyze into TRIS buffer to remove Imidazole.

4.6.5 NMR EXPERIMENTS

NMR samples were made with 100 μM of DnaK in a custom buffer containing 25 mM TRIS, 10 mM KCl, 15 mM $MgCl_2$, 5 mM EDTA, 5 μM PMSF and 20 mM ADP/Pi at pH 7.4. HSQC-TROSY [240] experiments were run at 30 C for 12 hrs (in 4 hr experiments) per sample on a Varian Inova 800 spectrometer equipped with a cold probe. The NMR data was processed with NMRPipe [212], and then converted to Sparky format [213]. A peak list based on assignments by Bertelsen and Zuiderweg [92] was used to assign peaks. It was verified that the addition of a C-terminal His-tag to the protein did not result in any significant chemical shift changes.

DnaJ R19C-MTSL was titrated in (equimolar to DnaK) in the same buffer (NMR Buffer A). HSQC-TROSY spectra were recorded for another 12 hours. Immediately

after this round of experiments, Ascorbic Acid was titrated in (to 750 μM) and 12 hours of HSQC-TROSY data was collected. The final round of experiments involved adding DTT (1 mM) to the NMR sample (DnaK + DnaJ-MTSL + Ascorbic Acid) and collecting another 12 hours of data. This sequence was repeated with all DnaJ mutants.

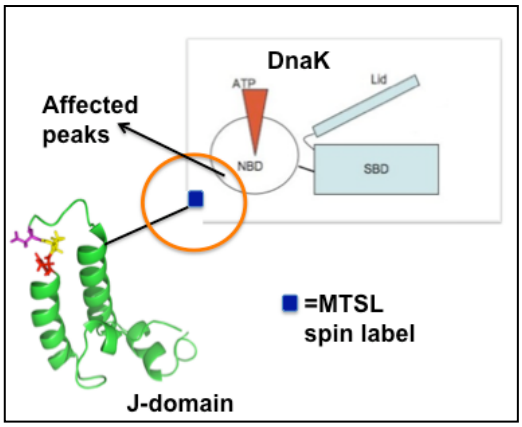

Figure 4.17: The effect of MTSL on neighbouring residues

Sparky overlays were generated with DnaK only, DnaK + DnaJ-MTSL, after addition of ascorbic acid, and after addition of DTT. Peaks that disappear upon addition of DnaJ-MTSL, but reappear upon addition of Ascorbic acid/DTT were listed. Residues corresponding to these peaks were identified on the DnaK-ADP.Pi structure. This provides the locus of interaction of DnaK and the J-domain. The characteristic signature of peaks that corresponds to possible interaction points is described below:

Experiment ⇒	DnaK only	DnaK + DnaJ-MTSL	DnaK + DnaJ-MTSL + ascorbic acid	DnaK + DnaJ-MTSL + ascorbic acid
Peak behaviour ⇒	Visible	Attenuated/invisible	Reappears	Remains

Table 4.7: Peak behaviour in NMR experiments

At each stage of experiments in Table 4.7, 5ul aliquots of sample were extracted from the NMR tube to run SDS-page gels. These gels indicated that the DnaK was intact (not subject to lysis at the linker), and the DnaJ did not dimerize (always a possibility with surface exposed Cys) at any stage in the experimental sequence. The latter was also an indication that the MTSL tag remained bound to DnaJ throughout.

4.6.6 BLANK TESTS WITH SPIN LABEL AND DnaK

We also tested the interaction of MTSL spin label with DnaK under the same conditions as all other NMR experiments described. No significant non-specific binding of MTSL to DnaK occurred, which would be discernable as line broadening of certain peaks.

4.7 RESULTS

The following peaks were recorded as displaying the expected signature (Table 4.7) from different DnaJ mutants:

DnaJ mutant	Peaks showing expected signature
R19C	T403, L411 and E430
M30C	N187, T189, I190, D211, G212, F216, E217, V218, L312, S330 and V331
K31C	none
K41C	A41, F67 and W102

Table 4.8: Peaks showing expected signature upon DnaJ-mtsl addition

Typical spectra for affected peaks is shown below:

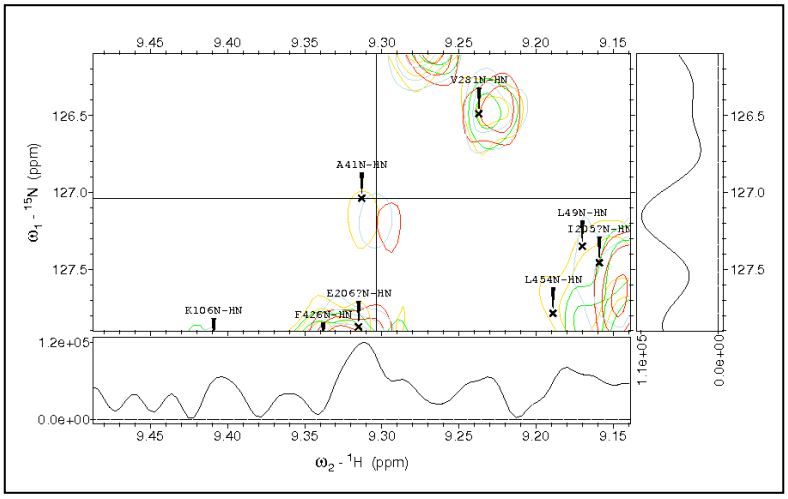

Figure 4.18: Sample spectra (peak A41) for DnaK ADP.Pi complexed with DnaJ K41C-MTSL

DnaK only – gold, DnaK + DnaJ K41C-mtsl – green, DnaJ K41C-mtsl + Ascorbic acid – red, DnaJ K41C-mtsl + Ascorbic acid + DTT – light blue

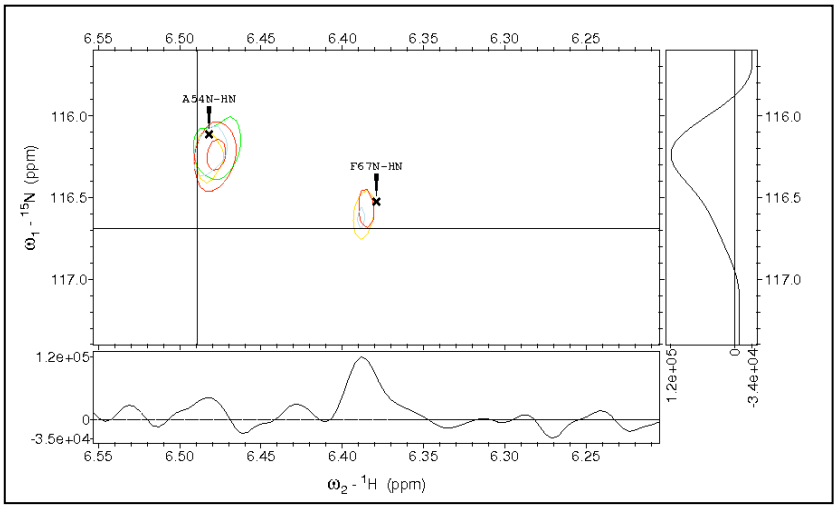

Figure 4.19: Sample spectra (peak F67) for DnaK ADP.Pi complexed with DnaJ K41C-MTSL

4.8 ASSEMBLING THE COMPLEX STRUCTURE

To assemble the structure, we used the affected residues as reference points, along with the additional constraint that the residue His33 on the J-domain from the HPD loop is also close to residue Arg167 on DnaK [95; 239]. We have generated a model for the complex structure, which has not yet been subject to energy minimization by a molecular dynamics approach. This model is shown below:

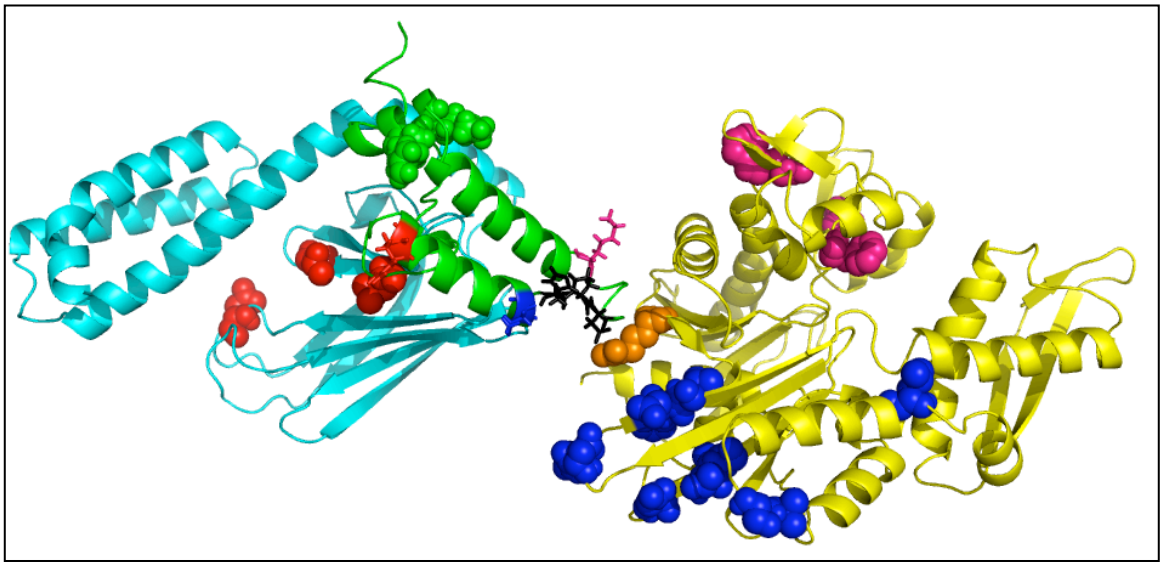

Figure 4.20: Model for the complex interaction of DnaK.ADP.Pi with DnaJ J-domain

R19 – red, M30 – blue, K41 – pink, HPD loop – black sticks, R167(DnaK) – orange spheres, NBD – yellow, SBD – cyan, J-domain – green

4.9 ATP HYDROLYSIS ASSAYS

In order to verify that the DnaJ mutants were functional, as attested to by stimulating ATP hydrolysis, we used a Malachite Green Assay. This series of experiments was carried out in consultation with Prof. Jason Gestwicki and Ms. Lyra Chang at the UM Life Sciences Institute. This is a phosphate release colorimetric assay based on the change of absorbance by Malachite Green working reagent (MGWR) at 620 nm. The colour change occurs upon complex formation when MGWR binds to phosphomolybdate under acidic conditions [243; 244]. The working reagent was made as described below.

The ingredients for the working reagent are:

1. Malachite green (0.0812% w/v)
2. Polyvinyl alcohol (2.32% w/v)
3. Ammonium heptamolybdate tetrahydrate (5.72% w/v)

Reagents 1,2, and 3 are mixed in a ratio of 2:1:1 with 2 parts of water. MG working reagent was always prepared fresh for assays. A calibration curve of absorbance vs. inorganic potassium phosphate (KH_2PO_4) concentration is also shown below:

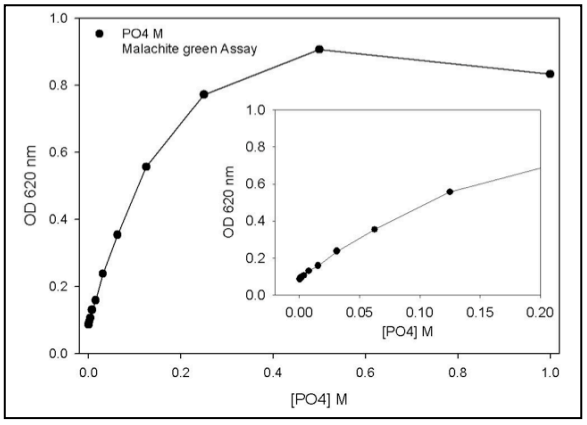

Figure 4.21: Malachite green absorbance calibration curve with potassium phosphate

Experiments were set up on 96 well plates. The protocol was as follows:

We dialyzed all proteins into a custom MG buffer – 100 mM TRIS, 20 mM KCl, 6mM $MgCl_2$, 5 μM PMSF, pH 7.4. EDTA was removed as it interferes with the assay. The absence of any nucleotide/inorganic phosphate in the buffer could potentially lead to instability problems for DnaK, but we observed that at the assay concentrations, this did not occur.

We tested a variety of assay conditions before picking the following set: 25 μl total reaction volume, 1 μM DnaK, 3 μM DnaJ, 0.25 mM ATP.

1. Add reaction ingredients in the following order: MG Buffer, ATP, DnaJ, and DnaK. The addition of DnaK starts the hydrolysis process.
2. Shake to mix, incubate at 37 C.
3. Add 80 μl of MG working reagent.
4. Immediately add 10 μl of sodium citrate (30% w/v) to quench the hydrolysis reaction. Shake to mix.
5. Incubate at 37 C for 15 minutes. Readout.

The result of this assay for the wild-type proteins is shown in Fig 4.22. The initial linear range of the graph (typically up to 3 hrs incubation time) was used to extract hydrolysis rates from slopes, shown in Figure 4.23.

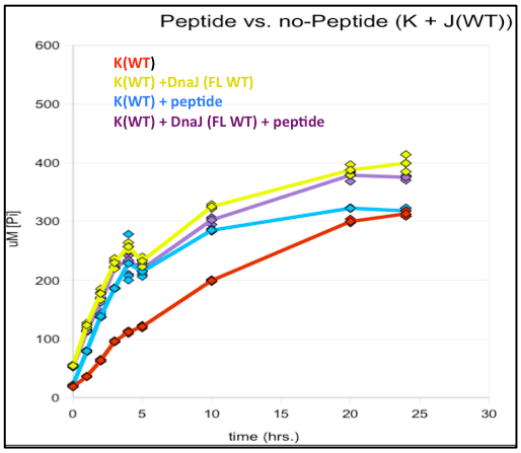

Figure 4.22: Stimulation of DnaK638WT (full length) by DnaJ (WT-FL), modulated by peptide

We expanded the scope of our assays to determine the role played by substrate peptide itself in the stimulation of ATP hydrolysis. We were particularly interested in this because our PRE-NMR experiments were carried out in the presence of peptide, to assure that the J-domain, its Gly-rich tail (residues 72-108) or the MSTL spinlabel could not bind into DnaK's substrate binding cleft. Such an interaction would seem non-physiological to us, given the function of DnaJ to recruit and deliver substrate to DnaK's substrate binding cleft. To the best of our knowledge, no DnaK/DnaJ ATP hydrolysis stimulation studies reported in the literature have been carried out in the presence of substrate peptide.

As Figures 4.22 and 4.23 show, the effect of peptide and DnaJ on the DnaK hydrolysis rate, approximately a 2.5-fold enhancement. The effect for both peptide and DnaJ is nearly identical and not additive. This suggests, but does not prove, that the enhancement is achieved by a similar mechanism for both interactions (see Discussion below).

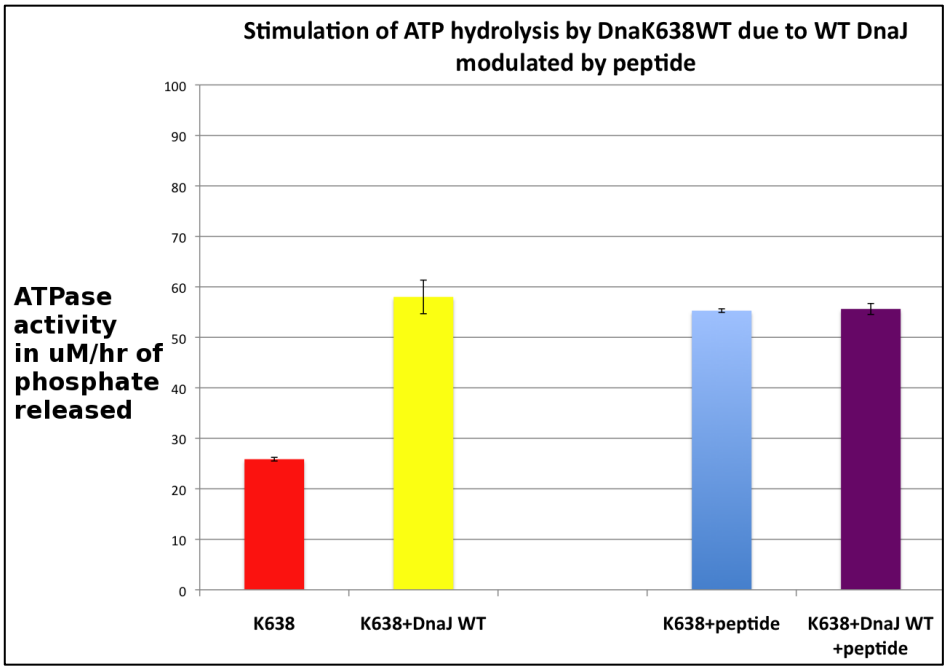

Figure 4.23: Stimulation of ATP hydrolysis rates of DnaK638 (WT-FL) by DnaJ (WT-FL), modulated by peptide

We tested the stimulation of WT DnaK (1-638, no His tag – sourced from Prof. Gestwicki) with DnaJ 1-108 M30C (with and without MTSL tag) and WT DnaJ (full length). These results are shown in Figs 4.24-4.27.

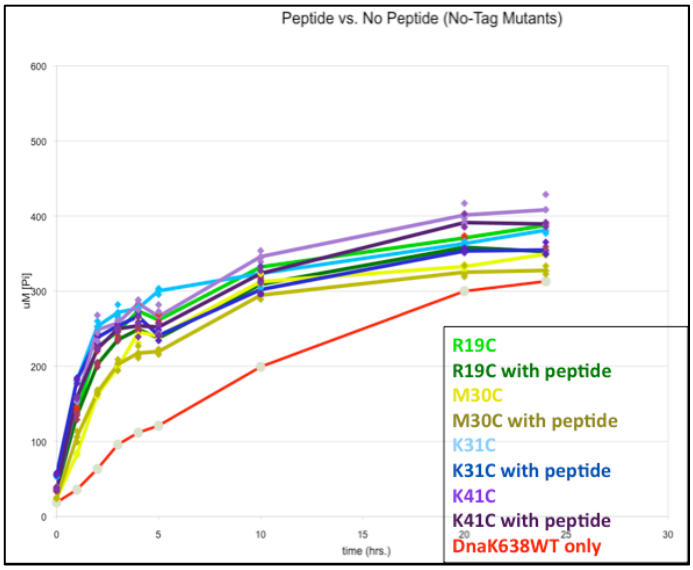

Figure 4.24: Stimulation of DnaK638(WT-FL) by untagged J-domain (1-108) mutants modulated by peptide

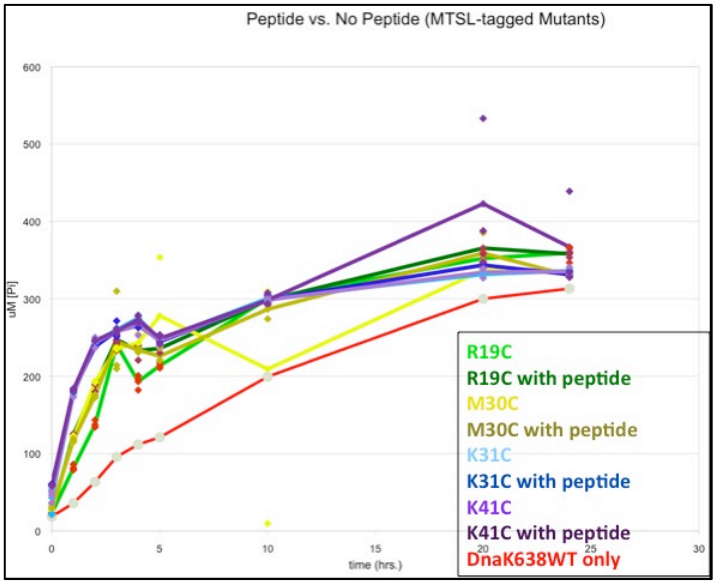

Figure 4.25: Stimulation of DnaK638(Wt-FL) by MTSL tagged J-domain (1-108) mutants modulated by peptide

As before, we have used the first three hours of data to extract slopes, which indicate the rate of ATP hydrolysis speedup by DnaK. These slopes are shown below:

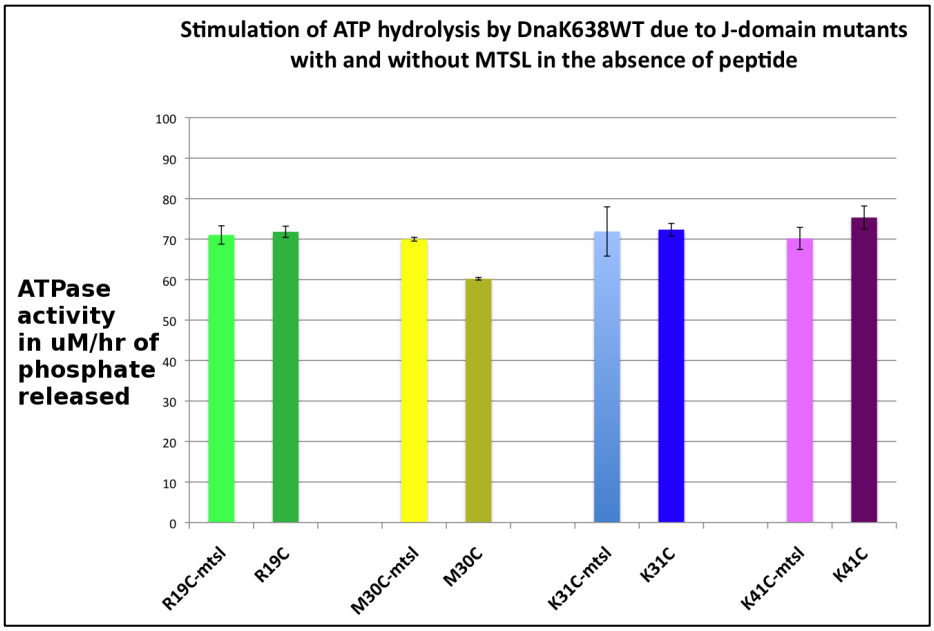

Figure 4.26: Stimulation of ATP hydrolysis rates of DnaK (WT-FL) by J-domain (1-108) mutants with and without MTSL tags in the absence of peptide

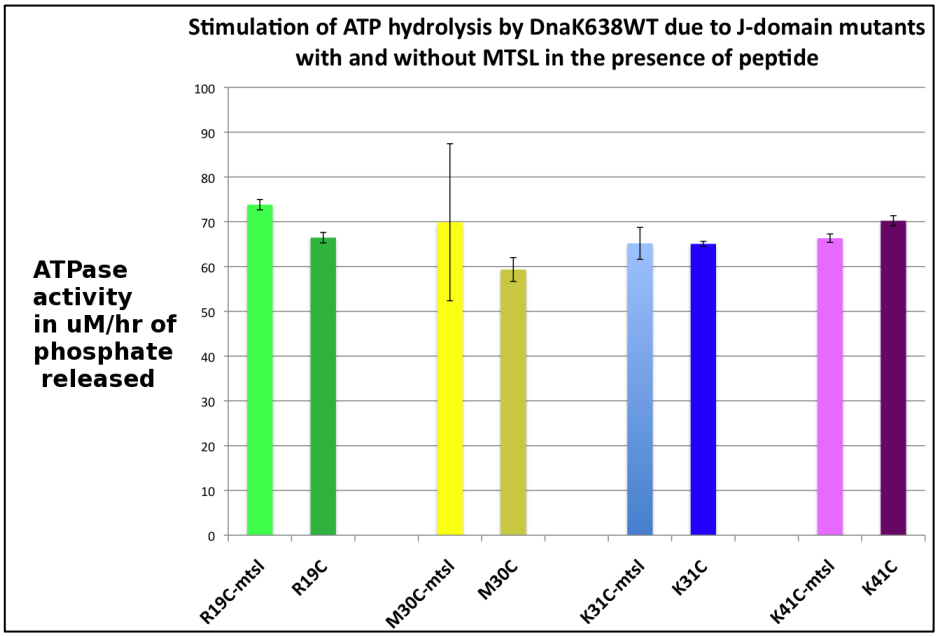

Figure 4.27: Stimulation of ATP hydrolysis rates of DnaK (WT-FL) by J-domain (1-108) mutants with and without MTSL tags in the presence of peptide

This set of experiments leads us to conclude, that the different DnaJ constructs that we have used are responsible for stimulation of DnaK's ATPase activity to more or less the same extent as WT-FL DnaJ (about 2-2.5 X). This is true for both the DnaJs

tagged with MTSL as well as untagged DnaJs. Thus, our J-domain mutants are biologically active. The deeper implications of the control experiments with peptide bound are discussed shortly.

4.10 DISCUSSION

The resultant complex model structure that we have arrived at does bear some similarities to the disulphide linked crystal structure published by Sousa and co-workers [94]. To compare and contrast the two structures, shown below is an annotated model of the disulphide-linked structure. The following residues on different species are equivalent:

Residues on E.coli DnaJ J-domain	H33	P34	D35	R19	M30	K41
Residues on bovine Auxilin J-domain	H874	P875	D876	Q860	L871	P882
Residues on E.coli DnaK	R167					
Residues on bovine Hsc70 NBD	R171					

Table 4.9: Equivalent residues on bovine Hsc70 NBD with Auxilin J domain compared to E.coli DnaK and J-domain

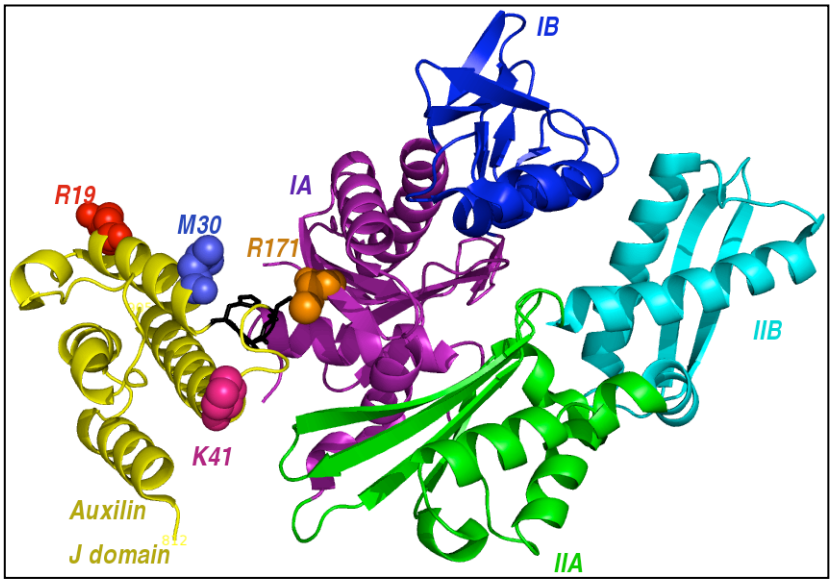

Figure 4.28: Annotated model of the disulphide linked 'complex' between Hsc70 NBD and the Auxilin J domain

R19 equivalent on Auxilin J-domain – red, M30 equivalent – blue, K41 equivalent – pink, HPD loop – black sticks, R171 (Hsc70) – orange spheres, Hsc70 NBD

Subdomain IA – purple, IB – blue, IIA – green, IIB – cyan, Auxilin J-domain – yellow [94]

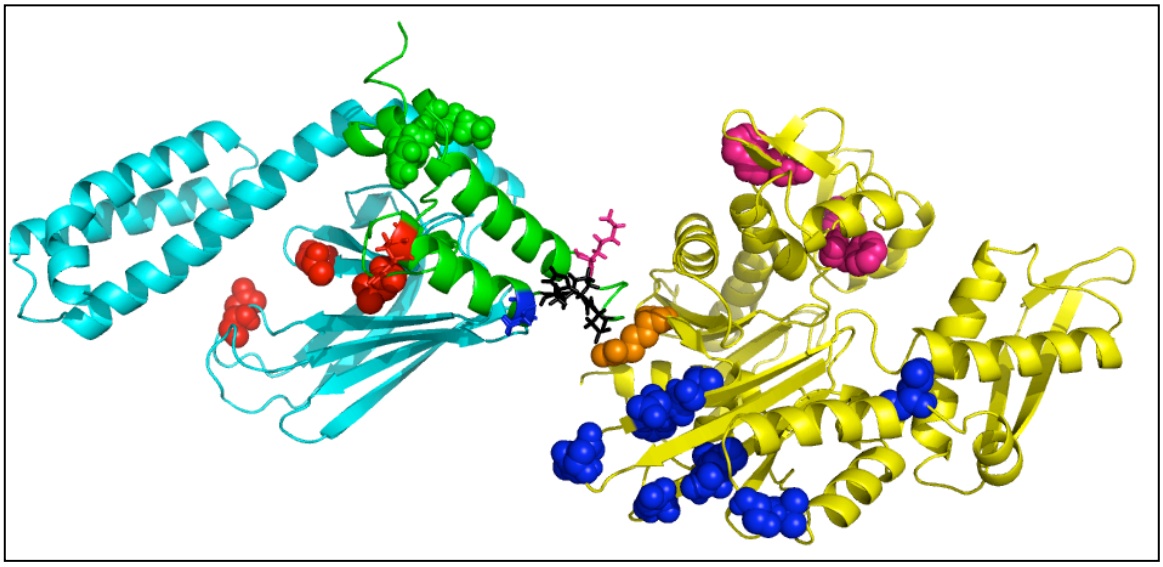

Figure 4.29: DnaK - DnaJ complex interaction model based on paramagnetic labeling NMR

R19 – red, M30 – blue, K41 – pink, HPD loop – black sticks, R167(DnaK) – orange spheres, NBD – yellow, SBD – cyan, J-domain – green

It can be seen from a comparison of our result (Fig 4.34/35) with the crystal structure of the disulphide-linked complex (Fig 4.33) that the overall placement of the J-domain w.r.t the NBD is more or less the same. This is unsurprising, given the evidence from Surface Plasmon resonance experiments carried out by Gross and co-workers [95; 239]. Indeed, those very experiments were the motivation for the disulphide linking strategy adopted by Sousa and co-workers.

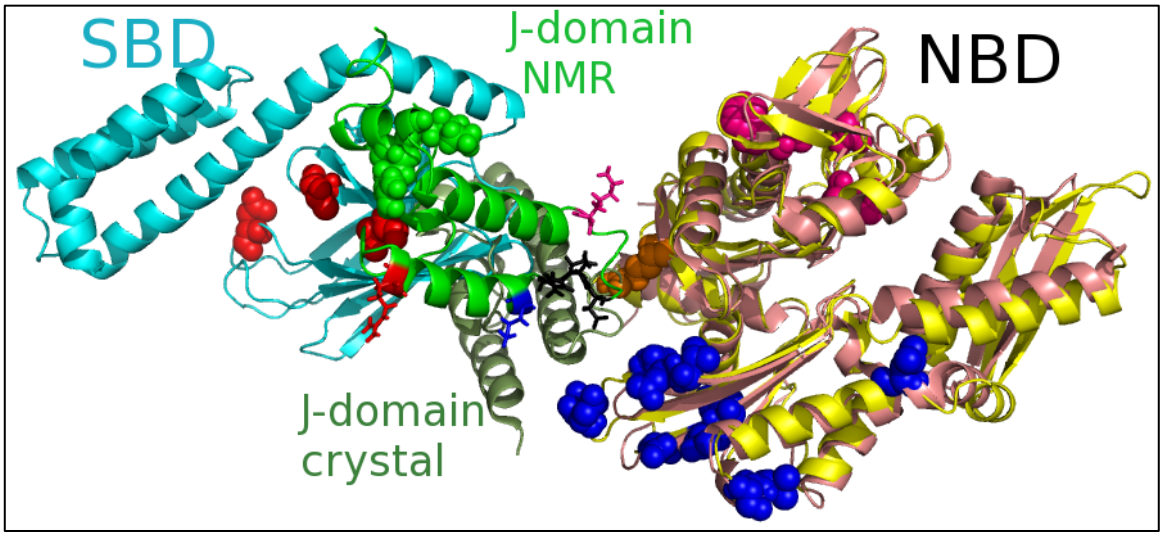

Figure 4.30: Peptide backbone overlay of NMR based interaction complex of DnaK-J-domain and the X-ray structure of the disulphide linked adduct of Hsc70 NBD the Auxilin J-domain

However, the relative orientations of the two J-domains are different by almost 90 degrees as is shown in Figure 4.30/31. In the crystal structure, the Auxilin J-domain (dark green) is oriented with its long axis (defined by helices 2 and 3) pointed behind and to the left of subdomain IA (as shown in this overlay). In our model, K41 on the J-domain is closer to NBD subdomain IB as compared to M30. In the crystal structure, we have the exact opposite situation. Thus, while the X-ray structure has one 'tether-point' between Hsc70 NBD and the J-domain, being defined by the disulphide bond between R171 (DnaK) and D876 (J domain), the similarity of the J-domain/NBD orientation with our model does not lend itself to further interpretation.

4.11 A POSSIBLE MODEL FOR ALLOSTERY IN THE Hsp70 SYSTEM

It has not escaped our notice that the model of DnaK/J-domain that we propose also sheds further light on the mechanism of interdomain allostery in the Hsp70 system. As we have already discussed in the results presented in Chapters 2 and 3, the ADP.Pi state[92] of this protein is characterized by the NBD and the SBD being connected by a loose, flexible linker. There are indications[93] that the ATP state of

DnaK has a structured linker[91] which docks in the inter-sub-domain interface between NBD subdomains IA and IIA [203]. This allows for the SBD to be 'reeled' in and dock with the NBD. From an inspection of the model presented here in Fig 4.30, it is seen that the J-domain lies transversely across the NBD and the SBD. It makes contacts with both domains of DnaK. This is in agreement with the SPR results where it was shown that the isolated DnaK NBD or SBD do not bind to DnaJ [95; 239]. We are also aware that ATP hydrolysis is required for DnaJ based activation of DnaK to bind to denatured/native protein surfaces for processing [220]. Thus, it seems plausible that the interaction of the J-domain (or rather, DnaJ) with DnaK mediates an undocking of the NBD and the SBD. Or in more concrete terms, it is possible that the J-domain pushes the NBD and the SBD apart, releases the linker from the NDB, allows larger flexibility in the IA / IIA interface and which stimulates ATP hydrolysis.

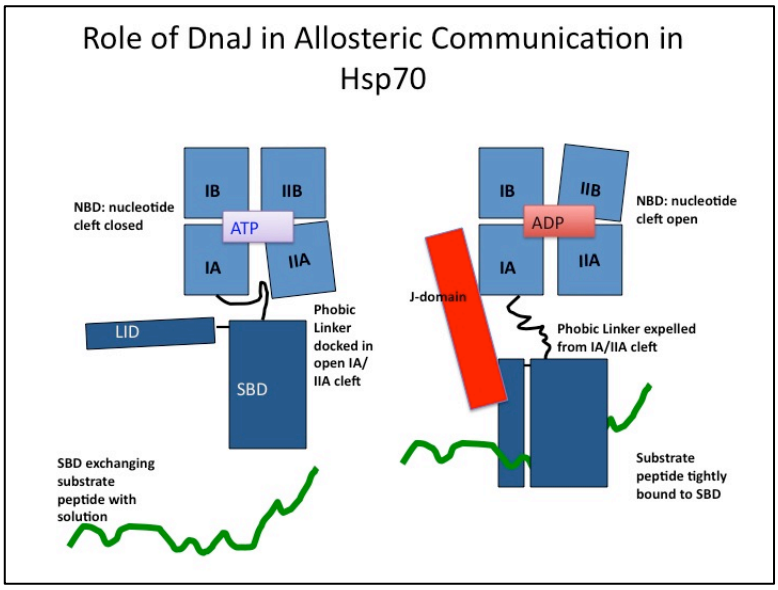

Figure 4.31: The role of the J-domain in allosteric behaviour in the Hsp70 system

The role played by the substrate peptide is also of great import. It is known that in the absence of substrate peptide, the α-helical lid domain is capable of becoming unstructured (especially in a truncated construct of the SBD from 383 to 540 residues, mammalian count [66]) and bind to the substrate binding pocket. In that sense, the substrate binding pocket in Hsp70s is somewhat nondiscriminating [220]. The

bidirectional allostery seen in Hsp70s is inferred from the stimulation of ATPase activity by addition of peptide. We surmise that while DnaJ performs the dual roles of loading substrate peptide onto the SBD and stimulating ATPase activity, these roles cannot necessarily be decoupled. Peptide binding induces the same undocking of the NBD and SBD that the addition of DnaJ is associated with. Further studies with other substrate peptides is needed to shed more light on any possibly competitive interaction of DnaK and peptide vs. DnaK and DnaJ.

4.12 LIMITATIONS OF THIS ANALYSIS

1. This is not a high-resolution structure as is usually determined by NOEs in NMR. Thus, this model should be considered just that: a model, similar to our RDC based eigenstructures of the TTh-NBD as reported in Chapters 2 and 3 [203].

2. There were three residues (R19, M30 and K41) on the J-domain where attached MTSL tags showed changes in the DnaK spectrum. The fourth residue tested (K31) did not show any changes. We have interpreted this result in terms of the K31 sidechain being oriented away from DnaK. As such, three points in space (along with the H33-R167 co-mutation result [95; 239]) are sufficient to orient the J-domain w.r.t DnaK. But further mutations on DnaJ could be made and tested. In fact, our lab possesses other DnaJ mutant plasmids: K51C and S60C. This would improve the precision of the model.

3. The NMR fingerprinting was performed by two dimensional HSQC-TROSY experiments [240]. The central portion of the spectrum was densely crowded. Although we used an existing peaklist assigned for DnaK in the ADP.Pi form, many residue assignments in the central region of the spectrum were highly ambiguous. We chose to be conservative in our analysis and only picked residues that were distinct and unambiguous. The application of 3-dimensional NMR experiments (HNCA/HNCO)[113] would provide another frequency dimension (^{13}C) and help in resolving these cluttered peak. It is then possible that

careful data mining would lead to a higher precision structure, and possibly a rigorous determination on the basis of Eq 4.1, as is embedded in XPLOR-NIH.

4. Our model agrees to a certain extent with the crystal disulphide linked complex of bovine Hsc70 NBD and the Auxilin J-domain. While the crystal structure is of high resolution (2.4 Angstrom), we have argued that the fundamental premise of the disulphide linkage is flawed. However, with the insight gained from our model, further efforts in crystallography/NMR could certainly refine our model.

4.13 REFERENCES

[1] J. Jiang, E.G. Maes, A.B. Taylor, L. Wang, A.P. Hinck, E.M. Lafer, and R. Sousa, Structural basis of J cochaperone binding and regulation of Hsp70. Mol Cell 28 (2007) 422-33.

[2] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-366.

[3] B. Bukau, J. Weissman, and A. Horwich, Molecular chaperones and protein quality control. Cell 125 (2006) 443-51.

[4] Q. Liu, and W.A. Hendrickson, Insights into hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-120.

[5] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-5.

[6] D.S. Williamson, J. Borgognoni, A. Clay, Z. Daniels, P. Dokurno, M.J. Drysdale, N. Foloppe, G.L. Francis, C.J. Graham, R. Howes, A.T. Macias, J.B. Murray, R. Parsons, T. Shaw, A.E. Surgenor, L. Terry, Y. Wang, M. Wood, and A.J. Massey, Novel adenosine-derived inhibitors of 70 kDa heat shock protein, discovered through structure-based design. J Med Chem 52 (2009) 1510-3.

[7] H. Sondermann, C. Scheufler, C. Schneider, J. Hohfeld, F.U. Hartl, and I. Moarefi, Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. Science 291 (2001) 1553-7.

[8] A. Bhattacharya, A.V. Kurochkin, G.N. Yip, Y. Zhang, E.B. Bertelsen, and E.R. Zuiderweg, Allostery in Hsp70 chaperones is transduced by subdomain rotations. J Mol Biol 388 (2009) 475-90.

[9] J.P. Hendrick, T. Langer, T.A. Davis, F.U. Hartl, and M. Wiedmann, Control of folding and membrane translocation by binding of the chaperone DnaJ to nascent polypeptides. Proc Natl Acad Sci U S A 90 (1993) 10216-20.

[10] A. Wawrzynow, B. Banecki, D. Wall, K. Liberek, C. Georgopoulos, and M. Zylicz, ATP hydrolysis is required for the DnaJ-dependent activation of DnaK chaperone for

binding to both native and denatured protein substrates. J Biol Chem 270 (1995) 19307-11.
[11] D.M. Cyr, T. Langer, and M.G. Douglas, DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. Trends Biochem Sci 19 (1994) 176-81.
[12] M. Zylicz, D. Ang, K. Liberek, and C. Georgopoulos, Initiation of lambda DNA replication with purified host- and bacteriophage-encoded proteins: the role of the dnaK, dnaJ and grpE heat shock proteins. EMBO J 8 (1989) 1601-8.
[13] S. Wickner, J. Hoskins, and K. McKenney, Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA. Nature 350 (1991) 165-7.
[14] M.K. Greene, K. Maskos, and S.J. Landry, Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proc Natl Acad Sci U S A 95 (1998) 6108-13.
[15] D.M. Cyr, X. Lu, and M.G. Douglas, Regulation of Hsp70 function by a eukaryotic DnaJ homolog. J Biol Chem 267 (1992) 20927-31.
[16] D.P. Atencio, and M.P. Yaffe, MAS5, a yeast homolog of DnaJ involved in mitochondrial protein import. Mol Cell Biol 12 (1992) 283-91.
[17] N. Rowley, C. Prip-Buus, B. Westermann, C. Brown, E. Schwarz, B. Barrell, and W. Neupert, Mdj1p, a novel chaperone of the DnaJ family, is involved in mitochondrial biogenesis and protein folding. Cell 77 (1994) 249-59.
[18] T. Szyperski, M. Pellecchia, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure determination of the Escherichia coli DnaJ molecular chaperone: secondary structure and backbone fold of the N-terminal region (residues 2-108) containing the highly conserved J domain. Proc Natl Acad Sci U S A 91 (1994) 11343-7.
[19] M. Pellecchia, T. Szyperski, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure of the J-domain and the Gly/Phe-rich region of the Escherichia coli DnaJ chaperone. J. Mol. Biol. 260 (1996) 236-250.
[20] M. Martinez-Yamout, G.B. Legge, O. Zhang, P.E. Wright, and H.J. Dyson, Solution structure of the cysteine-rich domain of the Escherichia coli chaperone protein DnaJ. J Mol Biol 300 (2000) 805-18.
[21] D. Wall, M. Zylicz, and C. Georgopoulos, The NH2-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. J Biol Chem 269 (1994) 5446-51.
[22] C. Alfano, and R. McMacken, Heat shock protein-mediated disassembly of nucleoprotein structures is required for the initiation of bacteriophage lambda DNA replication. J Biol Chem 264 (1989) 10709-18.
[23] M. Zylicz, T. Yamamoto, N. McKittrick, S. Sell, and C. Georgopoulos, Purification and properties of the dnaJ replication protein of Escherichia coli. J Biol Chem 260 (1985) 7591-8.
[24] J. Gamer, H. Bujard, and B. Bukau, Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32. Cell 69 (1992) 833-42.
[25] T. Langer, C. Lu, H. Echols, J. Flanagan, M.K. Hayer, and F.U. Hartl, Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356 (1992) 683-9.

[26] M. Revington, Y. Zhang, G.N. Yip, A.V. Kurochkin, and E.R. Zuiderweg, NMR investigations of allosteric processes in a two-domain Thermus thermophilus Hsp70 molecular chaperone. J Mol Biol 349 (2005) 163-83.
[27] Y. Zhang, and E.R. Zuiderweg, The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. Proc Natl Acad Sci U S A 101 (2004) 10272-7.
[28] J.F. Swain, G. Dinler, R. Sivendran, D.L. Montgomery, M. Stotz, and L.M. Gierasch, Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol Cell 26 (2007) 27-39.
[29] M. Vogel, M.P. Mayer, and B. Bukau, Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. J Biol Chem 281 (2006) 38705-11.
[30] M.P. Mayer, and B. Bukau, Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci 62 (2005) 670-84.
[31] S. Rudiger, A. Buchberger, and B. Bukau, Interaction of Hsp70 chaperones with substrates. Nat Struct Biol 4 (1997) 342-9.
[32] S. Ahle, and E. Ungewickell, Auxilin, a newly identified clathrin-associated protein in coated vesicles from bovine brain. J Cell Biol 111 (1990) 19-29.
[33] S.L. Schmid, W.A. Braell, D.M. Schlossman, and J.E. Rothman, A role for clathrin light chains in the recognition of clathrin cages by 'uncoating ATPase'. Nature 311 (1984) 228-31.
[34] A. Umeda, A. Meyerholz, and E. Ungewickell, Identification of the universal cofactor (auxilin 2) in clathrin coat dissociation. Eur J Cell Biol 79 (2000) 336-42.
[35] D.T. Haynie, and C.P. Ponting, The N-terminal domains of tensin and auxilin are phosphatase homologues. Protein Sci 5 (1996) 2643-6.
[36] Y. Ma, T. Greener, M.E. Pacold, S. Kaushal, L.E. Greene, and E. Eisenberg, Identification of domain required for catalytic activity of auxilin in supporting clathrin uncoating by Hsc70. J Biol Chem 277 (2002) 49267-74.
[37] T. Greener, X. Zhao, H. Nojima, E. Eisenberg, and L.E. Greene, Role of cyclin G-associated kinase in uncoating clathrin-coated vesicles from non-neuronal cells. J Biol Chem 275 (2000) 1365-70.
[38] A. Fotin, Y. Cheng, N. Grigorieff, T. Walz, S.C. Harrison, and T. Kirchhausen, Structure of an auxilin-bound clathrin coat and its implications for the mechanism of uncoating. Nature 432 (2004) 649-53.
[39] E. Ungewickell, H. Ungewickell, S.E. Holstein, R. Lindner, K. Prasad, W. Barouch, B. Martin, L.E. Greene, and E. Eisenberg, Role of auxilin in uncoating clathrin-coated vesicles. Nature 378 (1995) 632-5.
[40] D. Wall, M. Zylicz, and C. Georgopoulos, The conserved G/F motif of the DnaJ chaperone is necessary for the activation of the substrate binding properties of the DnaK chaperone. J Biol Chem 270 (1995) 2139-44.
[41] W.L. Kelley, and C. Georgopoulos, The T/t common exon of simian virus 40, JC, and BK polyomavirus T antigens can functionally replace the J-domain of the Escherichia coli DnaJ molecular chaperone. Proc Natl Acad Sci U S A 94 (1997) 3679-84.

[42] W.C. Suh, W.F. Burkholder, C.Z. Lu, X. Zhao, M.E. Gottesman, and C.A. Gross, Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. Proc Natl Acad Sci U S A 95 (1998) 15223-8.
[43] W.C. Suh, C.Z. Lu, and C.A. Gross, Structural features required for the interaction of the Hsp70 molecular chaperone DnaK with its cochaperone DnaJ. J Biol Chem 274 (1999) 30534-9.
[44] X.T. Zhu, X. Zhao, W.F. Burkholder, A. Gragerov, C.M. Ogata, M.E. Gottesman, and W.A. Hendrickson, Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272 (1996) 1606-1614.
[45] S.Y. Stevens, S. Cai, M. Pellecchia, and E.R. Zuiderweg, The solution structure of the bacterial HSP70 chaperone protein domain DnaK(393-507) in complex with the peptide NRLLLTG. Protein Sci 12 (2003) 2588-96.
[46] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc Natl Acad Sci U S A 106 (2009) 8471-6.
[47] E. Ungewickell, H. Ungewickell, and S.E. Holstein, Functional interaction of the auxilin J domain with the nucleotide- and substrate-binding modules of Hsc70. J Biol Chem 272 (1997) 19594-600.
[48] K. Pervushin, G. Wider, and K. W¸thrich, Single transition-to-single transition polarization transfer (ST2-PT) in [15N, 1H]-TROSY. Journal of Biomolecular NMR 12 (1998) 345-348.
[49] W. Jahnke, Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy. Chembiochem 3 (2002) 167-73.
[50] G. Bertani, Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol 186 (2004) 595-600.
[51] K. Liberek, D. Wall, and C. Georgopoulos, The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. Proc Natl Acad Sci U S A 92 (1995) 6224-8.
[52] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6 (1995) 277-293.
[53] T.D. Goddard, and D.G. Kneller, SPARKY 3. University of California, San Francisco (2000).
[54] P.P. Van Veldhoven, and G.P. Mannaerts, Inorganic and organic phosphate measurements in the nanomolar range. Anal Biochem 161 (1987) 45-8.
[55] P. Ekman, and O. Jager, Quantification of subnanomolar amounts of phosphate bound to seryl and threonyl residues in phosphoproteins using alkaline hydrolysis and malachite green. Anal Biochem 214 (1993) 138-41.
[56] Q. Liu, and W.A. Hendrickson, Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-20.
[57] R.C. Morshauser, W. Hu, H. Wang, Y. Pang, G.C. Flynn, and E.R. Zuiderweg, High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. J Mol Biol 289 (1999) 1387-403.

[58] A. Bax, and M. Ikura, An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. J Biomol NMR 1 (1991) 99-104.

CHAPTER 5 EPILOGUE AND FUTURE OUTLOOK

5.1 ABSTRACT

This chapter is a discussion of possible research to extend the work presented in this dissertation. This is in the specific context of the work on Hsp70 structure, allostery and function as studied by our group. We also discuss avenues to explore in the DnaK-DnaJ interaction project. Finally, we expand the context to the larger question of studying allostery in sizeable proteins and touch upon the contemporary and future roles of NMR spectroscopy in structural biology.

5.2 ALLOSTERY IN THE Hsp70 SYSTEM

5.2.1 INITIAL WORK ON STRUCTURES

Much is known about the Hsp70 molecule and its functional cycle. This is based on more than two decades of work by crystallographers and NMR spectroscopists. McKay was the first to crystallize the nucleotide-binding domain (NBD) and define the four different subdomains on the basis of secondary structure continuity and spatial proximity [49; 50]. His lab went on to define the nucleotide binding cleft and identifying crucial residues for ATP hydrolysis [53; 55; 164]. This early work examined the structure of the NBD in detail and defined the exacting standards by which all later researchers have evaluated their work. However, this body of work was ultimately a study of structure only. Identification of the Hsp70 as an allosteric machine came later. But in the meantime, other structural biologists had been working on the substrate-binding domain. In terms of NMR spectroscopy, various constructs of the SBD (the 15 kDa β-basket, for example) are smaller than the NBD (45 kDa), and hence within the limits of NMR technology. Our lab had investigated the structure of the SBD from mammalian Hsc70 (Rat)[58; 66]. These studies, along with Hendrickson's [67] crystal

structure of the *E.coli* DnaK SBD revealed the β-basket architecture of the substrate binding domain and the α-helical structure of the lid. Prior to these works, the SBD was modeled on the basis of MHC domains. Hendrickson's work, in particular, was very much a pathfinder, because of the use of a heptapeptide (NRLLLTG), which mimics hydrophobic exposed surface residues on misfolded proteins and is a suitable substrate for the SBD to bind to. This heptapeptide is an inhibitor of the substrate-binding domain.

5.2.2 MOVING STRUCTURES AND ALLOSTERY

It was when structural biologists focused their efforts on the substrate-binding domain that it became evident that Hsp70 was an allosteric system. Hendrickson and co-workers had identified crucial hinge-points on the SBD where the lid could swing back and open up access to the substrate binding cleft (residues N508, N522). They also identified the substrate-binding pocket (defined as 'hsp70 site 0'). This has already been discussed in the Introduction. However, it was the suggestion that the lid was capable of opening up and allowing substrates access to hsp70 site 0 that set the ground for much of the allostery studies to follow. Hendrickson and co-workers also hypothesized that the ATP state of the molecule corresponded to the lid-open state. The work by Zuiderweg and co-workers soon after showed that the lid was not essential for peptide binding. The same peptide used by Hendrickson's lab (NRLLLTG) was found to bind to the β-basket element of the *E.coli* DnaK SBD (residues 393-507) without causing any change in secondary structure [159]. At this point, it was clear that the lid domain is not essential for peptide binding. But the question remains: is it necessary for peptide sequestering and reengineering? This is a much more difficult and subtle question to answer. Most assays of DnaK function tend to focus on the ATP hydrolysis activity of the NBD. With reference to the analogy of the NBD = engine, SBD = wheels introduced earlier, this is akin to measuring the power output of an engine. However, the real question is: how does the power generated by the engine get

transferred to the wheels. To that end, researchers returned to looking at the role played by the NBD, and the linker.

5.2.3 DETAILED WORK ON ALLOSTERY

Zuiderweg and co-workers worked on *T.thermophilus* DnaK nucleotide binding domain and observed chemical shift changes between the ADP.Pi and the ADP.ALFx state (the latter being an ATP mimic). This thermophilic protein is more amenable to NMR spectroscopy as it can be probed at higher temperatures, generating cleaner spectra (as the molecule tumbles faster – lineshapes get sharper). The chemical shift changes observed are indicative of conformational change, but do not define what the change is [211]. The first detailed NMR based solution structure of the nucleotide binding domain (for bovine Hsc70 1-386) was also published by this group, where it was noted that the two lobes of the protein (IA-IB and IIA-IIB) orient slightly differently as compared to various crystal structures. This RDC based study showed with direct structural evidence (from dipolar couplings, as opposed to chemical shifts which provide secondary, or inferred data on conformation change) for the first time that solution conditions allow for more conformations to be sampled by the Hsp70 system [183]. The Zuiderweg lab then went on to study a larger, 54 kDa two-domain construct of *T.thermophilus* DnaK (1-501 residues). It was observed that the interdomain interface between the NBD and the SBD was tightest in the NBD-ATP/SBD-apo state, followed by the NBD-ADP.Pi/SBD-apo state, followed by the NBD-ADP.Pi/SBD-peptide state [186]. This indicated that the allosteric machinery in the protein works bidirectionally. This was supported by work from the Gierasch group, which showed that Trp102 on subdomain IB in the NBD of DnaK is quenched, blue shifted, and protected from solvent in the presence of ATP – evidence indicating that the lid might dock with the NBD in the ATP state [245].

5.2.4 WORK ON THE LINKER

Gierasch and co-workers went on to investigate the role of the linker in more detail using CD thermal melting transition analysis along with solution NMR on two constructs of DnaK (1-388 and 1-394 residues, the latter containing the conserved VLLL sequence). They showed that the linker binds to a hydrophobic cleft between subdomains IA and IIA of the NBD and hypothesized that the linker acts as an allosteric mediator between the NBD and the SBD [91]. The disappearance of the SBD peaks in the TROSY (NMR fingerprint) spectrum in the ATP form of DnaK also led them to state that when the NBD is bound to ATP, the NBD and the SBD are docked together. This was in line with previous hypotheses about domain association in the Hsp70 system connected to nucleotide turnover [184; 185; 246]. But as yet, no definite structure had been solved to confirm all of these ideas.

5.2.5 THE FIRST FULL LENGTH DnaK STRUCTURE

Bertelsen and Zuiderweg solved the first full-length structure of DnaK (1-605 residues, excluding the floppy tail at the C-terminus) using residual dipolar coupling analysis. This work showed that, in the ADP, peptide-bound state, the SBD is an independent entity from the NBD, connected by a floppy linker and can sample multiple conformations w.r.t. the NBD. This work treated the NBD and the SBD as independent rigid bodies and used paramagnetic labels to solve degeneracies in the RDC principal angle solutions [92].

5.2.6 MORE DETAILED WORK ON ALLOSTERY IN THE NUCLEOTIDE BINDING DOMAIN

We have discussed in Chapters 2 and 3 of this dissertation detailed work, which shows that the different subdomains of the Tth-DnaK NBD move w.r.t. each other as the molecule turns over nucleotide. This elucidated the mechanism of interaction of Hsp70s with nucleotide exchange factors, and also shed further light on the mechanism of allosteric signaling between the NBD and the SBD: the linker-docking-model proposed by Gierasch and co-workers was better explained as the IA/IIA

interface itself opens in the AMPPNP (ATP mimic) state and closes in the ADP.Pi state. Thus the linker actually has less room to become structured in the ADP.Pi state as compared to the ATP state [203]. This was a result which was unanticipated by crystallography. But all of the work described so far converges toward one major research target – the ATP state of a full-length Hsp70 molecule. In the immediate context of this dissertation, it would be of great interest to see if the open/close motion of the NBD nucleotide-binding cleft is, in fact seen in the full-length molecule. As the ADP.Pi state of this molecule has already been investigated [92], what remains is the ATP state.

5.3 THE ATP STATE OF Hsp70

The main candidate for the Hsp70 ATP state structure in the literature is Hsp110/ATP. This bears a certain degree of homology to Hsp70 – the *S.cerevisiae* form of this protein (SSe1) is about 35% homologous by sequence to the NBD of Hsp70, and less so for the SBD (19.1% for the β-sandwich and 12.7% for the α-helical lid). This molecule actually functions as a nucleotide exchange factor and was proposed as an Hsp70 structural analog based on putative secondary structure. The crystal structure solved by Hendrickson and co-workers does indeed show that the NBD and the SBD are docked together [93]. But apart from the relatively low homology (to use Hendrickson's rather colourful phrase: "in the twilight zone by sequence") this structure is suspect for another crucial reason: the linker is not conserved. The *E.coli* DnaK linker sequence from residue 386 runs: VKD(VLLL)DVTP with the maximally conserved sequence in parentheses. The SSe1 (S.cerevisiae) sequence from residue 389 (chosen by secondary structure alignment) runs VRPFKFEDIHP. As can be seen, the hydrophobic nature of the linker is not conserved in this molecule. Thus, while it may show some of the expected characteristics of the ATP state of Hsp70, it is difficult to consider this as a completely viable model.

5.4 SOLUTION NMR STUDIES ON THE ATP STATE OF Hsp70

A former postdoc in this lab had done some preliminary work on the full length DnaK molecule in the ATP state. The problem is one of three-body orientation: RDCs can unambiguously define the orientation of one body with respect to another. For a multi-body problem such as the NBD subdomains (discussed in Chapters 2 and 3), if we know that the different rigid bodies (subdomains IA, IIA, IB and IIB) are still tethered to each other by short loops, then principal angle degeneracies of +/- 90 or 180 degrees can be easily resolved by rejecting structures, which are geometrically unreasonable. This is not the case for the NBD-SBD-Lid problem. Thus, even with RDCs on all three domains, overall orientation of one domain (here, the lid) leads to a degenerate solution. One avenue out of this situation is to use paramagnetic labels, as discussed in Chapter 4. We picked a few residues on the DnaK molecule and mutated them to cysteines. These were D97C, V210 C and K421C. We performed QuickchangeTM mutations on DnaK1-605 plasmid (in pET22b) to incorporate the cysteines. Unfortunately, expression levels in BL21 (DE3) *E.coli* cells were very low. Furthermore, protein purification using FFQ/ATP-agarose columns also led to very low yields, and in spite of having protease inhibitors in solution, we frequently saw the protein being lysed at the linker. Thus we turned our focus upon the DnaK-DnaJ interaction and did not proceed further with this project.

Now, however, since we have a C-terminal HIS tagged construct of DnaK1-605, which delivers good protein yields and is easy to purify (Ni-NTA/Sepharose S-100 columns) – it must be of benefit to revisit this project. Triple QuickchangeTM mutations on the DnaK1-605 plasmid would be done in stages. This would enable us to repeat the Cysteine emplacements on D97, V210 and K421, as well as effect the T199A mutation, which inhibits ATP hydrolysis by the NBD and finally V436F that prevents binding of substrate to the SBD. This would allow us to stabilize the NBD-ATP/SBD-apo form of the protein in an easily purified form. MTSL addition to the surface exposed Cys residues (as discussed in Chapter 4) would enable us to find the orientation of the

lid w.r.t the other two domains. W.r.t Fig 5.1, D97, V210 and K421 are rendered as spheres. If the α-Lid domain opens up in the ATP state and docks with the NBD, close to subdomain IB, then lid residues should be broadened in proximity to D97C-MTSL. If the SBD β-basket docks with the NBD, then V210C-MTSL on subdomain IIA should lead to peaks on the SBD being broadened. A similar effect would then also be seen in NBD residues close to K21C-MTSL (which is located in the β-basket region).

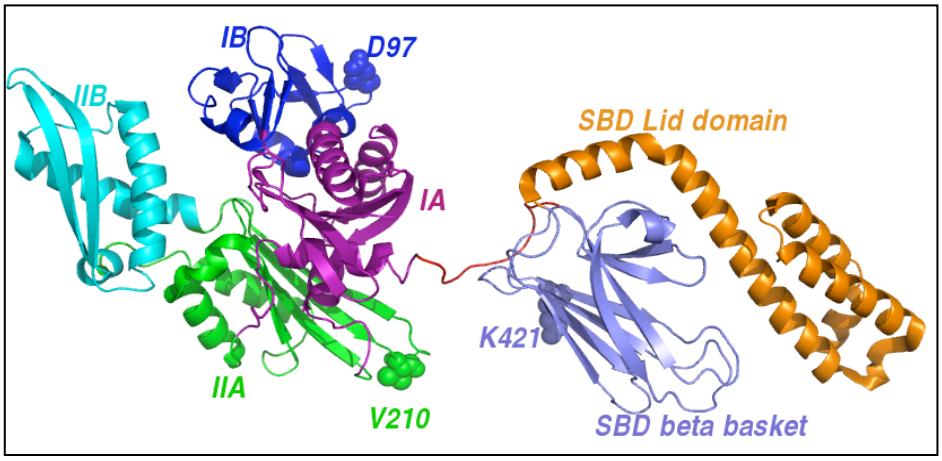

Figure 5.1: DnaK in the ADP.Pi state showing Cys mutation points

5.5 FURTHER STUDIES ON DnaK-DnaJ INTERACTION

There are three main avenues to be pursued here.

5.5.1 DOUBLE MUTANT STUDIES

The first is the question of double mutants. As discussed earlier in Chapter 4, the work of Suh and Gross showed via Surface Plasmon Resonance experiments that the dissociation constant of DnaK-DnaJ increases from 70 nM to 331 nM when R167 is mutated to A on DnaK. DnaK and DnaJ-D35N do not have any measurable interaction. But the double mutant analysis of DnaK R167H (or R167A) with DnaJ D35N shows substantial interaction with K_D=163 nM. It should be noted that no specific data was reported in this work which indicated that the double-mutant complex had substantially increased ATP hydrolysis rates. Our NMR analysis shows an interaction of the J-domain with Dnak ADP.Pi across both the NBD and the SBD, as discussed in Chapter

4. The next experiment to do would be to investigate whether this mutation also causes a loss of interaction at NMR concentrations (> 50 uM). This can be checked by seeing if changes occur in the spectrum of DnaK R167H/A upon titrating the J-domain and DnaJ-D35N into the sample, and whether the interaction of DnaK R167H/A with DnaJ-D35N is the same as the wild-type interaction. The experiments suggested are tabulated below (in all cases, "Cys mutant" refers to R19C/M30C/K31C/K41C):

DnaK mutant⇒ DnaJ mutant ⇓	1-605	1-605 R167H
1-108 (Cys mutant)	Done (Chapter 4)	(To be studied: expect interaction at NMR concentrations)
1-108 (Cys mutant) D35N	To be studied: expect no or little interaction at NMR concentrations	To be studied: expect interaction at NMR concentrations

Table 5.1: Suggested further experiments for DnaK-DnaJ interaction studies

5.5.2 TRUNCATED DnaJ STUDIES

The second approach to better understand this system is based on the fact that peptide addition appears to stimulate the rate of ATP hydrolysis by wt-FL DnaK638 about 2-2.5 times (see Section 4.9). This is the same as the speedup achieved by addition of either wt-FL DnaJ or J-domain (1-108) mutants. This might be due a competition for the hsp70 site 0 (as defined by Hendrickson and co-workers[67]). This can be tested by performing NMR paramagnetic studies and ATP hydrolysis assays with a smaller construct of the J-domain that does not contain the Gly-Phe region [237], which could potentially interact with the substrate binding cleft in the DnaK SBD. These studies could shed more light on any possible competition between substrate peptide and the Gly-Phe region of the J-domain for binding to the SBD.

5.5.3 FULL LENGTH DnaJ STUDIES

The third avenue to proceed along is to study the interaction of full length (386 residue) DnaJ with DnaK using the same techniques of paramagnetic labeling. As per

our current understanding of the allosteric mechanism of the Hsp70, the presence of peptide in the SBD stimulates ATP hydrolysis by the NBD, as does the addition of DnaJ/J-domain. There is the possibility that DnaJ 'competes' with peptide in terms of stimulation of the ATPase activity of the entire Hsp70 system. Thus, a more complete picture would emerge if we were able to study interaction interfaces of full length DnaK and DnaJ, and perhaps at a later stage, interfaces between Hsc70 and Auxilin. There are several techniques, which might be used to probe such complex structures. A few (chemical shift perturbation, H-D exchange and saturation transfer) are discussed below.

5.6 STUDYING INTERFACES BETWEEN LARGE PROTEINS: CHEMICAL SHIFT PERTURBATION NMR

Chemical shift perturbation is a very well exploited tool [247] to look at reaction interfaces between molecules or sites of ligand interactions with proteins. In fact the motivation to study nucleotide induced conformational changes in Tth-NBD was based on chemical shift changes in solution seen in the NBD spectrum between different nucleotide states (ADP.Pi vs AMPPNP) {Bhattacharya, 2009 #10860}. This technique can also be used to look at interactions between large proteins. The method is summarized as below:

1. Protein A (DnaK/Hsc70) is expressed in $^{15}N/^{13}C/^{2}H_20$ M9 minimal media – and is 'triple labeled' and suitable for multidimensional NMR spectroscopy.
2. Protein B (DnaJ/Auxilin) is expressed in unlabeled (no exotic isotopes) LB media [241] – this is NMR HSQC silent.
3. NMR fingerprint experiments on Protein A yield spectra which are assigned and then Protein B is titrated in. Shifting peaks/ changing linewidths is indicative of interaction interfaces on Protein A
4. Repeat the experiment series with isotope labeled Protein B and NMR silent Protein A. Identify interaction interfaces on Protein B.

5. Use results from 3 & 4 to dock Proteins A & B together.

The main problem with this technique is that for allosteric systems, there is always the possibility that an interaction between Proteins A and B at a particular interface causes long-range conformational changes which then induce substantial chemical shift perturbations (effectively on both proteins) far away from the binding site. Thus, chemical shift perturbation mapping can tell us that some kind of an interaction between proteins A and B occurs – but it may not be able to accurately tell us where it happens.

5.7 STUDYING INTERFACES BETWEEN LARGE PROTEINS: HYDROGEN-DEUTERIUM EXCHANGE NMR

Hydrogen-Deuterium exchange operates on the basic principle that backbone amide protons exchange with solvent protons when possible. This is especially fast when the amide group being monitored belongs to a solvent exposed region of the protein rather than being buried deep inside. Exchanging the solvent from H_20 to D_20 now allows for the exchange of surface exposed amide protons with solvent deuterons. This can be monitored by mass spectrometry [248] or multidimensional NMR [249]. In an NMR experiment, the replacement of an amide proton by a deuteron will lead to the loss of the ^1H signal from that proton to the residual water resonance. As a particular surface exposed proton exchanges with deuterium (this is, obviously, an ensemble/time averaged property), the NMR signal (i.e. cross-peak height) originating from that proton will decay exponentially with exchange time [250]. Thus, while monitoring H-D exchange for Protein A (DnaK/Hsc70: this is $^1H/^{13}C/^{15}N$ labeled) via 2D/3D NMR, we titrate in Protein B (DnaK/Auxilin: this is NMR silent) and observe the exchange rate for certain surface residues of Protein A to be reduced with increasing concentrations of Protein B, we conclude that the addition of Protein B sequesters these residues from solvent. Thus, we have identified the interaction interface.

Of course, there are several caveats to this strategy: it may not work very well, if

the interaction interface is not very tight and not very extensive – there is too little perturbation of solvent accessibility. Furthermore, other investigators have also found that amide proton exchange can also give false positives in allosteric systems [251].

5.8 STUDYING INTERFACES BETWEEN LARGE PROTEINS: CROSS SATURATION NMR

Cross saturation NMR spectroscopy relies on the transfer of magnetization from a perturbed member of a molecular complex to the unperturbed member. Protein A(DnaK/Hsc70) is prepared without any specific isotope labeling while Protein B (DnaK/Auxilin) is triple labeled (99% 2D/^{13}C /^{15}N). When both proteins form a complex in solution (90% D_20 buffer), the aliphatic ^{1}H resonances of protein A can be selectively irradiation using the proton channel. Spin diffusion [252] helps in transferring saturation to aromatic and amide resonances of Protein A very rapidly. The next step is the transfer of saturation from Protein A to Protein B by cross relaxation at the interaction surface. Protein B has 99% deuterons, but the remaining amide protons still enable us to record a 2-dimensional TROSY fingerprint spectrum [240]. The TROSY spectrum of Protein B recorded before and after saturation of Protein A will show some cross-peaks to be diminished. These will belong to interface residues, as the low density of protons on Protein B do not permit the cross-relaxation driven saturation transfer to propagate throughout all of Protein B. This technique was pioneered by Shimada and co-workers [253; 254]. The size limit to this method was about a 150 kDa complex. If Protein A and Protein B are in rapid exchange ($K_D > 10^{-5}$ M) and there is a large excess of Protein A over Protein B (10:1 ratio) then the transferred saturation can be detected on Protein B when Protein A is irradiated [255].

5.9 SUMMARIZING PROTEIN-PROTEIN INTERFACE ANALYSIS

The techniques summarized above are, but a sampling of the major interface

analysis techniques available. Supplementing these experimental tools with molecular dynamics simulations, a much better understanding of how DnaK and DnaJ interact can be found. This will help shed more light on the entire Hsp70 system and its partner molecules.

5.10 A FEW COMMENTS ON CURRENT TRENDS IN STRUCTURAL BIOLOGY

Crystallography has always been the overshadowing elder sibling of NMR spectroscopy when it comes to structural biology. The obvious question to ask is: what is the scope of NMR structural biology in particular and solution phase structural biology in general?

It is increasingly becoming apparent that the *ab-initio* structure determination by NMR involves a complicated and lengthy timeline – find stable solution conditions, assign the protein [111; 256], find distance constraints (NOEs [118; 257]), solve the structure [191; 258; 259; 260; 261]. The turnaround time involved in data analysis is quite long. Chemical shift indexing has been used to gain preliminary information about secondary structure [262]. New methods involve the use of chemical shift indexing to help Rosetta *ab initio* structure determinations [263] [264]. If this method becomes more reliable, NMR structure determinations for smaller proteins (< 100 kDa) may become sufficiently rapid as to compete with X-ray crystallography. There are automated assignment techniques in development [265], which may help in speeding up data analysis and structure computing [266; 267; 268; 269; 270]. As more NMR assignments come online and better data mining techniques, these programs will become more accurate. In spite of this, the degree of automation routinely available in X-ray crystallography will be difficult to achieve in NMR, at least in the foreseeable future.

There are specific avenues, however, where NMR spectroscopy can use crystal

structures as a starting point to gain greater insight into biophysical processes. Conformational changes in proteins, which can be identified in the solution phase; are frequently not seen in crystallography [203; 271]. The results of Chapters 2 and 3 of this thesis are other examples of this phenomenon. The strategy adopted in much of the work done in this dissertation (Chapters 2/3) has been as follows:

1. Start with a homology model of the protein of interest (based on a prior structure, usually but not necessarily solved by crystallography).
2. Solution NMR studies under different conditions (nucleotide change/interaction with other proteins) induce conformational changes.
3. Residual dipolar couplings detect conformational changes.
4. Solution 'eigenstructures' are generated based on the starting homology model and RDC analysis, which reflect actual changes in the molecular structure hitherto unobserved.

This strategy can be expanded to incorporate other tools as well:

1. Small Angle X-Ray Scattering can be used as additional data input to further refined RDC based structures [272; 273].
2. Fluorescence resonance energy transfer (FRET) spectroscopy is another 'molecular' ruler that can be used to measure distance changes between different fluorophores upon conformation change [274]. Distance changes between 10-100 Angstroms can potentially be measured, albeit with much lower resolution than by crystallography/NOEs. The range of applicability depends on the donor-acceptor pair used. In addition, even if a distance between the large donor ands acceptor fluorophores can be measured with precision, it is not at all straightforward to translate that distance to a distance between the amino acid residues in the protein the labels were attached to. Very large fluorophores (such as GFP – 26.9 kDa [275]) which themselves add on to the molecular weight of the system under study are not suitable for the systems like Hsp70s. Cyanine based dyes (Cy3/Cy5) are small molecules which are sometimes synthesized

with maleimide groups to enable conjugation with surface exposed Cys residues (much like the MTSL tagging procedure discussed in Chapter 4) [276]. Surface exposed intrinsic fluorophores (such as Trp102 in DnaK [245]) can also be used as either donors/acceptors.

3. Large scale high resolution molecular dynamics simulations of protein trajectories (AMBER [277]/CHARMM [278]) has the potential to bolster and refine experimental results. These methods, however tend to be computationally expensive and as yet, it is impossible to compute a MD trajectory for a large protein (~50 kDa) over the timescale of possible conformational exchange (~100 msec).

4. Coarse grained simulations of protein conformational changes [279] due to ligand binding/nucleotide turnover can also be run with a short turnaround time using Gaussian Network Modeling [280] (essentially a mass & springs representation of the protein). This can give a low resolution picture that evolves over a longer timescale than with all-atom high resolution simulations and can show trends such as clefts opening/closing upon ligand binding.

5.11 AFTERWORD

The techniques discussed above are increasingly used in conjunction with NMR spectroscopy to advance the state of the art in structural biology. Four decades after the seminal work on Haemoglobin [96] that illustrated that molecules are not static entities, structural biology has come of age. The way we will look at biomolecules in the future will be as movies, and not merely pictures. Al-Hashimi and colleagues have actually computed an NMR-movie for the TAR–RNA molecule [281] ; the first experimental protein movie has not yet appeared. NMR spectroscopy is uniquely poised to use an impressive arsenal of strategies to find out how large molecules and their complexes move and interact.

5.12 REFERENCES

[1] C.B. Anfinsen, Principles That Govern Folding of Protein Chains. Science 181 (1973) 223-230.

[2] P.E. Leopold, M. Montal, and J.N. Onuchic, Protein folding funnels: a kinetic approach to the sequence-structure relationship. Proc Natl Acad Sci U S A 89 (1992) 8721-5.

[3] R.K. Pathria, Statistical mechanics, 2nd ed. (1996).

[4] K.A. Dill, Polymer principles and protein folding. Protein Sci 8 (1999) 1166-80.

[5] J. Schonbrun, and K.A. Dill, Fast protein folding kinetics. Proc Natl Acad Sci U S A 100 (2003) 12678-82.

[6] A.L. Fink, Natively unfolded proteins. Curr Opin Struct Biol 15 (2005) 35-41.

[7] M. Karplus, and J.A. McCammon, The dynamics of proteins. Sci Am 254 (1986) 42-51.

[8] M. Karplus, and J.A. McCammon, Dynamics of proteins: elements and function. Annu Rev Biochem 52 (1983) 263-300.

[9] K. Henzler-Wildman, and D. Kern, Dynamic personalities of proteins. Nature 450 (2007) 964-72.

[10] J. Cavanagh, and R.A. Venters, Protein dynamic studies move to a new time slot. Nat Struct Biol 8 (2001) 912-4.

[11] D. Kern, and E.R. Zuiderweg, The role of dynamics in allosteric regulation. Curr Opin Struct Biol 13 (2003) 748-57.

[12] V.I. Abkevich, A.M. Gutin, and E.I. Shakhnovich, Specific nucleus as the transition state for protein folding: evidence from the lattice model. Biochemistry 33 (1994) 10026-36.

[13] H.S. Chan, S. Bromberg, and K.A. Dill, Models of cooperativity in protein folding. Philos Trans R Soc Lond B Biol Sci 348 (1995) 61-70.

[14] K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas, and H.S. Chan, Principles of protein folding--a perspective from simple exact models. Protein Sci 4 (1995) 561-602.

[15] A.R. Fersht, Nucleation mechanisms in protein folding. Curr Opin Struct Biol 7 (1997) 3-9.

[16] M.J. Gething, and J. Sambrook, Protein folding in the Cell. Nature 355 (1992) 33-45.

[17] S.W. Englander, L. Mayne, and M.M. Krishna, Protein folding and misfolding: mechanism and principles. Q Rev Biophys 40 (2007) 287-326.

[18] F.U. Hartl, and M. Hayer-Hartl, Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol 16 (2009) 574-81.

[19] V.I. Abkevich, A.M. Gutin, and E.I. Shakhnovich, Domains in folding of model proteins. Protein Sci 4 (1995) 1167-77.

[20] F.U. Hartl, and M. Hayer-Hartl, Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295 (2002) 1852-8.

[21] C.M. Dobson, Protein folding and misfolding. Nature 426 (2003) 884-90.

[22] C.M. Kaiser, H.C. Chang, V.R. Agashe, S.K. Lakshmipathy, S.A. Etchells, M. Hayer-Hartl, F.U. Hartl, and J.M. Barral, Real-time observation of trigger factor function on translating ribosomes. Nature 444 (2006) 455-60.
[23] A. Yonath, Large facilities and the evolving ribosome, the cellular machine for genetic-code translation. J R Soc Interface 6 Suppl 5 (2009) S575-85.
[24] L. Ferbitz, T. Maier, H. Patzelt, B. Bukau, E. Deuerling, and N. Ban, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. Nature 431 (2004) 590-6.
[25] A.V. Ludlam, B.A. Moore, and Z. Xu, The crystal structure of ribosomal chaperone trigger factor from Vibrio cholerae. Proc Natl Acad Sci U S A 101 (2004) 13436-41.
[26] S. Vorderwulbecke, G. Kramer, F. Merz, T.A. Kurz, T. Rauch, B. Zachmann-Brand, B. Bukau, and E. Deuerling, Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK. FEBS Lett 559 (2004) 181-7.
[27] A. Tissieres, H.K. Mitchell, and U.M. Tracy, Protein synthesis in salivary glands of Drosophila melanogaster: relation to chromosome puffs. J Mol Biol 84 (1974) 389-98.
[28] M.J. Schlesinger, Heat shock proteins. J Biol Chem 265 (1990) 12111-4.
[29] T. Hesterkamp, and B. Bukau, Identification of the prolyl isomerase domain of Escherichia coli trigger factor. FEBS Lett 385 (1996) 67-71.
[30] T. Hesterkamp, and B. Bukau, The Escherichia coli trigger factor. FEBS Lett 389 (1996) 32-4.
[31] G. Kramer, D. Boehringer, N. Ban, and B. Bukau, The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat Struct Mol Biol 16 (2009) 589-97.
[32] B. Bukau, J. Weissman, and A. Horwich, Molecular chaperones and protein quality control. Cell 125 (2006) 443-51.
[33] H. Schroder, T. Langer, F.U. Hartl, and B. Bukau, DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J 12 (1993) 4137-44.
[34] C.P. Georgopoulos, R.W. Hendrix, S.R. Casjens, and A.D. Kaiser, Host participation in bacteriophage lambda head assembly. J Mol Biol 76 (1973) 45-60.
[35] J. Zeilstra-Ryalls, O. Fayet, and C. Georgopoulos, The universally conserved GroE (Hsp60) chaperonins. Annu Rev Microbiol 45 (1991) 301-25.
[36] L. Ditzel, J. Lowe, D. Stock, K.O. Stetter, H. Huber, R. Huber, and S. Steinbacher, Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell 93 (1998) 125-38.
[37] T. Langer, C. Lu, H. Echols, J. Flanagan, M.K. Hayer, and F.U. Hartl, Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356 (1992) 683-689.

[38] J. Frydman, E. Nimmesgern, K. Ohtsuka, and F.U. Hartl, Folding of Nascent Polypeptide-Chains in a High-Molecular-Mass Assembly with Molecular Chaperones. Nature 370 (1994) 111-117.
[39] Z. Xu, A.L. Horwich, and P.B. Sigler, The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature 388 (1997) 741-50.
[40] W.A. Fenton, and A.L. Horwich, GroEL-mediated protein folding. Protein Sci. 6 (1997) 743-760.
[41] A.L. Horwich, K.B. Low, W.A. Fenton, I.N. Hirshfield, and K. Furtak, Folding in vivo of bacterial cytoplasmic proteins: role of GroEL. Cell 74 (1993) 909-917.
[42] P.B. Sigler, Z. Xu, H.S. Rye, S.G. Burston, W.A. Fenton, and A.L. Horwich, Structure and function in GroEL-mediated protein folding. Annu Rev Biochem 67 (1998) 581-608.
[43] Z. Xu, and P.B. Sigler, GroEL/GroES: structure and function of a two-stroke folding machine. J Struct Biol 124 (1998) 129-41.
[44] Y.C. Tang, H.C. Chang, A. Roeben, D. Wischnewski, N. Wischnewski, M.J. Kerner, F.U. Hartl, and M. Hayer-Hartl, Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. Cell 125 (2006) 903-14.
[45] J. Ma, P.B. Sigler, Z. Xu, and M. Karplus, A dynamic model for the allosteric mechanism of GroEL. J Mol Biol 302 (2000) 303-13.
[46] M.J. Kerner, D.J. Naylor, Y. Ishihama, T. Maier, H.C. Chang, A.P. Stines, C. Georgopoulos, D. Frishman, M. Hayer-Hartl, M. Mann, and F.U. Hartl, Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. Cell 122 (2005) 209-20.
[47] A. Brinker, G. Pfeifer, M.J. Kerner, D.J. Naylor, F.U. Hartl, and M. Hayer-Hartl, Dual function of protein confinement in chaperonin-assisted protein folding. Cell 107 (2001) 223-33.
[48] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-366.
[49] K.M. Flaherty, C. Delucaflaherty, and D.B. McKay, 3-DIMENSIONAL STRUCTURE OF THE ATPASE FRAGMENT OF A 70K HEAT-SHOCK COGNATE PROTEIN. Nature 346 (1990) 623-628.
[50] C. DeLuca-Flaherty, K.M. Flaherty, L.J. McIntosh, B. Bahrami, and D.B. McKay, Crystals of an ATPase fragment of bovine clathrin uncoating ATPase. J Mol Biol 200 (1988) 749-50.
[51] T.G. Chappell, B.B. Konforti, S.L. Schmid, and J.E. Rothman, The ATPase core of a clathrin uncoating protein. J Biol Chem 262 (1987) 746-51.
[52] T.G. Chappell, W.J. Welch, D.M. Schlossman, K.B. Palter, M.J. Schlesinger, and J.E. Rothman, Uncoating ATPase is a member of the 70 kilodalton family of stress proteins. Cell 45 (1986) 3-13.
[53] S.M. Wilbanks, C. DeLuca-Flaherty, and D.B. McKay, Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. I. Kinetic analyses of active site mutants. J Biol Chem 269 (1994) 12893-8.

[54] K.M. Flaherty, S.M. Wilbanks, C. DeLuca-Flaherty, and D.B. McKay, Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. J Biol Chem 269 (1994) 12899-907.
[55] M.C. O'Brien, K.M. Flaherty, and D.B. McKay, Lysine 71 of the chaperone protein Hsc70 Is essential for ATP hydrolysis. J. Biol. Chem. 271 (1996) 15874-15878.
[56] D.R. Palleros, K.L. Reid, L. Shi, W.J. Welch, and A.L. Fink, ATP-induced protein-Hsp70 complex dissociation requires K+ but not ATP hydrolysis. Nature 365 (1993) 664-6.
[57] B. Bukau, and A.L. Horwich, The Hsp70 and Hsp60 chaperone machines. Cell 92 (1998) 351-66.
[58] R.C. Morshauser, H. Wang, G.C. Flynn, and E.R. Zuiderweg, The peptide-binding domain of the chaperone protein Hsc70 has an unusual secondary structure topology. Biochemistry 34 (1995) 6261-6.
[59] G.M. Clore, and A.M. Gronenborn, Applications of three- and four-dimensional heteronuclear NMR spectroscopy to protein structure determination. Progr. Nucl. Magn. Reson. Spectrosc. 23 (1991) 43-92.
[60] H. Wang, and E.R. Zuiderweg, HCCH-TOCSY spectroscopy of 13C-labeled proteins in H2O using heteronuclear cross-polarization and pulsed-field gradients. J Biomol NMR 5 (1995) 207-11.
[61] S.W. Van Doren, A.V. Kurochkin, Q.Z. Ye, L.L. Johnson, D.J. Hupe, and E.R.P. Zuiderweg, Assignments for the main chain nuclear magnetic resonances and delineation of the secondary structure of the catalytic domain of human stromelysin-1 as obtained from triple resonance 3D NMR experiments. Biochemistry 32 (1993) 13109-13122.
[62] A. Bax, and S. Grzesiek, Methodological Advances in Protein NMR. Acc. Chem. Res. 26 (1993) 131-138.
[63] S. Rudiger, L. Germeroth, J. Schneider-Mergener, and B. Bukau, Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. Embo J 16 (1997) 1501-7.
[64] A. Gragerov, Z. Li, Z. Xun, W. Burkholder, and M.E. Gottesman, Specificity of DnaK Peptide Binding. J. Mol. Biol. 235 (1994) 848-854.
[65] A. Gragerov, and M.E. Gottesman, Different peptide binding specificities of hsp70 family members. J Mol Biol 241 (1994) 133-5.
[66] R.C. Morshauser, W. Hu, H. Wang, Y. Pang, G.C. Flynn, and E.R. Zuiderweg, High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. J Mol Biol 289 (1999) 1387-403.
[67] X.T. Zhu, X. Zhao, W.F. Burkholder, A. Gragerov, C.M. Ogata, M.E. Gottesman, and W.A. Hendrickson, Structural analysis of substrate binding by the molecular chaperone DnaK. Science 272 (1996) 1606-1614.
[68] S. Sadis, and L.E. Hightower, Unfolded proteins stimulate molecular chaperone Hsc70 ATPase by accelerating ADP/ATP exchange. Biochemistry 31 (1992) 9406-9412.

[69] K. Liberek, D. Skowyra, M. Zylicz, C. Johnson, and C. Georgopoulos, The Escherichia coli DnaK chaperone, the 70-kDa heat shock protein eukaryotic equivalent, changes conformation upon ATP hydrolysis, thus triggering its dissociation from a bound target protein. J Biol Chem 266 (1991) 14491-6.
[70] J.H. Ha, and D.B. McKay, ATPase kinetics of recombinant bovine 70 kDa heat shock cognate protein and its amino-terminal ATPase domain. Biochemistry 33 (1994) 14625-14635.
[71] H. Theyssen, H.P. Schuster, L. Packschies, B. Bukau, and J. Reinstein, The second step of ATP binding to DnaK induces peptide release. J. Mol. Biol. 263 (1996) 657-670.
[72] G.C. Flynn, T.G. Chappell, and J.E. Rothman, Peptide binding and release by proteins implicated as catalysts of protein assembly. Science 245 (1989) 385-90.
[73] B. Gao, L. Greene, and E. Eisenberg, Characterization of nucleotide-free uncoating ATPase and its binding to ATP, ADP, and ATP analogues. Biochemistry 33 (1994) 2048-54.
[74] K. Liberek, J. Marszalek, D. Ang, C. Georgopoulos, and M. Zylicz, Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. Proc Natl Acad Sci U S A 88 (1991) 2874-8.
[75] M. Zylicz, D. Ang, K. Liberek, and C. Georgopoulos, Initiation of lambda DNA replication with purified host- and bacteriophage-encoded proteins: the role of the dnaK, dnaJ and grpE heat shock proteins. Embo J 8 (1989) 1601-8.
[76] S. Wickner, J. Hoskins, and K. McKenney, Function of DnaJ and DnaK as chaperones in origin-specific DNA binding by RepA. Nature 350 (1991) 165-7.
[77] S. Wickner, J. Hoskins, and K. McKenney, Monomerization of RepA dimers by heat shock proteins activates binding to DNA replication origin. Proc Natl Acad Sci U S A 88 (1991) 7903-7.
[78] K. Liberek, and C. Georgopoulos, Autoregulation of the Escherichia coli heat shock response by the DnaK and DnaJ heat shock proteins. Proc Natl Acad Sci U S A 90 (1993) 11019-23.
[79] T. Szyperski, M. Pellecchia, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure determination of the Escherichia coli DnaJ molecular chaperone: secondary structure and backbone fold of the N-terminal region (residues 2-108) containing the highly conserved J domain. Proc Natl Acad Sci U S A 91 (1994) 11343-7.
[80] M. Pellecchia, T. Szyperski, D. Wall, C. Georgopoulos, and K. Wuthrich, NMR structure of the J-domain and the Gly/Phe-rich region of the Escherichia coli DnaJ chaperone. J. Mol. Biol. 260 (1996) 236-250.
[81] P.A. Silver, and J.C. Way, Eukaryotic DnaJ homologs and the specificity of Hsp70 activity. Cell 74 (1993) 5-6.
[82] C. Alfano, and R. McMacken, Heat shock protein-mediated disassembly of nucleoprotein structures is required for the initiation of bacteriophage lambda DNA replication. J Biol Chem 264 (1989) 10709-18.

[83] D. Wall, M. Zylicz, and C. Georgopoulos, The NH2-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. J Biol Chem 269 (1994) 5446-51.
[84] S. Ahle, and E. Ungewickell, Auxilin, a newly identified clathrin-associated protein in coated vesicles from bovine brain. J Cell Biol 111 (1990) 19-29.
[85] S.E. Holstein, H. Ungewickell, and E. Ungewickell, Mechanism of clathrin basket dissociation: separate functions of protein domains of the DnaJ homologue auxilin. J Cell Biol 135 (1996) 925-37.
[86] E. Ungewickell, H. Ungewickell, S.E. Holstein, R. Lindner, K. Prasad, W. Barouch, B. Martin, L.E. Greene, and E. Eisenberg, Role of auxilin in uncoating clathrin-coated vesicles. Nature 378 (1995) 632-5.
[87] E. Ungewickell, H. Ungewickell, and S.E. Holstein, Functional interaction of the auxilin J domain with the nucleotide- and substrate-binding modules of Hsc70. J Biol Chem 272 (1997) 19594-600.
[88] J. Jiang, A.B. Taylor, K. Prasad, Y. Ishikawa-Brush, P.J. Hart, E.M. Lafer, and R. Sousa, Structure-function analysis of the auxilin J-domain reveals an extended Hsc70 interaction interface. Biochemistry 42 (2003) 5748-53.
[89] L. Packschies, H. Theyssen, A. Buchberger, B. Bukau, R.S. Goody, and J. Reinstein, GrpE accelerates nucleotide exchange of the molecular chaperone DnaK with an associative displacement mechanism. Biochemistry 36 (1997) 3417-22.
[90] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-5.
[91] J.F. Swain, G. Dinler, R. Sivendran, D.L. Montgomery, M. Stotz, and L.M. Gierasch, Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. Mol Cell 26 (2007) 27-39.
[92] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. Proc Natl Acad Sci U S A 106 (2009) 8471-6.
[93] Q. Liu, and W.A. Hendrickson, Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-20.
[94] J. Jiang, E.G. Maes, A.B. Taylor, L. Wang, A.P. Hinck, E.M. Lafer, and R. Sousa, Structural basis of J cochaperone binding and regulation of Hsp70. Mol Cell 28 (2007) 422-33.
[95] W.C. Suh, W.F. Burkholder, C.Z. Lu, X. Zhao, M.E. Gottesman, and C.A. Gross, Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. Proc Natl Acad Sci U S A 95 (1998) 15223-8.
[96] J. Monod, J. Wyman, and J.P. Changeux, On the Nature of Allosteric Transitions: A Plausible Model. J Mol Biol 12 (1965) 88-118.
[97] E.E. Benson, and K. Linderstrom-Lang, Deuterium exchange between myoglobin and water. Biochim Biophys Acta 32 (1959) 579-81.
[98] D.E. Koshland, Jr., W.J. Ray, Jr., and M.J. Erwin, Protein structure and enzyme action. Fed Proc 17 (1958) 1145-50.

[99] D.E. Koshland, Application of a Theory of Enzyme Specificity to Protein Synthesis. Proc Natl Acad Sci U S A 44 (1958) 98-104.
[100] D.E. Koshland, Jr., G. Nemethy, and D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5 (1966) 365-85.
[101] P.J. Artymiuk, C.C. Blake, D.E. Grace, S.J. Oatley, D.C. Phillips, and M.J. Sternberg, Crystallographic studies of the dynamic properties of lysozyme. Nature 280 (1979) 563-8.
[102] M.J. Sternberg, D.E. Grace, and D.C. Phillips, Dynamic information from protein crystallography. An analysis of temperature factors from refinement of the hen egg-white lysozyme structure. J Mol Biol 130 (1979) 231-52.
[103] J.M. Yon, D. Perahia, and C. Ghelis, Conformational dynamics and enzyme activity. Biochimie 80 (1998) 33-42.
[104] H.J. Berendsen, and S. Hayward, Collective protein dynamics in relation to function. Curr Opin Struct Biol 10 (2000) 165-9.
[105] M.H. Levitt, Spin Dynamics: Basic Principles of NMR Spectroscopy. (2001).
[106] A.L. Ghatak, S, Quantum Mechanics: Theory And Applications. (2004).
[107] W.J.F. John Cavanagh, Arthur G. Palmer III, Nicholas J. Skelton, Protein NMR Spectroscopy: Principles and Practice. (2007).
[108] D.J. Griffiths, Introduction to Electrodynamics (3rd Edition). (1999).
[109] M. Ikura, L.E. Kay, and A. Bax, A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry 29 (1990) 4659-4667.
[110] M. Ikura, M. Krinks, D.A. Torchia, and A. Bax, An efficient NMR approach for obtaining sequence-specific resonance assignments of larger proteins based on multiple isotopic labeling. FEBS Lett 266 (1990) 155-8.
[111] M. Ikura, D. Marion, L.E. Kay, H. Shih, M. Krinks, C.B. Klee, and A. Bax, Heteronuclear 3D NMR and isotopic labeling of calmodulin. Towards the complete assignment of the 1H NMR spectrum. Biochem Pharmacol 40 (1990) 153-60.
[112] L.E. Kay, M. Ikura, R. Tschudin, and A. Bax, 3-Dimensional Triple-Resonance NMR-Spectroscopy of Isotopically Enriched Proteins. J. Magn. Reson. 89 (1990) 496-514.
[113] A. Bax, and M. Ikura, An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. J Biomol NMR 1 (1991) 99-104.
[114] A. Bax, M. Ikura, L.E. Kay, G. Barbato, and S. Spera, Multidimensional triple resonance NMR spectroscopy of isotopically uniformly enriched proteins: a powerful new strategy for structure determination. Ciba Found Symp 161 (1991) 108-19; discussion 119-35.

[115] G.T. Montelione, and G. Wagner, Conformation-Independent Sequential Nmr Connections in Isotope-Enriched Polypeptides by ^{1}H-^{13}C-^{15}N Triple-Resonance Experiments. J. Magn. Reson. 87 (1990) 183-188.
[116] R.T. Clubb, V. Thanabal, C. Osborne, and G. Wagner, 1H and 15N resonance assignments of oxidized flavodoxin from Anacystis nidulans with 3D NMR. Biochemistry 30 (1991) 7718-30.
[117] D.M. LeMaster, L.E. Kay, A.T. Brunger, and J.H. Prestegard, Protein dynamics and distance determination by NOE measurements. FEBS Lett 236 (1988) 71-6.
[118] A. Kumar, R.R. Ernst, and K. Wuthrich, A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. Biochem Biophys Res Commun 95 (1980) 1-6.
[119] G. Wagner, A novel application of nuclear Overhauser enhancement (NOE) in proteins: analysis of correlated events in the exchange of internal labile protons. Biochem Biophys Res Commun 97 (1980) 614-20.
[120] E.R.P. Zuiderweg, K. Hallenga, and E.T. Olejniczak, Improvement of 2D NOE spectra of biomacromolecules in H_2O solution by coherent suppression of the solvent resonance. J. Magn. Reson. 70 (1986) 336-343.
[121] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, and G.M. Clore, The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160 (2003) 65-73.
[122] G.M. Clore, M.R. Starich, and A.M. Gronenborn, Measurement of Residual Dipolar Couplings of Macromolecules Aligned in the Nematic Phase of a Colloidal Suspension of Rod-Shaped Viruses. J. Am. Chem. Soc. 120 (1998) 10571 -10572.
[123] M.R. Hansen, L. Mueller, and A. Pardi, Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol. 5 (1998) 1065-1074.
[124] M. Ottiger, F. Delaglio, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J Magn Reson 131 (1998) 373-8.
[125] M. Cai, H. Wang, E.T. Olejniczak, R.P. Meadows, A.H. Gunasekera, N. Xu, and S.W. Fesik, Accurate measurement of H^{N}-H^{α} residual dipolar couplings in proteins. J. Magn. Reson. 139 (1999) 451-453.
[126] J.J. Chou, S. Li, and A. Bax, Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. J Biomol NMR 18 (2000) 217-27.
[127] J. Meiler, N. Blomberg, M. Nilges, and C. Griesinger, A new approach for applying residual dipolar couplings as restraints in structure elucidation. J Biomol NMR 16 (2000) 245-52.
[128] J.H. Prestegard, H.M. al-Hashimi, and J.R. Tolman, NMR structures of biomolecules using field oriented media and residual dipolar couplings. Q. Rev. Biophys. 33 (2000) 371-424.

[129] A. Bax, G. Kontaxis, and N. Tjandra, Dipolar couplings in macromolecular structure determination. Methods Enzymol 339 (2001) 127-74.
[130] H.M. Al-Hashimi, H. Valafar, M. Terrell, E.R. Zartler, M.K. Eidsness, and J.H. Prestegard, Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. J Magn Reson 143 (2000) 402-6.
[131] H.M. Al-Hashimi, and D.J. Patel, Residual dipolar couplings: synergy between NMR and structural genomics. J Biomol NMR 22 (2002) 1-8.
[132] E. de Alba, and N. Tjandra, Residual dipolar couplings in protein structure determination. Methods Mol Biol 278 (2004) 89-106.
[133] J.R. Tolman, and K. Ruan, NMR residual dipolar couplings as probes of biomolecular dynamics. Chem Rev 106 (2006) 1720-36.
[134] W. Jahnke, Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy. Chembiochem 3 (2002) 167-73.
[135] J.L. Battiste, and G. Wagner, Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. Biochemistry 39 (2000) 5355-65.
[136] J. Iwahara, D.E. Anderson, E.C. Murphy, and G.M. Clore, EDTA-derivatized deoxythymidine as a tool for rapid determination of protein binding polarity to DNA by intermolecular paramagnetic relaxation enhancement. J Am Chem Soc 125 (2003) 6634-5.
[137] I. Bertini, and C. Luchinat, New applications of paramagnetic NMR in chemical biology. Curr Opin Chem Biol 3 (1999) 145-51.
[138] I. Bertini, C. Luchinat, and M. Piccioli, Paramagnetic probes in metalloproteins. Methods Enzymol 339 (2001) 314-40.
[139] P.R. Rosevear, Determination of Protein Structures Using Paramagnetic Labels, ACS-Midwest Sectional Meeting, Cleveland, OH, 1998.
[140] J. Iwahara, and G.M. Clore, Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature 440 (2006) 1227-1230.
[141] J. Iwahara, C. Tang, and G. Marius Clore, Practical aspects of 1H transverse paramagnetic relaxation enhancement measurements on macromolecules. Journal of Magnetic Resonance 184 (2007) 185-195.
[142] J. Iwahara, C.D. Schwieters, and G.M. Clore, Ensemble approach for NMR structure refinement against (1)H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. J Am Chem Soc 126 (2004) 5879-96.
[143] H. Schroder, T. Langer, F.U. Hartl, and B. Bukau, Dnak, Dnaj and Grpe Form a Cellular Chaperone Machinery Capable of Repairing Heat-Induced Protein Damage. Embo J. 12 (1993) 4137-4144.
[144] C. Garrido, M. Brunet, C. Didelot, Y. Zermati, E. Schmitt, and G. Kroemer, Heat shock proteins 27 and 70: anti-apoptotic proteins with tumorigenic properties. Cell Cycle 5 (2006) 2592-601.

[145] J. Nylandsted, K. Brand, and M. Jaattela, Heat shock protein 70 is required for the survival of cancer cells. Ann N Y Acad Sci 926 (2000) 122-5.
[146] R. Wadhwa, S.C. Kaul, and Y. Mitsui, Cellular mortality to immortalization: mortalin. Cell Struct Funct 19 (1994) 1-10.
[147] R. Wadhwa, S. Takano, K. Kaur, C.C. Deocaris, O.M. Pereira-Smith, R.R. Reddel, and S.C. Kaul, Upregulation of mortalin/mthsp70/Grp75 contributes to human carcinogenesis. Int J Cancer 118 (2006) 2973-80.
[148] J.E. Gestwicki, G.R. Crabtree, and I.A. Graef, Harnessing chaperones to generate small-molecule inhibitors of amyloid beta aggregation. Science 306 (2004) 865-9.
[149] K.K. Chung, and T.M. Dawson, Parkin and Hsp70 sacked by BAG5. Neuron 44 (2004) 899-901.
[150] S. Krobitsch, and S. Lindquist, Aggregation of huntingtin in yeast varies with the length of the polyglutamine expansion and the expression of chaperone proteins. Proc Natl Acad Sci U S A 97 (2000) 1589-94.
[151] T.V. Novoselova, B.A. Margulis, S.S. Novoselov, A.M. Sapozhnikov, J. van der Spuy, M.E. Cheetham, and I.V. Guzhova, Treatment with extracellular HSP70/HSC70 protein can reduce polyglutamine toxicity and aggregation. J. Neurochemistry 94 (2005) 597-606.
[152] J.L. Brodsky, and G. Chiosis, Hsp70 molecular chaperones: emerging roles in human disease and identification of small molecule modulators. Curr Top Med Chem 6 (2006) 1215-25.
[153] M.P. Mayer, S. Rudiger, and B. Bukau, Molecular basis for interactions of the DnaK chaperone with substrates. Biol Chem 381 (2000) 877-85.
[154] J.S. McCarty, A. Buchberger, J. Reinstein, and B. Bukau, The role of ATP in the functional cycle of the DnaK chaperone system. J. Mol. Biol. 249 (1995) 126-137.
[155] K.M. Flaherty, C. Deluca-Flaherty, and D.B. McKay, 3-Dimensional Structure of the ATPase Fragment of a 70k Heat-Shock Cognate Protein. Nature 346 (1990) 623-628.
[156] H. Sondermann, C. Scheufler, C. Schneider, J. Hohfeld, F.U. Hartl, and I. Moarefi, Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. Science 291 (2001) 1553-7.
[157] H. Wang, Y. Pang, A.V. Kurochkin, W. Hu, G.C. Flynn, and E.R.P. Zuiderweg, The solution structure of the 21 kDa chaperone protein DnaK substrate binding domain: a preview of chaperone - protein interaction. Biochemistry 37 (1998) 7929-7940.
[158] M. Pellecchia, D.L. Montgomery, S.Y. Stevens, C.W. Vander Kooi, H.P. Feng, L.M. Gierasch, and E.R. Zuiderweg, Structural insights into substrate binding by the molecular chaperone DnaK. Nat Struct Biol 7 (2000) 298-303.
[159] S.Y. Stevens, S. Cai, M. Pellecchia, and E.R. Zuiderweg, The solution structure of the bacterial HSP70 chaperone protein domain DnaK(393-507) in complex with the peptide NRLLLTG. Protein Sci 12 (2003) 2588-96.

[160] M. Vogel, M.P. Mayer, and B. Bukau, Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. J Biol Chem 281 (2006) 38705-11.
[161] E.B. Bertelsen, L. Chang, J.E. Gestwicki, and E.R. Zuiderweg, Solution conformation of wild-type E.coli Hsp70 (DnaK) chaperone
complexed with ADP and substrate.
. Proc. Natl. Acad. Sci. U. S. A. (2009) in the press.
[162] Q. Liu, and W.A. Hendrickson, Insights into hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. Cell 131 (2007) 106-120.
[163] S.M. Wilbanks, L.L. Chen, H. Tsuruta, K.O. Hodgson, and D.B. Mckay, Solution Small-Angle X-Ray-Scattering Study of the Molecular Chaperone Hsc70 and Its Subfragments. Biochemistry 34 (1995) 12095-12106.
[164] M.C. O'Brien, and D.B. McKay, Threonine 204 of the chaperone protein Hsc70 influences the structure of the active site, but is not essential for ATP hydrolysis. J. Biol. Chem. 268 (1993) 24323-24329.
[165] D.B. McKay, Structure and mechanism of 70-kDa heat-shock-related proteins. Adv. Protein Chem. 44 (1993) 67-98.
[166] S.M. Wilbanks, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. II. Potassium binds specifically in the ATPase active site. J Biol Chem 270 (1995) 2251-7.
[167] J. Jiang, K. Prasad, E.M. Lafer, and R. Sousa, Structural basis of interdomain communication in the Hsc70 chaperone. Mol. Cell 20 (2005) 513-524.
[168] M.C. O'Brien, and D.B. McKay, How potassium affects the activity of the molecular chaperone Hsc70. I. Potassium is required for optimal ATPase activity. J. Biol. Chem. 270 (1995) 2247-2250.
[169] D. Klostermeier, R. Seidel, and J. Reinstein, Functional properties of the molecular chaperone DnaK from Thermus thermophilus. J Mol Biol 279 (1998) 841-53.
[170] 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. Proc. Natl. Acad. Sci. USA 92 (1995) 9279-9283.
[171] M.W. Fischer, J.A. Losonczi, J.L. Weaver, and J.H. Prestegard, Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. Biochemistry 38 (1999) 9013-9022.
[172] M. Revington, and E.R. Zuiderweg, TROSY-driven NMR backbone assignments of the 381-residue nucleotide-binding domain of the Thermus Thermophilus DnaK molecular chaperone. J Biomol NMR 30 (2004) 113-4.
[173] N. Tjandra, J.G. Omichinski, A.M. Gronenborn, G.M. Clore, and A. Bax, Use of dipolar ^{1}H-^{15}N and ^{1}H-^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat. Struct. Biol. 4 (1997) 732-738.
[174] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, and J.H. Prestegard, Structural and dynamic analysis of residual dipolar coupling data for proteins. J. Am. Chem. Soc. 123 (2001) 1416-1424.

[175] D. Yang, J.R. Tolman, N.T. Goto, and L.E. Kay, An HNCO-based Pulse Scheme for the Measurement of $^{13}C[agr]$-$^{1}H[agr]$ One-bond Dipolar couplings in ^{15}N, ^{13}C Labeled Proteins. J. Biomol. NMR 12 (1998) 325-332.
[176] D. Yang, R.A. Venters, G.A. Mueller, W.Y. Choy, and L.E. Kay, TROSY-based HNCO pulse sequences for the measurement of ^{1}HN-^{15}N, ^{15}N-^{13}CO, ^{1}HN-^{13}CO, ^{13}CO-$^{13}C[agr]$ and ^{1}HN-$^{13}C[agr]$ dipolar couplings in ^{15}N, ^{13}C, ^{2}H-labeled proteins. J. Biomol. NMR 14 (1999) 333-343.
[177] H. Valafar, and J.H. Prestegard, REDCAT: a residual dipolar coupling analysis tool. J Magn Reson 167 (2004) 228-41.
[178] S.M. Wilbanks, L. Chen, H. Tsuruta, K.O. Hodgson, and D.B. McKay, Solution small-angle X-ray scattering study of the molecular chaperone Hsc70 and its subfragments. Biochemistry 34 (1995) 12095-12106.
[179] C.J. Harrison, M. Hayer-Hartl, M. Di Liberto, F. Hartl, and J. Kuriyan, Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. Science 276 (1997) 431-435.
[180] Z. Polier S Fau - Dragovic, F.U. Dragovic Z Fau - Hartl, A. Hartl Fu Fau - Bracher, and A. Bracher, Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. Cell 133 (2008).
[181] J. Schuermann Jp Fau - Jiang, J. Jiang J Fau - Cuellar, O. Cuellar J Fau - Llorca, L. Llorca O Fau - Wang, L.E. Wang L Fau - Gimenez, S. Gimenez Le Fau - Jin, A.B. Jin S Fau - Taylor, B. Taylor Ab Fau - Demeler, K.A. Demeler B Fau - Morano, P.J. Morano Ka Fau - Hart, J.M. Hart Pj Fau - Valpuesta, E.M. Valpuesta Jm Fau - Lafer, R. Lafer Em Fau - Sousa, and R. Sousa, Structure of the Hsp110:Hsc70 nucleotide exchange machine. Molecular Cell 31 (2008).
[182] H.M. Al-Hashimi, Y. Gosser, A. Gorin, W. Hu, A. Majumdar, and D.J. Patel, Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. J Mol Biol 315 (2002) 95-102.
[183] Y. Zhang, and E.R. Zuiderweg, The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. Proc Natl Acad Sci U S A 101 (2004) 10272-7.
[184] W. Rist, C. Graf, B. Bukau, and M.P. Mayer, Amide hydrogen exchange reveals conformational changes in hsp70 chaperones important for allosteric regulation. J Biol Chem 281 (2006) 16493-501.
[185] A. Buchberger, H. Theyssen, H. Schroder, J.S. McCarty, G. Virgallita, P. Milkereit, J. Reinstein, and B. Bukau, Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication. J. Biol. Chem. 270 (1995) 16903-16910.
[186] M. Revington, Y. Zhang, G.N. Yip, A.V. Kurochkin, and E.R. Zuiderweg, NMR investigations of allosteric processes in a two-domain Thermus thermophilus Hsp70 molecular chaperone. J Mol Biol 349 (2005) 163-83.

[187] F. Ottiger M Fau - Delaglio, A. Delaglio F Fau - Bax, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra.
[188] J.A. Lukin, G. Kontaxis, V. Simplaceanu, Y. Yuan, A. Bax, and C. Ho, Quaternary structure of hemoglobin in solution. Proc Natl Acad Sci U S A 100 (2003) 517-20.
[189] M. Ikura, L.E. Kay, and A. Bax, A novel approach for sequential assignment of 1H, 13C, and 15N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. Biochemistry 29 (1990) 4659-67.
[190] M. Ikura, L.E. Kay, M. Krinks, and A. Bax, Triple-resonance multidimensional NMR study of calmodulin complexed with the binding domain of skeletal muscle myosin light-chain kinase: indication of a conformational change in the central helix. Biochemistry 30 (1991) 5498-504.
[191] K. Wuthrich, Protein structure determination in solution by nuclear magnetic resonance spectroscopy. Science 243 (1989) 45-50.
[192] K. Wuthrich, Protein structure determination in solution by NMR spectroscopy. J Biol Chem 265 (1990) 22059-62.
[193] K. Wuthrich, Six years of protein structure determination by NMR spectroscopy: what have we learned? Ciba Found Symp 161 (1991) 136-45; discussion 145-9.
[194] C.D. Schwieters, and G.M. Clore, The VMD-XPLOR visualization package for NMR structure refinement. J Magn Reson 149 (2001) 239-44.
[195] P. Mandal, and A. Majumdar, A comprehensive discussion of HSQC and HMQC pulse sequences. Concepts in Magnetic Resonance Part A 20 (2004) 1-23.
[196] N. Tjandra, and A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278 (1997) 1111-1114.
[197] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, and J.H. Prestegard, NMR evidence for slow collective motions in cyanometmyoglobin. Nat. Struct. Biol. 4 (1997) 292-297.
[198] N. Tjandra, Establishing a degree of order: obtaining high-resolution NMR structures from molecular alignment. Structure 7 (1999) R205-11.
[199] M. Ottiger, F. Delaglio, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J. Magn. Reson. 131 (1998) 373-378.
[200] K. Pervushin, R. Riek, G. Wider, and K. Wuthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A 94 (1997) 12366-71.
[201] A. Sheth, M. Ravikumar, and R. Hosur, Application of J scaling in two-dimensional spin-echo-correlated spectroscopy to observation of small coupling correlations. Journal or Magnetic Resonance 74 (1969) 352-355.

[202] M. Revington, Y. Zhang, G.N. Yip, A.V. Kurochkin, and E.R. Zuiderweg, NMR investigations of allosteric processes in a two-domain Thermus thermophilus Hsp70 molecular chaperone. J Mol Biol 349 (2005) 163-83.
[203] A. Bhattacharya, A.V. Kurochkin, G.N. Yip, Y. Zhang, E.B. Bertelsen, and E.R. Zuiderweg, Allostery in Hsp70 chaperones is transduced by subdomain rotations. J Mol Biol 388 (2009) 475-90.
[204] D.S. Weaver, and E.R. Zuiderweg, Protein proton-proton dynamics from amide proton spin flip rates. J Biomol NMR 45 (2009) 99-119.
[205] M.P. Mayer, and B. Bukau, Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci 62 (2005) 670-84.
[206] M.C. Sousa, and D.B. McKay, The hydroxyl of threonine 13 of the bovine 70-kDa heat shock cognate protein is essential for transducing the ATP-induced conformational change. Biochemistry 37 (1998) 15392-9.
[207] M.W.F. Fischer, A. Majumdar, and E.R.P. Zuiderweg, Protein NMR relaxation: theory, applications and outlook. Prog. NMR Spectrosc. 33 (1998) 207-272.
[208] T. Schulte-Herbruggen, and O.W. Sorensen, Clean TROSY: compensation for relaxation-induced artifacts. J Magn Reson 144 (2000) 123-8.
[209] A. Abragam, The Principles of Nuclear Magnetism, Clarenden Press, Oxford, 1961.
[210] A. Carrington, and A. MacLachlan, Introduction to Magnetic Resonance with Applications to Chemistry and Chemical Physics, Harper & Row, New York, 1967.
[211] M. Revington, T.M. Holder, and E.R. Zuiderweg, NMR study of nucleotide-induced changes in the nucleotide binding domain of Thermus thermophilus Hsp70 chaperone DnaK: implications for the allosteric mechanism. J Biol Chem 279 (2004) 33958-67.
[212] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6 (1995) 277-293.
[213] T.D. Goddard, and D.G. Kneller, SPARKY 3. University of California, San Francisco (2000).
[214] A.L. Hansen, and H.M. Al-Hashimi, Insight into the CSA tensors of nucleobase carbons in RNA polynucleotides from solution measurements of residual CSA: towards new long-range orientational constraints. J Magn Reson 179 (2006) 299-307.
[215] J.A. Losonczi, M. Andrec, M.W. Fischer, and J.H. Prestegard, Order matrix analysis of residual dipolar couplings using singular value decomposition. J. Magn. Reson. 138 (1999) 334-342.
[216] C.A. McLellan, D.A. Raynes, and V. Guerriero, HspBP1, an Hsp70 cochaperone, has two structural domains and is capable of altering the conformation of the Hsp70 ATPase domain. J Biol Chem 278 (2003) 19017-22.
[217] M. Zweckstetter, G. Hummer, and A. Bax, Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. Biophys J 86 (2004) 3444-60.

[218] D.S. Williamson, J. Borgognoni, A. Clay, Z. Daniels, P. Dokurno, M.J. Drysdale, N. Foloppe, G.L. Francis, C.J. Graham, R. Howes, A.T. Macias, J.B. Murray, R. Parsons, T. Shaw, A.E. Surgenor, L. Terry, Y. Wang, M. Wood, and A.J. Massey, Novel adenosine-derived inhibitors of 70 kDa heat shock protein, discovered through structure-based design. J Med Chem 52 (2009) 1510-3.
[219] J.P. Hendrick, T. Langer, T.A. Davis, F.U. Hartl, and M. Wiedmann, Control of folding and membrane translocation by binding of the chaperone DnaJ to nascent polypeptides. Proc Natl Acad Sci U S A 90 (1993) 10216-20.
[220] A. Wawrzynow, B. Banecki, D. Wall, K. Liberek, C. Georgopoulos, and M. Zylicz, ATP hydrolysis is required for the DnaJ-dependent activation of DnaK chaperone for binding to both native and denatured protein substrates. J Biol Chem 270 (1995) 19307-11.
[221] D.M. Cyr, T. Langer, and M.G. Douglas, DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. Trends Biochem Sci 19 (1994) 176-81.
[222] M.K. Greene, K. Maskos, and S.J. Landry, Role of the J-domain in the cooperation of Hsp40 with Hsp70. Proc Natl Acad Sci U S A 95 (1998) 6108-13.
[223] D.M. Cyr, X. Lu, and M.G. Douglas, Regulation of Hsp70 function by a eukaryotic DnaJ homolog. J Biol Chem 267 (1992) 20927-31.
[224] D.P. Atencio, and M.P. Yaffe, MAS5, a yeast homolog of DnaJ involved in mitochondrial protein import. Mol Cell Biol 12 (1992) 283-91.
[225] N. Rowley, C. Prip-Buus, B. Westermann, C. Brown, E. Schwarz, B. Barrell, and W. Neupert, Mdj1p, a novel chaperone of the DnaJ family, is involved in mitochondrial biogenesis and protein folding. Cell 77 (1994) 249-59.
[226] M. Martinez-Yamout, G.B. Legge, O. Zhang, P.E. Wright, and H.J. Dyson, Solution structure of the cysteine-rich domain of the Escherichia coli chaperone protein DnaJ. J Mol Biol 300 (2000) 805-18.
[227] M. Zylicz, T. Yamamoto, N. McKittrick, S. Sell, and C. Georgopoulos, Purification and properties of the dnaJ replication protein of Escherichia coli. J Biol Chem 260 (1985) 7591-8.
[228] J. Gamer, H. Bujard, and B. Bukau, Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32. Cell 69 (1992) 833-42.
[229] T. Langer, C. Lu, H. Echols, J. Flanagan, M.K. Hayer, and F.U. Hartl, Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356 (1992) 683-9.
[230] S. Rudiger, A. Buchberger, and B. Bukau, Interaction of Hsp70 chaperones with substrates. Nat Struct Biol 4 (1997) 342-9.
[231] S.L. Schmid, W.A. Braell, D.M. Schlossman, and J.E. Rothman, A role for clathrin light chains in the recognition of clathrin cages by 'uncoating ATPase'. Nature 311 (1984) 228-31.

[232] A. Umeda, A. Meyerholz, and E. Ungewickell, Identification of the universal cofactor (auxilin 2) in clathrin coat dissociation. Eur J Cell Biol 79 (2000) 336-42.
[233] D.T. Haynie, and C.P. Ponting, The N-terminal domains of tensin and auxilin are phosphatase homologues. Protein Sci 5 (1996) 2643-6.
[234] Y. Ma, T. Greener, M.E. Pacold, S. Kaushal, L.E. Greene, and E. Eisenberg, Identification of domain required for catalytic activity of auxilin in supporting clathrin uncoating by Hsc70. J Biol Chem 277 (2002) 49267-74.
[235] T. Greener, X. Zhao, H. Nojima, E. Eisenberg, and L.E. Greene, Role of cyclin G-associated kinase in uncoating clathrin-coated vesicles from non-neuronal cells. J Biol Chem 275 (2000) 1365-70.
[236] A. Fotin, Y. Cheng, N. Grigorieff, T. Walz, S.C. Harrison, and T. Kirchhausen, Structure of an auxilin-bound clathrin coat and its implications for the mechanism of uncoating. Nature 432 (2004) 649-53.
[237] D. Wall, M. Zylicz, and C. Georgopoulos, The conserved G/F motif of the DnaJ chaperone is necessary for the activation of the substrate binding properties of the DnaK chaperone. J Biol Chem 270 (1995) 2139-44.
[238] W.L. Kelley, and C. Georgopoulos, The T/t common exon of simian virus 40, JC, and BK polyomavirus T antigens can functionally replace the J-domain of the Escherichia coli DnaJ molecular chaperone. Proc Natl Acad Sci U S A 94 (1997) 3679-84.
[239] W.C. Suh, C.Z. Lu, and C.A. Gross, Structural features required for the interaction of the Hsp70 molecular chaperone DnaK with its cochaperone DnaJ. J Biol Chem 274 (1999) 30534-9.
[240] K. Pervushin, G. Wider, and K. W¸thrich, Single transition-to-single transition polarization transfer (ST2-PT) in [15N, 1H]-TROSY. Journal of Biomolecular NMR 12 (1998) 345-348.
[241] G. Bertani, Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol 186 (2004) 595-600.
[242] K. Liberek, D. Wall, and C. Georgopoulos, The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. Proc Natl Acad Sci U S A 92 (1995) 6224-8.
[243] P.P. Van Veldhoven, and G.P. Mannaerts, Inorganic and organic phosphate measurements in the nanomolar range. Anal Biochem 161 (1987) 45-8.
[244] P. Ekman, and O. Jager, Quantification of subnanomolar amounts of phosphate bound to seryl and threonyl residues in phosphoproteins using alkaline hydrolysis and malachite green. Anal Biochem 214 (1993) 138-41.
[245] J.F. Swain, E.G. Schulz, and L.M. Gierasch, Direct comparison of a stable isolated Hsp70 substrate-binding domain in the empty and substrate-bound states. J Biol Chem 281 (2006) 1605-11.
[246] F. Moro, V. Fernandez-Saiz, and A. Muga, The allosteric transition in DnaK probed by infrared difference spectroscopy. Concerted ATP-induced rearrangement of the substrate binding domain. Protein Sci 15 (2006) 223-33.

[247] M.P. Foster, D.S. Wuttke, K.R. Clemens, W. Jahnke, I. Radhakrishnan, L. Tennant, M. Reymond, J. Chung, and P.E. Wright, Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA. J Biomol NMR 12 (1998) 51-71.
[248] V. Katta, and B.T. Chait, Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. Rapid Commun Mass Spectrom 5 (1991) 214-7.
[249] S.W. Englander, and N.R. Kallenbach, Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q Rev Biophys 16 (1983) 521-655.
[250] Y. Paterson, S.W. Englander, and H. Roder, An antibody binding site on cytochrome c defined by hydrogen exchange and two-dimensional NMR. Science 249 (1990) 755-9.
[251] S. Arumugam, C.L. Hemme, N. Yoshida, K. Suzuki, H. Nagase, M. Berjanskii, B. Wu, and S.R. Van Doren, TIMP-1 contact sites and perturbations of stromelysin 1 mapped by NMR and a paramagnetic surface probe. Biochemistry 37 (1998) 9650-9657.
[252] A. Kalk, and H. Berendsen, Proton magnetic relaxation and spin diffusion in proteins. J. Magn. Reson 24 (1976) 343ñ366.
[253] H. Takahashi, T. Nakanishi, K. Kami, Y. Arata, and I. Shimada, A novel NMR method for determining the interfaces of large protein-protein complexes. Nat. Struct. Biol. 7 (2000) 220-223.
[254] I. Shimada, NMR techniques for identifying the interface of a larger protein-protein complex: cross-saturation and transferred cross-saturation experiments. Methods Enzymol 394 (2005) 483-506.
[255] T. Nakanishi, M. Miyazawa, M. Sakakura, H. Terasawa, H. Takahashi, and I. Shimada, Determination of the interface of a large protein complex by transferred cross-saturation measurements. J Mol Biol 318 (2002) 245-9.
[256] S. Grzesiek, and A. Bax, Improved 3d Triple-Resonance Nmr Techniques Applied to a 31-Kda Protein. J. Magn. Reson. 96 (1992) 432-440.
[257] G. Wagner, A. Kumar, and K. Wuthrich, Systematic application of two-dimensional 1H nuclear-magnetic-resonance techniques for studies of proteins. 2. Combined use of correlated spectroscopy and nuclear Overhauser spectroscopy for sequential assignments of backbone resonances and elucidation of polypeptide secondary structures. Eur J Biochem 114 (1981) 375-84.
[258] E.R.P. Zuiderweg, R.M. Scheek, R. Boelens, W.F. Vangunsteren, and R. Kaptein, Determination of Protein Structures from Nuclear Magnetic-Resonance Data Using a Restrained Molecular-Dynamics Approach - the Lac Repressor DNA-Binding Domain. Biochim. 67 (1985) 707-715.
[259] R.M. Cooke, and I.D. Campbell, Protein structure determination by nuclear magnetic resonance. Bioessays 8 (1988) 52-6.
[260] G.M. Clore, A.T. Brunger, M. Karplus, and A.M. Gronenborn, Application of molecular dynamics with interproton distance restraints to three-

dimensional protein structure determination. A model study of crambin. J Mol Biol 191 (1986) 523-51.
[261] K. Wuthrich, Determination of three-dimensional protein structures in solution by nuclear magnetic resonance: an overview. Methods Enzymol 177 (1989) 125-31.
[262] D.S. Wishart, B.D. Sykes, and F.M. Richards, The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry 31 (1992) 1647-51.
[263] R.W. Montalvao, A. Cavalli, X. Salvatella, T.L. Blundell, and M. Vendruscolo, Structure determination of protein-protein complexes using NMR chemical shifts: case of an endonuclease colicin-immunity protein complex. J Am Chem Soc 130 (2008) 15990-6.
[264] Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G. Liu, A. Eletsky, Y. Wu, K.K. Singarapu, A. Lemak, A. Ignatchenko, C.H. Arrowsmith, T. Szyperski, G.T. Montelione, D. Baker, and A. Bax, Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci U S A 105 (2008) 4685-90.
[265] H.N. Moseley, and G.T. Montelione, Automated analysis of NMR assignments and structures for proteins. Curr Opin Struct Biol 9 (1999) 635-42.
[266] J.P. Linge, S.I. O'Donoghue, and M. Nilges, Automated assignment of ambiguous nuclear overhauser effects with ARIA. Methods Enzymol 339 (2001) 71-90.
[267] J.P. Linge, M. Habeck, W. Rieping, and M. Nilges, ARIA: automated NOE assignment and NMR structure calculation. Bioinformatics 19 (2003) 315-6.
[268] M. Habeck, W. Rieping, J.P. Linge, and M. Nilges, NOE assignment with ARIA 2.0: the nuts and bolts. Methods Mol Biol 278 (2004) 379-402.
[269] T. Herrmann, P. Guntert, and K. Wuthrich, Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 319 (2002) 209-27.
[270] T. Herrmann, P. Guntert, and K. Wuthrich, Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR 24 (2002) 171-89.
[271] H.J. Woo, J. Jiang, E.M. Lafer, and R. Sousa, ATP Induced Conformational Changes in Hsp70: Molecular Dynamics and Experimental Validation of an In Silico Predicted Conformation. Biochemistry (2009).
[272] A. Grishaev, J. Ying, M.D. Canny, A. Pardi, and A. Bax, Solution structure of tRNAVal from refinement of homology model against residual dipolar coupling and SAXS data. J Biomol NMR 42 (2008) 99-109.
[273] A. Grishaev, V. Tugarinov, L.E. Kay, J. Trewhella, and A. Bax, Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints. J Biomol NMR 40 (2008) 95-106.
[274] C.G. dos Remedios, and P.D. Moens, Fluorescence resonance energy transfer spectroscopy is a reliable "ruler" for measuring structural changes in

proteins. Dispelling the problem of the unknown orientation factor. J Struct Biol 115 (1995) 175-85.
[275] F.G. Prendergast, and K.G. Mann, Chemical and physical properties of aequorin and the green fluorescent protein isolated from Aequorea forskalea. Biochemistry 17 (1978) 3448-53.
[276] L.A. Ernst, R.K. Gupta, R.B. Mujumdar, and A.S. Waggoner, Cyanine dye labeling reagents for sulfhydryl groups. Cytometry 10 (1989) 3-10.
[277] W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell, and P. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. Journal of the American Chemical Society 117 (1995) 5179-5197.
[278] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J. Comput. Chem. 4 (1983) 187-217.
[279] M.M. Tirion, Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. Phys Rev Lett 77 (1996) 1905-1908.
[280] L.W. Yang, A.J. Rader, X. Liu, C.J. Jursa, S.C. Chen, H.A. Karimi, and I. Bahar, oGNM: online computation of structural dynamics using the Gaussian Network Model. Nucleic Acids Res 34 (2006) W24-31.
[281] Q. Zhang, A.C. Stelzer, C.K. Fisher, and H.M. Al-Hashimi, Visualizing spatially correlated dynamics that directs RNA conformational transitions. Nature 450 (2007) 1263-7.