

**Protein NMR studies of two systems
involved in bacterial pathogenicity**

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

vorgelegt von
Sigrun Rumpel
aus Göttingen

Göttingen 2006

D7

Referent: Prof. Dr. Christian Griesinger

Korreferent: Prof. Dr. Axel Zeck

Tag der mündlichen Prüfung: 01.11.2006

Zusammenfassung

Die vorliegende Arbeit beinhaltet neue Strukturdaten über bakterielle Proteine, die an der Transkriptionsregulation und Proteinsekretion beteiligt sind, sowie einen Beitrag zur Verbesserung der Strukturaufklärung von homodimeren Proteinen mittels NMR-Spektroskopie. Ein Hauptaugenmerk liegt auf der Untersuchung von zwei homodimeren bakteriellen Proteinen: Dem Transkriptionsfaktor CylR2 aus *Enterococcus faecalis* und dem Chaperon CesT aus enteropathogenem *Escherichia coli*.

Enterococcus faecalis hat sich zu einer Hauptursache von Infektionen in Krankenhäusern aufgrund von Antibiotikaresistenzen entwickelt, wobei die schwere der Infektion mit einem von *Enterococcus faecalis* sekretierten Protein, dem Cytolysin, zusammenhängt. Die Produktion dieses Cytolysins wird von einem zwei-Komponenten CylR1/CylR2 System über einen autoinduzierten Quorum-Sensing Mechanismus reguliert. In dieser Arbeit wird gezeigt, dass das regulatorische Protein CylR2 ein rigides Dimer formt. Die Kristall- und NMR-Struktur sind im wesentlichen identisch. Jedes Monomer besteht aus einem fünf Helix-Bündel, das ein Helix-Turn-Helix DNA-Bindingsmotiv enthält und durch ein zweisträngiges antiparalleles β -Faltblatt verlängert ist. Die Lösung der NMR-Struktur erfolgt mit Hilfe einer neuartigen Strukturbestimmungsmethode für homodimere Proteine. Um die beiden Untereinheiten als rigide Körper aneinander zu docken, werden experimentelle residuale dipolaren Kopplungen der Rückgrat-Amide und intermolekulare, langreichweitige Distanzen, die über paramagnetische Relaxationsverstärkung ermittelt werden, verwendet. Eine Modellstruktur für CylR2 im Komplex mit seiner spezifischen palindromischen DNA innerhalb des Cytolysinpromotors wird basierend auf der Veränderung der chemischen Verschiebungen von CylR2 bei DNA-Bindung erstellt. Diese Ergebnisse deuten auf eine Rolle von CylR2 als Repressor der Cytolysintranskription hin.

Enteropathogenes *Escherichia coli* ist eine Hauptursache von Durchfallerkrankungen, die durch Sekretion bakterieller Proteine über ein Type III Sekretionssystem in die menschliche Wirtszelle ausgelöst werden. CesT ist an dem Sekretionsmechanismus beteiligt indem es Effektorproteine vor der Sekretion spezifisch im bakteriellen Cytoplasma bindet und dadurch den Effektor in einem sekretionsbereiten Zustand hält. Die genaue Funktion des Chaperons und damit der Mechanismus der bakteriellen Type III Proteinsekretion ist noch ungeklärt. Mittels dipolarer Kopplungen wird gezeigt, dass CesT in Lösung, anders als im Kristall, eine den homologen Proteinen ähnliche Struktur ausbildet. Viele hydrophobe Interaktionen werden für die Erkennung von Effektoren durch CesT als wichtig identifiziert und abschliessend wird ein Modell für die Erkennung des Chaperon/Effektor Komplexes durch das Type III Sekretionssystem vorgeschlagen.

Abstract

The present work provides new structural information about bacterial proteins involved in transcription regulation and protein secretion, as well as a contribution to the improvement of structure determination of homodimeric proteins by NMR spectroscopy. A main focus is placed on two homodimeric bacterial proteins: the transcription factor CylR2 from *Enterococcus faecalis* and the chaperone CesT from enteropathogenic *Escherichia coli*.

Enterococcus faecalis has emerged as a leading agent of hospital-acquired antibiotic-resistant infections and the severity of its infections has been linked to a secreted protein called cytolysin. Production of this enterococcal cytolysin is regulated by the two-component CylR1/CylR2 system through an autoinduction quorum-sensing mechanism. Here, the regulatory protein CylR2 is found to form a rigid dimer with an essentially identical crystal and solution NMR structure. Each monomer contains a helix-turn-helix DNA-binding motif as part of a five helix-bundle, which is extended by an antiparallel β -sheet. The determination of the solution NMR structure involves the development of a novel method for homodimeric proteins. This method applies rigid-body docking driven by backbone amide residual dipolar couplings and intermolecular long-range distances from paramagnetic relaxation enhancement. A model structure for CylR2 in complex with its specific palindromic DNA within the cytolysin promoter region is derived based on NMR chemical shift perturbation experiments. These results suggest that CylR2 acts as a repressor of cytolysin transcription.

The other studied homodimer, the chaperone CesT, originates from enteropathogenic *Escherichia coli* which is a chief cause of diarrhea and involves secretion of bacterial proteins via a type three secretion system into the human host cell. CesT participates in the secretion mechanism by specifically binding effector proteins in the bacterial cytoplasm prior to their secretion; the effector is thereby kept in a secretion-competent form. However, the exact function of the chaperone is still open and remains a key question for understanding the bacterial type three protein secretion mechanism. A novel application of NMR dipolar couplings is used to elucidate the structure of CesT in solution. In this way CesT is shown to form a structure similar to its homologues. Many hydrophobic interactions are identified to be important for the complex formation with at least two of its effectors. Finally, a model is suggested for the recognition of the chaperone/effector complex by the type three secretion system.

Acknowledgements

During the years of this thesis many people have contributed in many different ways to the work. I have been very fortunate to work under a good atmosphere in a department with great scientists and with the most modern equipment. So I am deeply grateful and would like to express thanks.

Most importantly, my direct supervisor Dr. Markus Zweckstetter, for the interesting projects, his good ideas, patience and advice.

To Prof. Christian Griesinger for giving me the opportunity to do my PhD. in his department with the very stimulating scientific environment.

Dr. Stefan Becker for helpful guidance concerning protein expression, purification and crystallization.

To Prof. Axel ZeecK for accepting me as an external PhD. student in the Chemistry Faculty of the Georg-August University, Göttingen.

Vinesh Vijayan for the company during the time of this thesis and for a lot of help with the measurement of NMR spectra.

All technical assistants and especially Karin Giller for her many technical tricks. Kerstin Overkamp for the breaths of fresh air, the organization of volleyball games and other events and of course all the HPLC runs and mass spectra.

All members of the department “NMR based structural biology”: in particular, Dirk Lennartz, Monika Bayrhuber and Daniela Fischer, my officemates during different years of this work, for their cordiality, helpfulness and a lot of useful discussions. Pierre Montaville and Fernando Rodriguez-Castañeda for company and advice in the lab. Young-Sang Jung, Jochen Junker, Karel Kubíček, Marcel Reese and Jegannath Korukottu for help with different computational aspects. My trainees, Ekaterina Torbashevich and Hai-Young Kim, who have contributed to the work about the type three secretion system.

All my friends, from tennis, school and studies, who have made these years special and well-balanced between a lot of scientific work and “real” life.

Last but not least I would like to thank my parents for their endless support, my brother for all the joint adventures, and Christophe for all the special moments that gave me the strength to finish this thesis.

Publication list

This thesis is based on the following papers:

Chapter 4

- [1] S. Rumpel, A. Razeto, C. M. Pillar, V. Vijayan, A. Taylor, K. Giller, M. S. Gilmore, S. Becker and M. Zweckstetter. Structure and DNA-binding properties of the cytolysin regulator CylR2 from *Enterococcus faecalis*. *EMBO Journal*, 23(18):3632-3642, 2004.

Chapter 6

- [2] S. Rumpel, H. Y. Kim, V. Vijayan, S. Becker and M. Zweckstetter. Backbone resonance assignment of the homodimeric, 35 kDa chaperone CesT from enteropathogenic *Escherichia coli*. *Journal of Biomolecular NMR*, 31(4):377-378, 2005.

Chapter 5

- [3] S. Rumpel, S. Becker and M. Zeckstetter. High-resolution structure determination of the CylR2 homodimer using intermonomer distances from paramagnetic relaxation enhancement and NMR dipolar couplings. *submitted*.

Table of contents

Abbreviations	xvii
1 Introduction	1
1.1 Bacterial pathogenesis	1
1.2 NMR spectroscopy in structural biology	2
1.3 Macromolecular complexes	3
1.4 Protein symmetry and function	4
1.5 Rationale and Outline	4
2 Materials	7
2.1 Bacterial strains and plasmids	7
2.2 Oligodesoxyribonucleotides and peptides	9
2.3 Culture media and antibiotics	10
2.4 Buffers and solutions	11
2.5 Enzymes and Chemicals	13
2.6 Equipment	14
2.7 Software	15
3 Methods	17
3.1 Molecular biology methods	17
3.1.1 Agarose gel electrophoresis	17
3.1.2 Isolation and purification of DNA	17

3.1.3	Enzymatic modifications of DNA	18
3.1.4	Polymerase chain reaction	18
3.1.5	Site-directed mutagenesis	19
3.1.6	DNA sequencing	19
3.1.7	Concentration and purity of DNA	20
3.1.8	Transformation of <i>E. coli</i>	20
3.2	Protein expression	20
3.2.1	Cultivation and storage of <i>E. coli</i>	20
3.2.2	Cell lysis	21
3.3	Protein methods	21
3.3.1	Nickel-nitrilotriacetic acid agarose affinity chromatography	21
3.3.2	Ion exchange chromatography	22
3.3.3	Gel filtration	22
3.3.4	Reversed phase-high performance liquid chromatography	22
3.3.5	Cleavage with TEV-protease	22
3.3.6	Concentrating of proteins	23
3.3.7	Protein storage	23
3.3.8	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis	23
3.3.9	Edman degradation	24
3.3.10	Limited proteolysis	24
3.3.11	Peptide synthesis	24
3.3.12	Introduction of a paramagnetic center	25
3.4	Spectroscopic methods	25
3.4.1	Determination of protein concentration	25
3.4.2	Circular dichroism spectroscopy	26
3.4.3	Electrospray ionization mass spectrometry	26
3.5	NMR spectroscopy	26
3.5.1	Chemical shift assignment	27
3.5.2	Secondary structure determination	27
3.5.3	Residual dipolar couplings	28

3.5.4	NMR relaxation data	29
3.5.5	Mapping binding surfaces	31
3.5.6	Determination of protein-ligand dissociation constants	31
3.5.7	Long-range distances from paramagnetic relaxation enhancement	32
3.6	Structure calculation	34
3.6.1	Monomer structure calculation	34
3.6.2	Dimer structure calculation	34
3.6.3	<i>Ab initio</i> docking	34
3.6.4	Protein/DNA docking	35
3.7	Structure analysis	35
4	Structure and DNA-binding properties of CylR2	37
4.1	Introduction	37
4.2	Materials and methods	40
4.2.1	Expression and purification	40
4.2.2	Preparation of ds DNA	40
4.2.3	NMR experiments	41
4.2.4	DNA titration	41
4.2.5	Model of the CylR2/DNA complex	42
4.3	Results and discussion	43
4.3.1	Three dimensional structure of CylR2	43
4.3.2	Dimerisation	45
4.3.3	Solution NMR studies of CylR2 and CylR2/DNA complex	47
4.3.4	Mapping of the DNA-binding site	50
4.3.5	Model for the CylR2/DNA complex	53
4.3.6	Specificity of CylR2/DNA interactions	56
4.3.7	Interaction between CylR2 and CylR1	58
4.4	Conclusions	59
5	Structure of a homodimer using RDC and PRE data	61
5.1	Introduction	61

5.2	Materials and methods	63
5.2.1	Construction of CylR2-mutants	63
5.2.2	NMR sample preparation	63
5.2.3	NMR spectroscopy	64
5.2.4	Monomer structure calculation	64
5.2.5	Intermonomer distances	65
5.2.6	Homodimer structure calculation	65
5.2.7	Ranking of homodimer models from <i>ab initio</i> docking	66
5.3	Results	66
5.3.1	Monomer solution structure	66
5.3.2	Mutagenesis and spin-labeling	68
5.3.3	Long-range distances from PRE	72
5.3.4	Homodimeric solution structure	74
5.3.5	The ATCUN-motif as a paramagnetic position	78
5.3.6	Ranking of homodimer structures from <i>ab initio</i> docking	79
5.4	Discussion	80
5.5	Conclusions	82
6	Function of the chaperone CesT in protein secretion	85
6.1	Introduction	85
6.2	Materials and methods	93
6.2.1	Cloning	93
6.2.2	Protein expression	93
6.2.3	Purification	94
6.2.4	NMR experiments	96
6.3	Results and discussion	96
6.3.1	Experimental difficulties	96
6.3.2	Extent of the assignment of CesT	98
6.3.3	Secondary structure and dynamics of CesT in solution	99
6.3.4	Solution structure of CesT	100

6.3.5	CesT _{V116R} aggregates	102
6.3.6	Structural characterisation of Tir and Map	103
6.3.7	Crystallisation of the CesT/Map complex	105
6.3.8	Backbone assignment and secondary structure of Tir108	105
6.3.9	Effects of CesT binding on Tir and Map	107
6.3.10	Effect of CesT on effector regions outside the CBD	110
6.3.11	Effects of Tir and Map binding on CesT	110
6.3.12	Properties of the CesT/Map84 complex in solution	117
6.3.13	Conclusions	120
7	Summary and future considerations	123
	References	128
A	Appendix: Sample preparation	149
A.1	CylR2	149
A.2	CesT	150
A.3	Map	151
A.4	Tir	151
A.5	CesT/effector-complexes	153
B	Appendix: NMR experiments and chemical shift assignments	155
B.1	NMR experiments	155
B.2	Chemical shift assignments	157
C	Appendix: NMR restraints	165
C.1	Experimental residual dipolar couplings	165
C.2	Distance restraints for CylR2	170
C.3	Dihedral angle restraints for CylR2	180
D	Appendix: Xplor input files	181
D.1	Water refinement	181
D.2	Rigid body docking	188

D.3 Simulated annealing	195
Curriculum Vitae	203
Lebenslauf	204

List of figures

2.1	Plasmid map of pET16bTEV and pETDuet1	8
4.1	Model of cytolysin maturation and regulation of cytolysin production	38
4.2	Cytolysin promotor region	39
4.3	The 22-mer DNA sequence of IR1 used for the NMR experiments. . .	41
4.4	Sequence of CylR2 and stereo view of the crystal structure of CylR2 .	44
4.5	Superposition of the CylR2 monomer with SinR and 434 repressor . .	44
4.6	Dimer interface of CylR2	45
4.7	Dimer interfaces of CylR2 and GerE superposed	46
4.8	Superposition of one subunit of CylR2 to 434 repressor	47
4.9	Correlation between experimental and back-calculated HN-RDC for free CylR2 and DNA-bound CylR2	48
4.10	^{15}N - ^1H -NOE of CylR2 and CylR2 in complex with its DNA	50
4.11	Overlay of ^{15}N -HSQC spectra of free and DNA-bound CylR2	51
4.12	Chemical shift changes upon CylR2/DNA complex formation	52
4.13	Surface representation of CylR2	53
4.14	Model of the CylR2/DNA complex structure	55
4.15	Overlay of NMR titration ^{15}N -HSQC experiments of CylR2 with IR2	57
4.16	Chemical shift changes upon binding of CylR2 to unspecific DNA . .	58
4.17	Prediction of transmembrane helices for CylR1	59
5.1	Monomer solution structure of CylR2	68

5.2	Reaction of the MTSL spin label with the -SH group of a protein . . .	69
5.3	Chemical shift changes due to single cysteine mutations	69
5.4	Overall strategy to derive intermonomer distances from PRE in ho- modimers.	70
5.5	^{15}N -HSQC spectra of paramagnetic and diamagnetic CylR2 _{T55C} . . .	71
5.6	Accuracy of the intermolecular distances determined from PRE . . .	73
5.7	Homodimeric solution structure of CylR2	75
5.8	Comparison of the solution and crystal structure of CylR2	77
5.9	Overlay of ^{15}N -HSQC spectra of CylR2 and ATCUN-CylR2	79
5.10	NMR-based ranking of structural models from <i>ab initio</i> docking . . .	80
6.1	Schematic representation of a type three secretion system	86
6.2	Structures of TTSS chaperones	89
6.3	Structure of CesT	90
6.4	Structures of chaperone/effector complexes	91
6.5	^{15}N -HSQC spectrum of His ₆ -CesT	97
6.6	^{15}N -TROSY-HSQC spectrum of CesT	98
6.7	Secondary structure of CesT	99
6.8	Surface representation of the model and X-ray structure of CesT . . .	100
6.9	Correlation between experimental and predicted or back-calculated RDCs for CesT	101
6.10	ESI-MS spectrum of coexpressed and copurified CesT/Map101	104
6.11	Far-UV-CD spectra of His ₇ -Map84 and Tir108	104
6.12	Secondary chemical shifts of Tir108	106
6.13	Overlay of ^{15}N -HSQC spectra of free and bound Tir108	107
6.14	Chemical shift changes of Tir108 upon CesT/Tir108 complex formation	108
6.15	Overlay of ^{15}N -HSQC spectra of free His ₇ -Map84 and bound Map84 .	109
6.16	Overlay of ^{15}N -HSQC spectra of TirN and TirN bound to CesT . . .	111
6.17	Overlay of ^{15}N -TROSY-HSQC spectra of free and effector-bound CesT	112
6.18	Secondary structure of CesT bound to Map84	113

6.19	Chemical shift changes of CesT upon formation of the CesT/Map84 complex	113
6.20	Structure-based sequence alignment of TTSS chaperones	114
6.21	Mapping of shift changes upon Map84 binding to the surface of CesT	115
6.22	Chemical shift changes of CesT upon formation of the CesT/Tir108 complex	116
6.23	^{15}N - ^1H -NOE of CesT and Map84 in complex with each other	117
6.24	^{15}N -TROSY-HSQC spectrum of ^{15}N -labeled CesT/Map84 complex . .	119
6.25	Model of effector targeting to the TTSS through CesT	121
A.1	Expression and purification of CylR2	149
A.2	ESI-MS spectrum of CylR2 _{T55C}	150
A.3	ESI-MS spectrum of perdeuterated ^2H -His ₇ -CesT	150
A.4	Expression and purification of His ₇ -Map84	151
A.5	Expression and purification of Tir108	151
A.6	ESI-MS spectrum of Tir108	152
A.7	ESI-MS spectrum of TirN	152
A.8	Coexpression and copurification of CesT/Map84	153
A.9	ESI-MS spectrum of coexpressed and copurified CesT/Map84	153
A.10	ESI-MS spectrum of coexpressed and copurified CesT/Tir108	154
B.1	Dimerization of CylR2	155

List of tables

2.1	Name, genotype and reference for the used <i>E. coli</i> strains.	7
2.2	Plasmids used in this work.	8
2.3	Name, sequence and application of the applied oligonucleotides. . . .	9
2.4	Name, sequence and residues of CylR1 of the applied peptides.	10
2.5	Culture media and antibiotics	10
2.6	Buffers and solutions	11
2.7	Enzymes and chemicals	13
2.8	Equipment and supplier	14
2.9	Software	15
3.1	Composition of the 17.5 % SDS-gels	24
4.1	Best-fit of HN-RDCs for DNA-bound CylR2 to the X-ray structure of free CylR2.	49
5.1	Statistics for CylR2 monomer solution structure determination.	67
5.2	Statistics for CylR2 dimer solution structure determination.	75
B.1	NMR spectra for assignment and structure determination of CylR2	156
B.2	NMR spectra for backbone assignment of CesT	156
B.3	NMR spectra for backbone assignment of Tir108	156
B.4	NMR spectra for backbone assignment of CesT in complex with Map84	157
B.5	NMR spectra for backbone assignment of CesT in complex with Tir108	157
B.6	NMR spectra for backbone assignment of Map84 in complex with CesT	157

B.7	Assigned chemical shifts of CylR2 in complex with its DNA	158
B.8	Assigned chemical shifts of CesT bound to Map84	159
B.9	Assigned chemical shifts of CesT bound to Tir108	160
B.10	Assigned chemical shifts of Tir108	161
B.11	Assigned chemical shifts of Map84 in complex with CesT	162
C.1	HN-RDCs of CylR2	165
C.2	HN-RDCs of DNA-bound CylR2	166
C.3	HN-RDCs of CesT	167
C.4	NC'-RDCs of CesT	167
C.5	C α C'-RDCs of CesT	168
C.6	HN-RDCs of Tir108	169
C.7	Intramolecular NOE distances of CylR2	170
C.8	Intermolecular distances of CylR2 from PRE	178
C.9	Intermolecular NOE distances of CylR2	179
C.10	Dihedral angle restraints for CylR2	180

Abbreviations

A

A	<u>a</u> denine
A_λ	absorption at wavelength λ
Ala	<u>a</u> lanine (A)
Ap	<u>a</u> mpicillin
Arg	<u>a</u> rginine (R)
Asn	<u>a</u> sparagine (N)
Asp	<u>a</u> spartic acid (D)
ATCUN	<u>a</u> mino <u>t</u> erminal <u>Cu</u> (II)- and <u>Ni</u> (II)-binding
ATP/ADP	<u>a</u> denosine 5'- <u>t</u> ri <u>p</u> hosphate / <u>a</u> denosine 5'- <u>d</u> i <u>p</u> hosphate

B

bp	<u>b</u> ase <u>p</u> air
BMRB	<u>B</u> iological <u>M</u> agnetic <u>R</u> esonance <u>B</u> ank
BSA	<u>b</u> ovine <u>s</u> erum <u>a</u> lbumin

C

CBD	<u>c</u> haperone <u>b</u> inding <u>d</u> omain
CD	<u>c</u> ircular <u>d</u> ichroism
CesT	<u>c</u> haperone for <u>e</u> fficient <u>s</u> ecretion of <u>T</u> ir
CIAP	<u>c</u> alf <u>i</u> ntestine <u>a</u> lkaline <u>p</u> hosphatase
CyIR1	<u>c</u> ytolysin <u>r</u> egulator <u>1</u>
CyIR2	<u>c</u> ytolysin <u>r</u> egulator <u>2</u>

Cys	<u>c</u> ysteine (C)
D	
D_a	magnitude of the alignment tensor
DNA	<u>d</u> eoxyribo <u>n</u> ucleic <u>a</u> cid
DNase	<u>d</u> eoxyribo <u>n</u> uclease
dNTP	<u>d</u> eoxyribo <u>n</u> ucleotide tri <u>p</u> hos <u>p</u> hate
ds	<u>d</u> ouble- <u>s</u> tranded
DTT	<u>d</u> ithio <u>t</u> hreit <u>o</u> l
δ	chemical shift
E	
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. faecalis</i>	<i>Enterococcus faecalis</i>
<i>EHEC</i>	entero <u>h</u> emorrhagic <i>Escherichia coli</i>
ELISA	enzyme- <u>l</u> inked immunosorbent <u>a</u> ssay
<i>EPEC</i>	enteropathogenic <i>Escherichia coli</i>
EDTA	ethylene <u>d</u> iamine <u>t</u> etraacetic <u>a</u> cid
ESI	<u>e</u> lectrospray ionization
F	
FPLC	<u>f</u> ast <u>p</u> rotein <u>l</u> iquid <u>c</u> hromatography
G	
G	<u>g</u> uanine
Gln	<u>g</u> lutamine (Q)
Glu	<u>g</u> lutamic acid (E)
Gly	<u>g</u> lycine (G)
γ	nuclear gyromagnetic ratio
H	
h	Planck's constant (=6.6260693 x 10 ⁻³⁴ Js)
\hbar	reduced Planck's constant (=1.05457168 x 10 ⁻³⁴ Js)
HEPES	2-[4-(2- <u>h</u> ydroxyethyl)1-1 <u>p</u> iperaziny] <u>e</u> thansulfonic acid
His	<u>h</u> istidine (H)

HPLC	<u>h</u> igh <u>p</u> erformance <u>l</u> iquid <u>c</u> hromatography
HSQC	<u>h</u> eteronuclear <u>s</u> ingle- <u>q</u> uantum <u>c</u> oherence
HTH	<u>h</u> elix- <u>t</u> urn- <u>h</u> elix
I	
Ile	iso <u>l</u> eu <u>c</u> ine (I)
INEPT	<u>i</u> nsensitive <u>n</u> uclei <u>e</u> nhancement by <u>p</u> olarization <u>t</u> ransfer
IPAP	<u>i</u> n- <u>p</u> hase / <u>a</u> nti- <u>p</u> hase
IPTG	iso <u>p</u> ropyl- β -D- <u>t</u> hiogalactopyranoside
IR	<u>i</u> nverted <u>r</u> ep <u>e</u> at
K	
K_d	dissociation constant
kDa	<u>k</u> ilo- <u>D</u> alton (=10 ³ g/mol)
L	
LB	<u>L</u> uria <u>B</u> ertani
Leu	<u>l</u> eu <u>c</u> ine (L)
Lys	<u>l</u> ysine (K)
M	
M	mol/l
Map	<u>m</u> itochondrial <u>a</u> ssociated <u>p</u> rotein
MES	2-(N- <u>m</u> orpholino) <u>e</u> thane- <u>s</u> ulfonic acid
MCS	<u>m</u> ultiple <u>c</u> loning <u>s</u> ite
MeOH	methanol
Met	<u>m</u> ethionine (M)
MOPS	3-(N- <u>m</u> orpholino) <u>p</u> ropane- <u>s</u> ulfonic acid
MS	<u>m</u> ass <u>s</u> pectrometry
MTSL	(1-oxy-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)- <u>m</u> ethanethiosulfonate spin- <u>l</u> abel
MW	<u>m</u> olecular <u>w</u> eight
MWCO	<u>m</u> olecular <u>w</u> eight <u>c</u> utoff
μ_0	permeability of vacuum (=1.25663706144 x 10 ⁻⁶ NA ⁻²)

N

Ni-NTA	<u>n</u> ickel- <u>n</u> itri <u>l</u> o <u>t</u> riacetic <u>a</u> cid
NMR	<u>n</u> uclear <u>m</u> agnetic <u>r</u> esonance
NOE	<u>n</u> uclear <u>O</u> verhauser <u>e</u> ffect
NOESY	<u>n</u> uclear <u>O</u> verhauser <u>e</u> ffect <u>s</u> pectroscopy

O

OD	<u>o</u> ptical <u>d</u> ensity
----	---------------------------------

P

PAGE	<u>p</u> oly <u>a</u> crylamide <u>g</u> el <u>e</u> lectrophoresis
PCR	<u>p</u> olymerase <u>c</u> hain <u>r</u> eaction
PCS	<u>p</u> seudo <u>c</u> ontact <u>s</u> hift
PDB	<u>P</u> rotein <u>D</u> ata <u>B</u> ank
Pf1	<u>f</u> ilamentous <u>p</u> hage
Phe	<u>p</u> henylalanine (F)
PMSF	<u>p</u> henyl <u>m</u> ethylsulphonyl <u>f</u> luoride
ppm	<u>p</u> arts <u>p</u> er <u>m</u> illion (=10 ⁻⁶)
PRE	<u>p</u> aramagnetic <u>r</u> elaxation <u>e</u> nhancement
Pro	<u>p</u> roline (P)

R

R	rhom <u>b</u> icity of the alignment tensor
R ₁	longitudinal or spin-lattice relaxation rate
R ₂	transversal or spin-spin relaxation rate
RDC	<u>r</u> esidual <u>d</u> ipolar <u>c</u> oupling
rmsd	<u>r</u> oot <u>m</u> ean <u>s</u> quare <u>d</u> eviation
RP-HPLC	<u>r</u> eversed <u>p</u> hase- <u>h</u> igh <u>p</u> erformance <u>l</u> iquid <u>c</u> hromatography

S

SDS	<u>s</u> odium <u>d</u> odecyl <u>s</u> ulfate
SDSL	<u>s</u> ite- <u>d</u> irected <u>s</u> pin- <u>l</u> abeling
Ser	<u>s</u> erine (S)
ss	<u>s</u> ingle- <u>s</u> tranded

SVD	<u>s</u> ingular <u>v</u> alue <u>d</u> ecomposition
T	
T	<u>t</u> hymine
T₁	longitudinal or spin-lattice relaxation time
T₂	transversal or spin-spin relaxation time
TEMED	N,N,N',N'- <u>t</u> etramethylethylenediamine
TEV	<u>t</u> obacco <u>e</u> tch <u>v</u> irus
Tir	<u>t</u> ranslocated <u>i</u> ntimin <u>r</u> eceptor
TFA	<u>t</u> ri <u>f</u> luoroacetic <u>a</u> cid
Thr	<u>t</u> hreonine (T)
TOCSY	<u>t</u> otal <u>c</u> orrelation <u>s</u> pectroscopy
Tris	tris(hydroxymethyl)-amino-methan
TROSY	<u>t</u> ransverse <u>r</u> elaxation <u>o</u> ptimised <u>s</u> pectroscopy
Trp	<u>t</u> ryptophan (W)
TTSS	<u>t</u> ype <u>t</u> hree <u>s</u> ecretion <u>s</u> ystem
Tyr	<u>t</u> yrosine (Y)
τ	correlation time for the electron-nuclear interaction
τ_c	global or rotational correlation time
V	
Val	<u>v</u> aline (V)
Y	
YT	<u>y</u> east/ <u>t</u> ryptone



1

Introduction

Proteins (greek: protas = of primary importance) are essential to all life processes. Association of proteins with each other or with DNA leads to the formation of enzymatic and regulatory complexes that are integral to cellular control processes such as transcription, signal transduction and protein translocation. Numerous of these processes are involved in bacterial pathogenicity.

1.1 Bacterial pathogenesis

Bacterial pathogenicity is defined as the sum of biochemical mechanisms whereby micro-organisms cause disease in a host organism [1]. These infectious diseases are a main cause of death worldwide. Emerging new infectious diseases, re-emerging deadly infectious diseases and increasing prevalence of antimicrobial resistant strains mark a formidable threat to public health and welfare. Pathogenic bacteria utilize a versatile and flexible repertoire of mechanisms by which they exert influence over their hosts. Yet, as these mechanisms become better understood, there is significant new evidence that very different microbial pathogens use common strategies. Two general abilities underlie bacterial pathogenicity. Firstly, the capacity of tissue invasion: this encompasses mechanisms for adherence and initial multiplication, the ability to bypass or overcome host defence mechanisms and the production of extracellular substances which facilitate invasion. Secondly, the competence to produce toxins. There are two types of toxins. The first type, the exotoxins, are released from bacterial cells and may act at tissue sites remote from the site of bacterial growth while the second type,

the endotoxins, are cell-associated substances that are structural components of the cell walls of Gram-negative bacteria [2, 3].

Our understanding of biological systems and their interaction networks in general and of bacterial pathogenesis in particular, relies largely on the three-dimensional (3D) structures of proteins as well as protein/DNA- and protein/protein-complexes. Unraveling their mechanisms and functions may render them feasible for control and correction by drugs. Moreover, 3D structures play an important role in rational drug design by enabling the examination of interactions between proteins and potential drug compounds at atomic detail.

1.2 NMR spectroscopy in structural biology

Currently, NMR spectroscopy and X-ray crystallography are the most prominent techniques capable of determining 3D structures of macromolecules at atomic level. The most important technique used in the present study is NMR spectroscopy. For a detailed treatment of NMR spectroscopy, the reader is invited to consult the well-established books [4, 5, 6, 7].

The basic phenomenon of NMR, that energy levels of atomic nuclear spins are split by a magnetic field and that transitions between these energy levels can be induced by irradiating with a frequency equal to the energy difference between the two levels, was discovered in 1946 [8, 9]. In the last two decades, NMR spectroscopy has evolved enormously and become a powerful alternative structural elucidation tool to X-ray crystallography. In addition, NMR spectroscopy provides at least two important advantages compared to X-ray crystallography. First, the 3D structures are solved in aqueous solution and thus under physiologically-relevant conditions. Second, NMR spectroscopy offers a unique perspective on dynamics. Advances in NMR spectroscopy were acknowledged by Nobel prizes in 1991 [10] and in 2002 [11] and were accompanied by an ever-growing number of structures solved by NMR and deposited in the protein data bank (PDB, <http://www.rcsb.org/pdb>) [12]. In spite of the success of NMR spectroscopy, limits on the size of the studied systems remain. Yet, the symbiotic development of deuteration strategies [13], transverse relaxation

optimized spectroscopy (TROSY) [14, 15] and new computational methods [16, 17] have expanded the maximum molecular weight amenable to high-resolution structure determination by NMR spectroscopy from 10-15 kDa up to 30-40 kDa. However, the development of such novel methods remains a main research focus in NMR spectroscopy.

1.3 Macromolecular complexes

The vast majority of the reported structures in the PDB were solved by X-ray crystallography and often represent only monomeric proteins. This is due to the difficulty in production and crystallization of macromolecular assemblies compared to individual proteins. It is foreseeable that this gap between solved monomeric and complex structures will even widen as structural genomics efforts increase and mainly contribute monomeric structures. Hence, complex structures will continue to represent an important bottleneck in our understanding of cellular processes [18]. Improvement of existing methods as well as development of new tools for macromolecular structure determination are thus required. Most promising is the application of hybrid methods to improve efficiency, accuracy and completeness [19]. A prime example of combined approaches to complex structure determination is the combination of NMR, crystallography and computation. First, the structures of the free components of a macromolecular assembly can be solved by NMR or crystallography. Then, in the context of complex assemblies, efficient NMR tools can be combined with computational methods. Contact surfaces between the individual monomers can be identified by chemical shift perturbation or slight conformational changes upon complex formation can be identified by residual dipolar couplings (RDCs). Reviews covering the manifold theoretical, experimental and computational aspects of RDCs are available [20, 21, 22, 23].

1.4 Protein symmetry and function

A particular case of macromolecular assemblies are homooligomers which are presumed to form symmetrical complexes. They have evolved due to functional genetic and physicochemical needs and are favored because of stability [24]. In fact, most of the soluble and membrane-bound proteins form homooligomers. However, structural characterization of oligomeric states of proteins still remains a challenge. This is illustrated by the fact, that although 66 % of the structures in the PDB in 1993 were monomeric, only about one fifth of the proteins annotated in the SWISS-PROT for *E. coli* is monomeric. Out of the remaining four fifth of oligomeric proteins, 79 % form homooligomers [24]. In general, oligomerization plays a key role in protein function. These functions include the possibility of allosteric control, higher local concentration of active sites, larger binding surfaces and new active sites at subunit interfaces [25]. The over-representation of monomers and lack of homooligomers in the PDB is one reason for the lack of knowledge about the assembly of oligomers. One mechanism involved to form a dimer or higher oligomer is 3D domain swapping [25]. This 3D domain swapping is defined as the replacement of a portion of the tertiary structure of a protein with an identical structural element from a second protein molecule [26]. 3D domain swapping has gained recent interest because it can serve as a mechanism for reversible oligomerization and is implicated in amyloidosis [27].

1.5 Rationale and Outline

The work presented in this thesis brings new insights into two mechanisms involved in bacterial pathogenicity with a focus on advances in solution NMR structure determination of homodimeric proteins. Overall, an emphasis is placed on the interpretation of the complementary data from X-ray crystallography and NMR spectroscopy combined with essential *in vivo* data. Thus, the work attempts to achieve an in-depth understanding of molecular functions and mechanisms through structure determination or identification of interaction surfaces between biomolecules with very recent tools for liquid state NMR.

The work at hand is organized as follows: Chapter 2 specifies commercially available and self-made materials followed by the methods applied to gain experimental results given in Chapter 3. The subsequent three Chapters 4-6 provide experimental details, results and conclusions. Specifically, two different mechanisms of bacterial pathogenicity which are involved in the virulence of *Enterococcus faecalis* (*E. faecalis*) and enteropathogenic *Escherichia coli* (EPEC) are studied in the Chapters 6 and 4, respectively. Chapter 6 presents aspects of chaperone/substrate interactions in the bacterial cytoplasm preceding protein translocation via a type three secretion system. In the first part of Chapter 6, structural details of the homodimeric chaperone are studied using experimental measurement of residual dipolar couplings and a novel application for their prediction to unravel the potential functional relevance of a 3D domain swap. The second part of this Chapter 6 deals with the characterization of the interactions of the chaperone with its substrate. In Chapter 4 a structural model for a protein/DNA complex is presented based on X-ray crystallographic and NMR spectroscopic data. This model is interpreted together with *in vivo* data to propose a role for the protein in the transcriptional regulation of the exotoxin cytolysin which is of major importance for the pathogenicity of *E. faecalis*. Using the same protein as in Chapter 4, a novel method for homooligomeric protein structure determination by NMR is demonstrated in Chapter 5. Finally the results are summarized and future experiments are suggested in Chapter 7.



2

Materials

In this chapter, a general description is provided for the most important materials used in this work.

2.1 Bacterial strains and plasmids

Bacterial strains used in this work are summarized in Table 2.1.

Table 2.1: Name, genotype and reference for the used *E. coli* strains.

STRAIN	GENOTYPE	REFERENCES
BL21	F ⁻ , <i>ompT</i> , <i>hsdS_B</i> , (<i>r_B-</i> , <i>m_B-</i>), <i>dcm</i> , <i>gal</i>	[28]
BL21(DE3)	F ⁻ , <i>ompT</i> , <i>hsdS_B</i> , (<i>r_B-</i> , <i>m_B-</i>), <i>dcm</i> , <i>gal</i> , λ (DE3)	[28]
XL2-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac</i> [F' <i>proAB</i> <i>lacI^qZ</i> Δ M15Tn10 (Tet ^r) Amy Cam ^r]	Stratagene

Plasmids containing the desired gene were transformed into the *Escherichia coli* (*E. coli*) expression strains BL21 or BL21(DE3). For expression of CylR2, the pET32a vector was used and a modified pET16b vector, referred to as pET16bTEV (Figure 2.1A), was used for cloning the coding sequences of the TTSS chaperone CesT and its effectors Tir and Map. For coexpression of protein-protein complexes, the genes were cloned into the pETDuetI vector (Figure 2.1B). All plasmids used in this work are listed in Table 2.2.

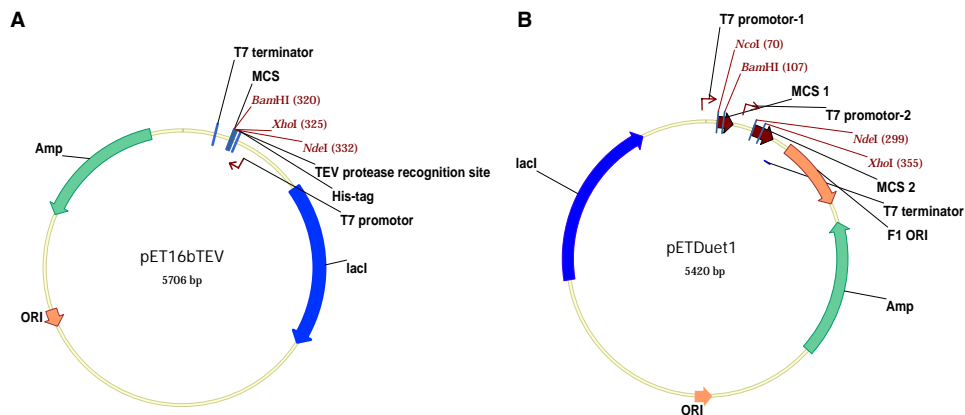


Figure 2.1: Plasmid map of pET16bTEV (A) and pETDuet1 (B)

Table 2.2: Plasmids used in this work.

PLASMID	RESISTENCE	PROPERTIES	REFERENCE/ORIGIN
pET32a	Ap	expression vector	Novagen, Madison, USA
pET32a-CylR2	Ap	pET32a with inserted CylR2-gene	K. Giller, Max Planck Institute Göttingen, Chapters 4 and 5
pET32a-MGSHG-CylR2	Ap	pET32a with inserted MGSHG-CylR2-gene	this work, Chapter 5
pET32a-MGHG-CylR2	Ap	pET32a with inserted MGHG-CylR2-gene	this work, Chapter 5
pQE30-His ₆ CesT	Ap	pQE30 with inserted His ₆ -CesT-gene	[29]
pET28a-Map	Km	pET28a with inserted Map-gene	[30]
pET28a-Tir	Km	pET28a with inserted Tir-gene	[31]
pET28a-TirN	Km	pET28a with inserted Tir-gene	[32]
pET16bTEV	Ap	modified pET16b with an additional TEV-protease recognition site for removal of the N-terminal His-tag	K. Giller, Max Planck Institute Göttingen, Chapter 6
pET16bTEV-CesT	Ap	pET16bTEV with inserted CesT-gene	this work, Chapter 6
pET16bTEV-Map	Ap	pET16bTEV with inserted Map-gene	this work, Chapter 6
pET16bTEV-Map101	Ap	pET16bTEV with inserted Map101-gene	this work, Chapter 6
pET16bTEV-Map84	Ap	pET16bTEV with inserted Map84-gene	this work, Chapter 6
pET16bTEV-Tir	Ap	pET16bTEV with inserted Tir-gene	this work, Chapter 6
pET16bTEV-TirN	Ap	pET16bTEV with inserted TirN-gene	this work, Chapter 6
pET16bTEV-Tir110	Ap	pET16bTEV with inserted Tir110-gene	this work, Chapter 6
pETDuet1	Ap	expression vector	Novagen, Madison, USA
pETDuet1-CesT-Map101	Ap	pETDuet1 with inserted CesT- and Map101-gene	this work, Chapter 6
pETDuet1-CesT-Map84	Ap	pETDuet1 with inserted CesT- and Map84-gene	this work, Chapter 6
pETDuet1-CesT ₁₋₁₃₅ -Map84	Ap	pETDuet1 with inserted CesT ₁₋₁₃₅ - and Map84-gene	this work, Chapter 6
pETDuet1-CesT-Tir108	Ap	pETDuet1 with inserted CesT- and Tir108-gene	this work, Chapter 6

2.2 Oligodesoxyribonucleotides and peptides

All oligonucleotides were ordered from Invitrogen (Karlsruhe, Germany), IBA (Göttingen, Germany) or Eurogentec (Liège, Belgium) and are presented in Table 2.3.

Table 2.3: Name, sequence and application of the applied oligonucleotides.

NAME	SEQUENCE 5' to 3'	APPLICATION
ir1.5'	ttgtcaagtgcacttgacaaa	oligonucleotides to construct IR1 for NMR studies
ir1.3'	ttgtcaagtgcacttgacaaa	
ir1short.5'	gtcaagtgcacttgaca	oligonucleotides to construct the central 18-bp region of IR1
ir1short.3'	gtcaagtgcacttgact	
ir2.5'	ttgtaacatattatggtatgca	oligonucleotides to construct IR2 for NMR studies
ir2.3'	tgcataacataatattgtacaa	
ACTG.5'	actgactgactgactgactgac	oligonucleotides to construct a random dsDNA for NMR studies
ACTG.3'	gtcagtcagtcagtcagtcagt	
CylR2.T55C5'	gcttactactgaattgtccactagaagatatttttcaatggc	mutagenesis of the gene for CylR2 to CylR2 _{T55C}
CylR2.T55C3'	gccattgaaaaatattcttagtgacaattcaggtagtaagc	
CylR2.N40C5'	ggtatagaaaaataaatattgcccttctttacagtttagcattaaaaattgc	mutagenesis of the gene for CylR2 to CylR2 _{N40C}
CylR2.N40C3'	gcaatttttaagtctaactgtaagaaggcaatattttttttctataacc	
MGSHG-CylR2.5'	ggaggacatatggcgagccatggcatgataatcaataactaaaattaattag	attaching the ATCUN-motif to CylR2
MGSHG-CylR2.3'	ggaggactcgagtttattcaggttg	
ATCUN-CylR2.5'	gaaggagatatacatatgggcatggcatgataatcaataac	mutagenesis of the gene for MGSHG-CylR2 to MGHG-CylR2
ATCUN-CylR2.3'	gttattgattatcatgccatggccatattgatatctccttc	
CesT5'	ggaggacatatgctgcgcatctgaacttttattagataggtttgcgg	amplification of the gene for CesT
CesT3'	ggaggaggatcctttatcttccggcg	
CesT _{V116R} 5'	gagaatgaaatagaagtcccgcttaagtcgatgg	mutagenesis of the gene for CesT to CesT _{V116R}
CesT _{V116R} 3'	ccatcgacttaacgcgacttctatttcattctc	
CesT _{E135} 3'	ggaggaggatccttattcgttttctaatggtattccctgattatg	amplification of the gene for CesT ₁₋₁₃₅
Map5'	ggaggacatatgtagtccaacggcaatggtag	amplification of the gene for Map
Map3'	ggaggactcgagttacagccgagtatcctgcacattg	
Map101.3'	ggaggaggatccttatgccgaattattactcatgcaatagtc	amplification of the gene for Map101
Map84.3'	ggaggactcgagttagtagataaaaagtaatctgctcttgcttg	amplification of the gene for Map84
Tir5'	ggaggacatatgcctattggtaaccttgtaataatg	amplification of the gene for Tir
Tir108.3'	ggaggactcgagttaaagggtccaatttgcgtattgag	amplification of the gene for Tir108

The peptides comprising predicted intracellular loops of CylR1 were synthesized by Kerstin Overkamp. The names, sequences and residues of the CylR1 peptides are summarized in Table 2.4.

Table 2.4: Name, sequence and residues of CylR1 of the applied peptides.

NAME	SEQUENCE	RESIDUES of CylR1
CylR1A	KKDERILYTISL	residue 24 to 35
CylR1B	DSKKDERILYTISL	residue 22 to 35
CylR1C	KKDERILYTISLTS	residue 24 to 37
CylR1D	FFNKRTS	residue 88 to 94
CylR1E	LLFFNKRTS	residue 86 to 94

2.3 Culture media and antibiotics

All culture media were prepared according to Table 2.5 followed by autoclave sterilisation. Before use, 100 $\mu\text{g}/\text{l}$ of ampicillin or 70 $\mu\text{g}/\text{l}$ of kanamycin was added to every medium. For agar plates 15 g agar per 1 l medium were added before autoclaving.

Table 2.5: Culture media and antibiotics

NAME	AMOUNT	CHEMICAL	REMARK
2 x YT-medium	16 g	tryptone	
	10 g	yeast extract	
	5 g	NaCl	
	up to 1 l	H ₂ O	
LB-medium	10 g	tryptone	
	5 g	yeast extract	
	10 g	NaCl	
	up to 1 l	H ₂ O	
M9-minimalmedium	6.8 g	Na ₂ HPO ₄	
	3 g	KH ₂ PO ₄	
	0.5 g	NaCl	
	1 g	NH ₄ Cl or ¹⁵ NH ₄ Cl	
	4 g	glucose or ¹³ C ₆ -glucose or D ₈ -glycerol	
	2 ml, 1 M	MgSO ₄	
	50 μl , 2 M	CaCl ₂	
	0.03 g	thiaminechloride hydrochloride	
	10 ml	trace elements	
up to 1 l	H ₂ O or 99.9 % D ₂ O		
trace elements	0.6 g	FeSO ₄ · 7 H ₂ O	
	0.094 g	MnCl ₂ · 2 H ₂ O	
	0.08 g	CoCl ₂ · 6 H ₂ O	
	0.07 g	ZnSO ₄ · 7 H ₂ O	
	0.03 g	CuCl ₂ · 2 H ₂ O	
	0.002 g	H ₃ BO ₃	
	0.025 g	(NH ₄) ₆ Mo ₇ O ₂₄ · 4 H ₂ O	
	up to 100 ml	H ₂ O	stir 10 min
0.5 g	EDTA	stir over night	
ampicillin stock solution	100 mg/ml	ampicillin sodium salt	steril filtrated, stored at -20 °C
kanamycin stock solution	70 mg/ml	kanamycin	steril filtrated, stored at -20 °C
IPTG stock solution	1 M	IPTG	steril filtrated, stored at -20 °C

2.4 Buffers and solutions

The buffers and solutions used for the methods and protein preparations described in this work are subsumed in Table 2.6.

Table 2.6: Buffers and solutions

METHOD	NAME	AMOUNT	CHEMICAL	
Affinity chromatography	wash buffer 1	20 mM	Tris/HCl pH 7.9	
		500 mM	NaCl	
		10 mM	imidazole	
	wash buffer 2	20 mM	Tris/HCl pH 7.9	
		500 mM	NaCl	
		40 mM	imidazole	
	elution buffer 1	20 mM	Tris/HCl pH 7.9	
		500 mM	NaCl	
elution buffer 2	100 mM	imidazole		
	20 mM	Tris/HCl pH 7.9		
	500 mM	NaCl		
Agarose gel	5 x DNA loading buffer	500 mM	imidazole	
		12.5 g	Ficoll 400	
		2.5 ml, 0.5 M	EDTA pH 8.0	
		0.5 g	SDS	
		25 mg	bromphenol blue	
		25 mg	xylene cyanol FF	
	10 x TBE buffer	up to 50 ml	H ₂ O	
		108 g	Tris	
		55 g	boric acid	
		40 ml, 0.5 M	EDTA pH 8.0	
Cell lysis	denaturing lysis buffer	up to 1 l	H ₂ O	
		8 M	urea	
		20 mM	Tris/HCl pH 7.9	
		500 mM	NaCl	
	native lysis buffer	10 mM	imidazole	
		20 mM	Tris/HCl pH 7.9	
		500 mM	NaCl	
	lysis buffer for CylR2	10 mM	imidazole	
		50 mM	HEPES pH 6.0	
		100 mM	NaCl	
		1 mM	EDTA	
		1 mM	MgCl ₂	
5 mM		DTT		
Gel filtration	gel filtration buffer	0.1 mg/1	DNase I	
		20 mM	HEPES pH 7.0	
Ion exchange	anion exchange buffer A for CesT	150 mM	NaCl	
		20 mM	Tris/HCl pH 8.2	
		0.1 M	NaCl	
		0.5 mM	EDTA	
	anion exchange buffer B for CesT	1 mM	DTT	
		20 mM	Tris/HCl pH 8.2	
		1 M	NaCl	
		0.5 mM	EDTA	
	ion exchange buffer A for CylR2	1 mM	DTT	
		50 mM	HEPES pH 6.0	
		0.1 M	NaCl	
		1 mM	EDTA	
		1 mM	MgCl ₂	
		5 mM	DTT	
		ion exchange buffer B for CylR2	50 mM	HEPES pH 6.0
			1 M	NaCl
1 mM	EDTA			
1 mM	MgCl ₂			
		5 mM	DTT	

continued on next page

METHOD	NAME	AMOUNT	CHEMICAL
Ligation	T4-DNA ligase buffer	40 mM	Tris/HCl pH 7.8
		10 mM	MgCl ₂
		1 mM	DTT
		0.5 mM	ATP
NMR	NMR buffer for CylR2	50 mM	HEPES pH7.0
		250 - 600 mM	NaCl
		5 - 10 %	D ₂ O (v/v)
	NMR buffer for CesT	50 mM	potassium phosphate pH 6.8
		100 mM	NaCl
		1 mM	DTT
		0.5 mM	EDTA
		5 - 10 %	D ₂ O (v/v)
	NMR buffer for Tir108	50 mM	potassium phosphate pH 6.8
100 mM		NaCl	
5 mM		DTT	
5 %		D ₂ O (v/v)	
PCR	Cloned Pfu buffer	20 mM	Tris/HCl pH 8.8
		2 mM	MgSO ₄
		10 mM	KCl
		10 mM	(NH ₄) ₂ SO ₄
		0.1 %	Triton X-100
		0.1 mg/ml	BSA
RP-HPLC	RP-HPLC buffer A	99.9 %	H ₂ O
		0.1 %	TFA
	RP-HPLC buffer B	99.9 %	acetonitrile
		0.1 %	TFA
SDS-PAGE	APS	10 %	(w/v) in H ₂ O
	destaining solution	100 ml	acetic acid
		900 ml	H ₂ O
	4 x protein loading buffer	1.7 g	SDS
		7.5 ml, 1 M	Tris/HCl pH 6.8
		23 ml	glycerol
		50 mg	bromphenol blue
		0.5 ml	β -mercaptoethanol
	running buffer	up to 50 ml	H ₂ O
		1 g	SDS
3.03 g		Tris	
14.4 g		glycine	
staining solution	up to 1 l	H ₂ O	
	2.2 g	Coomassie Brilliant blue G250	
	100 ml	acetic acid	
	250 ml	isopropanol	
TEV-cleavage	TEV-buffer	650 ml	H ₂ O
		50 mM	Tris/HCl pH 8.0
		1 mM	DTT
		0.5 mM	EDTA

2.5 Enzymes and Chemicals

Enzymes and chemicals used in this thesis can be found in Table 2.7.

Table 2.7: Enzymes and chemicals

CHEMICALS/ENZYMES	COMPANY
Agar, DNase I, IPTG	AppliChem, Darmstadt, Germany
Pf1 NMR co-solvent	ASLA ^{biotech} , Riga, Latvia
ammonium chloride (>98 % ¹⁵ N)	Cambridge Isotope Laboratories, Andover, USA
¹³ C ₆ -D-glucose (>98 % ¹³ C)	Spectra Stable Isotopes, Columbia, USA
Wizard Screen TM I/II	Emerald BioSystems, Bainbridge Island, USA
99.9 % D ₂ O, 98 % D ₈ -glycerol	Eurisotop, Gif-sur-Yvette, Frankreich
<i>Bam</i> HI, CIAP, λ <i>Hind</i> III, <i>Nco</i> I, <i>Nde</i> I, ϕ X174DNA/ <i>Bsu</i> RI(<i>Hae</i> III), T4-DNA ligase	Fermentas, St. Leon-Rot, Germany
ammonium molybdate tetrahydrate, ascorbic acid, Coomassie Brilliant Blue R-250, copper chloride dihydrate, iron (II) sulfate heptahydrate	Fluka, Neu-Ulm, Germany
DTT	Gerbu, Gaiberg, Germany
agarose, kanamycin sulphate	GibcoBRL, Karlsruhe, Germany
Crystal Screen TM I/II, Index Screen, PEG/Ion Screen	Hampton Research, Aliso Viejo, USA
Ficoll 400	ICN Biomedicals Inc, Costa Mesa, USA
Bench Mark protein ladder	Invitrogen, Karlsruhe, Germany
acetic acid, α -D(+)-glucose monohydrate, ammonium acetate, ammonium chloride, ammonium hydrocarbonate, boric acid, disodium hydrogen phosphate, ethanol, glycerol, glycine, hydrochloric acid, manganese chloride dihydrate, magnesium sulfate heptahydrate, sodium chloride, sodium dihydrogen phosphate, sodium hydroxide, TFA, thiaminechloride hydrochloride, Tris, urea	Merck, Darmstadt, Germany
BSA	New England Biolabs, Ipswich, USA
Hot Star Taq, Ni-NTA Agarose, QIAGEN Plasmid Midi Kit, QI-Aprep Spin Miniprep Kit, QIAquick Gel Extraction Kit, QIAquick PCR Purification Kit	Qiagen, Hilden, Germany
cobalt chloride hexahydrate	Riedel-de Haën, Seelze, Germany
Complete TM protease inhibitors, Complete TM protease inhibitors EDTA-free, elastase, protease K, trypsin	Roche Diagnostics, Mannheim, Germany
acetonitrile, ampicillin sodium salt, APS, dipotassium hydrogen phosphate, ethidium bromide, EDTA, HEPES, imidazole, magnesium chloride hexahydrate, MES, MOPS, potassium dihydrogen phosphate, Rotiphorese Gel 30, sodium acetate, TEMED, tryptone, yeast extract	Roth, Karlsruhe, Germany
SDS	Serva, Heidelberg, Germany
ATP, chymotrypsin	Sigma, Deisenhofen, Germany
Turbo <i>Pfu</i> DNA polymerase, dNTP	Stratagene, La Jolla, USA
MTSL	Toronto Research Chemicals, Toronto, Canada

2.6 Equipment

In the following, laboratory instruments and consumables are tabulated.

Table 2.8: Equipment and supplier

COMMON NAME	IDENTIFIER/COMPANY
Balances	Sartorius B 3100 S, Sartorius, Göttingen, Germany Sartorius AC 210 S, Sartorius, Göttingen, Germany
Centrifuges	Beckmann-Coulter Avanti J-20 and J-301, rotors: JLA 8.100, JLA 9.100, JLA 16.250, JA 25.50 Ti, JA 30.50 Ti, Krefeld, Germany Eppendorf Centrifuge 5415D, Wesseling-Berzdorf, Germany Eppendorf Centrifuge 5804, Wesseling-Berzdorf, Germany Heraeus Biofuge primo, Kendro, Hanau, Germany
Concentrators	Microcon, YM-3 and YM-10, Amicon, Bedford, USA Centricon, YM-3 and YM-10, Amicon, Bedford, USA Centriplus, YM-3 and YM-10, Amicon, Bedford, USA Vivaspin 2 ml MWCO 10.000 PES, Vivascience, Hannover, Germany
Crystallization	TECAN Genesis RSP 150, Crailsheim, Germany Mosquito Crystal, TTP LabTech Ltd., Melbourn, UK
Desalting	NAP TM -10, Amersham Pharmacia Biotech, Freiburg, Germany PD TM -10, Amersham Pharmacia Biotech, Freiburg, Germany
Dialysis	Slide-A-Lyzer Dialysis Cassettes, MWCO 3500, 0.1-0.5 ml Capacity, Pierce Biotechnology, Inc., Rockford, IL, USA Slide-A-Lyzer Dialysis Cassettes, MWCO 10000, 0.1-0.5 ml Capacity, Pierce Biotechnology, Inc., Rockford, IL, USA Spectra Por membranes, MWCO 10000, Roth, Karlsruhe, Germany Spectra Por membranes, MWCO 3500, Roth, Karlsruhe, Germany
Electrophoresis	Kodak Electrophoresis documentation and analysis system 120, Eastman Kodak Co., New York, NY, USA Power Pac 300, BioRad, München, Germany Polyacrylamide gel electrophoresis: Mini-PROTEAN 3 Cell, BioRad, München, Germany Agarose gel electrophoresis: Mini-Sub Cell GT, BioRad, München, Germany
-80 °C freezer	MDF-U71V Ultra-low temperature freezer, SANYO Electric Co., Ltd, Osaka, Japan
Filtering	sterile filter 0,20 µm, Sartorius, Göttingen, Germany
FPLC	Äkta prime, Amersham Pharmacia Biotech, Freiburg, Germany Äkta basic, Amersham Pharmacia Biotech, Freiburg, Germany Disposable 10 ml polypropylene columns, Pierce Biotechnology, Inc., Rockford, IL, USA Frac-100, Amersham Pharmacia Biotech, Freiburg, Germany HiLoad Superdex 75 prep grade 16/60, Amersham Pharmacia Biotech, Freiburg, Germany HiTrap Q-Sepharose, Amersham Pharmacia Biotech, Freiburg, Germany HiTrap S-Sepharose, Amersham Pharmacia Biotech, Freiburg, Germany
HPLC	system 1: UV-975, PU-980, LG-980-02, DG-980-50, AS-2055Plus, CO-200, JASCO International, Groß-Umstadt, Germany system 2: MD-910, PU-980, LG-1580-04, DG-1580-54, AS-950-10, CO-200, JASCO International, Groß-Umstadt, Germany system 3: MD-2010Plus, PU-2080Plus, LG-2080-04, DG-2080-54, AS-2055Plus, CO-200, JASCO International, Groß-Umstadt, Germany Vydac C4 8x250 mm, Hesperia, CA, USA Vydac C18 8x250 mm, Hesperia, CA, USA
Incubator	Infors Multitron HT, Einsbach, Germany Certomat R, B. Braun Biotech International, Melsungen, Germany
Lyophilisation	Christ Alpha 2-4, B. Braun Biotech International, Melsungen, Germany
NMR	AVANCE 400, Bruker, Karlsruhe, Germany AVANCE 600, Bruker, Karlsruhe, Germany DRX 600, Bruker, Karlsruhe, Germany AVANCE 700, Bruker, Karlsruhe, Germany DRX 800, Bruker, Karlsruhe, Germany AVANCE 900, Bruker, Karlsruhe, Germany Quality NMR Sample Tubes 5 mm, Norell, Inc., Landisville, NJ, USA Shigemi NMR tube 5 mm, Shigemi Corp., Tokyo, Japan

continued on next page

COMMON NAME	IDENTIFIER/COMPANY
Peptide synthesis	433A Peptide Synthesizer, Applied Biosystems, Darmstadt, Germany
pH-Meter	PB11 PY-P10, Sartorius, Göttingen, Germany
Spectroscopy	UV/VIS-Spectrophotometer, Hewlett-Packard 8453, Böblingen, Germany CD spectropolarimeter J-720, JASCO International, Groß-Umstadt, Germany Mass spectrometer, Water Micromass ZQ single quadrupole, Waters, Saint-Quentin, France
Thermocycler	HYBAID PCR sprint, Perkin Elmer, Wellesley, MA, USA
Ultrasound	SONOPLUS HD 2200, Bandelin, Berlin, Germany

2.7 Software

The software used for processing and analysing NMR spectra, for calculating and analysing structures and for preparing figures is summarized in Table 2.9.

Table 2.9: Software

PROGRAM	REFERENCE/ORIGIN
AQUA 2.0	[33]
BOBSRIPT	[34]
CCP4	[35, 36]
ClusPro	http://nrc.bu.edu/cluster , [37, 38, 39]
CYANA 2.0	[40]
DINO	[41]
InsightII	MSI 2000 release, San Diego, CA, USA
MARS	[42]
MOLMOL	[43]
MOLSCRIPT	[44]
MSMS	[45]
NMRPipe/NMRDraw	[46]
NMRView 5.0.4	[47]
PALES	[48]
PROCHECK	[49]
PROCHECK-NMR	[33]
PyMOL	http://www.pymol.org , [50]
RASTER3D	[51]
Sparky 3	T. D. Goddard and D. G. Kneller, University of California, San Francisco
TALOS	[52]
XPLOR-NIH 2.9.7	[53]
X-WINNMR 3.5	Bruker, Karlsruhe, Germany



3

Methods

Here, all general methods applied in this thesis are presented. The detailed purification strategies for the studied proteins are given in Chapters 4 and 6.

3.1 Molecular biology methods

3.1.1 Agarose gel electrophoresis

DNA fragments were separated on horizontal agarose gels. The gels were prepared by melting 1 % (w/v) agarose in 1 x TBE buffer and adding 25 μ l/l ethidium bromide. Each DNA sample was mixed with 1/5 volumes of 5 x DNA-load. Gels were run at a constant voltage of 100 V, imaged under UV-light and digitized for documentation.

3.1.2 Isolation and purification of DNA

10 ml *E. coli* XL2-blue overnight cultures were used for purification of up to 20 μ g plasmid DNA with the help of the QIAprep Spin Miniprep Kit and 200 ml *E. coli* XL2-blue overnight cultures were used for purification of up to 100 μ g plasmid DNA with the help of the QIAGEN Plasmid Midi Kit. Plasmid DNA was isolated from the cell pellets according to the producer's instructions.

For removal of enzymes, salts and buffers, columns with a silica gel membrane (QIAquick spin columns) were used according to the user manual.

DNA fragments were purified from 1 % agarose gels with the QIAquick Gel Extraction Kit.

3.1.3 Enzymatic modifications of DNA

3.1.3.1 Fragmentation with restriction enzymes

DNA was digested with 2-10 units restriction enzyme per μg DNA under buffer conditions recommended by the manufacturer. The reaction was performed at 37 °C for 1-2 h with *Bam*HI or overnight for all other restriction enzymes.

3.1.3.2 Dephosphorylation

To avoid self-ligation of restricted vectors, 5'-overhanging phosphate groups were cleaved by adding 1 unit CIAP (calf intestine alkaline phosphatase) per 1 μg DNA and incubating for 1 h at 37 °C.

3.1.3.3 Ligation

To ligate DNA-fragments with overhanging ends, 50-100 ng vector DNA were mixed with insert-DNA at a ratio of 1 : 3 to 1 : 25 in T4-DNA ligase buffer and incubated overnight at 16 °C with 4 units T4-DNA ligase.

3.1.4 Polymerase chain reaction (PCR)

For amplification of DNA fragments by PCR [54], 25 μl reaction mixture containing 20-30 ng plasmid DNA template, 10 pmol of each primer, 100 nmol dNTPs and 2.5 units *Pfu* Turbo DNA polymerase in Cloned *Pfu* buffer were prepared on ice. The PCR was carried out after an initial denaturing step of 2 min at 95 °C with 30 cycles comprising the following steps: 1.) denaturing for 30 s at 95 °C, 2.) annealing for 30 s at 55 °C and 3.) elongation for 1 min per kb of fragment length at 72 °C. The final elongation step was extended by 10 min at 72 °C.

Colony PCR was used to verify a successful cloning directly from single colonies without prior preparation of plasmid DNA. Each 25 μl PCR reaction contained 10 pmol of each primer, 50 nmol dNTPs and 0.3 units HotStar Taq polymerase. Bacterial cells from single colonies were transferred into PCR reaction mixtures with the help of sterile pipette tips, which were afterwards used to inoculate 10 ml LB medium containing ampicillin. The LB cultures were incubated overnight at 37 °C. The PCR

reaction mixtures were initially heated to 95 °C for 15 min to activate the HotStar Taq polymerase. Afterwards, the PCR was performed with 25 cycles comprised of: 1.) denaturing for 30 s at 95 °C, 2.) annealing for 1 min at 55 °C and 3.) elongation for 2 min per kb of fragment length at 72 °C. The final elongation step was extended by 10 min at 72 °C. The PCR products were analysed by agarose gel electrophoresis (see 3.1.1). Plasmid DNA isolated from LB cultures (see 3.1.2) with a positive PCR result was confirmed by DNA sequencing (see 3.1.6).

3.1.5 Site-directed mutagenesis

Single amino acid changes were generated using the QuikChange site-directed mutagenesis kit. Primers used for mutagenesis were designed according to the instruction manual and are included in Table 2.3. The 50 μ l PCR reaction mixtures contained 125 ng of each primer, 10 ng plasmid DNA template, 6.25 μ mol dNTPs and 2.5 units Turbo *Pfu* DNA polymerase in Cloned Pfu buffer. The cycling after an initial step at 95 °C for 30 s was performed as follows: 1.) denaturing for 30 s at 95 °C, 2.) annealing for 1 min at 55 °C and 3.) elongation for 2 min per kb of plasmid length at 68 °C. For creating single nucleotide changes and for creating double or triple nucleotide changes the number of cycles was 12 and 16, respectively. Following temperature cycling, 10 units of *Dpn* I restriction enzyme were directly added to each PCR reaction mixture and incubated for 1 h at 37 °C. 1 μ l of the reaction mix was transformed into *E. coli* XL2-blue (see 3.1.8), plasmid DNA was purified (see 3.1.2) and correct introduction of the desired mutation was confirmed by DNA sequencing of the insert region of the plasmid DNA (see 3.1.6).

3.1.6 DNA sequencing

Sequencing of purified plasmid DNA was performed via the extended Hot Shot DNA sequencing service of SeqLab (Göttingen, Germany). 200 μ l PCR tubes with a flat lid were loaded with a total volume of 7 μ l containing 20 pmol of primer and 0.6-0.7 μ g plasmid DNA in H₂O.

3.1.7 Concentration and purity of DNA

The absorption at 260 nm (A_{260}) was measured to determine the DNA concentration. An A_{260} of 1 equals 50 $\mu\text{g}/\text{ml}$ dsDNA [55]. The ratio of A_{260}/A_{280} was calculated. A ratio between 1.8 and 2 indicates reasonably pure DNA, suitable for DNA sequencing.

3.1.8 Transformation of *E. coli*

1 μl plasmid DNA or 5 μl ligation assay were added to 40 μl competent cells and incubated for 30 min on ice. Cells were subjected to a heat shock of 42 °C for 60 s and supplemented with 1 ml of 2 x YT medium. After incubation for 1 h at 37 °C and centrifugation for 1 min at 7500 x g the cell pellet was resuspended in $\sim 50 \mu\text{l}$ medium and plated onto LB-agar plates containing ampicillin.

3.2 Protein expression

3.2.1 Cultivation and storage of *E. coli*

All proteins were expressed in the *E. coli* strain BL21(DE3). Agar plates and all liquid nutrient media contained the appropriate antibiotic. Cells were grown at 37 °C if not denoted differently. A single colony of *E. coli* BL21(DE3) harbouring the appropriate plasmid was grown on an agar plate and used to inoculate a 1 ml LB overday preculture. 20 μl of this preculture were transferred to a 20 ml overnight preculture. The 20 ml overnight preculture was then used to inoculate a 1 l LB culture. For producing labeled protein, the 20 ml LB and the 1 l LB culture were replaced by a 50 ml and a 1 l M9-minimal medium culture containing the required nitrogen and carbon source. In the case of deuterium labeling, M9-minimal medium culture was prepared with 99.9 % D_2O . For perdeuteration, 2 ml M9-minimal medium precultures with the deuterium concentration increasing from 33 %, 50 %, 75 % to 99.9 % and 4 g/l D_8 -glycerol as the sole carbon source were grown for 10-22 h each. Each preculture was centrifuged at 7500 x g and the pellet was resuspended in the medium with the next deuterium concentration. Cell growth was monitored

by measuring the OD₆₀₀. At an OD₆₀₀ of 0.6-0.8 the protein expression was induced by addition of 1 M IPTG to a final concentration of 1 mM. The cells were harvested at an OD₆₀₀ of 1.2-1.6 by centrifugation at 7500 x g and 4 °C for 15 min. The cell pellets were stored at -80 °C.

For long term storage of the *E. coli* strains, 350 µl stationary LB culture were mixed with 150 µl sterile 50 % glycerol and stored at -80 °C.

3.2.2 Cell lysis

For cell lysis under native conditions, cell pellets from 1 l cultures were thawed on ice and suspended in 40-50 ml lysis buffer containing 0.5 mM PMSF and half a tablet of CompleteTM protease inhibitors. The suspension was sonicated 9 x 20 s on ice. The lysate was centrifuged for 30-45 min at 48000 x g and 4 °C.

Cell lysis under denaturing conditions was performed in 40-50 ml lysis buffer containing 8 M urea. Pellets were suspended and stirred for 1 h or overnight at room temperature. The lysate was centrifuged for 1 h at 14 °C and 48000 x g.

3.3 Protein methods

3.3.1 Nickel-nitrilotriacetic acid (Ni-NTA) agarose affinity chromatography

Components of the required buffers can be found in Table 2.6. For purification of sevenfold histidine (His₇)-tagged proteins under denaturing conditions, the wash and elution buffers (see 2.6) contained 8 M urea in addition.

Proteins with a N-terminal His₇-tag were purified via a Ni-NTA agarose column. 3 ml of resin were used per protein solution from 1 l of expression culture. The resin, equilibrated with wash buffer 1, was incubated with the supernatant from the cell lysis (see 3.2.2) for 1 h at 4 °C under native conditions or at room temperature under denaturing conditions. After washing the column with 10 column volumes of wash buffer 1, bound proteins were eluted by increasing the imidazole concentration in three steps: 1.) 5 column volumes of wash buffer 2, 2.) 5 column volumes of elution buffer 1 and 3.) 5 column volumes of elution buffer 2.

3.3.2 Ion exchange chromatography

Ion exchange chromatography was performed on an ÄKTA prime low pressure liquid chromatography system with HiTrap columns. Before each run, the HiTrap column was prepared by washing at first with 5 bed volumes of buffer A and then with 5 bed volumes of buffer B. Finally, the column was equilibrated with 5-10 bed volumes of buffer A. Samples were loaded at a flow rate of 1 ml/min onto the equilibrated column. Unbound protein was washed with 5-10 column volumes of buffer A. The elution was performed applying a linear gradient of 0-100 % buffer B at a flow rate of 1-3 ml/min and collecting 2 or 3 ml fractions. Protein elution was detected by measuring the absorption at 280 nm (A_{280}).

3.3.3 Gel filtration

An ÄKTA basic medium pressure liquid chromatography system was used for gel filtration with a HiLoad Superdex 75 prep grade 16/60 column. 2-5 ml of the concentrated protein solution containing 10-20 mg protein were loaded onto the column, equilibrated with gel filtration buffer. The gel filtration was performed at a constant flow rate of 1 ml/min. 2 ml fractions were collected.

3.3.4 Reversed phase-high performance liquid chromatography (RP-HPLC)

RP-HPLC systems were operated by Kerstin Overkamp or Gerhard Wolf.

A semi-preparative C_{18} -HPLC column was used for final purification of all CylR2 samples and a semi-preparative C_4 -HPLC column was used for final purification of all Tir samples. The column was pre-equilibrated with RP-HPLC buffer A (Table 2.6). The elution was performed at a flow rate of 3 ml/min with a linear gradient of 0-100 % RP-HPLC buffer B in 50 min for Tir and in 30 min for CylR2.

3.3.5 Cleavage with TEV-protease

N-terminal His-tags were cleaved with the TEV-protease. At first, His-tagged proteins after affinity chromatography were dialysed at 4 °C against two times 5 l TEV-buffer

(see 3.3.1). Then the protein concentration was determined (see 3.4.1) and 0.5-2 μg of TEV-protease per 100 mg of protein were added. The cleavage reaction was performed for 16-48 h at room temperature. As the TEV-protease contains a His-tag it was removed via a second affinity chromatography step. After completed cleavage, the protein solution was dialysed against 5 l of wash buffer 1 and applied to a 1-ml Ni-NTA Agarose column. The flow through and the first 5 ml from wash step 1 were collected for further purification.

3.3.6 Concentrating of proteins

Protein solutions were concentrated by ultrafiltration using 0.5, 2 or 20 ml concentrators with the appropriate molecular weight cutoff (MWCO). The centrifugation was performed at 4 °C as recommended by the supplier (Table 2.8). The progress of the ultrafiltration was monitored by measuring the A_{280} .

3.3.7 Protein storage

Proteins in H_2O were shock-frozen in liquid nitrogen and dried by lyophilisation. Lyophilised proteins were stored at -20 °C. Proteins in buffer solutions were shock-frozen in liquid nitrogen and stored at -80 °C.

3.3.8 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The discontinuous Tris-glycine buffer system as described by Laemmli [56] was used to separate proteins by molecular weight. The components of the stacking and separating gel are given in Table 3.1. All gels had a separating gel with an acrylamide concentration of 17.5 % and a 3 % acrylamide stacking gel. For detection of proteins with a molecular weight smaller than 8-10 kDa, 2 M Tris/HCl pH 8.8 was used as the separation gel buffer instead of the 1 M buffer [57]. Protein samples were mixed with 4 x protein loading buffer and denatured at 100 °C for 5 min. Gels were run at 25 mA and stained with Coomassie blue R-250 by heating for 90 s in the microwave and 10-30 min incubation. For destaining, several 90 s heating steps in the microwave

were performed and gels were completely destained by shaking overnight. Finally, the SDS-gels were imaged and digitized for documentation.

Table 3.1: Composition of the 17.5 % SDS-gels

	Seperating gel	Stacking gel
Rotiphorese Gel 30	3 ml	250 μ l
1 M or 2 M Tris/HCl pH 8.8	1.88 ml	-
1 M Tris/HCl pH 6.8	-	313 μ l
H ₂ O	30 μ l	1.88 ml
10 % SDS	50 μ l	25 μ l
TEMED	2.5 μ l	2 μ l
10 % APS	50 μ l	25 μ l

3.3.9 Edman degradation

N-terminal protein sequencing was performed via Edman degradation [58] by U. Pleßmann in the group of Prof. Dr. Klaus Weber.

3.3.10 Limited proteolysis

Limited proteolysis was used to identify the rigid core of Map in complex with CesT [59]. Proteolysis was performed with the proteases trypsin, proteaseK, elastase and chymotrypsin. A \sim 30 μ M protein solution in 20 mM HEPES pH 7.0 and 150 mM NaCl was digested at protein:protease ratios of 50:1, 100:1, 1000:1 and 10000:1 at room temperature. The reaction was stopped after 5, 10, 20, 30, 60 and 120 min by addition of SDS-PAGE loading buffer and boiling for \sim 5 min at 100 °C. All samples were analysed by SDS-PAGE.

3.3.11 Peptide synthesis

Solid-phase peptide synthesis (SPPS) was performed by Kerstin Overkamp. Raw peptides were purified via RP-HPLC.

3.3.12 Introduction of a paramagnetic center

3.3.12.1 Site-directed spin-labeling (SDSL)

SDSL is a technique to introduce a spin label into a protein based on a specific reaction between the spin label and an amino acid. In the most commonly used procedure, a unique sulfhydryl group is selectively modified with a nitroxide reagent [60]. Single cysteine mutants for the attachment of a spin label were generated by site-directed mutagenesis (see 3.1.5) and modified with the thiol-specific nitroxide spin label reagent (1-oxy-2,2,5,5-tetramethyl-3-pyrroline-3-methyl)-methanethiosulfonate (MTSL). 100 $\mu\text{g}/\mu\text{l}$ MTSL in cold acetone was added with a 3 to 5-fold molar excess and incubated overnight at room temperature. Unreacted spin label was removed by size exclusion chromatography via a PD-10 column. Complete labeling with MTSL was verified by ESI-MS (see 3.4.3).

3.3.12.2 Copper-binding

Introduction of paramagnetic Cu^{2+} was achieved by prepending the amino terminal Cu(II)- and Ni(II)-binding (ATCUN) motif to the N-terminus of CylR2 via recombinant cloning. Cu^{2+} was bound by adding a 3 to 5-fold excess of CuCl_2 dissolved in H_2O . Excessive Cu^{2+} was removed by dialysis against MOPS buffer. MOPS buffer was used to minimize the interaction between buffer molecules and Cu^{2+} [61].

3.4 Spectroscopic methods

3.4.1 Determination of protein concentration

The concentration of proteins in solution was determined by measuring the absorption at 280 nm (A_{280}).

$$A_{\lambda} = \epsilon_{\lambda} \cdot c \cdot d \quad (3.1)$$

A_{λ} : absorption at wavelength λ

ϵ_{λ} : molar extinction coefficient at wavelength λ in $\text{M}^{-1}\text{cm}^{-1}$

c : protein concentration (M)

d : thickness of the cuvette (cm)

3.4.2 Circular dichroism (CD) spectroscopy

CD spectra were recorded from 200 to 260 nm in a 0.1 cm cuvette at a scan-rate of 20 nm/min and with 0.5 nm resolution for 200 μ l of a 20-30 μ M sample in the NMR buffer. The temperature was kept constant at 25 °C. The mean molar ellipticity per amino acid was calculated from the measured ellipticity according to the formular [62]:

$$[\Theta]_{MRW} = \frac{10^5 \cdot \Theta}{c \cdot d \cdot N} \quad (3.2)$$

$[\Theta]_{MRW}$: mean molar ellipticity per amino acid ($\text{deg}\cdot\text{cm}^2\cdot\text{dmol}^{-1}$)

Θ : measured ellipticity

c: protein concentration (μM)

d: thickness of the cuvette (cm)

N: number of amino acids

3.4.3 Electrospray ionization mass spectrometry (ESI-MS)

ESI-MS was performed after the final purification of the first preparation of each protein sample to confirm the identity by the molecular weight and verify the purity. Fractions from RP-HPLC (see 3.3.4) were analysed directly by ESI-MS. All other samples were transferred into 20 mM NH_4OAc by dialysis or via a NAPTM-10-column.

3.5 NMR spectroscopy

The preparation of the NMR samples is described under materials and methods of Chapters 4, 5 and 6.

All NMR experiments were acquired at 600, 700, 800 or 900 MHz on Bruker AVANCE 600, 700, and 900 or DRX 600 and 800 spectrometers running under X-WINNMR 3.5. All NMR spectra were processed using X-WINNMR 3.5 or NMRPipe/NMRDraw [46] and analysed using NMRView 5.0.4 [47] or Sparky 3 (T. D. Goddard and D. G. Kneller, University of California, San Francisco).

3.5.1 Chemical shift assignment

All NMR experiments performed for chemical shift assignments of the different samples are summarized in the Appendix B.1.

3.5.1.1 Backbone chemical shift assignment

Backbone C_α , C' , N and H_N assignments were obtained using standard triple resonance experiments for proteins smaller than 30 kDa [63]. The assignments were derived from 3D HNCA, HN(CO)CA, CBCA(CO)NH or HN(CO)CACB, HNCACB, HNCO and HN(CA)CO experiments. For proteins and protein complexes bigger than 30 kDa, TROSY versions of the experiments were measured with deuterium decoupling [64, 65, 66]. In addition, the magnetization was completely transferred from C_α to C_β in HN(COCA)CB and HN(CA)CB experiments to detect C_β carbons for molecules bigger than 30 kDa [67]. The manual assignment of backbone resonances was assisted by using the automatic assignment program MARS [42].

3.5.1.2 Side chain chemical shift assignment

Aliphatic side chain resonances were derived from 3D HCCH-TOCSY, CCONH-TOCSY and HCCONH-TOCSY spectra. Refocusing and mixing of ^{13}C magnetization was performed using adiabatic pulses [68, 69]. Additional data provided by 3D [$^{15}\text{N}, ^1\text{H}$]-NOESY-HSQC and [$^{13}\text{C}, ^1\text{H}$]-NOESY-HSQC experiments were used for further assignment as well as confirmation of the through-bond data. Aromatic ring proton and carbon resonances were assigned using 2D [$^{13}\text{C}, ^1\text{H}$]-HSQC, (H β)C β (C γ C δ)H δ and (H β)C β (C γ C δ C ϵ)H ϵ experiments [70].

3.5.2 Secondary structure determination

Regions of secondary structure can be identified by calculating the difference between the experimental chemical shift (C^{exp}) and the random coil chemical shift (C^{coil}). Tabulated random coil values were used for folded proteins. For unfolded proteins, sequence corrected random coil values as empirically determined at pH 3.0 by Schwarzingger *et al.* were used for all residues with the exception of His, Glu and

Asp [71]. Since His, Glu and Asp are particularly sensitive to pH, the experimental random coil chemical shifts at pH 5.0 from Wishart *et al.* were applied for these residues [72]. Secondary chemical shifts ($\Delta\delta C$) for C_α , C_β and C' were calculated as the difference between measured ^{13}C chemical shift (δC^{exp}) and random coil ^{13}C chemical shift (δC^{coil}) [73]:

$$\Delta\delta C = \delta C^{exp} - \delta C^{coil}. \quad (3.3)$$

To derive secondary structure information, the secondary chemical shifts were combined using the following formular:

$$\Delta\delta(C_\alpha C_\beta C') = \Delta\delta C_\alpha - \Delta\delta C_\beta + \Delta\delta C' \quad (3.4)$$

or by adding $\Delta\delta C_\alpha$ and $\Delta\delta C'$ if C_β chemical shifts were not available. C_β chemical shifts were not used for unfolded proteins.

3.5.3 Residual dipolar couplings (RDCs)

Anisotropic media for measurement of RDCs were prepared by addition of filamentous phages (Pf1) to a concentration of 7-12 mg ml⁻¹ [74, 75]. RDCs were back-calculated from X-ray or NMR structures using singular value decomposition (SVD) as implemented in PALES [48]. The correlation coefficient and the quality or Q-factor were used to evaluate the agreement between a structure and the observed RDCs. The Q-factor was determined as $\text{rms}(^1D_{HN}^{measured} - ^1D_{HN}^{calculated})/\text{rms}(^1D_{HN}^{measured})$ [76].

3.5.3.1 HN-RDCs

^{15}N - ^1H splittings were measured under isotropic and anisotropic conditions using 2D IPAP- ^{15}N - ^1H HSQC experiments [77]. HN-RDCs were extracted by subtraction of the $^1J_{NH}$ scalar coupling measured for the isotropic sample [78]. Alternatively, HN-RDCs were determined from the ^{15}N - ^1H splittings of a modified 3D TROSY-HNCO experiment for strongly overlapped peaks in large molecules [79].

3.5.3.2 C'N- and C'C_α-RDCs

The quantitative J correlation method was used to determine one-bond dipolar ¹³C'-¹⁵N couplings and one-bond dipolar ¹³C'-¹³C_α couplings. ¹J_{C'N} or ¹J_{C'C_α} values were measured under isotropic and anisotropic conditions for ²H,¹³C,¹⁵N-labeled CesT using a 3D TROSY-HNCO [80] or a 3D TROSY-HN(CO)CA-QJ experiment [81], respectively.

3.5.3.3 Prediction of RDCs

The electrostatic module of the software PALES was used to predict RDCs from the surface charge distribution of proteins in filamentous phage medium [48, 82] in order to validate structures or to rank results from *ab initio* docking (see 3.6.3). Predictions were performed for the experimental Pf1 concentration, for different NaCl concentrations (0.005, 0.01, 0.015, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45 and 0.50 M) and with the default charge attached to all ionizable residues. The correlation between predicted and experimental couplings was finally multiplied with the correlation obtained from a best-fit of experimental values to the structure using singular value decomposition.

3.5.4 NMR relaxation data

3.5.4.1 Approximation of the molecular weight

The rotational correlation time (τ_c) is proportional to the molecular weight (MW) and thus the aggregation or oligomerization state of a macromolecule:

$$\tau_c[ns] \sim \frac{1}{2} MW[kDa]. \quad (3.5)$$

An approximate value of τ_c can be derived from the T₂ times of the backbone amide protons

$$\tau_c[ns] \approx \frac{1}{5T_2[s]} \quad (3.6)$$

or from the T_{1ρ} times of the backbone amide nitrogens:

$$\tau_c[s] = \frac{5r_{HN}^6 16\pi^2}{\gamma_H^2 \gamma_N^2 \hbar^2 \mu_0^2 T_{1\rho}}, \quad (3.7)$$

where r_{HN} is the distance between amide proton and nitrogen, γ_i is the gyromagnetic ratio of nucleus i , \hbar is the reduced Planck's constant and μ_0 is the permeability of vacuum. For unlabeled samples transverse relaxation times (T_2) of backbone protons were estimated from two 1D one-one echo experiments [83] measured with a relaxation delay of 100 μ s (Δ_A) and a relaxation delay of 3.1 ms (Δ_B) using the following formular:

$$T_2 = \frac{2(\Delta_A - \Delta_B)}{\ln(I_B/I_A)} \quad (3.8)$$

For ^{15}N -labeled samples 1D $T_{1\rho}$ experiments with a spin-lock power of 2.5 kHz. Two 1D $T_{1\rho}$ experiments were acquired. One with a spin-lock pulse duration of 2 ms (Δ_A) and a second with a spin-lock pulse duration Δ_B which corresponds to ~ 1.3 times the relaxation time calculated from the theoretical MW. The average relaxation time of all backbone nitrogens was calculated using the following formular:

$$T_{1\rho} = \frac{\Delta_A - \Delta_B}{\ln(I_B/I_A)}. \quad (3.9)$$

3.5.4.2 ^{15}N transverse relaxation times (T_2)

^{15}N relaxation data were acquired with modified versions of pulse sequences described earlier [84], incorporating pulsed field gradients and a WATERGATE for solvent suppression. T_2 times were sampled using seven different ^{15}N relaxation delays (Δ_{relax}): 7.6, 50, 90, 130, 160, 190 and 220 ms. To extract ^{15}N T_2 relaxation times by measuring the peak volumes (I) for all spectra recorded with different relaxation delays and fitting the volumes to a single-exponential function:

$$y = I(O) \exp\left(-\frac{\Delta_{relax}}{R_2}\right), \quad (3.10)$$

where R_2 is the transverse relaxation rate ($T_2 = 1/R_2$).

3.5.4.3 ^{15}N - ^1H steady state NOE

The heteronuclear nuclear Overhauser effects (NOEs) of ^{15}N nuclei in NH groups were measured using the Bruker standard pulse program `invinoef3gpsi`. ^{15}N - ^1H NOE

values were calculated as the intensity ratios of the ^{15}N - ^1H correlation peaks from pairs of interleaved spectra acquired with and without ^1H presaturation during the recycle time of 5 s.

3.5.5 Mapping binding surfaces

An easy and rapid way to gain qualitative information about the interaction between two molecules is to study the chemical shift values. Chemical shifts depend on the local environment. The chemical shifts of ^1H and ^{15}N are particularly sensitive to changes of the environment. Therefore ^{15}N and ^1H chemical shift perturbations recorded via ^{15}N - ^1H -HSQC spectra throughout an NMR titration are commonly combined in order to map a binding site on a protein according to the following equation [85]:

$$\Delta\delta_{HN} = \sqrt{\frac{(\Delta\delta_N/5)^2 + (\Delta\delta_H)^2}{2}} \quad (3.11)$$

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

$\Delta\delta_N$: amide nitrogen chemical shift perturbation

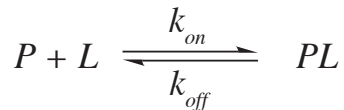
$\Delta\delta_H$: amide proton chemical shift perturbation

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

3.5.6 Determination of protein-ligand dissociation constants

Chemical shift changes observed in a series of ^{15}N - ^1H -HSQC spectra during the titration of a ligand to a protein can be used to calculate the dissociation constant K_d if the exchange rate is much faster than the difference between the chemical shifts of the free and bound protein. In this case of fast exchange on the NMR time scale, the K_d is larger than 10^{-3} M.

Considering a complex with 1:1 stoichiometry composed of a protein (P) which binds a DNA ligand (L), the complex formation and dissociation is described as follows: and K_d is defined as:



$$K_d = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} = \frac{([P]_0 - X_{PL}[P]_0)([L]_0 - X_{PL}[P]_0)}{X_{PL}[P]_0} \quad (3.12)$$

with $[P]_0 = [P] + [PL]$ and $X_{PL} + X_P = 1$.

k_{off} : rate of dissociation

k_{on} : rate of association

$[P]_0$: total or starting concentration of the protein

$[P]$: concentration of the protein

$[L]_0$: total or starting concentration of the ligand

$[L]$: concentration of the ligand

$[PL]$: concentration of the complex

X_P : fraction of the free protein

X_{PL} : fraction of the bound protein

The observed chemical shift δ_{obs} can be written as:

$$\delta_{obs} = X_P\delta_P + X_{PL}\delta_{PL} \quad (3.13)$$

and with $r = \frac{[L]_0}{[P]_0}$ the following equation can be derived:

$$\delta_{obs} = \delta_P + (\delta_{PL} - \delta_P) \frac{(K_d + (1+r)[P]_0) - \sqrt{(K_d + (1+r)[P]_0)^2 - 4[P]_0^2 r}}{2[P]_0} \quad (3.14)$$

δ_P : chemical shift of the protein

δ_{PL} : chemical shift of the complex

Fitting the experimental data to equation 3.14 yields the K_d [86].

3.5.7 Long-range distances from paramagnetic relaxation enhancement (PRE)

The distance-dependent transverse relaxation enhancements of resonances in protein samples containing paramagnetic centers were used to derive long-range distances.

Peak intensities of cross-peaks for amide proton-nitrogen pairs in the ^{15}N -HSQC spectrum of the paramagnetic (I_{para}) and of the diamagnetic (I_{dia}) state of a protein were measured. Peak intensity ratios I_{para}/I_{dia} were calculated to linearly fit for the enhancement of the transverse relaxation rate by the unpaired electron (R_2^{para}) [87]:

$$\frac{I_{para}}{I_{dia}} = \frac{R_2 \exp(-R_2^{para}t)}{R_2 + R_2^{para}}, \quad (3.15)$$

where t is the total INEPT evolution time of the ^{15}N -HSQC (~ 11.3 ms) and amide proton R_2 values were approximated by experimental amide nitrogen R_2 values (see 3.5.4.2). The distances r between the unpaired electron and the amide protons have been determined according to:

$$r = \left[\frac{K}{R_2^{para}} \left(4\tau + \frac{3\tau}{1 + \omega_h^2 \tau^2} \right) \right]^{1/6}, \quad (3.16)$$

in which K is $1.23 \times 10^{-32} \text{ cm}^6\text{s}^{-2}$ [88], τ is the correlation time for the electron-nuclear interaction and ω_h is the Larmor frequency of the proton. For calculating the distances, τ was assumed to be equal to the global correlation time τ_c of the protein. τ_c was determined using the Stokes' law [5]:

$$\tau_c = \frac{4\pi\eta_W r_H^3}{3k_B T}. \quad (3.17)$$

- η_W : viscosity of the solvent
- r_H : effective hydrodynamic radius
- k_B : Boltzmann's constant
- T : temperature

Under the assumption that the specific volume of the protein is $\bar{V} = 0.73 \text{ cm}^3\text{g}^{-1}$ and that a hydration layer of $r_W = 1.6\text{-}3.2 \text{ \AA}$ surrounds the protein, a rough estimate of r_H can be made [5]:

$$r_H = \left[\frac{3\bar{V}M_r}{4\pi N_A} \right]^{1/3} + r_W. \quad (3.18)$$

- N_A : Avogadro's number

3.6 Structure calculation

3.6.1 Monomer structure calculation

The monomer structure of CylR2 was calculated using simulated annealing and torsion angle dynamics within the program CYANA [40]. Structure calculation was combined with automated NOE assignment for seven cycles calculating 100 structures each and using the 20 structures with the lowest target function for the next cycle. The final 20 structures with the lowest target function were refined with HN-RDCs in explicit solvent using a XPLOR-NIH protocol [53] which is provided in the Appendix D.

3.6.2 Dimer structure calculation

The CylR2 homodimer structure was calculated on the basis of HN-RDCs and intermolecular distances using a XPLOR-NIH protocol [89]. This rigid-body docking protocol is composed of three rounds of rigid-body minimization and a final simulated annealing. In the first round of the rigid-body minimization the force constant for dipolar couplings k_{dip} was set to $0.001 \text{ kcal mol}^{-1} \text{ Hz}^{-2}$ and the force constant for the intermonomer distances k_{inter} was set to $0.01 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. In the second round k_{dip} was slowly increased from $0.001 \text{ kcal mol}^{-1} \text{ Hz}^{-2}$ to $0.1 \text{ kcal mol}^{-1} \text{ Hz}^{-2}$ and k_{inter} was slowly increased from $0.01 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ to $30 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and in the third round k_{dip} and k_{inter} were set to $0.1 \text{ kcal mol}^{-1} \text{ Hz}^{-2}$ and $60 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$, respectively. 100 structures were calculated during the final simulated annealing applying a standard protocol with inter- and intramonomer distances and with all coordinates fixed with the exception of the side chain atoms of residues contributing to the dimer interface. The 20 lowest energy structures after the simulated annealing were refined with the same protocol as the monomer ensemble.

3.6.3 *Ab initio* docking

For *ab initio* docking with restricted symmetry the ClusPro protein-protein docking Web server (<http://nrc.bu.edu/cluster>) was used [37, 38]. Refined monomer struc-

tures were docked with the standard settings, but with ZDOCK chosen as the docking program and the symmetry restricted to C_2 [39]. 15 structures were received as the output.

3.6.4 Protein/DNA docking

The structure of the CylR2/DNA complex was modeled by Monte Carlo docking using the program Monty [90, 91]. DNA bending in Monty is achieved by a two-step procedure. Firstly a regular helix is fitted through the B-DNA backbone which is randomly over- or underwound. The straight starting double helix is subsequently bent by shifting each base pair, which is not in van der Waals contact with the protein. These base pairs are shifted along an intermolecular vector toward the protein by a distance smaller or equal to the so-called bending parameter. The outer base pairs are shifted by a larger amount than the inner ones, which is determined by a quadratic function of base pair numbering. A larger value for the bending parameter thus will result in a stronger shifting of base pairs toward the protein, i.e., a stronger bending of the DNA [90, 91]. Here, docking simulations were performed with DNA bending parameters varying from 0.0 to 0.6. Corresponding to NMR titration results, side chains were randomised and allowed to rotate freely during the entire simulation and energy bonuses of 30 and 40 kcal/mol were given to complex structures where the backbone amides are in close contact with the DNA. The square-well depths for hydrogen bonds and van der Waals interactions were 5 and 0.5 kcal/mol, respectively. The value of kT was set to 7.7 kcal/mol. For each DNA bending parameter 64 Monte Carlo simulations of 200000 attempted moves were performed.

3.7 Structure analysis

References for the utilized softwares can be found in Table 2.9. For the analyses of the X-ray structure of CylR2 PROCHECK and the CCP4 suite of programs were employed. Structure figures were prepared with MOLSCRIPT, BOBSCRIPT, RASTER3D, MSMS, DINO, PyMOL and MOLMOL. NMR structures were analysed with MOLMOL, AQUA and PROCHECK-NMR. Energies were evaluated with

XPLOR-NIH.



4

Structure and DNA-binding properties of the cytolysin regulator CylR2

4.1 Introduction

Enterococcus faecalis is one of the major causes for hospital-acquired antibiotic-resistant infections [92, 93]. It produces an exotoxin, called cytolysin [94], which is lethal for a wide range of Gram-positive bacteria and is toxic to higher organisms. The toxicity of strains expressing cytolysin was found to be at least 10-times higher than for isogenic, non-cytolytic mutants [95, 96] and the virulence of enterococcal infections is five-fold increased because of cytolysin secretion [97]. Cytolysin is a heterodimer consisting of a large subunit Cyl_L and a small subunit Cyl_S, which are post-translationally modified by CylM. CylM introduces the modified amino acids lanthionine, β -methylanthionine, dehydroalanine and dehydrobutyrine which are characteristic of the lantibiotic class of bacteriocins. After modification, both subunits interact with CylB. This ATP-binding cassette (ABC) transporter has a dual function: it removes a leader sequence from the modified subunits and secreted them into the extracellular space [94, 98]. In the extracellular space the subtilisin-class serine protease, CylA, activates cytolysin by cleaving the precursors and releasing the mature subunits Cyl_L and Cyl_S [99]. These subunits interact to cause lysis of erythrocytes, to destroy neural tissue and to kill a wide range of Gram-positive bacteria [100, 101]. The cytolysin-producing cell is protected from the bactericidal activity by the specific immunity protein CylI [102]. The genes for all these proteins

including those of cytolysin were found on a collinear gene cluster [103](Figure 4.1). As cytolysin is lethal for strains that lack CylII [105], regulation of cytolysin ex-

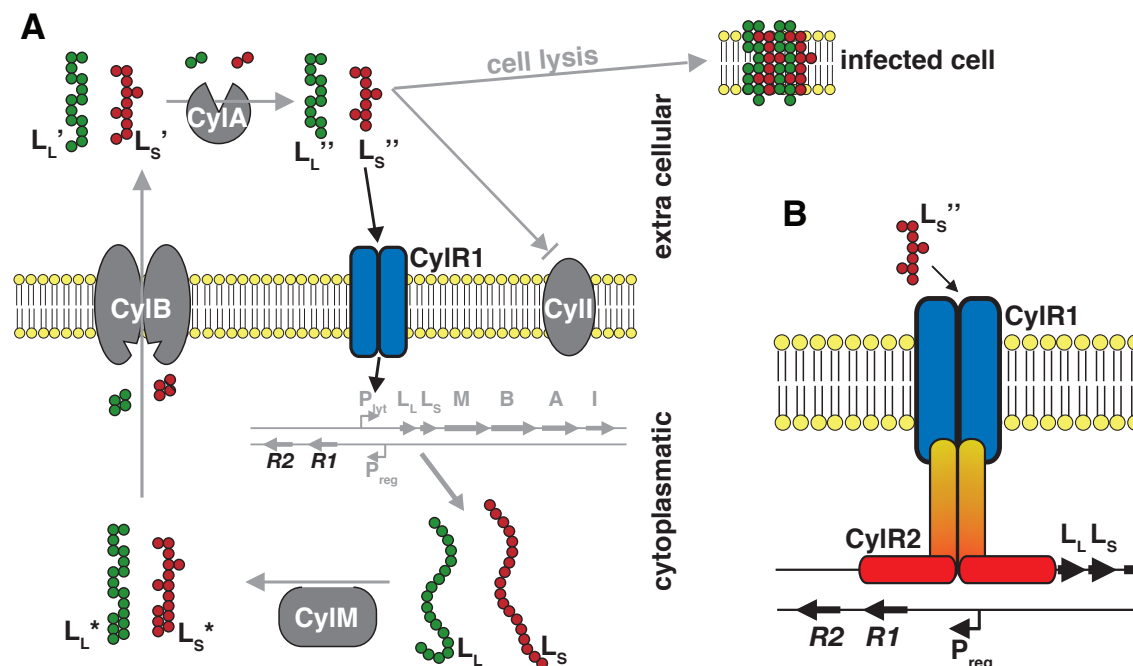


Figure 4.1: (A) Model of cytolysin maturation. The subunits $CylL_L$ and $CylL_S$ are ribosomally synthesized, modified post-translationally by CyIM, proteolytically cleaved, secreted and extracellularly activated by CyIA. The fully active subunits $CylL_L''$ and $CylL_S''$ interact to cause cell lysis. The cytolysin-producing cell is protected by CylII. The box indicates the part of the scheme which is enlarged in (B): Model of the regulation of cytolysin production. CylR1 and CylR2 act in concert to repress expression of the cytolysin operon with CylR2 binding to the promoter P_{lyt} . Interaction between $CylL_S''$ and CylR1 induces transcription from P_{lyt} [104].

pression has to be carefully orchestrated with regulation of immunity. Regulation of cytolysin production involves two proteins, CylR1 and CylR2, whose genes were identified upstream of the collinear gene cluster in opposite orientation [104]. Both, CylR1 and CylR2, were shown to be essential for the regulation. This CylR1/CylR2 two-component regulatory system was found to be connected to autoinduction by a quorum sensing mechanism [104]. Quorum sensing is generally a bacterial mechanism of bacteria for communication and coordinated behaviour at the level of cell density and the signalling molecules are also referred to as autoinducers [106]. In most cases of cell-cell communication between bacteria, binding of a signalling peptide to a histidine kinase receptor on the cell surface results in autophosphorylation of the

kinase. The phosphate is then transferred to a response regulator converting it to an active transcription factor. Alternatively, the peptide signal is imported into the cell by an oligopeptide permease where it interacts with an intracellular receptor [107]. However, the quorum sensing mechanism involving the CylR1/CylR2 two-component regulatory system is different from the previously described two-component quorum sensing signal transduction systems and neither CylR1 nor CylR2 are related to the family of bacterial two-component regulators [104]. Repression of the cytolysin gene is overcome in a cell density-dependent fashion as a result of the accumulation of one of the posttranslationally modified subunits of the toxin, CylL_S”, but the mechanism by which this accumulation is sensed by the cell is presently unknown. It is predicted that CylR2 may contain a helix-turn-helix motif, and that CylR1 would likely be a membrane-bound protein. Thus, a model was proposed by Haas *et al.* [104] in which CylR2, interacting with membrane-bound CylR1, binds and represses the cytolysin promoter until CylL_S” is accumulated to a threshold level to trigger a change that alters or releases the promoter, allowing for transcription (Figure 4.1).

Aside from the demonstration that mutations in either CylR1 or CylR2 lead to derepression of the operon [104], precise details of the mechanism of how quorum sensing autoinduction of cytolysin expression is achieved remain to be determined. The binding of CylR2 to either of the two inverted repeats (IR1 and IR2) within the cytolysin promoter (Figure 4.2) has been determined by gel shift analysis [108]. The absence or

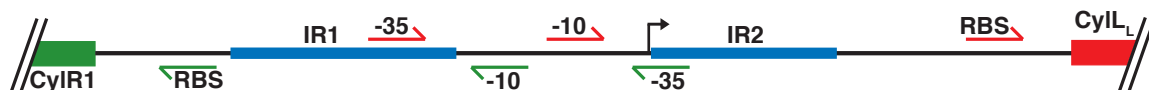


Figure 4.2: Intergenic region between *cylL_L* and *cylR1*, containing divergent promoters, P_{reg} and P_{lyt} , involved in cytolysin expression and regulation. The two palindromic sequences, IR1 and IR2, within the promoter region are indicated and the transcriptional start site is indicated as a bent arrow. Promotor elements (-10: -10 box, -35: -35 box) for P_{reg} and P_{lyt} are shown in green and red, respectively.

presence of band shifts revealed that CylR2 binds only to the 25-bp double-stranded (ds) IR1 and neither to single-stranded (ss) IR1, nor ds IR2 or ss IR2. Furthermore, the CylR2-binding was unchanged for 25-bp ds DNA with up to 4 nucleotides missing

at the 5' or 3' end of IR1. Having this information on-hand, it is of interest to determine the high-resolution structure of CylR2 and its DNA-binding properties as an important step towards unravelling the mechanism of regulation of the enterococcal virulence factor cytolysin and towards approving the proposed mechanism.

4.2 Materials and methods

4.2.1 Expression and purification

Cell cultures for the expression of CylR2, wildtype or mutants, were switched to 28 °C at an OD₆₀₀ of 0.4-0.5 and were harvested 6-7 h after induction (see 3.2.1). The cell lysis was performed under native conditions with DNase I supplemented in order to degrade DNA (see 3.2.2). After the native cell lysis, the supernatant was filtrated through a 0.20 μm sterile filter and loaded on a 5-ml anion exchange column to remove the DNA (see 3.3.2). The flow-through and 5 column volumes of wash were pooled and loaded on a 10-ml cation exchange column equilibrated with lysis buffer. Elution was performed at a flow rate of 3.0 ml/min with a linear gradient of 0-100 % of ion exchange buffer B for CylR2 (see 3.3.2). Fractions containing CylR2 were pooled, shock-frozen in liquid nitrogen and dried by lyophilisation. The lyophilised protein was dissolved in ~3 ml water and further purified by RP-HPLC (see 3.3.4). The purity of the eluted CylR2 was confirmed by ESI-MS (Figure A.2). The purified CylR2 was stored in lyophilised form at -20 °C. The yield is ~25 mg CylR2 per liter of expression culture.

4.2.2 Preparation of ds DNA

Complementary oligonucleotides (Table 2.3) were resuspended in 50 mM HEPES pH 7.0 and 350 mM NaCl, mixed to a final 1:1 molar ratio and annealed by heating to 100 °C for 5 min in a heatblock and cooling down slowly to room temperature. The annealed DNA was used for NMR titration experiments or in the case of the ds DNA comprising IR1 (Figure 4.3) for the preparation of CylR2/DNA complex samples.

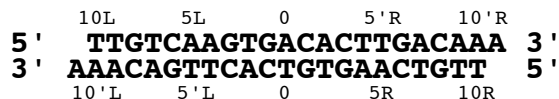


Figure 4.3: The 22-mer DNA sequence of IR1 used for the NMR experiments.

4.2.3 NMR experiments

About 1 mM CylR2 samples for resonance assignment were prepared by dissolving lyophilised CylR2 directly in 250 μ l NMR buffer with 600 mM NaCl. At lower salt concentrations CylR2 is less soluble and tends to crystallize. Therefore, for NMR samples with lower salt concentrations CylR2 was first dissolved in 250 μ l 50 mM HEPES pH 7.0, 1.5 M NaCl followed by dialysis against 50 mM HEPES pH 7.0 with 250 to 350 mM NaCl. Finally 5 % (v/v) D₂O were added.

For CylR2/DNA complex samples, the protein in 50 mM HEPES pH 7.0 and 1.5 M NaCl was slowly mixed with the DNA (see 4.2.2) up to a 1:1 molar ratio. Afterwards the complex was dialysed against 50 mM HEPES pH 7.0 and 350 mM NaCl and 5 % (v/v) D₂O were added. NMR samples contained 1 mM protein/DNA complex.

Anisotropic samples for the measurement of HN-RDCs contained 12 mg/ml Pf1. To lower the electrostatic attraction between CylR2 and the highly negatively charged Pf1, HN-RDCs for free CylR2 were measured at a NaCl concentration of 0.5 M. IPAP-¹⁵N-HSQC spectra of CylR2 in complex with the DNA under anisotropic conditions were acquired at NaCl concentrations of 0 M, 0.025 M, 0.05 M, 0.1 M, 0.2 M, 0.3 M and 0.5 M. HN-RDCs for DNA-bound CylR2 at 0.5 M NaCl were measured for 12 mg/ml and 14 mg/ml Pf1.

All NMR experiments were carried out at 298 K at a field strength of 600, 700 or 800 MHz and are summarized in the Appendix Table B.1.

4.2.4 DNA titration

The CylR2/DNA complex formation was monitored by recording a series of 2D ¹⁵N-¹H-HSQC spectra of a 100 μ M ¹⁵N-labeled CylR2 NMR sample with a stepwise increasing DNA concentration. DNA was added from a 1.4 M stock solution and final DNA concentrations were 10, 20, 30, 50, 60 and 100 μ M. The combined chemical shift

changes for a particular residue upon DNA-binding were calculated according to:

$$\Delta\delta_{tot} = \sqrt{(\Delta\delta_{HN})^2 + (f_{C\alpha}\Delta\delta_{C\alpha})^2 + (f_{C\beta}\Delta\delta_{C\beta})^2 + (f_{C'}\Delta\delta_{C'})^2} \quad (4.1)$$

with $\Delta\delta_{HN}$ calculated following equation 3.11.

$$\Delta\delta_{C\alpha,\beta} = \sqrt{(f_{C\alpha}\Delta\delta_{C\alpha})^2 + (f_{C\beta}\Delta\delta_{C\beta})^2} \quad (4.2)$$

$$\Delta\delta_{C'\star} = \sqrt{(f_{C'}\Delta\delta_{C'})^2} \quad (4.3)$$

and for amide side chains:

$$\Delta\delta_{side} = \sqrt{(f_{H\epsilon}\Delta\delta_{H\epsilon})^2 + (f_{N\epsilon}\Delta\delta_{N\epsilon})^2 + (f_{C\gamma/\delta}\Delta\delta_{C\gamma/\delta})^2} \quad (4.4)$$

$$\Delta\delta_{sideC} = \sqrt{(f_{C\gamma/\delta}\Delta\delta_{C\gamma/\delta})^2} \quad (4.5)$$

$\Delta\delta_i$ are the differences in chemical shifts between the free and bound protein and f_i denotes the weight factor of nucleus i ; $f_{H\epsilon} = 1$, $f_{N\epsilon} = 0.154$, $f_{C\alpha} = f_{C\beta} = 0.256$ and $f_{C\gamma/\delta} = f_{C'} = 0.341$ [109].

4.2.5 Model of the CylR2/DNA complex

Initially, a starting model for the CylR2/DNA complex was built by superimposing the backbone atoms of residues 1 to 66 of the CylR2 dimer on corresponding atoms of residues 1 to 66 of the N-terminal domain of repressor of phage 434 in complex with its DNA operator (PDB code: 1rpe) [110]. The 434 repressor DNA was replaced by a 22 bp B-DNA, which was adjusted to the IR1 inverted repeat sequence using InsightII, by positioning bp 9L to 9R of the B-DNA on bp 2R to 3L of the 434 repressor DNA. The program Monty was used for the refinement of this start structure (see 3.6.4). Further docking trials were performed starting from initial models, where the DNA operator in complex with the 434 repressor (PDB code: 1rpe) or the lambda repressor (PDB code: 1lmb) [111] was used instead of regular B-DNA. No DNA flexibility was allowed in these cases and the DNA sequence was adjusted prior to docking to IR1 using InsightII. Prolonged simulations with 600000 Monte Carlo moves showed virtually no improvement. From the 64 final solutions, those in which the DNA was

detached from the protein were removed, leaving a total of 41 CylR2/DNA complexes. Since the number of possible side chain conformers is extremely large and sampling is computationally limited, it is unlikely that one can obtain complexes with all correct protein-DNA interactions simultaneously present. Therefore, important CylR2/DNA interactions were identified based on their presence in more than 20 % (van der Waals) or more than 30 % (H-bonds) of the successful docking solutions.

4.3 Results and discussion

4.3.1 Three dimensional structure of CylR2

The crystal structure of CylR2 was determined at 1.9 Å resolution in the group of Stefan Becker by Adelia Razeto. The model of CylR2 consists of 132 residues, corresponding to two molecules in the asymmetric unit, 15 iodide ions, a sodium ion and 194 water molecules and is deposited with the PDB (accession code: 1utx). The structure of CylR2 consists of five helices ($\alpha 1$ - $\alpha 5$) and an antiparallel β -sheet, which is formed by strands $\beta 1$ and $\beta 2$, at the N- and C-terminus of CylR2, respectively (Figure 4.4). Helices $\alpha 2$ and $\alpha 3$ make up the HTH DNA-binding motif and are defined as the ‘scaffold’ and ‘recognition’ helix, respectively [112]. Helices $\alpha 1$, $\alpha 2$, and $\alpha 3$ pack against each other and the architecture of their fold is supported by hydrophobic and hydrophilic interactions, reflecting their amphipathic nature. The hydrophobic residues are oriented towards the centre of the molecule, while hydrophilic residues, on the surface of the molecule, protect the hydrophobic core. A buried salt-bridge between Arg10 and Glu35 connects helix $\alpha 1$ to $\alpha 3$. The salt-bridge has already been indicated as an important feature in HTH DNA-binding proteins with the function of stabilising the fold and determining the relative orientation of the scaffold and recognition helices [113]. In HTH DNA-binding proteins the tight turns between the helices are usually sterically favoured by glycine [114]. However, in CylR2 other residues replace these glycines in sharp bends: Lys14 ($\alpha 1$ - $\alpha 2$), Glu25 ($\alpha 2$ - $\alpha 3$), Asn37 ($\alpha 3$ - $\alpha 4$) and Asn54 ($\alpha 4$ - $\alpha 5$). All of these have main-chain dihedral angles lying in the left-handed α -helical region of the Ramachandran plot, which is usually occupied by

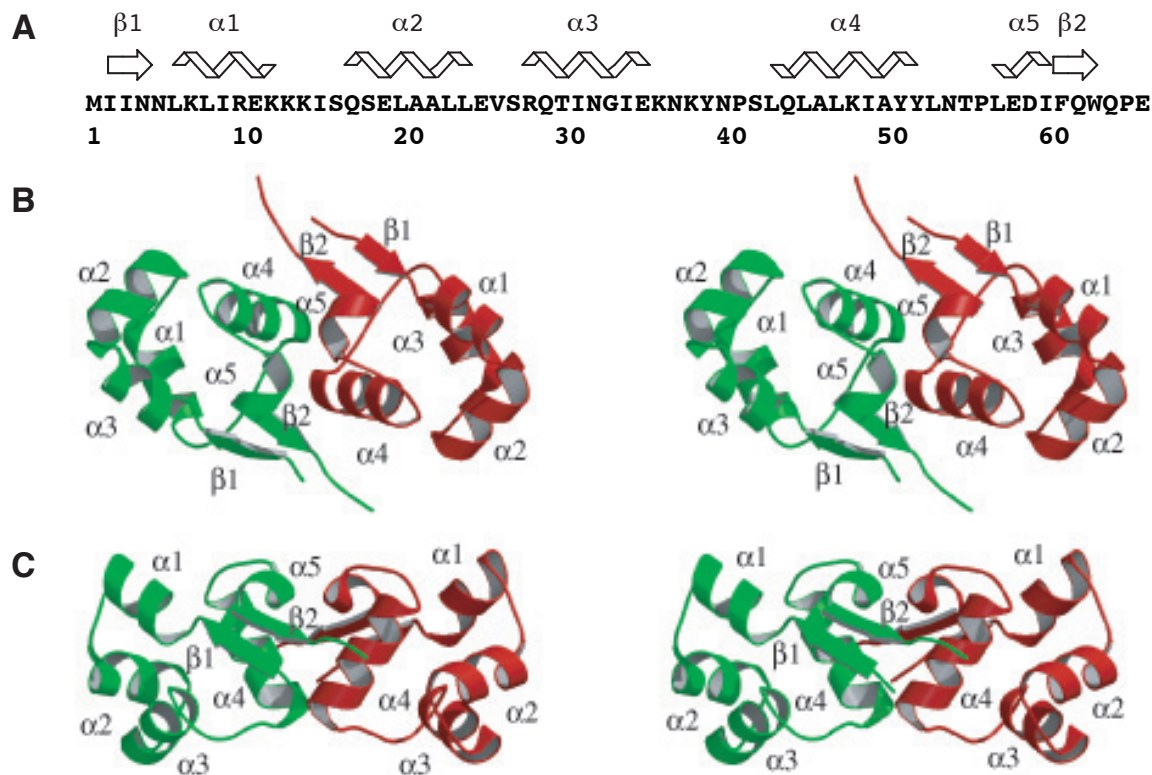


Figure 4.4: Sequence of CylR2 (A) and stereo view of the crystal structure of CylR2 (B,C). Helices are labeled with $\alpha 1$ - $\alpha 5$ and β -strands are labeled with $\beta 1$ and $\beta 2$. (B,C) Subunit A and B are shown in red and green. (B) Stereo view of the homodimer looking down the dimer twofold axis. (C) CylR2 is rotated by 90° about the horizontal axis.

glycines. Their main-chain conformations are mostly stabilised by H-bonds between their amino groups and carbonyl groups of residues (n-3), a typical feature of β -turns. Helices $\alpha 4$ and $\alpha 5$ are predominantly hydrophobic, contribute to the hydrophobic core of the molecule and, more importantly, mediate the dimeric contacts (see 4.3.2).

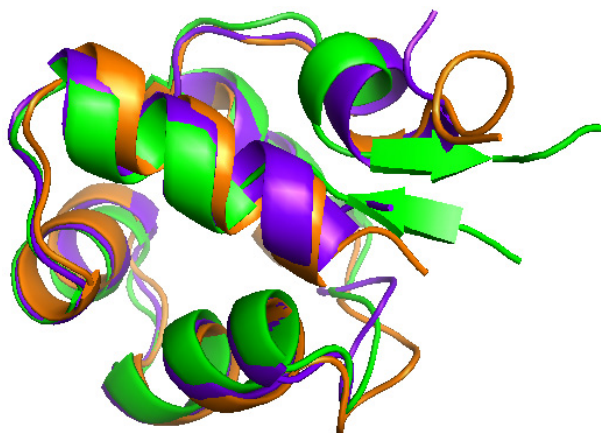


Figure 4.5: Superposition of the CylR2 monomer with SinR and 434 repressor. CylR2 is shown in green, SinR in orange and 434 repressor in violet.

A DALI [115] search identified *Bacillus subtilis* SinR [116] and the N-terminal domain of 434 repressor as the structures having the closest structural similarity to CylR2 dimer with rmsd values of 2.2 and 1.7 Å, respectively (alignment of 62 C α atoms of SinR and 59 C α atoms of 434 repressor with structurally equivalent atoms of CylR2). The α -helical portion of the structure superimposes quite well on both, SinR [116] and 434 repressor [114]. The main divergences are in the turn between α 3 and α 4 and in the N- and C-terminal regions (Figure 4.5). In CylR2 the N- and C-terminal helices α 1 and α 5 are shorter than in SinR and 434 repressor and are preceded and followed by β -strands, which form an antiparallel β -sheet (Figure 4.5). SinR and 434 repressor lack this β -sheet.

4.3.2 Dimerisation

In the crystal structure CylR2 forms a dimer with overall dimensions of about 18 Å x 36 Å x 44 Å. The two subunits, which are related by a twofold non-crystallographic symmetry, are very similar: the rmsd is 0.7 Å by superposition of C α atoms (Figure 4.4). The N- and C-terminal residues have the highest rms deviations of about 2 Å. Helices α 4 and α 5 contribute extensively to the dimeric interface (Figure 4.6). Helices α 4 are very close to one another at their N-termini (the distance between

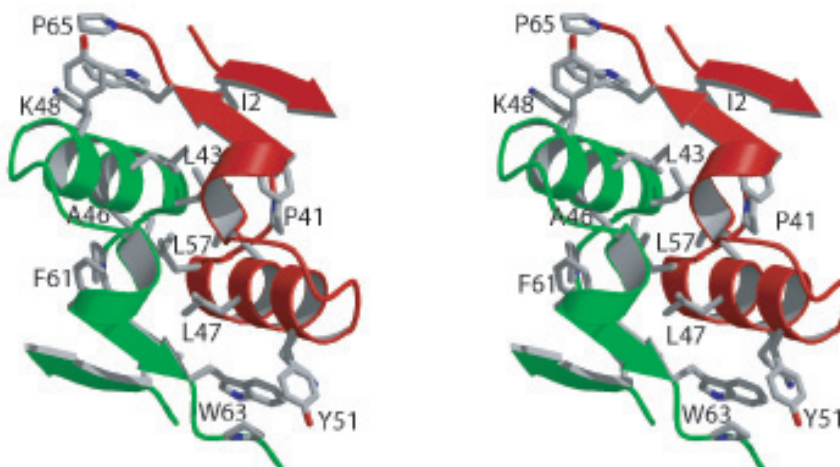


Figure 4.6: Dimer interface of CylR2. Residues 1 to 5 and 40 to 65 of subunit A and B are colored red and green. Side chains of the predominantly hydrophobic residues Ile2, Pro41, Leu43, Ala46, Leu47, Lys48, Tyr51, Leu57, Phe61, Trp63 and Pro65 are shown. The orientation corresponds to Figure 4.4A.

the Leu43 Cas is 4.5 Å) and diverge at their C-termini. C-terminally of $\alpha 4$ the two CylR2 subunits come again closer by means of helices $\alpha 5$, whose N-termini contact one another. Strands $\beta 1$ and $\beta 2$ mediate also intersubunit contacts by sandwiching helices $\alpha 4$ and $\alpha 5$. In general, the side chains along the dimeric interface have the same conformations and superimpose well. The interface is predominantly hydrophobic, consisting of residues: Ile2, Pro41, Leu43, Ala46, Leu47 and the aliphatic part of Lys48, Tyr51, Leu57, Phe61, Trp63, Pro65 (Figure 4.6). The hydrophilic interactions between the subunits are mostly water-mediated. There are only four direct hydrogen bonds from the side chains of Gln44 and Glu58 of one subunit to the main chain nitrogens of Met1 and Leu57 of the other subunit, respectively. The side chains of the two Glu58 come close to one another. A sodium ion between the Glu58 side chains neutralises the local negative potential.

A similar dimerisation mode is also observed for the HTH DNA-binding protein GerE [117]. However, helices $\alpha 5$ and strands $\beta 1$ and $\beta 2$ are missing and the angle between helices $\alpha 4$ is 20° narrower than in CylR2 (Figure 4.7). In CylR2 helices

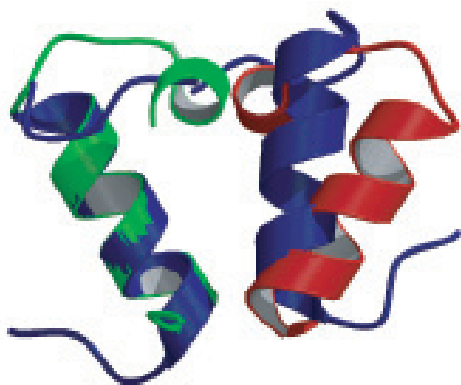


Figure 4.7: Residues 40 to 60 of CylR2 (red/green) are superposed on residues 55 to 74 of GerE (blue). The orientation corresponds to Figure 4.4B.

$\alpha 5$ are pushing helices $\alpha 4$ apart (Figure 4.6). Dimerisation of CylR2 is achieved by formation of an extensive hydrophobic interface, which is shielded from the solvent (Figure 4.6). It buries a total of 1600 Å² of the solvent accessible surface, accounting for about 18 % of the total accessible area of each subunit. This is significantly more than in GerE and the N-terminal domain of 434 repressor, where 12 % [117] and 9 % [118] of the accessible area of each subunit are buried upon dimerisation, respectively. Especially for the 434 repressor, dimerisation is energetically more favourable

in the DNA complex structure as charged residues at the dimer interface are neutralised upon DNA-binding [118]. The relative orientations of the two monomers in CylR2 and the N-terminal domain of 434 repressor bound to DNA are quite different. If only one monomer of CylR2 is superimposed on one 434 repressor monomer, the other monomers are related by a rotation of about 20° around the axis connecting the CylR2 Ser42 C α s (Figure 4.8). In addition, the two recognition helices are about

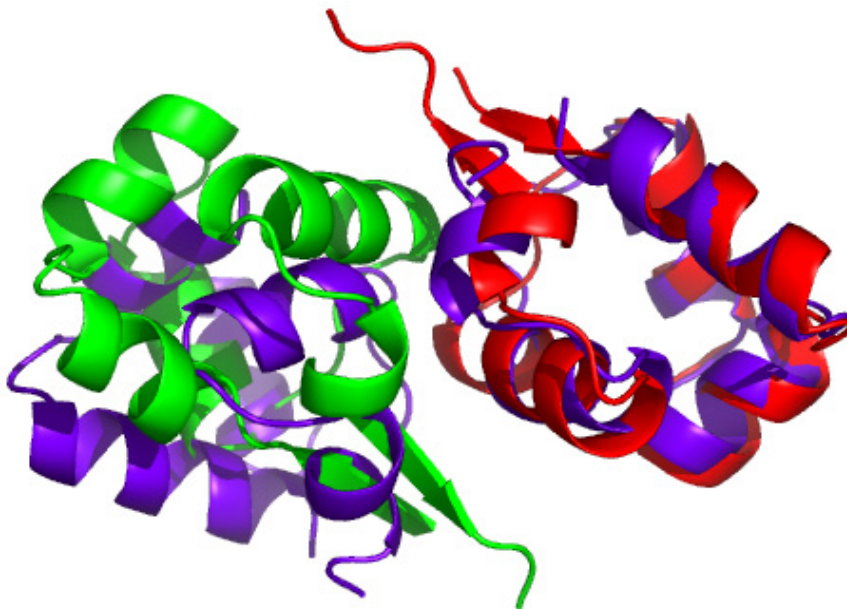


Figure 4.8: Superposition of one subunit of CylR2 dimer to 434 repressor dimer. The subunits of CylR2 are colored red/green and 434 repressor is colored violet. The orientation corresponds to Figure 4.4A.

2 Å further apart in CylR2 than in the N-terminal domain of 434 repressor bound to DNA.

4.3.3 Solution NMR studies of CylR2 and CylR2/DNA complex

Analysis of triple resonance NMR spectra allowed identification and sequential assignment of all backbone amide resonances and of all C α and C β chemical shifts (see 3.5.1). The assignment of C' (except of Pro65), H α and the aliphatic proton and carbon side chain resonances is complete for all residues with exception of the carboxyterminal Glu66. Protonated ^{15}N resonances have been found for the side chains of all five glutamine residues, of five out of six asparagines and of the tryptophan side

chain. More than 80 % of the aromatic proton and carbon ^1H and ^{13}C side chain resonances have been unambiguously assigned. Assigned ^1H , ^{13}C and ^{15}N chemical shifts have been deposited in the BMRB (<http://www.bmrb.wisc.edu>), accession code 6317. This 98.8 % complete assignment, was used in Chapter 5 for calculating the monomer solution structure of CylR2.

NMR relaxation times indicate that CylR2 is dimeric in solution for concentrations between 10 μM and 2 mM (Appendix Figure B.1). In order to detect differences between crystal and solution structure of CylR2 HN-RDCs were measured (see 3.5.3.1). Comparison of experimental $^1\text{D}_{\text{HN}}$ couplings measured at 0.5 M NaCl with values back-calculated from the CylR2 crystal structure using SVD resulted in a correlation coefficient of 0.97 and a Q-factor of 0.19 for 57 out of 62 backbone amide protons (Figure 4.9A). The experimental HN-RDCs of free CylR2 can be found in the Appendix Table C.1. This proves that the backbone structure of CylR2 in solution

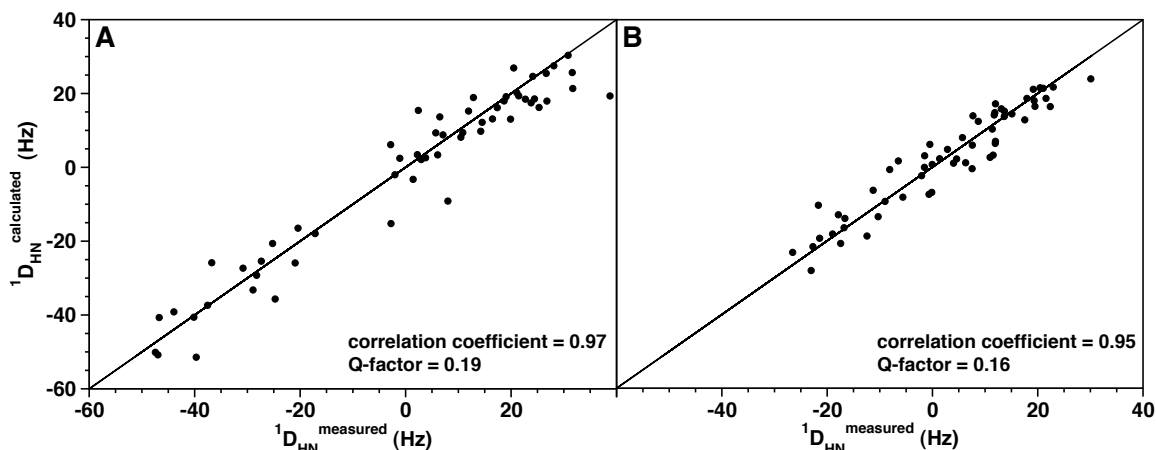


Figure 4.9: HN-RDC of CylR2. Correlation between experimental backbone $^1\text{D}_{\text{HN}}$ couplings and values back-calculated from the crystal structure. (A) Correlation for free CylR2 at 0.5 M NaCl and (B) correlation for DNA-bound CylR2 at 0 M NaCl.

agrees well with that in the crystalline state. The magnitude D_a and the rhombicity R of the alignment tensor were back-calculated as $D_a = -26.0$ Hz and $R = 0.16$.

In order to evaluate if DNA-binding is accompanied by changes in the backbone conformation of CylR2, HN-RDCs were measured for CylR2 when bound to DNA (see 3.5.3.1). Best-fitting $^1\text{D}_{\text{HN}}$ couplings observed at 0 M NaCl for CylR2 in complex with DNA to the crystal structure of free CylR2 results in a correlation coefficient

of 0.94 and a Q-factor of 0.16 between experimental and back-calculated RDCs for 57 out of 62 backbone amide protons (Figure 4.9B). The back-calculated alignment tensor is characterized by $D_a = -22.1$ Hz and $R = 0.07$ for the protein/DNA complex at 0 M NaCl. Due to a significantly reduced signal-to-noise ratio of the IPAP- ^{15}N -HSQC spectrum acquired at 0.5 M NaCl for DNA-bound CylR2, experimental HN-RDCs could only be determined accurately for 49 out of 62 backbone amides. These HN-RDCs resulted in a correlation coefficient of 0.94 and a Q-factor of 0.19. The SVD results for DNA-bound CylR2 at the different NaCl concentrations are summarized in Table 4.1 and the experimental HN-RDCs are listed in the Appendix Table C.2. The

Table 4.1: Best-fit of HN-RDCs for DNA-bound CylR2 to the X-ray structure of free CylR2.

NaCl	Pfl	correlation coefficient	Q-factor	D_a	R	number of HN-RDCs
0 M	12 mg/ml	0.951	0.158	-22.1 Hz	0.067	57
0.025 M	12 mg/ml	0.941	0.177	-20.5 Hz	0.057	57
0.05 M	12 mg/ml	0.937	0.182	-19.2 Hz	0.075	57
0.1 M	12 mg/ml	0.939	0.184	-17.6 Hz	0.184	57
0.2 M	12 mg/ml	0.942	0.176	-15.1 Hz	0.176	57
0.3 M	12 mg/ml	0.941	0.189	-13.4 Hz	0.121	52
0.5 M	14 mg/ml	0.943	0.189	-19.2 Hz	0.189	49

very good correlation between the experimental HN-RDCs for DNA-bound CylR2 and the X-ray structure of free CylR2 at all salt concentrations shows that there are no major changes in the backbone structure of CylR2 upon binding to DNA.

Dynamics of free and DNA-bound CylR2 in solution were probed by heteronuclear NMR spectroscopy to reveal flexible regions and mobility changes upon DNA-binding. Steady state heteronuclear ^{15}N - ^1H -NOE values (see 3.5.4.3) for most of the CylR2 backbone amides in the free and bound form are above 0.7 indicating that the backbone of CylR2 is well ordered in solution (Figure 4.10). Only Ser42 in the loop connecting α -helices 3 and 4 is flexible with a ^{15}N - ^1H -NOE of 0.4 and the four N- and C-terminal residues are slightly affected by motion. Fluctuations of ^{15}N - ^1H -NOE values in the β -strands and α -helix 2 for bound CylR2 are due to a bad signal to

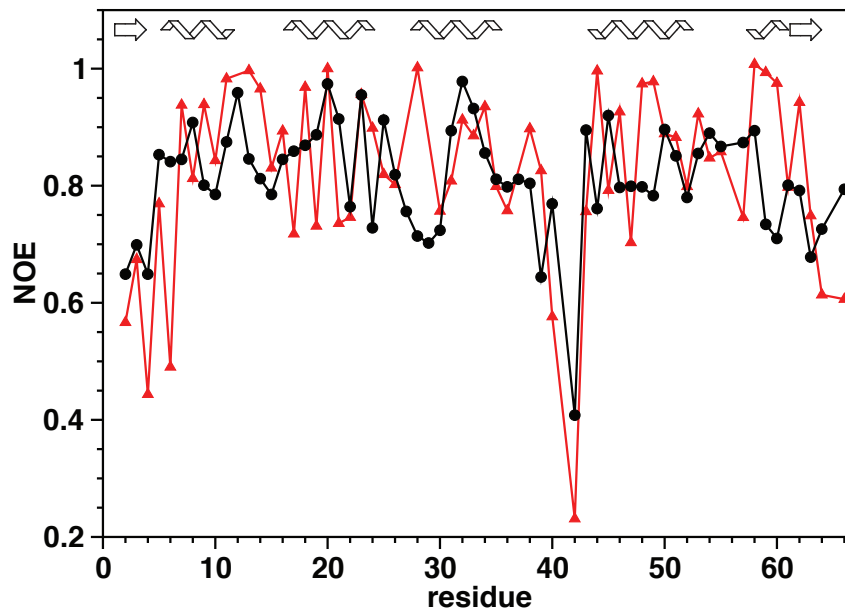


Figure 4.10: The ^{15}N - ^1H -NOE plotted against the residue number. Values for free CylR2 are indicated as black circles and values for CylR2 in complex with its DNA are shown as red triangles. Secondary structure elements are indicated.

noise ratio.

4.3.4 Mapping of the DNA-binding site

The DNA interaction surface of CylR2 was determined using the NMR chemical shift perturbation method (see 3.5.5), in which ^{15}N -HSQC spectra were recorded with successive addition of the 22-bp DNA comprising IR1 (see 4.2.4). The spectra for the free CylR2 and the DNA-bound CylR2 are shown in Figure 4.11. Upon DNA-binding, resonances of unbound CylR2 weaken continuously and new resonances appear with significant chemical shift differences. This indicates that the free protein is in slow exchange on the NMR time scale with its DNA-bound form. The apparent K_d of 2.2 nM was determined by gel shift analysis by Chris M. Pillar [108]. A similarly strong DNA-binding with a K_d of about 3.3 nM was observed for the repressor from bacteriophage 434 [119, 120]. For both, free and bound protein, only a single set of resonances was observed. In order to allow quantification of the observed chemical shift changes, backbone resonances of CylR2 in complex with its DNA were reassigned. Assigned chemical shifts of CylR2 in the complex can be found in the Appendix B.7.

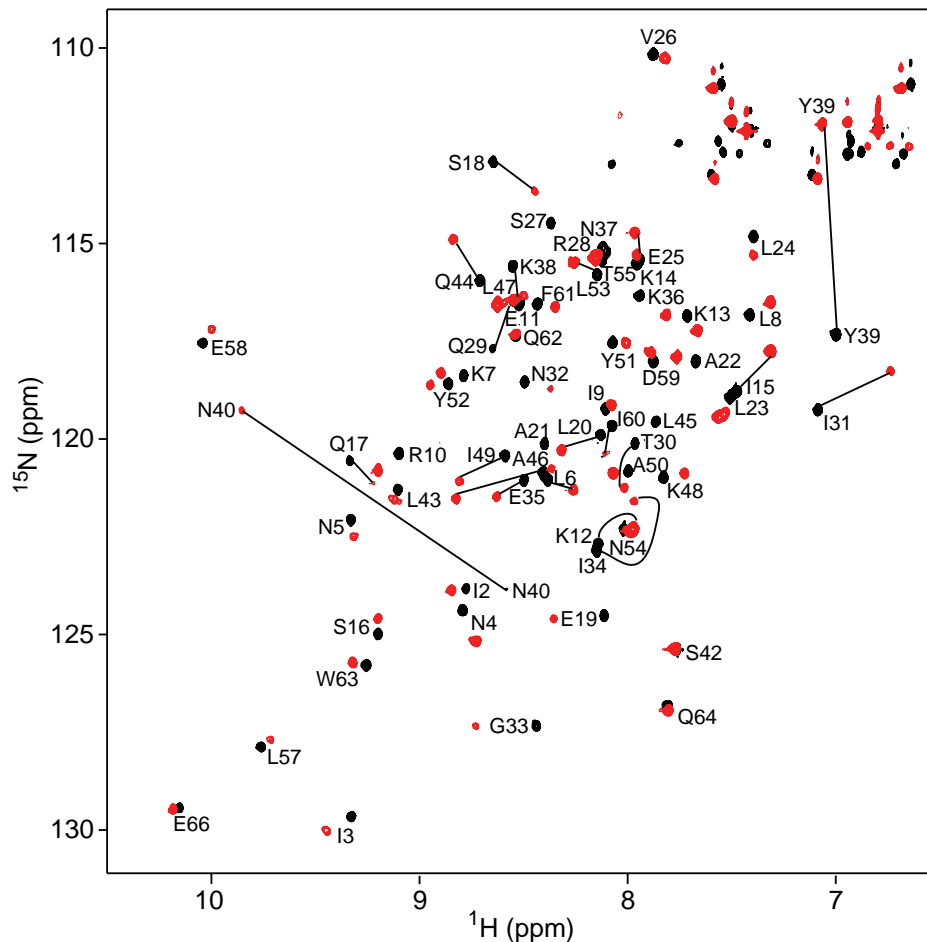


Figure 4.11: Overlay of ^{15}N -HSQC spectra of CylR2 (black) and CylR2/DNA complex (red). Black line indicate strongly shifted resonances of free and complexed CylR2.

The result of the mapping of the DNA-binding surface of CylR2 for the amide resonances is shown in Figure 4.12A. Largest chemical shift changes were observed for Tyr39 and Asn40. In addition to the amide chemical shift changes, the chemical shift perturbation was analysed for C^α , C^β , C' and the side chain amides (see 4.2.4) to give more detailed information about the DNA-binding. The side chain amides of Gln29, Asn40 and Gln44 are strongly affected and large chemical shift changes are found for residue 16-20, 28-34 and 44-46. This indicates that especially the recognition helix ($\alpha 3$), the loop connecting helices $\alpha 3$ and $\alpha 4$ and the N-terminal halves of helices $\alpha 2$ and $\alpha 4$ are involved in IR1 binding. Residues of CylR2 form a positively charged surface for DNA-binding and residues strongly affected by DNA-binding are colored on the surface of CylR2 (Figure 4.13). A similar DNA-binding mode was found for

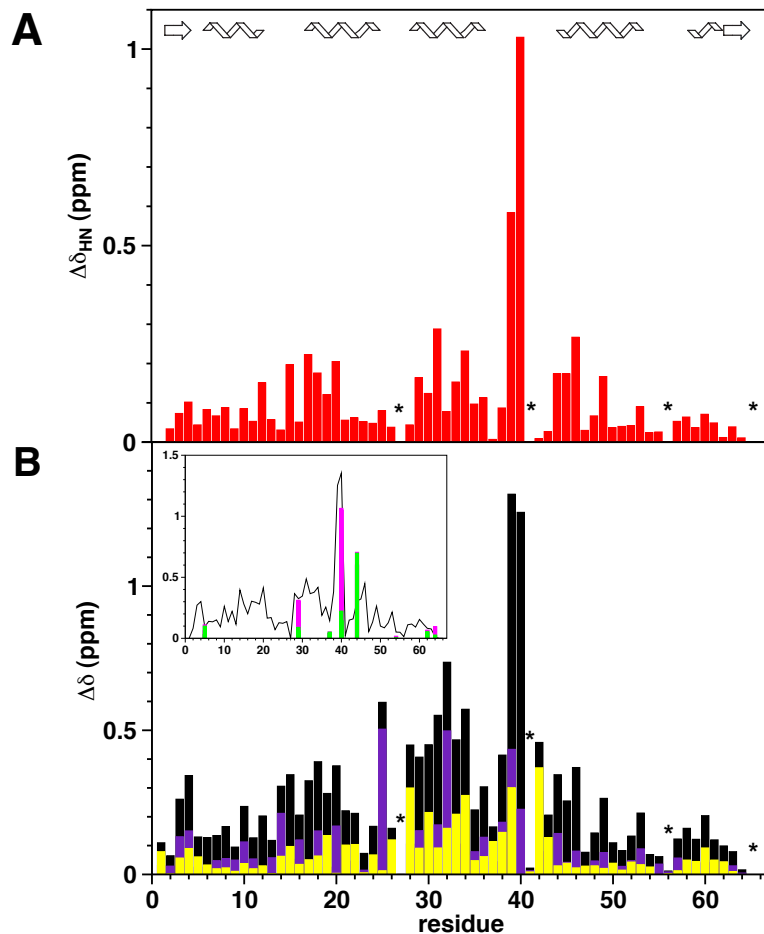


Figure 4.12: Chemical shift changes upon CylR2/DNA complex formation. The for DNA-bound CylR2 unassigned residue Ser27 and proline residues are labeled with an asterisk. (A) Amide chemical shift changes and (B) Combined chemical shift changes ($\Delta\delta_{tot}$ in black, $\Delta\delta_{C\alpha C\beta}$ in violet and $\Delta\delta_{C'\ast}$ in yellow). The inset displays $\Delta\delta_{tot}$ as a solid line and the chemical shift changes for the complete amide side chain $\Delta\delta_{side}$ in magenta and for the amide C^γ or C^δ $\Delta\delta_{sideC}$ in green.

the N-terminal domain of the 434 repressor [110, 118]. The DNA titration was in addition performed for an 18-bp DNA comprising IR1, but no chemical shift changes could be observed for CylR2. This proves that CylR2/DNA interactions also occur outside of the 18-bp core and the minimum length of an IR1 fragment required for binding to CylR2 is 20 or 22 bp long.

The NMR chemical shift perturbation studies implicated an important role of Asn40 of CylR2 in binding the cytolysin promoter. Based on this result, the consequence on cytolysin expression upon mutating this residue to Ala was examined by Chris

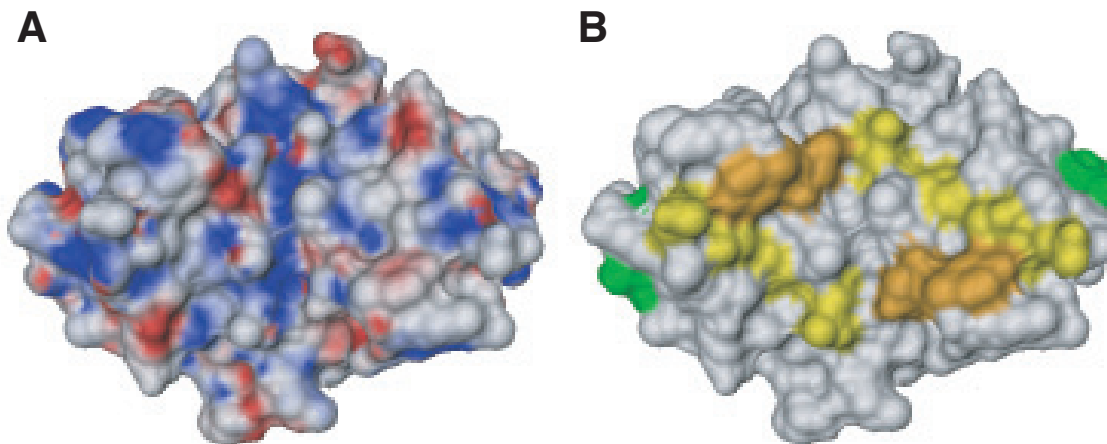


Figure 4.13: (A) Electrostatic surface of CylR2 with positive and negative potentials colored blue and red, respectively. (B) Chemical shift changes upon DNA-binding mapped to the CylR2 surface. Residues with very strong (Tyr39, Asn40), strong (Gln29, Thr30, Ile31, Gln44, Leu45, Ala46) and intermediate chemical shift changes (Gln17, Ser18, Glu19, Leu20) are colored orange, yellow and green, respectively.

M. Pillar *in vivo* with a LacZ reporter system, in which the LacZ transcription is driven by the cytolysin promoter. The LacZ activity was repressed in the presence of wild-type CylR2 and was 10-fold increased only after addition of the autoinducer CylL_S". However, for an identical construct in which the Asn40 of CylR2 was mutated to an Ala the LacZ activity was high whether the operon was induced or not, revealing a derepression of the cytolysin promoter [108]. These results obtained from experiments, which were performed due to the NMR chemical shift perturbation, are consistent with the proposed role of CylR2 as a repressor of cytolysin expression from P_{lyt} via its interaction with cytolysin promoter DNA (see 4.1).

4.3.5 Model for the CylR2/DNA complex

Chris M. Pillar has shown that CylR2 binds specifically to the 25-bp IR1 inverted repeat that exists within the cytolysin promoter [108]. The NMR shift perturbation results of CylR2 with a 22-bp fragment of the IR1 repeat sequence (Figures 4.11, 4.12) further demonstrated that similar regions in dimeric CylR2 are interacting with DNA as compared to the N-terminal domain of 434 repressor. Furthermore, HN-RDCs for CylR2 when bound to DNA showed that there are no major changes in the backbone structure of CylR2 upon binding to DNA (Figure 4.9).

Based on the combined information from crystallographic (see 4.3.1) and solution NMR studies (see 4.3.3), a model of the CylR2/DNA complex was constructed. In agreement with the absence of major conformational changes in CylR2 upon binding to DNA, a starting structure for the complex was built by superposition of the structure of the unbound CylR2 dimer onto the N-terminal domain of 434 repressor and replacement of the 434 operator by standard linear B-DNA. This initial model was refined using the protein-DNA docking program Monty [90, 91] guided by information obtained from the NMR chemical shift perturbation studies: Side chains of residues 10, 16-20, 24-32 and 34-45 were allowed to rotate freely in accordance with the chemical shift differences (Figure 4.12). In addition, energy bonuses were given during the simulation to complex structures where the backbone amides of Tyr39, Asn40, Gln44, Leu45 and Ala46 are in close contact to the DNA and when the NH₂ of Gln29 contacts the DNA reflecting the strong chemical shift changes upon addition of DNA. Figure 4.14 shows the model for the CylR2/DNA complex obtained from this docking approach. It agrees with established dimeric HTH/DNA interactions [112, 121]; binding occurs with dyad symmetry on two adjacent major grooves through their recognition helix. In a similar manner as observed for the 434 repressor/O_R1 complex structure, side chains of Ser27 and Gln29 form a van der Waals pocket to receive the methyl group of thymine 4' (Figures 4.12 and 4.14B). Moreover, complex models indicate a bidentate hydrogen bonding for the NH₂ of Arg28 to either G8, T7 or C6 and a hydrogen bond between the side chain of Gln29 and bp 4. The hydrogen bond of Gln29 to bp 4 is likely to be formed with O^ε of Gln29 as the hydrogen bond acceptor, because the C^δ resonance showed a strong downfield shift of ~0.9 ppm while N^ε and H^ε remained nearly unchanged (Figures 4.12 and 4.14B). For the side chain of Arg28 only C^α and C^β could be assigned, whereas both backbone and side chain resonances were missing for Ser27 in the complex. The disappearance or strong perturbation of these resonances upon binding to DNA supports their importance for complex formation. At the N-terminus of helix 2, Gln17 forms van der Waals contacts and a hydrogen bond with the phosphate of T9, in agreement with the disappearance of the side chain resonances of the Gln17 carbonyl group upon DNA-binding. Docking re-

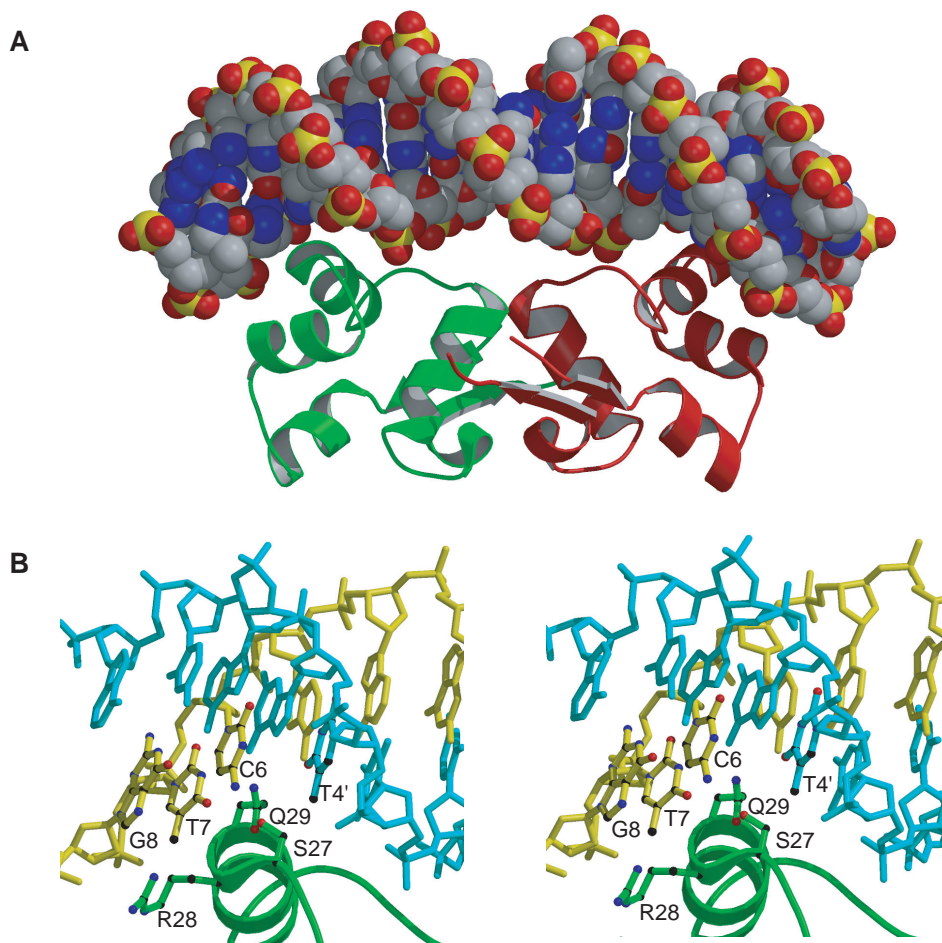


Figure 4.14: (A) Overall model of the CylR2/DNA complex structure. (B) Detailed stereo view indicating important protein-DNA interactions in the major groove. The DNA is shown with the sense strand in yellow and the antisense strand in cyan. Side chains of CylR2 and nucleobases interacting by van der Waals contacts and/or hydrogen bonds are indicated.

sults and chemical shift changes also suggest that bp 6 accepts a hydrogen bond from Asn32. NMR resonances from the carbonyl side chain of Asn32 were neither found for free nor complexed CylR2, but the C^β and C^α of Asn32 were shifted by 2.8 ppm and 1.1 ppm downfield, respectively (Figure 4.14B). Very important for stabilizing the CylR2/DNA complex is the (α 3- α 4) loop: for both CylR2 molecules a hydrogen bond is formed from the main chain amide group of Asn40 to the phosphate of C1 in more than 30 % of docking solutions, in agreement with a downfield shift of 1.1 ppm of its amide proton resonance.

The best-fit of DNA to the dimeric structure of CylR2 was obtained for a DNA

bending parameter of 0.4. For a more extended DNA, formation of protein/DNA interactions, which were expected on the basis of the NMR chemical shift perturbation studies, was not possible for both molecules of the CylR2 dimer simultaneously. When the DNA was more strongly bent, on the other hand, no docking solutions could be obtained that provided the necessary space to accommodate both CylR2 molecules in the orientation and position that were observed for the uncomplexed CylR2 structure. This orientation and position were experimentally indicated by the very good agreement of the experimental HN-RDCs for DNA-bound CylR2 with the X-ray structure of free CylR2 (Figure 4.9B). Similar to the 434 repressor/DNA complex, the DNA is relatively straight in the middle of the operator and bends towards the ends to enable interactions with the side chains of Arg28 and Gln29 (Figure 4.14B). The helical axes of individual bp, when projected onto the mean plane of bending, lie on a circle with a radius of 65 Å (Figure 4.14A). A specific feature of the 434/DNA complex is a strong compression of the minor groove due to insertion of the side chains of Arg43 of both molecules of the dimeric 434 repressor: the two positively charged side chains allow a closer approach of the negatively charged phosphates [118]. In CylR2, on the other hand, Arg43 is replaced by a serine; hence, the positive side chains are absent within the minor groove and no compression of the minor groove is required to dock DNA onto CylR2.

4.3.6 Specificity of CylR2/DNA interactions

The DNA-binding site of CylR2 had been revealed by gel shift analysis as the inverted repeat IR1 of the intergenic region of the cytolysin operon (Figure 4.2). In contrast, no DNA band shift was observed for the inverted repeat IR2 in the presence of CylR2 [108]. Nevertheless, NMR titration experiments were performed for IR2. ^{15}N -HSQC spectra recorded at different protein-to-IR2 ratios showed chemical shift changes for residues which were already found to be affected by IR1-binding (Figure 4.15). However, CylR2 signals shift continuously until a ~ 7 -fold excess of IR2 is reached. This indicates fast exchange between CylR2 and IR2 and agrees with the absence of retardation in the gel shift analysis. ^{15}N -HSQC spectra measured for

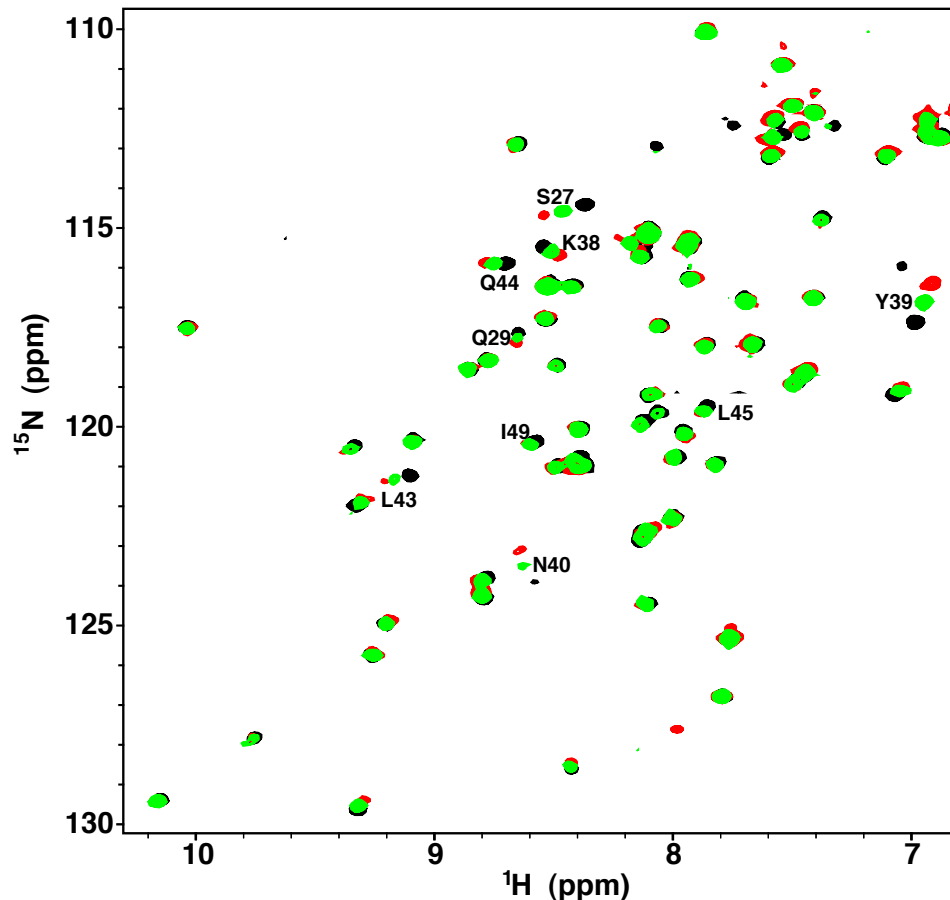


Figure 4.15: Overlay of ^{15}N -HSQC spectra of CylR2 at a molar protein to IR2 ratio of 1 to 0 (black), of 1 to 1 (green) and of 1 to 5 (red). Shifted resonances are labeled with the residue number.

a titration of CylR2 with a random ds DNA comprising a 22-bp long ACTG-repeat displayed almost identical continuous shifts as IR2. From these data a K_d of ~ 2 mM was calculated (see 3.5.6). Thus, the affinity of CylR2 for a non-specific DNA is by six orders of magnitude reduced. Figure 4.16 shows the amide chemical shift changes upon binding of CylR2 to unspecific DNA. At a ~ 7 -fold excess of the unspecific DNA the chemical shift changes for the mainly affected region around residue 40 are very similar to the chemical shift changes of CylR2 upon binding to its specific IR1 DNA while the regions around residue 17 and 45 are less affected by the unspecific DNA. This suggests an important function of the regions around residue 17 and 45 of CylR2 for the specificity to IR1 DNA.

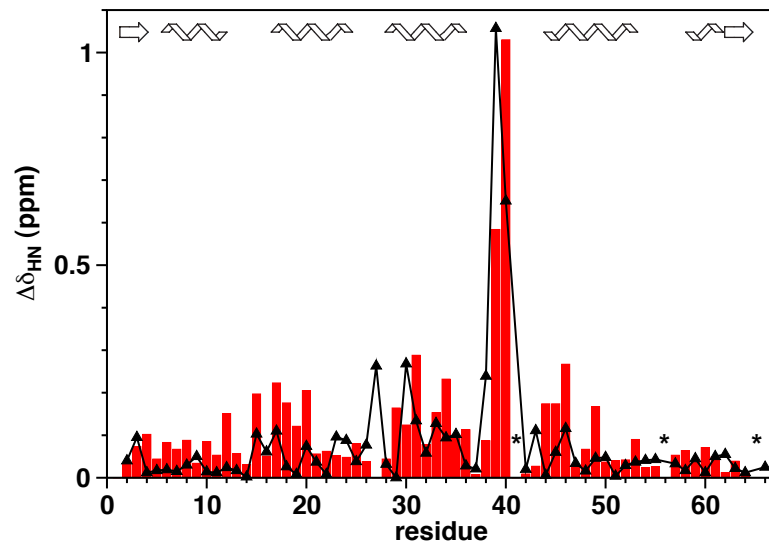


Figure 4.16: Amide chemical shift changes upon binding of CylR2 to unspecific DNA (black triangles). The red bars show the amide chemical shift changes upon binding CylR2 to its specific IR1 DNA for comparison. Proline residues are labeled with an asterisk.

4.3.7 Interaction between CylR2 and CylR1

In addition to CylR2, the 94-residue protein CylR1 has been demonstrated to be essential for repression of cytolysin production [104]. Using TMHMM Server v. 2.0 for the prediction of transmembrane helices, 20 N-terminal residues, residues 40 to 60 and residues 68 to 89 of CylR1 were found to probably form transmembrane helices [122](Figure 4.17). Furthermore, intracellular loop regions were predicted for residues 20 to 40 and the C-terminal residues therewith suggesting an interaction of these residues with CylR2. As all trials to recombinantly prepare CylR1 failed, unlabeled peptides comprising the predicted intracellular loop regions were synthesized (see 3.3.11 and Table 2.4) for NMR-titration experiments with ^{15}N -CylR2. Even at a 20-fold excess of the peptide, the recorded ^{15}N -HSQC spectra did not show chemical shift changes so that the interaction between CylR2 and CylR1 cannot be investigated with the peptides constituting the predicted intracellular loop regions. These data propose the absence of a direct interaction between CylR2 and CylR1, the binding of CylR2 to two or all three intracellular loops in the correct 3D arrangement or the interaction of CylR2 with transmembrane regions of CylR1. To distinguish these

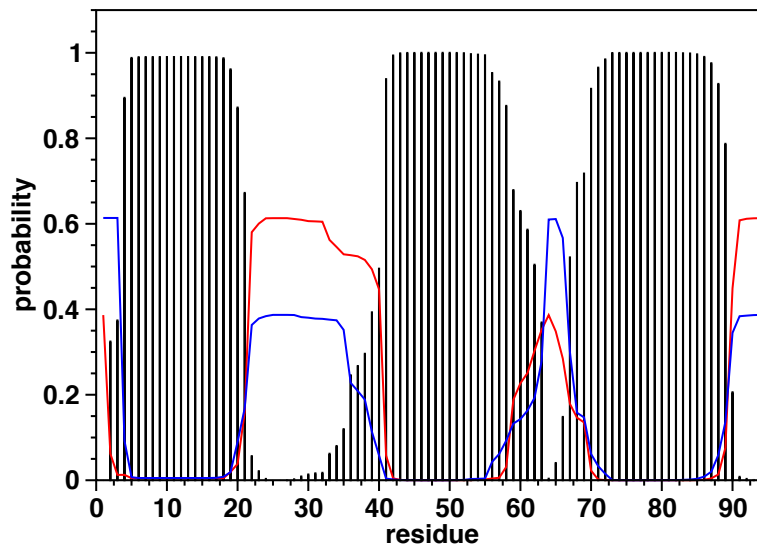


Figure 4.17: Prediction of transmembrane helices for CylR1. Probabilities for helical, inside and outside regions are indicated in black, red and blue, respectively.

possibilities intact CylR1 is required. However, this is a demanding task as reconstitution of CylR1 in micelles is necessary and so far there is no generally recommended refolding protocol for α -helical membrane proteins available [15].

4.4 Conclusions

The performed studies show that CylR2 binds as a preformed dimer to a 22-bp fragment of the cytolysin promoter region. The promoter binding is highly specific and is mediated through a complex network of CylR2/DNA interactions. In particular, the very strong chemical shift change of Asn40 upon DNA-binding motivated an *in vivo* activity test in the group of Michael S. Gilmore. This showed significantly reduced strength for the CylR2-operator binding when Asn40 was mutated to Ala. Based on these results a mechanism was proposed where repression is achieved by CylR2 obstruction of the promoter preventing transcription of the cytolysin operon.

CylR2 works together with CylR1 in order to repress the transcription of cytolysin genes [104]. To get an idea on how CylR1 influences the CylR2-operator interaction, comparison with the repressor of phage 434 is particularly interesting. In the case of the 434 repressor, it is the intricate combination of direct protein-DNA interactions and DNA conformational effects that enable differential binding to different oper-

ator binding sites. This very well orchestrated, differential binding determines the developmental fate of phage 434. By analogy, it can therefore be speculated that in the presence of the autoinducer CylL_S”, CylR1 shifts the DNA-binding specificity of CylR2 to sequences adjacent to the IR1 repeat. The three-dimensional structure of CylR2 and its DNA-binding properties presented here provide first steps in understanding the molecular basis of cytolysin repression by the CylR1/CylR2 two-component regulatory system.



5

Using dipolar couplings and intermonomer paramagnetic relaxation enhancement to determine the structure of the CylR2 homodimer

5.1 Introduction

The majority of structures in the PDB represents monomers, but most of the soluble and membrane proteins form homodimers or higher oligomers. Therefore oligomerization plays an important role in protein function, and improved methods for structure determination of homooligomers are required.

Solving the NMR solution structure of homodimeric proteins is especially complicated as inter- and intramonomer contacts are difficult to distinguish. The standard method to determine intermonomer distances in homodimeric proteins combines isotope-filtered and -edited NOE experiments conducted on samples formed by labeled and unlabeled protein [123]. An alternative method uses a dimer mixed from a completely deuterated ^{15}N -labeled and an unlabeled protonated protein to measure a 3D $[^{15}\text{N}-^1\text{H}]$ NOESY-HSQC, which shows in the aliphatic region only intermonomer contacts between HN and aliphatic protons [124]. In addition, computational methods exist which restrain the symmetry during structure calculation with ambiguous NOE data [125]. All of these methods require experimental observation of a significant number of side chain to side chain or HN to side chain intermolecular NOEs. How-

ever, deuteration is necessary for large homodimers to reduce the linewidths which increase with molecular weight. Due to deuteration intermolecular NOE information is very limited.

As an alternative or a complement to NOE restraints, which are limited to distances up to ~ 6 Å, long-range distance information may be derived from paramagnetic relaxation enhancement (PRE) [88]. As paramagnetic metalloproteins represent only a small class of proteins in nature, approaches to introduce paramagnetic centers into proteins have been developed to determine PRE. These approaches include ions bound to engineered metal-binding sites [126, 127], chelating tags [128, 129] and nitroxide groups [60].

The interaction between a specifically attached paramagnetic center and protons within a distance of up to 25-30 Å causes broadening of their NMR signal due to an increase in transverse relaxation rate [88]. This effect has an r^{-6} dependence on the electron-proton distance and thus allows the detection of long-range interactions in proteins. PRE-derived distances have already been used in monomer structure calculations [128, 130, 131] and are particularly useful for the characterization of highly flexible systems, such as intrinsically disordered proteins [132]. Moreover, they can be used to detect transient intermediates in macromolecular binding processes [133]. PRE-derived distances might also be used to overcome the difficulties in detection of intermolecular NOEs in large homodimers. By adding a paramagnetic probe substoichiometrically, the homodimer symmetry is broken and monomer specific RDCs and pseudocontact shifts (PCSs) are measured [134]. However, due to the broken symmetry and the pseudocontact shifts, the number of signals is increased significantly, limiting the use of this approach especially for larger proteins.

This chapter introduces a method for structure determination of homodimeric proteins, which is based on intermonomer distances derived from paramagnetic relaxation enhancement and residual dipolar couplings. Using this method, the high-resolution solution structure of the 15.4 kDa homodimer of CylR2 was determined. In order to derive intermonomer distances, two paramagnetic labels have been applied. Firstly, site-directed spin-labeling (SDSL) of a single cysteine residue with the nitroxide spin

label methanethiosulfonate (MTSL) [60] and secondly, prepending the three-residue amino terminal Cu^{2+} (Ni^{2+})-binding (ATCUN) motif to the N-terminus [130]. In addition, it is shown that only HN-RDCs can be used to rank homodimeric models of CylR2 obtained from *ab initio* docking.

5.2 Materials and methods

5.2.1 Construction of CylR2-mutants

The plasmid pET32a-CylR2 provided by Karin Giller (Max Planck Institute for Biophysical Chemistry, Göttingen) was the template for mutagenesis using the QuikChangeTM Site-Directed Mutagenesis Kit (see 3.1.5). The plasmid was amplified with the mutagenic primers CylR2_N40C5' and CylR2_N40C3' to generate pET32a-CylR2_{N40C} and with the mutagenic primers CylR2_T55C5' and CylR2_T55C3' to generate pET32a-CylR2_{T55C} (see Tables 2.3 and 2.2).

To attach an ATCUN-motif to CylR2, pET32a-CylR2 was amplified by PCR (see 3.1.4) with the primers MGS HG-CylR2_5' and MGS HG-CylR2_3' (see Table 2.3). The purified PCR-product was ligated between the *Nde*I and *Xho*I sites of the pET32a expression vector to construct pET32a-MGS HG-CylR2 (see Table 2.2). The codon for serine within MGS HG was deleted by mutagenesis (see 3.1.5) with the primers ATCUN-CylR2_5' and ATCUN-CylR2_3' to generate pET32a-MGHG-CylR2 (see Tables 2.3 and 2.2).

5.2.2 NMR sample preparation

Wild type CylR2 and CylR2 mutants were purified as described in 4.2.1. Samples to determine intermonomer distances were prepared by dissolving ¹⁵N-labeled wild type (¹⁵N-wt) and unlabeled mutant CylR2 (mut) in 8 M urea and mixing them in a 1:1 molar ratio. Samples containing single cysteine mutants were refolded by dialysis against NMR buffer with 600 mM NaCl and 5 mM DTT whereas samples containing the ATCUN-motif were refolded by dialysis against 50 mM MOPS pH 7.0 and 600 mM NaCl. Directly before labeling with MTSL (see 3.3.12.1), DTT was removed via a

PD-10 column from the buffer for single cysteine mutants. MTSL-labeled samples to measure intermonomer distances are referred to as 1:1-mixed ^{15}N -wt/mut* sample. Following NMR analysis of the paramagnetic oxidized state of the proteins, a 2-3 molar excess of 200 mM ascorbic acid dissolved in NMR buffer was added and NMR spectra were acquired for the resulting diamagnetic reduced state of the proteins (1:1-mixed ^{15}N -wt/mut). For ATCUN-CylR2 (see 3.3.12.2), NMR spectra were measured in the absence and presence of Cu^{2+} .

5.2.3 NMR spectroscopy

2D ^{15}N - ^1H -HSQC and ^{15}N T_2 experiments (see 3.5.4.2) for site-directed spin-labeling studies were performed at 298 K and 700 MHz with NMR samples containing 0.4-0.8 mM CylR2. For structure determination, a 3D [^{15}N , ^1H] NOESY-HSQC and a [^{13}C , ^1H] NOESY-HSQC with a mixing time of 120 ms were measured at 600 MHz. All experiments for the solution structure determination of CylR2 are summarized in the Appendix Table B.1.

5.2.4 Monomer structure calculation

The monomer NMR structure ensemble was calculated as described in 3.6.1. The 98.8 % complete resonance assignment (see 4.3.3), torsion angle restraints as predicted from chemical shifts with TALOS [52] and the peaklists of the 3D [^{15}N , ^1H] and the [^{13}C , ^1H] NOESY-HSQC were used as input for CYANA.

A monomeric MTSL-modified CylR2-mutant structure was calculated as the wild type monomer structure. Thereby the wild type structure was used as the starting structure and intramolecular distances derived from PRE plus 5 Å (see 3.5.7) were used as additional upper limit restraints. For peaks broadened beyond detection for the paramagnetic sample, the upper distance limit was assigned to 12 Å. To avoid inclusion of wrong distances into the structure calculation, intramolecular PRE distances have only been used for those residues without an intensity difference in the experiments to determine intermonomer distances.

5.2.5 Intermonomer distances

Intermonomer distances were calculated from PRE (see 3.5.7). For peaks broadened beyond detection for the paramagnetic sample, distances were assigned to 7 ± 5 Å. To calculate intermonomer distances for peaks that did not show a chemical shift difference between homodimeric wild type and heterodimeric wild type/paramagnetic-mutant, intensities of the paramagnetic heterodimeric molecule I_{para} have been calculated using:

$$I_{para} = 2(I_{para\star} - \frac{I_{dia}}{2}), \quad (5.1)$$

where $I_{para\star}$ is the intensity in the spectrum of the paramagnetic state and I_{dia} is the intensity in the spectrum of the diamagnetic state. Upper and lower long-range distances were set to ± 5 Å of the calculated distances. For residues neighboring residues without a PRE effect, no upper limit was used and for residues without a PRE effect, the lower limit was set to 25 Å.

Intermolecular NOEs have been manually assigned for 20 peaks of the [^{13}C , ^1H] NOESY-HSQC spectrum which have not been automatically assigned during the monomer structure calculation with CYANA (Appendix Table C.9). For NOE data upper and lower distances were set to ± 2 Å of the calculated distances.

5.2.6 Homodimer structure calculation

All homodimer calculations (see 3.6.2) were performed with the lowest energy structure of the monomer ensemble. In the starting coordinates, one monomer was translated by 50 Å and oriented differently with respect to the other monomer. Seven different sets of restraints have been used: (a) NOE and RDCs, (b) PRE and RDCs, (c) PRE, NOE and RDCs, (d) PRE, RDCs and symmetry (e) PRE, NOE, RDCs and symmetry, (f) PRE and symmetry and (g) PRE, RDCs and symmetry. Twofold symmetry was implied by distance difference restraints [135]. With the exception of (a) all rigid-body dockings were performed with two monomers which contain MTSL at position N40C and T55C. The intermonomer PRE distances in the rigid-body docking were restrained from the nitrogen of the MTSL ring in one monomer to the amide

protons of the other monomer. Each monomer in the homodimer structure resulting from the rigid-body minimization was replaced by the refined wild type monomer structure. Intermolecular HN-HN distances ± 2 Å from N40 and Thr55 to the other subunit were extracted using MOLMOL and used for the simulated annealing. The rigid-body docking as well as the simulated annealing protocol are provided in the Appendix D.

5.2.7 Ranking of homodimer models from *ab initio* docking

The lowest energy monomer structure was docked to itself using the ClusPro Web server (see 3.6.3). Ranking of the docking results without experimental data about the dimer interface was carried out via prediction of RDCs with PALES (see 3.5.3.3).

5.3 Results

5.3.1 Monomer solution structure

The solution monomer structure of CylR2 was solved based on a 98.8 % complete chemical shift assignment, 987 interproton distances, 86 dihedral angle restraints and 57 HN-RDCs. The chemical shift assignments can be found in the BMRB under the accession code 6317 and the structural restraints are provided in the Appendix C. Nine medium- and seven long-range ^{13}C -distances have been manually assigned to avoid wrong assignments during the automated NOE assignment within CYANA. Thereby convergence was ensured even in the presence of intermolecular NOEs in the peak list. The ensemble of 20 structures determined by CYANA has a backbone rmsd (residues 3-63) of 0.45 Å and an average CYANA target function of 0.84 Å². After the final water refinement with XPLOR-NIH including 57 HN-RDCs, the backbone rmsd (residues 3-63) is slightly increased to 0.60 Å, but all restraints are well satisfied (Table 5.1). The structural characteristics have been described in the previous chapter (see 4.3.1). Highest rmsd values within the 20 structure bundle were found for the N- and C-terminal residues (Figure 5.1B), which also have the highest

Table 5.1: Statistics for CylR2 monomer solution structure determination.

NOE distance restraints:	
total	987
short range ($ i-j \leq 1$)	594
medium range ($1 \leq i-j \leq 5$)	180
long range ($ i-j \geq 5$)	213
Dihedral angle restraints	86
^{15}N - ^1H residual dipolar couplings	57
Mean rmsd from experimental restraints	
NOE	0.0048 Å
dihedral angles	0.743 deg
Average number of	
NOE violations ≥ 0.5 Å	0
Dihedral angle violations $\geq 5^\circ$	0
Mean rmsd from idealized covalent geometry	
Bonds	0.0097 Å
Angles	1.35 deg
Impropers	1.62 deg
Ramachandran analysis	
Most-favored region	88.5 %
Additionally allowed region	10.3 %
Generously allowed region	1.0 %
Disfavoured region	0.3 %
Rmsd from the mean ^a	
Backbone atoms	0.60 ± 0.14 Å
All heavy atoms	1.06 ± 0.12 Å
Rmsd from the X-ray structure ^a	
Backbone atoms	0.89 Å
All heavy atoms	1.88 Å

^a determined for residues 3-63

rmsd values when superimposing the two subunits of the X-ray structure. Overall, there is a good correlation between the rmsd values within the ensemble of 20 NMR structures and the ^{15}N transverse relaxation times (T_2) (Figure 5.1C). There are no major differences between the X-ray structure (Figure 4.4) and the NMR structure ensemble. A superposition of the mean NMR structure and the X-ray structure is shown in Figure 5.1A. The average backbone rmsd (residues 3-63) between the mean NMR-structure and the X-ray structure is 0.89 Å. Differences within the ensemble and to the X-ray structure are similar with the exception of the region around residue Ser42 and residue Glu11, where the backbone rmsd to the X-ray structure is around 1 Å higher (Figure 5.1B). For the region around residue Ser42, this can be explained by the results from the steady-state heteronuclear ^{15}N - ^1H NOE (Figure 4.10) and by the T_2 times (Figure 5.1C), which reveal Ser42 as the most flexible residue.

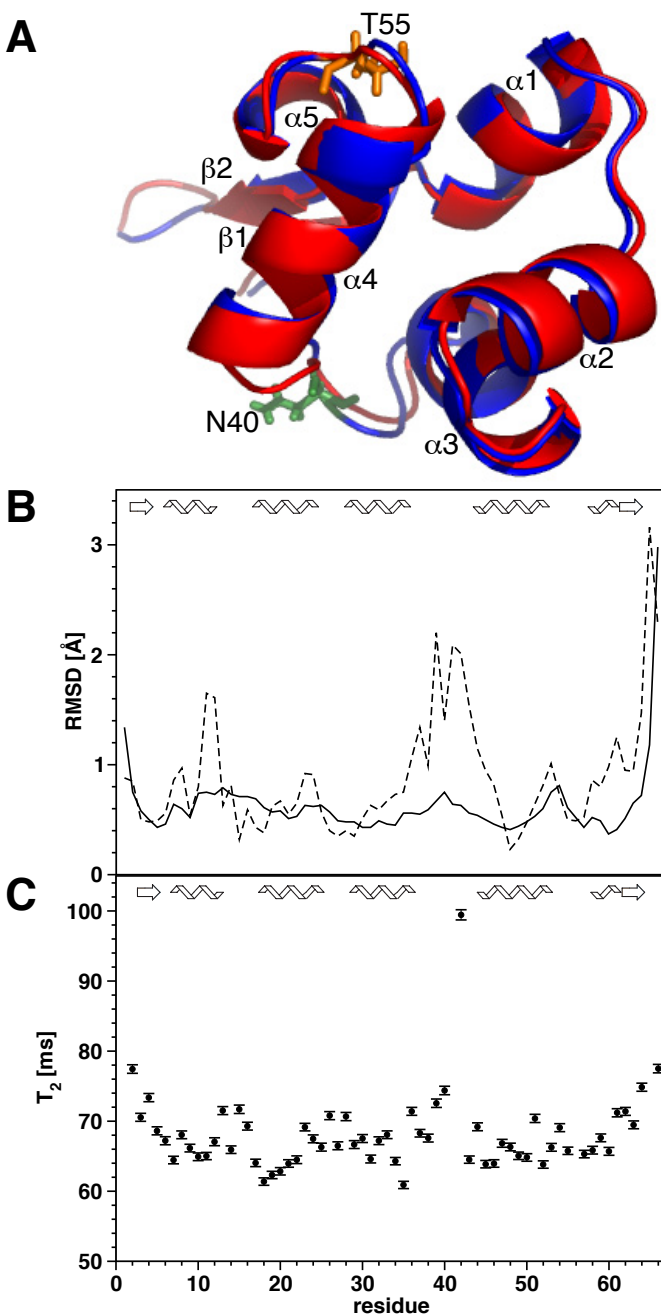


Figure 5.1: Monomer solution structure of CylR2. (A) Mean structure of the NMR ensemble (blue) superimposed on a monomer from the X-ray structure (red). Side chains of Asn40 and Thr55 which are mutated to cysteine for dimer structure determination are shown in green and orange, respectively. (B) Average backbone rmsd per residue between the mean NMR-structure and the 20 NMR-structures (solid line) and the X-ray structure (dashed line). (C) Measured T_2 values versus the residue number. Secondary structure elements are indicated.

5.3.2 Mutagenesis and spin-labeling

SDSL with MTSL was applied to determine distances from one subunit of CylR2 to the other (Figure 5.2). Therefore single-cysteine residues were introduced into the cysteine-free wild type CylR2 (see 5.2.1). Conservative mutation sites were chosen at position Asn40 and Thr55 in the middle of loop regions on opposite sites of the monomer structure. By comparing amide proton and nitrogen chemical shifts as well

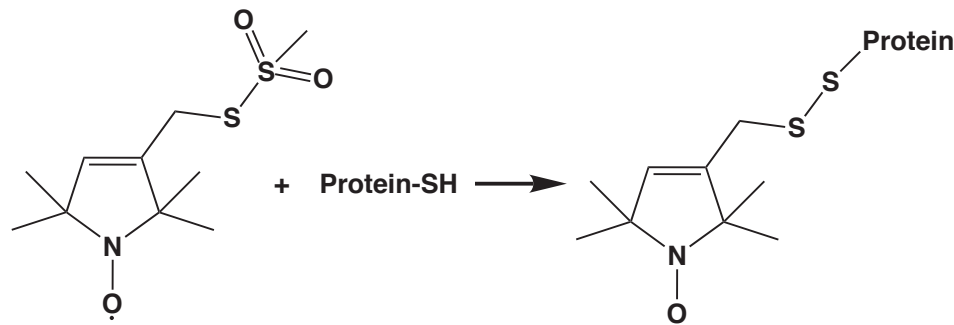


Figure 5.2: Structure of the MTSL spin label and reaction with the -SH group of a protein.

as rotating frame relaxation times ($T_{1\rho}$) between wild type and mutant, it was confirmed that introduction of a Cys residue at these two positions does not significantly affect the protein structure (Figure 5.3). Chemical shift differences are smaller than

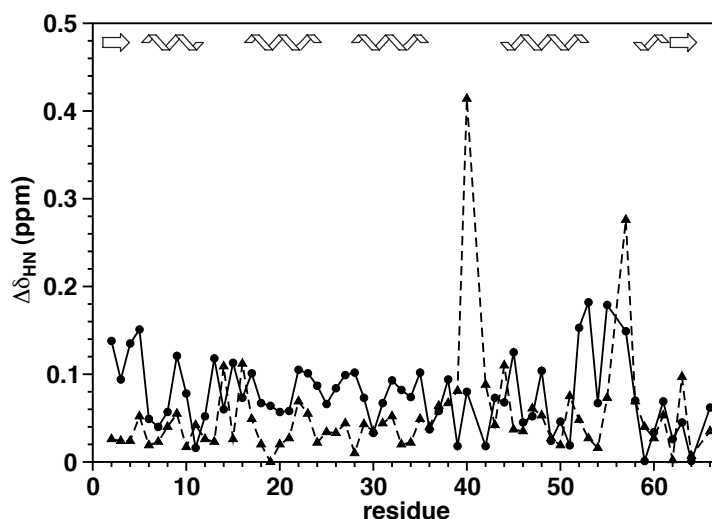


Figure 5.3: Chemical shift changes of backbone amides due to mutation of Asn40 (triangles and dashed line) and Thr55 (circles and solid line) to cysteine. Secondary structure elements are indicated.

0.1 ppm with the exception of Leu57 for CylR2_{N40C} and residues around the mutation site. Residues in the N-terminal region of CylR2_{T55C} show $\Delta\delta_{HN}$ values between 0.1 and 0.16 ppm. Additionally, unchanged $T_{1\rho}$ values indicate an unaltered oligomerization state. Subsequent modification of the single cysteine mutants with the spin label MTSL (Figure 5.2) allows to measure distances within a range of $\leq 25\text{-}30$ Å from the spin label.

^{15}N -HSQC spectra have been recorded for the para- and diamagnetic form of each CylR2-mutant. These spectra were measured for two different samples of each CylR2-mutant: a pure ^{15}N - and spin-labeled mutant (^{15}N -mut(*)) and a 1:1 mixture of

^{15}N -labeled wild type and spin-labeled mutant (1:1-mixed ^{15}N -wt/mut(*)). Paramagnetic states are indicated by a star. Each 1:1-mixed ^{15}N -wt/mut(*) sample is composed of three different dimers: 50 % ^{15}N -wild type/spin-labeled mutant (^{15}N -wt/mut(*)), 25 % ^{15}N -wild type (^{15}N -wt/ ^{15}N -wt) and 25 % spin-labeled mutant (mut(*)/mut(*))(Figure 5.4). Both, ^{15}N -wt/mut* heterodimer and ^{15}N -wt/ ^{15}N -wt

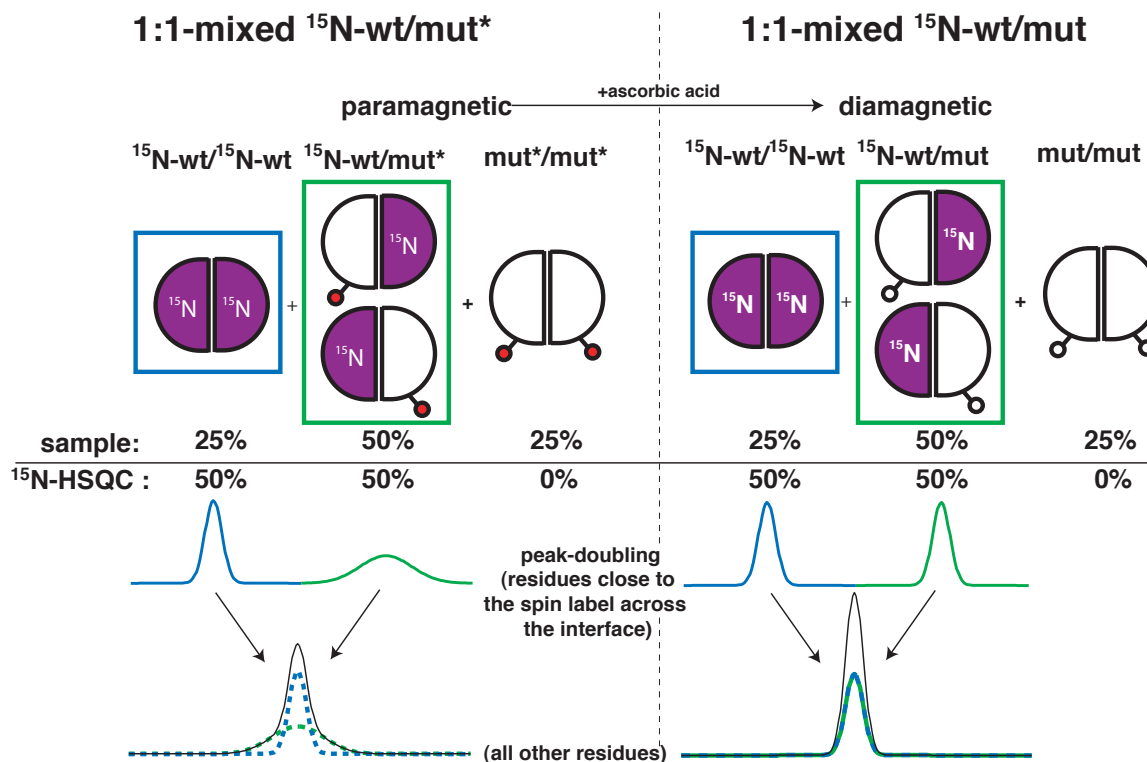


Figure 5.4: Overall strategy to derive intermonomer distances from PRE in homodimers. The paramagnetic (*) and the diamagnetic sample (no *) are shown on the left and right hand side of the figure, respectively. The 1:1-mixed samples are composed of equal amounts of ^{15}N -labeled wild type (^{15}N -wt, violet) and of paramagnetic mutant (mut(*), white) monomers. The monomers combine into three distinct dimerization pairs: 25 % ^{15}N -wt/ ^{15}N -wt (blue), 50 % ^{15}N -wt/mut(*) (green) and 25 % mut(*)/mut(*). The former two species contribute equally to the NMR signal while the latter is undetected. For a few residues close to the para- or diamagnetic tag across the dimer interface, the chemical shift can be distinguished (peak-doubling), while for all other residues, the ^{15}N -wt/ ^{15}N -wt and ^{15}N -wt/mut(*) peaks overlap. The PRE distance is derived from the peak intensity ratio (I_{para}/I_{dia}) obtained from the paramagnetic and diamagnetic lines (green lines). For the overlapped case, I_{para} can be obtained by subtracting $I_{dia}/2$ according to formula 5.1. For MTSL, the diamagnetic sample can easily be obtained from the paramagnetic sample by ascorbic acid reduction.

homodimer, contribute 50 % to the NMR signal. For CylR2_{T55C}, an overlay of the ^{15}N -HSQC spectra of the para- and diamagnetic state for ^{15}N -mut(*) and for 1:1-

mixed ^{15}N -wt/mut(*) are shown in Figure 5.5A and in Figure 5.5B, respectively. ^{15}N -HSQC experiments measured for ^{15}N -mut(*) were used to accurately define the

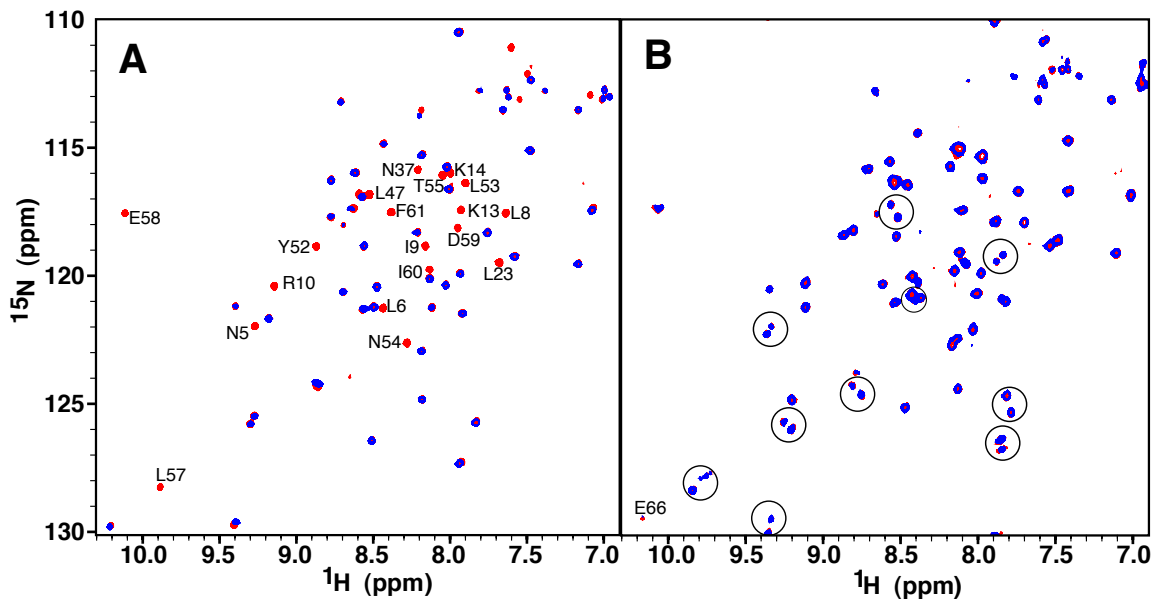


Figure 5.5: Overlay of ^{15}N -HSQC spectra of paramagnetic (blue) and diamagnetic (red) forms of ^{15}N -CylR2 $T_{55}C$ (A) and the 1:1 mixture of ^{15}N -CylR2 and CylR2 $T_{55}C$ (B). Residues that disappeared in the paramagnetic state are labeled and doubled peaks in (B) are indicated by ellipses.

two positions of the spin label MTSL within the monomer structure.

Intermonomer distances are derived from peak intensities of HSQC spectra for the paramagnetic and diamagnetic 1:1-mixed ^{15}N -wt/mut(*) sample. Overall reduction of signal intensities for the 1:1-mixed ^{15}N -wt/mut* sample is significantly less important than for the ^{15}N -mut* sample due to larger distances of the backbone amide HN group to the spin label. Indeed, whereas around one third of all backbone amide signals disappeared for the ^{15}N -mut* sample, only a single signal (Glu66) was not observed for the 1:1-mixed ^{15}N -wt/mut* sample. For the 1:1-mixed ^{15}N -wt/mut(*) sample, peak-doubling is observed for 10 of the 62 backbone amide signals because of slightly different chemical shifts for the wild type monomer residues within the homodimer and the wild type monomer residues that are close to the mutation site within the heterodimer (Figure 5.5B).

5.3.3 Long-range distances from PRE

The approach by Battiste and Wagner was used to derive long-range distances from PRE [87] and has been described in 3.5.7 and 5.2.5. For CylR2, the correlation time τ for the electron-nuclear interaction was estimated as 6 ns from Stokes' law (3.17) under the assumption that τ is equal to τ_c . Intramonomer distances were determined for residues whose signals were not broadened for the 1:1-mixed ^{15}N -wt/mut* sample. The resulting 24 and 31 intramonomer distances for CylR2_{N40C} and CylR2_{T55C}, respectively, were used to calculate the average position of the spin label MTSL at position N40C and T55C within the monomer structure. These distances were added as restraints to the protocol described in paragraph 5.2.4 and the monomer structure containing MTSL at position N40C and T55C was calculated using the wild type structure as the start structure. After the water refinement, the mean monomer structure of the 20 lowest energy structures was calculated using MOLMOL and used for the PRE-based rigid-body docking (see 5.3.4).

Intermonomer distances for the residues with peak-doubling for the 1:1-mixed ^{15}N -wt/mut(*) sample were derived like the intramonomer distances in the ^{15}N -mut(*) sample. The peak corresponding to the heterodimeric ^{15}N -wt/mut(*) mutant was identified as the signal which is shifted compared to the ^{15}N -HSQC of wild type CylR2. For the residues without peak-doubling, the intensity corresponding to PRE across the dimer interface (I_{para}) for all molecules of the 1:1-mixed ^{15}N -wt/mut(*) sample was calculated according to formula 5.1. This formula is valid under the assumption that the sample contains 50 % ^{15}N -wt/mut(*) and 25 % ^{15}N -wt/ ^{15}N -wt. From the ^{15}N -HSQC peak intensities (I_{para*}) is subtracted half the reference ^{15}N -HSQC peak intensity (I_{dia}), the latter corresponding to the contribution of the ^{15}N -wt/ ^{15}N -wt (Figure 5.4). The calculated intermonomer distances for the spin label at residue 40 and 55 are listed in the Appendix Table C.8.

The accuracy of the intermonomer PRE distances was assessed by comparing the calculated distances with the theoretically expected distances for the X-ray structure (Figure 5.6A). Only 6 out of 103 experimental PRE distances for each subunit are

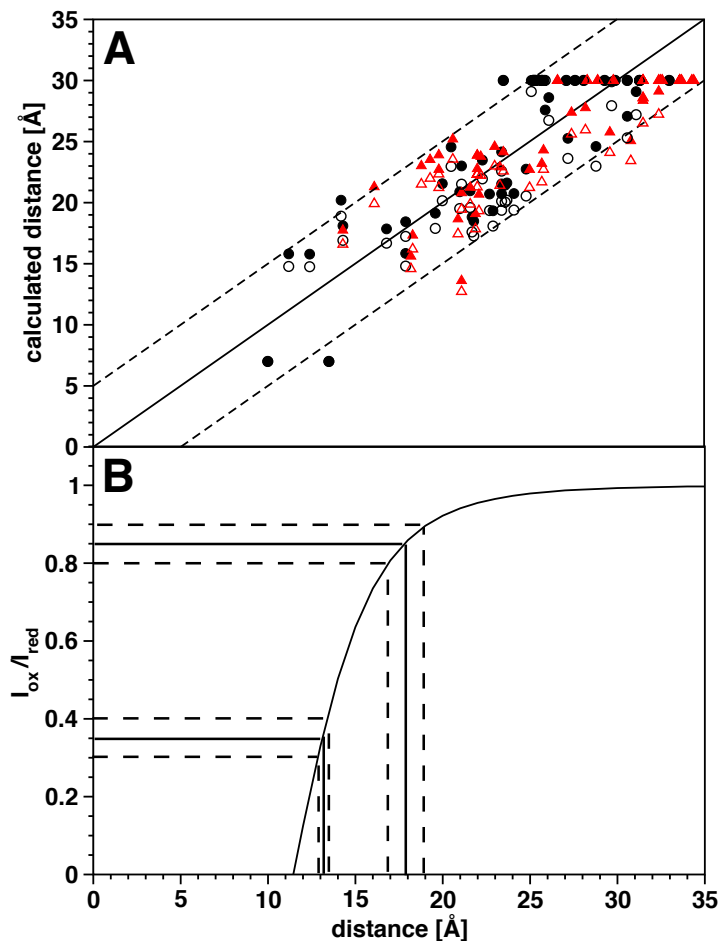


Figure 5.6: (A) Theoretically expected distances from the X-ray structure versus distances calculated from PRE data. The solid line indicates optimal correlation between experimental and expected distances and the dashed lines mark the ± 5 Å error bounds. Distances calculated with the spin label at position N40C and T55C are shown as black circles and red triangles, respectively. Distances calculated with a τ of 6 ns and 4 ns are indicated as filled and empty symbols, respectively. (B) Measured intensity ratio plotted as a function of the calculated distance. The dashed lines show that for an intensity ratio of 0.85 ± 0.05 the uncertainty of the distance is approximately four times larger than for an intensity ratio of 0.35 ± 0.05 .

found to deviate by more than ± 5 Å from the theoretically expected distances. The ± 5 Å error bounds are based on different assumptions and sources of experimental errors. Amide proton T_2 relaxation times were approximated by the experimental amide nitrogen T_2 relaxation times (Figure 5.1C) and τ was assumed as essentially equal to τ_c . Because of the r^{-6} dependence of the PRE, errors in the T_2 relaxation time and τ (Figure 5.6A) have a very small effect on the calculated distance. Positional averaging of the flexible nitroxide side chain of MTSL is another source of errors. However, these errors were minimized by calculating the average position of the spin label (see 5.2.4). Errors in the determination of protein concentration and interference of MTSL during refolding the dimer from 8 M urea (see 5.2.2) may result in a deviation from the 50 % contribution of the $^{15}\text{N-wt/mut}(\ast)$ and the $^{15}\text{N-wt}/^{15}\text{N-wt}$ to the $^{15}\text{N-HSQC}$ signal which was assumed to calculate intermonomer distances

for non-doubled peaks. In addition, the error is larger for intermonomer than for intramonomer distances as the measured distances are in general larger (Figure 5.6B). MTSL attached to N40C is slightly closer to the dimer interface than MTSL attached to T55C (Figure 5.7). The average theoretical distance from the nitrogen of the MTSL ring to amide protons across the interface is ~ 3.5 Å smaller for the spin-labeled N40C than for spin-labeled T55C.

5.3.4 Homodimeric solution structure

Accurate docking of two proteins requires both orientational and translational information. Measurement of HN-RDCs in a single alignment medium, leads to four possible orientations of one subunit of CylR2 with respect to the second subunit, when the alignment tensor is not axially symmetric. The four-fold degeneracy is resolved by intermolecular PRE distances or intermolecular NOE data. PRE distances have the same distance dependence as NOE data and so the XPLOR-protocol for rapid protein-protein docking on the basis of RDCs and intermolecular NOEs and PRE data can be used with the same standard energy function [89].

The performance of this PRE-based rigid-body docking has been tested for the 57 HN-RDCs for each monomer in combination with two different sets of experimental intermolecular restraints: PRE data (Table 5.2a) and PRE together with NOE data (Table 5.2b). In addition, the docking was tested with symmetry distance restraints to imply twofold symmetry [135]. These symmetry distance restraints were either used together with RDC data (Table 5.2c and d) or used to replace RDC data (Table 5.2e and f). For comparison, the same rigid-body protocol was run with RDCs and intermolecular NOEs as the only experimental restraints (Table 5.2g). These seven rigid-body dockings with different sets of restraints resulted in structures with the same monomer-to-monomer orientation. After simulated annealing with all coordinates fixed with the exception of the side-chain atoms of residues contributing to the dimer interface and refinement in explicit solvent, the seven ensembles, composed of 20 structures, have a backbone rmsd (residues 3-63) between 0.54 Å and

Table 5.2: Statistics for CylR2 dimer solution structure determination.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
intermolecular distance restraints:							
symmetry ^a	-	-	yes	yes	yes	yes	-
PRE	206	206	206	206	206	206	-
NOE	-	48	-	48	-	48	48
HN-RDCs	114	114	-	-	114	114	114
Ramachandran analysis:							
Most-favored region (%)	88.3	88.0	88.0	88.2	88.5	88.7	87.7
Additionally allowed region (%)	10.4	10.8	10.5	10.3	9.4	9.9	11.2
Generously allowed region (%)	1.0	1.0	1.2	1.3	2.0	1.3	1.0
Disfavoured region (%)	0.2	0.2	0.3	0.2	0.2	0.2	0.1
Rmsd from the mean ^b :							
Backbone atoms (Å)	0.59±	0.59±	0.59±	0.61±	0.54±	0.59±	0.65±
0.11		0.08	0.11	0.07	0.09	0.10	0.13
All heavy atoms (Å)	0.90±	0.89±	0.88±	0.91±	0.86±	0.89±	0.94±
0.10		0.09	0.11	0.09	0.08	0.10	0.13
Rmsd from the X-ray structure ^b :							
Backbone atoms (Å)	1.15	1.14	1.18	1.22	1.99	1.23	2.14
All heavy atoms (Å)	2.08	2.01	2.10	2.07	2.79	2.10	2.92
intermonomer rmsd:							
mean NMR structure:							
Backbone atoms (Å)	0.29	0.37	0.34	0.25	0.52	0.33	0.33
All heavy atoms (Å)	0.72	0.75	0.66	0.88	0.81	0.78	0.76
NMR ensemble (20 structures):							
Backbone atoms (Å)	0.88±	0.79±	0.84±	0.94±	0.73±	0.92±	0.83±
0.28		0.27	0.36	0.28	0.22	0.34	0.42
All heavy atoms (Å)	1.49±	1.44±	1.47±	1.54±	1.34±	1.54±	1.43±
0.25		0.26	0.4	0.3	0.3	0.32	0.4
intermolecular energy (kcal/mol)	-1609.6	-1343.5	-1662.2	-1309.7	-1527.9	-1422.1	-1436.4

^a Restrains distance differences to imply twofold symmetry [135]

^b determined for residues 3-63 of each monomer

0.65 Å (Table 5.2). The 10 lowest energy structures of the NMR ensemble calculated with HN-RDCs and PREs is shown in Figure 5.7. The small backbone intermonomer

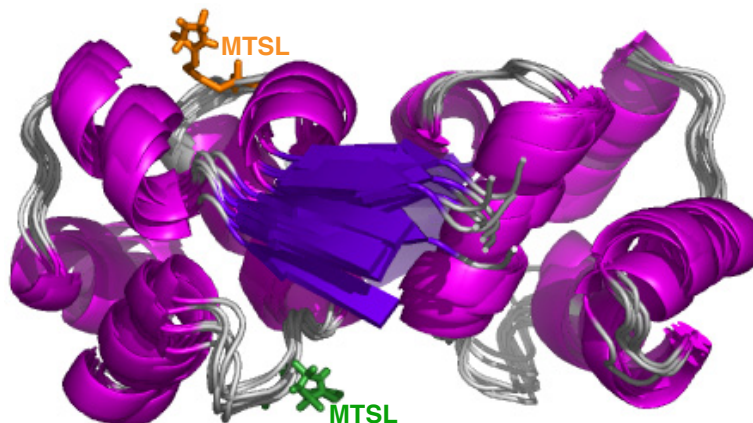


Figure 5.7: Homodimeric solution structure of CylR2. Superposition of the 10 NMR structures with the lowest energy. Helices and β -strands are shown in magenta and violet, respectively. The calculated average position of MTSL attached to either position N40C (green) or position T55C (orange) is indicated for the left subunit.

rmsd values for the mean structure ($< 0.6 \text{ \AA}$) and all structures of the NMR ensemble ($< 1 \text{ \AA}$) indicate the high symmetry of the homodimers and are similar to the backbone intermonomer rmsd value of 0.75 \AA between the monomers of the X-ray structures. As expected, the orientation of the monomers within the homodimer can be restrained by experimental RDCs or a symmetry function [135] (Table 5.2). Therefore, in this case, the experimental RDCs were not absolutely required. However, considering the backbone rmsd (residues 3-63) to the X-ray structure, the experimental HN-RDCs perform significantly better in the absence of intermolecular NOEs (1.15 \AA versus 1.99 \AA) and slightly better in the presence of intermolecular NOEs (1.14 \AA versus 1.22 \AA). This can be explained by the fact that with the HN-RDCs, one of the principal axis of the alignment tensor must be parallel and the other two orthogonal to the twofold axis [136] while the symmetry function restrains symmetry-related intermonomer distance differences to 0 without restraining the two axis. As the HN-RDCs restrain the symmetry more stringently than the symmetry function, the docking results are not improved when HN-RDCs are used together with the symmetry function. This is demonstrated by an almost unchanged backbone rmsd (residues 3-63) to the X-ray structure independent of the use of intermolecular NOEs for the calculations (Table 5.2a,b,e and f).

Decreasing the error bounds to $\pm 4 \text{ \AA}$ resulted in an increased rmsd and in a larger number of violated intermolecular restraints. Changing τ from 4 ns to 6 ns did not change the docking results because of the small difference in the calculated distances (Figure 5.6). Due to the different sources of errors for the PRE-derived distances (see 5.3.3), the docking did not converge to a unique solution when only intermolecular distances for one spin label position together with the HN-RDCs were used.

The backbone rmsd (residues 3-63) between the mean NMR dimer structure generated using intermolecular PRE distances as the only intermolecular restraints together with HN-RDCs and the X-ray structure is 1.15 \AA (Table 5.2a). The precision of the structure was only insignificantly improved by addition of 24 manually assigned intermolecular NOEs as indicated by the backbone rmsd (residues 3-63) to the X-ray structure of 1.14 \AA (Table 5.2b). Moreover, the combination of HN-RDCs and inter-

molecular NOEs is not sufficient to define the dimer structure as accurately as the intermolecular PREs in combination with HN-RDCs. This is indicated by a backbone rmsd of 2.14 Å relative to the X-ray structure (Table 5.2g).

There are small differences between the orientation of the two monomers in the NMR and the X-ray structure as the rmsd values are overall higher for the dimer than for the monomer (Figures 5.1B and 5.8B). Most notable are the differences for the

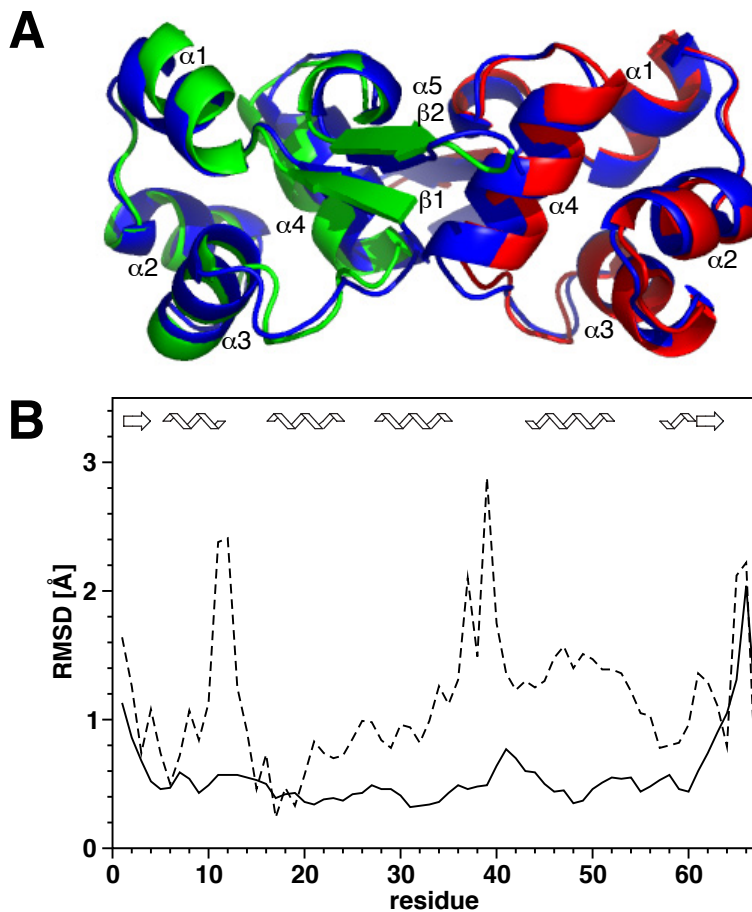


Figure 5.8: Comparison of the solution and crystal structure of CylR2. Secondary structure elements are indicated. (A) Mean structure of the NMR ensemble (blue) superimposed on the X-ray structure (green/red). (B) Average backbone rmsd per residue between the mean NMR-structure and the 20 NMR-structures (solid line) and the X-ray structure (dashed line).

longest helix $\alpha 4$ (residues 43-52) that contributes strongly to the dimer interface and the loop connecting helix $\alpha 3$ and $\alpha 4$ involved in the DNA binding. Within this loop the flexible residue Ser42 is found. Structural flexibility in this region is probably important for DNA binding (see 4.3.4 and 4.3.5).

5.3.5 The ATCUN-motif as a paramagnetic position

In order to obtain a paramagnetic probe at a third position and to use another paramagnetic probe, the three-residue Cu^{2+} -binding ATCUN-motif was attached via a glycine-linker to the N-terminal residue of wild type CylR2 (see 5.2.1) [130]. Because a histidine residue at position three is essential for the high affinity Cu^{2+} binding, the endogenous cleavage of the N-terminal translation initiator methionine in *E. coli* is determinant in choosing the length of the ATCUN-motif insert for CylR2. This N-terminal methionine has been shown to be cleaved more efficiently by the methionyl-aminopeptidase in *E. coli* with small side chains at the second residue position [137]. Since MGSHG-CylR2 contains glycine at the second position, it was most reasonable to produce GSHG-CylR2 with the required histidine at position three from *E. coli* BL21(DE3)pET32a-MGSHG-CylR2. But ESI-MS (see 3.4.3) revealed the N-terminal methionine to be present in 100 % of the purified protein. Therefore, the serine was deleted and the expected molecular weight for MGHG-CylR2 (ATCUN-CylR2) was confirmed by ESI-MS.

Intermonomer distances derived with the help of this third paramagnetic position were additionally included in the docking or replaced intermonomer distances for N40C or T55C. However, the rmsd to the X-ray was increased by ~ 1.5 Å when intermonomer distances for all three paramagnetic positions were used. When intermonomer distances for N40C or T55C were replaced by the intermonomer distances derived with the help of the ATCUN-motif, solutions from rigid-body docking did not converge. Hence, the intermonomer distances determined via the ATCUN-motif are too imprecise. The reason for the inaccuracy may be the presence of different conformations for the terminal regions of ATCUN-CylR2 as suggested by the ^{15}N -HSQC spectra (Figure 5.9). The ^{15}N -HSQC spectrum exhibited peak-doubling for N- and C-terminal residues and Leu24, Ala46 and Leu47 (Figure 5.9A). After Cu^{2+} -binding to ATCUN-CylR2 (see 5.2.2), paramagnetic peak broadening was observed and in addition, the extent of peak-doubling was increased compared to ATCUN-CylR2 (Figure 5.9B). Most noticeable is the effect on Leu57 and Glu58 which have

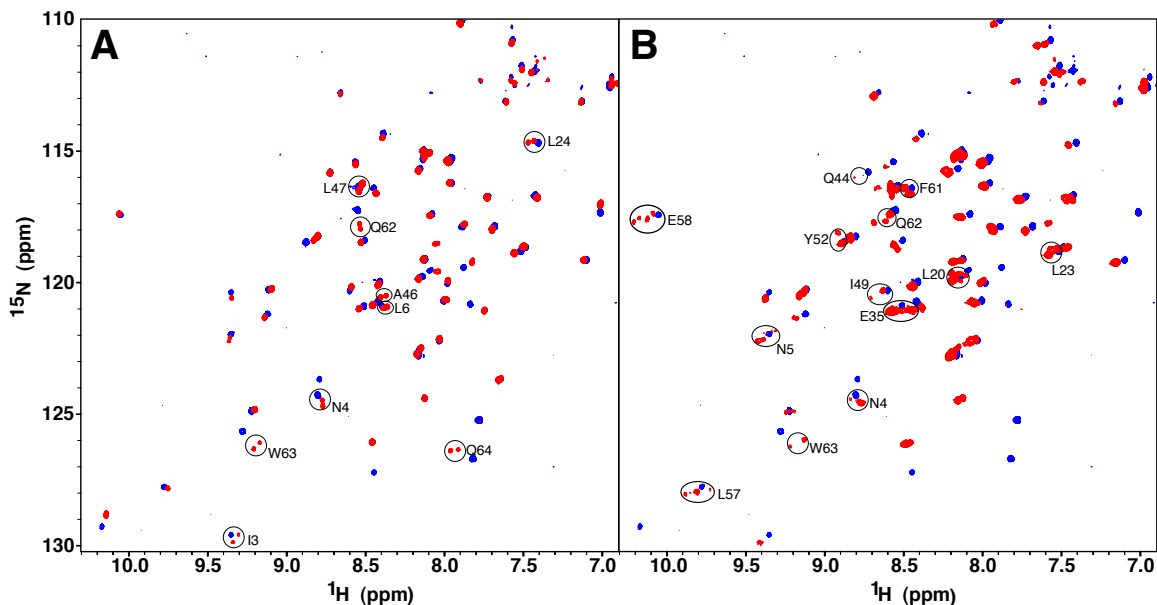


Figure 5.9: Overlay of the ^{15}N -HSQC spectrum of CylR2 (blue) on the ^{15}N -HSQC spectrum of ^{15}N -ATCUN-CylR2 (red) (A) and on the ^{15}N -HSQC spectrum of ^{15}N - Cu^{2+} -ATCUN-CylR2 (red) (B). Doubled or quadrupled peaks are indicated by ellipses.

quadrupled peaks. Hence, the distances derived with the help of the ATCUN-motif were not used for the solution NMR structure determination of CylR2.

5.3.6 Ranking of homodimer structures from *ab initio* docking

Good progress has been made in the *ab initio* docking of protein complexes and homooligomeric proteins [18]. For many systems, an ensemble of low energy docking solutions can be calculated, which contains a structural model with an rmsd of 2-5 Å to the real structure. RDCs are highly useful to evaluate the quality of structures of biomolecules [78]. Previously, it was shown that it is possible to filter docking solutions using RDCs, chemical shift perturbation data and one or two distances from biological data [138]. To evaluate the feasibility of a structure determination of homodimers with HN-RDCs as the only experimental information, *ab initio* docking was performed for the lowest energy structure of the monomer NMR ensemble with the ClusPro Web server (<http://nrc.bu.edu/cluster>). The obtained *ab initio* docking results were then ranked according to the agreement of experimental RDCs with values predicted by the electrostatic module of PALES (see 3.5.3.3). The highest

correlation between predicted and experimental RDCs was 0.73 and was found for the homodimer structure with the lowest rmsd to the X-ray structure (Figure 5.10). This docking solution has a backbone rmsd (residues 3-63) to the X-ray structure

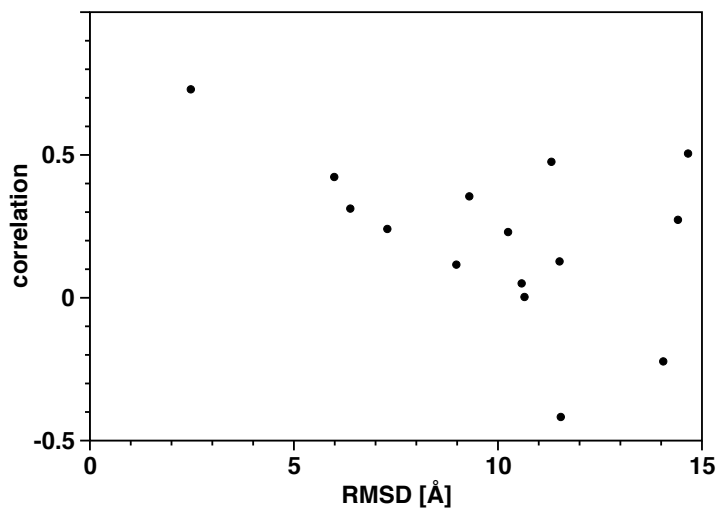


Figure 5.10: NMR-based ranking of structural models obtained from *ab initio* docking. Correlation between measured and charge-shape predicted RDCs multiplied by the RDC correlation obtained from best-fitting RDCs to the docking solution. Values are compared to the backbone rmsd (residues 3-63) of the *ab initio* model to the X-ray structure.

of 2.47 Å and is thus only slightly worse than the NOE-only structure of CylR2, but ~ 1.3 Å worse than the structure obtained from intermonomer PREs and NOEs. For all other docking solutions, the correlation was at least 0.22 worse (Figure 5.10). Thus, the correct monomer-to-monomer orientation was identified without any experimental data about the dimer interface only using a small number of easily accessible HN-RDCs.

5.4 Discussion

The structure determination of homooligomeric proteins by NMR is problematic because of the difficulty to distinguish inter- and intramonomer contacts. To overcome this problem, backbone HN-RDCs have been combined with long-range distances derived from PRE in a rigid-body docking protocol. Application of this method to CylR2, the regulator of cytolysin production from *Enterococcus faecalis*, resulted in a homodimeric structure that deviated by 1.15 Å from the X-ray structure (Figure 5.8). The distance information derived from PREs has four advantages: (i) it is not limited to the dimer interface, (ii) it is long-range, (iii) it can be used in the case of

fully deuterated proteins or for proteins for which no side chain assignment can be obtained and (iv) the number of long-range distances can be increased as required by using additional paramagnetic positions. The quality of structures determined using distances from PRE depends on the number and accuracy of the calculated distances. Compared to intramolecular long-range distances, the ratio of residues with strong and medium PRE to residues with weak PRE is decreased (Figures 5.5 and 5.6). PRE-derived distances from two mutation sites were required to unambiguously define the monomer-to-monomer orientation of CylR2. Due to the fact that hydrophobic interactions determine the dimer interface of CylR2 almost exclusively, it is especially difficult to unambiguously define it. The small differences between the NMR and the X-ray structure of CylR2 may be due to crystal packing artefacts.

For the measurement of intermolecular PRE distances, sample preparation by re-folding or mixing spin-labeled and ^{15}N -labeled protein is required. HSQC spectra of the 1:1-mixed ^{15}N -wt/mut(*) sample may show two peaks for residues close to the interface (Figure 5.5B), but peak-doubling is limited to residues close to the mutation site. Hence, the number of residues with doubled peaks will not depend on the molecular weight. This method is complementary to a previously reported approach, in which Co^{2+} was introduced as a paramagnetic probe sub-stoichiometrically into a homodimer to break the symmetry and measure RDCs and PCSs [134]. The approach by Gaponenko *et al.* abrogates degeneracy more efficiently than the approach reported here. Due to the PCSs, paramagnetic effects across the dimer interface can be determined directly from NMR spectra and must not be calculated. However, the triplicated signals and thus the increased peak overlap required a 3D HNCOC to identify the signals of a medium sized protein (~ 30 kDa). Therefore, a ^{13}C - and ^{15}N -labeled sample is necessary to determine the intermolecular distances. In contrast, the presented approach needs ^{13}C - and ^{15}N -labeling only for the sequential assignment. This promises to render the novel method superior for larger homodimers.

For proteins with multiple structurally important cysteines, PRE distances may be derived using other paramagnetic probes, e.g. the ATCUN-motif [130]. Application of PRE distances determined with the help of the ATCUN-motif was not possible

for CylR2 because an attached ATCUN-motif resulted in multiple conformations of CylR2 (Figure 5.9).

For successful rigid-body minimization, the position of the spin label has to be defined by intramolecular PRE distances prior to the docking. In addition, the intramolecular PRE distances represent an easy check on the specificity of MTSL binding. By comparing experimental intramolecular PRE distances with the theoretical distances within the monomer, unspecific binding of MTSL, which is a problem for some proteins [139], can be excluded.

Experimental data about the dimer interface may not be needed at all to identify the correct monomer-to-monomer orientation when *ab initio* docking results are filtered according to the correlation between experimental RDCs and values predicted from the 3D charge distribution and shape of the *ab initio* structural models. In case of CylR2, RDCs predicted from the docking solution, which was closest to the native structure, correlated best with experimental values (Figure 5.10). Previously, Zweckstetter *et al.* have used charge-shape predicted RDCs to distinguish between parallel and antiparallel orientations of helices in coiled-coil proteins [140]. In that case, the two possible structural assemblies had very different electrostatic properties, and distinction between the two assemblies was unambiguous. On the other hand, *ab initio* docking of protein complexes produces in general many different solutions. Some of these solutions are quite similar, at least with respect to their global electrostatic properties. In addition, PALES is based on a strongly simplified electrostatic model. Therefore, the 15-Å dimer model of CylR2 had a similar RDC correlation as a 6 Å model (Figure 5.10B). Thus, charge-shape predicted RDCs may fail to rank *ab initio* docking solutions in other cases, but they have the advantage that no mutations have to be introduced into the protein and experimental data about the dimer interface are not required.

5.5 Conclusions

In this chapter, it was demonstrated that a combination of long-range distances derived from paramagnetic relaxation enhancement induced by a spin label and a small

number of ^1H - ^{15}N residual dipolar couplings enables a high-resolution structure determination of a symmetric homodimer in solution. Here, the structure of the CylR2 monomer was determined by conventional NOE-based methods. For larger homodimers, which require deuteration, the structure determination of the monomeric unit will be more difficult. However, assignment of side chain resonances is not required as the distances and RDCs are solely determined from HSQC spectra. Therefore, the method is particularly useful for large homodimers or higher oligomers, for which a backbone assignment can be obtained, but for which the essential deuteration limits the availability of NOE data. In these cases, the intramolecular PREs, which were only used for positioning the spin label in case of CylR2, will also be highly valuable for the structure determination of the monomer [87]. In addition, the NMR-based rigid-body docking might be combined with X-ray crystallography for complexes, for which the dimer interface is not certain from the electron density map. Only based on a backbone assignment and a small number of HN-RDCs revealing similar solution and crystal structures, the dimer interface can be defined by measurement of PRE broadenings for a complex composed of spin-labeled and ^{15}N -labeled protein. Alternatively, charge-shape predicted RDCs, whose value for ranking *ab initio* docking solutions has been demonstrated in paragraph 5.3.6 of this chapter, can be applied to unravel uncertainties from X-ray crystallography. An example for this application of charge-shape predicted RDCs is presented for the 35.4 kDa homodimer CesT in paragraph 6.3.4 of the following chapter.



6

Structural studies of the function of the chaperone CesT in protein secretion via the type three secretion machinery of enteropathogenic *E. coli*

6.1 Introduction

A large number of Gram-negative bacterial pathogens of plants and animals utilize a type three secretion system (TTSS) to deliver virulence effector proteins into eukaryotic cells. These TTSSs are composed of ~25 proteins, which assemble to form a needle-like secretion complex spanning the bacterial inner and outer membrane and comprising an inner pore with a diameter of 2-3 nm (Figure 6.1). Through this pore, bacterial proteins are secreted and translocated into the eukaryotic cell.

Amongst the Gram-negative bacteria, which utilize a TTSS, are the human pathogens enteropathogenic *Escherichia coli* (EPEC) and enterohemorrhagic *Escherichia coli* (EHEC) O157:H7. They have evolved from commensal *E. coli*, the most abundant facultative anaerobic Gram-negative bacterium of the intestinal microflora, by horizontal gene transfer. These pathogenic bacteria form lesions on the intestinal surface by attaching and effacing (A/E) on host epithelial cells. Both, EPEC and EHEC, use a related translocation mechanism [141]. EPEC is the major cause of diarrhoea in young children and EHEC causes hemorrhagic colitis [142]. Figure 6.1 depicts the

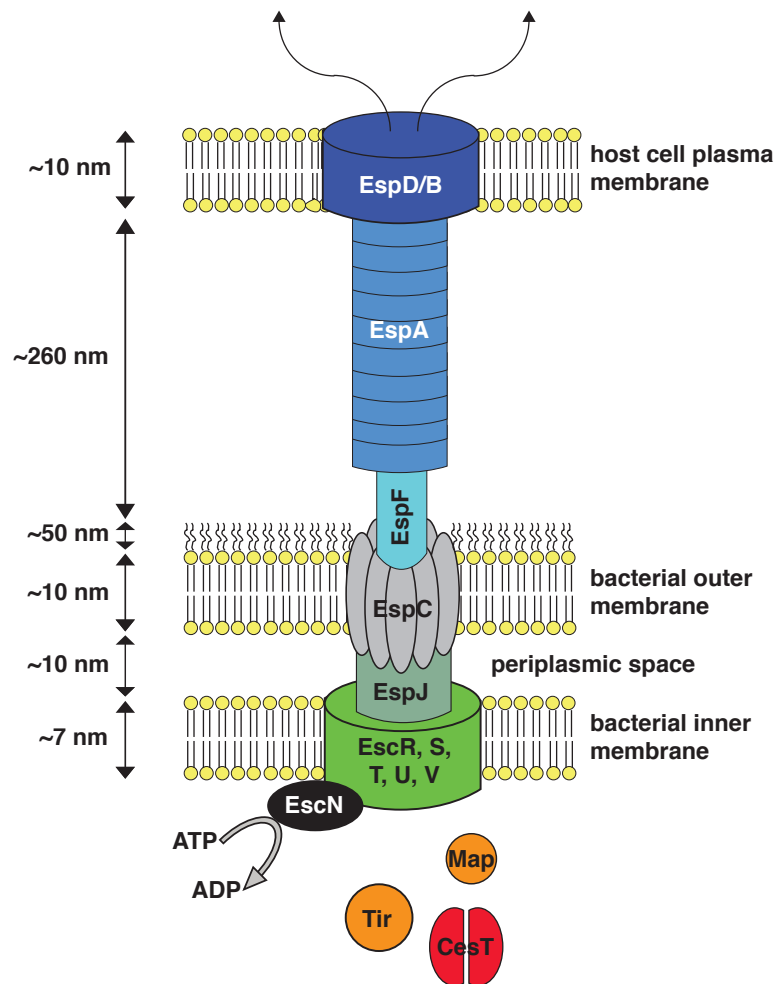


Figure 6.1: Schematic representation of the EPEC/EHEC type three secretion system. The basal body of the TTSS is composed of the outer membrane protein EscN, the inner membrane proteins EscR, EscS, EscT, EscU and EscV and the lipoprotein EscJ, which connects the inner and outer membrane ring structures. The needle structure is composed of EspF and EspA filaments. The translocation pore in the host cell plasma membrane is formed by EspB and EspD. By ATP hydrolyses, the ATPase EscN provides energy to the system. The chaperone CesT and the effectors Tir and Map are indicated as cytoplasmic components of the TTSS [141].

TTSS of EPEC and EHEC including the proteins composing the needle complex, proteins involved in the secretion process in the bacterial cytoplasm and proteins translocated from the bacterial cytoplasm into the eukaryotic host cell. The translocated proteins are called TTSS effectors and have a variety of activities that damage host cells. Two of the TTSS effectors of EPEC and EHEC are the translocated intimin receptor (Tir) and the mitochondrial associated protein (Map).

Map was given the name mitochondrial associated protein because it targets mitochondria via an N-terminal targeting sequence. The N-terminal 44 residues were shown to be sufficient to target Map to mitochondria where it is imported into the mitochondrial matrix [143]. Map mediates filopodia formation, causes mitochondrial dysfunction and may have an anti-apoptotic function [144], [145]. Recently, the biological function of Map was demonstrated to be the subversion of small G protein signaling pathways through molecular mimicry of the activated forms of distinct Rho family small G proteins [146]. However, the relevance of Map to the pathogenesis of EPEC-induced disease is uncertain.

Tir is inserted into the host cell plasma membrane forming a hairpin-like structure composed of two transmembrane domains. The extracellular loop between the two transmembrane domains interacts with intimin, a bacterial adhesion molecule. The C- and N-terminus, which are located within the host cell, interact with focal adhesion and cytoskeletal proteins. Thus, the bacterium is linked to the host cell cytoskeleton and actin-rich pedestals beneath adherent bacteria are formed [141].

As many other TTSS effectors, Tir and Map have a specific chaperone which is important for their secretion. The TTSS chaperones are a family of proteins which share only limited sequence similarity, tend to form homodimers, are highly acidic and are predicted to contain a C-terminal amphipathic α -helix [147]. A classification scheme of these chaperones has been proposed based on the function of the substrates [148]: Class I are effector binding chaperones whereas class II are chaperones binding the translocators, which form the TTSS channel. Class I was divided into two subclasses based on the number of substrates and the location of the genes encoding the chaperone and its substrates: class IA chaperones bind one effector and are encoded in the vicinity of their substrate gene, whilst class IB chaperones bind several effectors and are encoded within the operons for the TTSS components. Class III is composed of the chaperones of the flagellar export system which is the evolutionary predecessor of TTSSs [148, 149]. In general, the function of chaperones is to assist other proteins to achieve a proper folding. Though the function of TTSS chaperones is not clearly understood and seems to be versatile. Most widely accepted is their role in prevent-

ing aggregation or degradation of effector molecules. There is also evidence for a regulatory role in the release of effectors [150] and an involvement in the feedback regulation of the TTSS-associated gene expression [151].

The TTSS chaperone specific for Tir and Map is called CesT. CesT stands for chaperone for efficient secretion of Tir as it was originally discovered to bind only Tir [30] and was grouped to class IA of TTSS chaperones [148]. CesT is not absolutely required for the secretion of Tir [152]. In 2003, it was found to also bind Map [30] and in 2005 its binding to five effectors, NleA, NleH, NleF, EspH and SepZ, in addition to Tir and Map was revealed [153]. Therefore, Thomas *et al.* [153] redefined CesT as a chaperone with broad specificity for multiple TTSS effectors.

At present, nine X-ray structures of class IA TTSS chaperones are solved, including the structure of CesT from EHEC O157:H7 [154], SycE from *Yersinia pseudotuberculosis* [155], SycH from *Yersinia pestis* [156], SycT from *Yersinia enterocolitica* [157], [158], SicP [159] and SigE [154] from *Salmonella enterica* as well as the AvrPphFORf1 from the plant pathogen *Pseudomonas syringae* (Figure 6.2). The X-ray structures of Spa15 from *Shigella flexneri* [160] and InvB from *Salmonella typhimurium* [161] are the only solved structures of a class IB chaperone. In spite of their limited sequence similarity, the TTSS chaperones exhibit similar structures. A structure-based sequence alignment of the nine TTSS chaperones with known X-ray structures is displayed in Figure 6.20. Each monomer is composed of five β strands and three α helices. CesT contains an additional β strand while SycT lacks the dimerization helix $\alpha 2$ (Figure 6.2). The hydrophobic side of the predicted amphipathic C-terminal helix $\alpha 3$ is buried in the core of the chaperone. Each homodimeric chaperone possesses two pairs of hydrophobic patches on the surface [162].

For CesT, the structure shown in Figure 6.2E is a model constructed due to structural homology. However, the crystal structure of CesT from EHEC O157:H7 reveals a domain-swapped homodimer (Figure 6.3A) [154], where the N- and C-terminal regions lie on the adjacent molecule. As all other solved crystal structures of TTSS chaperones do not show a 3D domain swap, the question arises whether it is a crystallographic artefact or facilitates the chaperone-effector interaction [154].

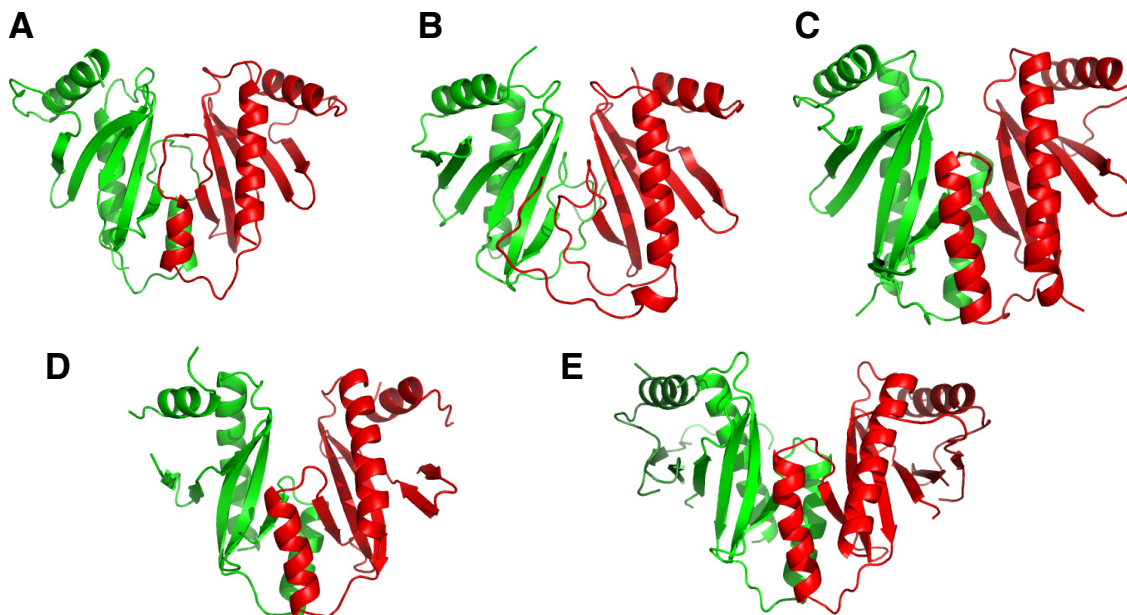


Figure 6.2: X-ray structures of TTSS chaperones as ribbon models with one monomer shown in red and the other in green: (A) SycE (PDB access code: 1L2W), (B) SycT (PDB access code: 2BSJ), (C) SicP (PDB access code: 1JYO), (D) SigE (PDB access code: 1K3S) and (E) the *CesT* model structure (PDB access code: 1K3E) with an analogous dimerization interface and the amino- and carboxy-terminal regions of each monomer in dark red and dark green.

With the help of recent structural data, progress has been made towards understanding the key elements of this chaperone-effector interaction. Four crystal structures of complexes of the N-terminal chaperone binding domain (CBD) of the effector and its cognate class IA chaperone have been solved: the *Salmonella typhimurium* chaperone SicP with its virulence factor SptP [159], the *Yersinia pseudotuberculosis* chaperone SycE with its effector YopE [162](Figure 6.4), the *Yersinia pestis* chaperone SycH and YSCM2, a negative regulator of type three secretion [156] and the heterodimeric chaperone SycN-YscB from *Yersinia pestis* with its substrate YopN [163]. Very recently, the X-ray structure of the class IB chaperone InvB with SipA from *Salmonella typhimurium* was published [161]. This structure is an exception as the interaction between InvB and SipA occurs only with one molecule of the homodimeric InvB [161]. All other known chaperone/effector structures show effector interactions with both chaperone monomers. These structures have revealed the N-terminal CBDs of the effectors as extended, nonglobular polypeptides which are wrapped around the

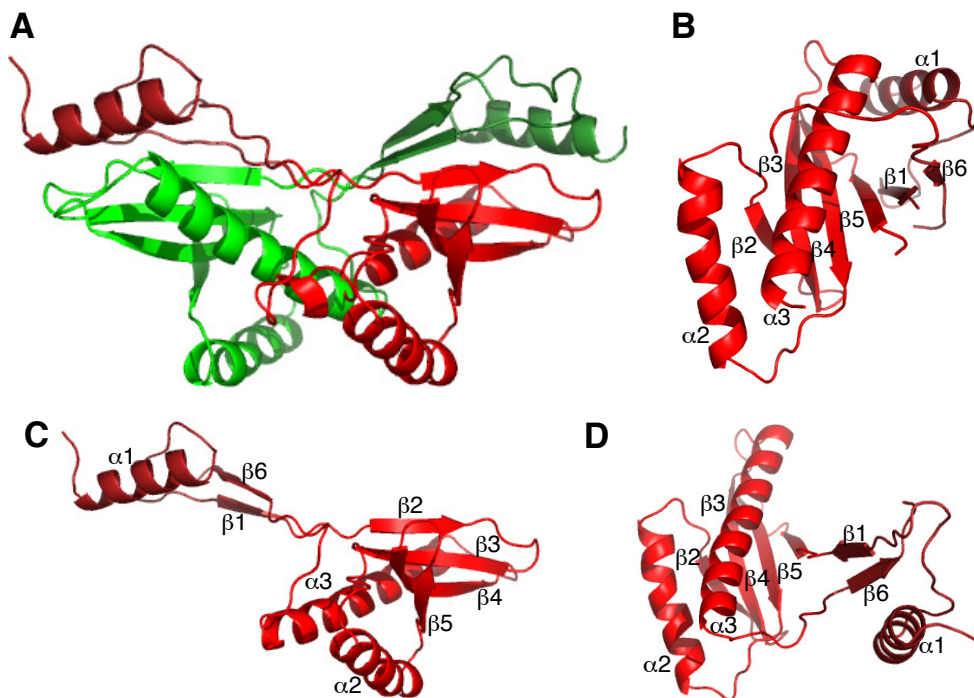


Figure 6.3: Structure of CesT. Dimeric X-ray structure (A), one monomeric subunit of the model structure (B) and one monomeric subunit of the X-ray structure (C,D) of CesT. The orientation of the monomeric subunit of the X-ray structure in (C) and (D) corresponds to the orientation of (A) and (B), respectively. One subunit is colored red and the other green. The small domain of each monomer (residues 1-34 and 139-146 forming $\alpha 1$, $\beta 1$ and $\beta 6$) are colored in dark red and dark green.

dimeric chaperones (Figure 6.4). Chaperone/effector interactions occur mainly between the hydrophobic patches on the chaperone surface and secondary structure elements of the effector. As different parts of the effector bind to identical regions of the chaperone, the specificity of the chaperone may arise from two complementary contributions [148]. Based on the X-ray structures of chaperone/effector complexes (Figure 6.4), the N-terminal CBDs can be roughly superimposed indicating a general three-dimensional fold [164]. The only conserved structural element in all five solved complex structures is a β -motif and hence an adaptor-targeting model was proposed by Lilic *et al.* [161]: the β -motif of the CBD functions as an adaptor element, which targets the effector to the chaperone. Then, the chaperone docks the effector to the membrane-associated ATPase of the TTS machinery. Only a homology-based non-swapped model structure of CesT is consistent with the key binding elements of the chaperone/effector interaction [164].

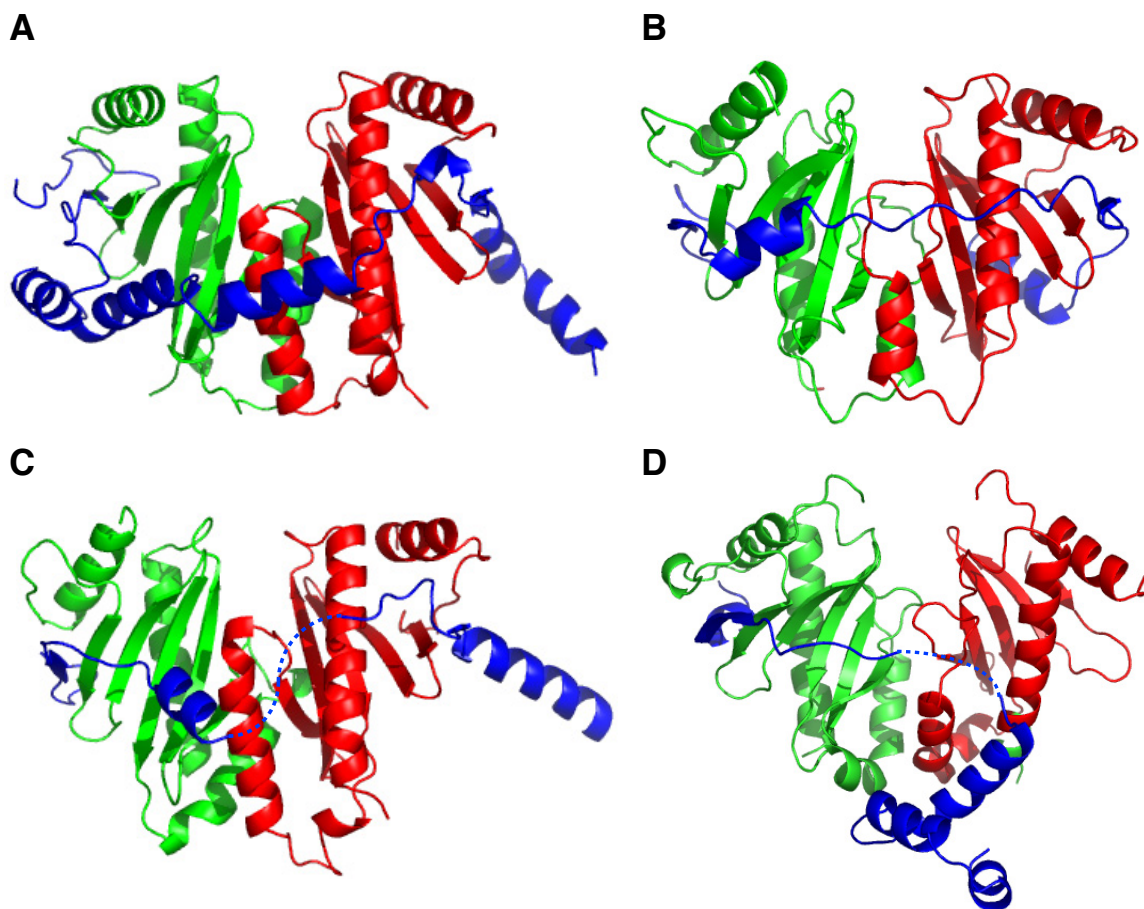


Figure 6.4: Ribbon diagrams of crystal structures of chaperone/effector complexes. The two subunits of the chaperones are shown in red and green and the effectors in blue with disordered regions as dotted lines. (A) The chaperone SicP with residues 35-139 of SptP (PDB access code:1JYO), (B) the chaperone SycE with residues 17-85 of YopE (PDB access code: 1L2W), (C) the heterodimeric chaperone SycN-YscB with residues 32-92 of YopN (PDB access code:1XKP) and (D) the chaperone InvB with residues 23-93 of SipA (PDB access code:2FM8).

Little is known about the structure of the free CBDs of TTSS effectors. The X-ray structure of the CBD of SipA comprising residues 48-158 is the only known structure of a CBD of a TTSS effector. However, SipA⁴⁸⁻²⁶⁴ is not able to form a stable complex and SipA²²⁻²⁶⁴ was used to solve the complex structure of the InvB/SipA complex [161].

At this stage, different experimental data exist which render a global unfolding activity of the chaperone on the effector unlikely. Specifically, experimental data about the TTSS effector, e.g. for Tir and SigD, demonstrate similar physical or enzymatic

properties of the effector in the absence and presence of their chaperone [154]. On the other hand, an unfolding activity of the TTSS-associated ATPase InvC in the presence of ATP for the chaperone/effector complex SicP/SptP from *Salmonella enterica* was demonstrated [165].

Overall, the secretion mechanism is only partly understood. Especially intriguing is the question how the passage of the effector proteins through the ~ 3 nm narrow pore is achieved and what the role of the specific chaperones is. Disassembly of the chaperone/effector complex prior to and during secretion may occur partly or entirely through domain swapping [164]. Further structural insights, particularly of the chaperone and chaperone/effector complexes are prerequisite for a better understanding of the TTSS mechanism which is crucial for many medically important bacterial pathogens. To answer these questions, the chaperone CesT and its complexes may prove to be extremely valuable. Firstly, as the CesT X-ray structure differs from the other X-ray structures of other TTSS chaperones, further structural studies of CesT may improve our understanding of the TTSS mechanism. Secondly, in recent experiments five effector proteins in addition to Map and Tir, which bind CesT, were found and CesT was shown to specifically deliver multiple effectors to the TTSS [153]. Furthermore, it was demonstrated that CesT interacts not only with TTSS effectors but also with the membrane-associated ATPase EscN [166].

During this project, a contribution to the understanding of how the TTSS chaperone CesT from EPEC E2348/69 aids in the injection of the TTSS effector proteins Tir and Map was made. In the first part of this chapter, the backbone assignment of the 35.4 kDa homodimer CesT and its properties in solution are depicted as a fundamental step for all further NMR studies. Without calculating a NMR solution structure of CesT experimental RDCs and charge-shape predicted RDCs with PALES were applied to investigate whether the solution structure is the domain-swapped (Figure 6.3) or the non swapped structure (Figure 6.2E).

In the second part of this chapter, steps towards a CesT/effector complex structure are presented. These structural studies are the starting point to further understand the mechanism of the TTS machinery in general and in particular the TTSS of EPEC

and EHEC. This knowledge is important to answer the question how the translocation and unfolding of TTSS effectors is triggered. An answer to these questions may render the introduction of engineered proteins into almost any cultured cell via a TTSS possible. Eventually the design of drugs that can combat a broad range of diseases should be possible.

6.2 Materials and methods

6.2.1 Cloning

Initially, the *NdeI* cleavage site within the gene for *CesT* from EPEC E2348/69 in pQE30-His₆*CesT* [29] was eliminated through a silent mutation using the QuikChangeTM Site-directed mutagenesis Kit (see 3.1.5). Afterwards the gene coding for *CesT* from EPEC E2348/69 was amplified by PCR and cloned into the *NdeI* and *BamHI* restriction sites of pET16bTEV.

Template plasmids containing genes coding for Tir and Map are specified in Table 2.2. The genes coding for Tir and Map as well as different C-terminally truncated genes were cloned into the *NdeI* and *XhoI* restriction sites of pET16bTEV (Figure 2.1A). For coexpression of *CesT*/effector complexes, the gene coding for the effector was inserted into the *NdeI* and *XhoI* restriction sites of the multiple cloning site 2 (MCS2) of pETDuet1 and the gene coding for *CesT* was inserted into the *NcoI* and *BamHI* restriction sites of the MCS1 of pETDuet1 (Figure 2.1B). The different constructs are summarized in Table 2.2.

6.2.2 Protein expression

Details of protein expression are described in 3.2.1. For coexpression of *CesT* and of *CesT*/Map complexes in M9-based minimal medium the temperature was reduced to 17 °C at an OD₆₀₀ of 0.4 - 0.5 and cells were harvested 16 - 20 h after induction. Cells producing *CesT* and growing in M9-based minimal medium prepared with 99.9 % D₂O were harvested 40 - 48 h after induction.

6.2.3 Purification

Images of SDS-PAGEs documenting the different protein purifications as well as ESI-MS spectra confirming the purity and molecular weight of the proteins can be found in Appendix A.

6.2.3.1 Purification of CesT

The cell lysis was performed under native conditions for cells grown in LB and under denaturing conditions for cells grown in M9-based minimal medium (see 3.2.2). Recombinant, His-tagged CesT was initially purified via affinity chromatography on a Ni-NTA Agarose resin under native or denaturing conditions (see 3.3.1). After removal of the His-tag by cleavage with the TEV-protease (see 3.3.5), CesT was further purified by anion exchange chromatography on a 5-ml HiTrap Q-Sepharose column (see 3.3.2). The protein solution was diluted to 120 ml with anion exchange buffer A and was supplemented with DTT and EDTA to a final concentration of 1 mM and 0.5 mM, respectively. The protein was eluted by linearly increasing the percentage of anion exchange buffer B from 0 - 100 % over a volume of 100 ml at a flow rate of 3.0 ml/min. Fractions containing CesT were pooled, dialysed extensively against H₂O and stored in lyophilised form at -20 °C. One litre of culture medium yielded ~40 mg of unlabeled CesT, ~30 mg of ¹⁵N or ¹⁵N/¹³C labeled CesT and ~20 mg of ²H, ²H/¹⁵N or ²H/¹⁵N/¹³C labeled CesT.

6.2.3.2 Purification of Map

At first, cells were ruptured under native conditions. After centrifugation the pellet was applied to cell lysis under denaturing conditions (see 3.2.2). Then the His-tagged Map was purified via a Ni²⁺-column (see 3.3.1). Finally the His-tagged Map was used for the preparation of CesT/Map complexes or dialysed directly against NMR buffer.

6.2.3.3 Purification of Tir

Cell pellets containing recombinantly expressed Tir were ruptured under native conditions (see 3.2.2), the His-tagged Tir was purified by affinity chromatography on a Ni-NTA Agarose resin (see 3.3.1) and the His-tag was cleaved with the TEV-protease (see 3.3.5). After removal of the TEV-protease Tir was further purified by RP-HPLC (see 3.3.4). The purity of Tir was confirmed by ESI-MS (see 3.4.3). The purified Tir was stored in lyophilised form at -20 °C. The yield is ~5 mg/l.

6.2.3.4 Purification of CesT/effector complexes

After native cell rupture (see 3.2.2) coexpressed CesT/effector complexes were bound via the N-terminal His-tag of CesT to Ni-NTA Agarose resin and eluted by increasing the imidazole concentration. For the preparation of selectively labeled CesT/Map complexes, both complex components were purified under denaturing conditions via a Ni²⁺-column (see 3.3.1). Protein concentrations were determined by measuring the A₂₈₀ (see 3.4.1). Sample volumes of CesT and Map were mixed with a three- to fourfold molar excess of CesT. The CesT/Map complex was formed during dialysis against 2 times 5 l TEV-buffer.

All His-tags were cleaved using the TEV-protease (see 3.3.5). For further purification and removal of the excess of CesT, CesT/Map complexes were applied to anion exchange chromatography on a 5-ml HiTrap Q-Sepharose column performed as described for CesT (see 3.3.2 and 6.2.3.1). To achieve an optimal separation of CesT and the CesT/Map complex, the flow rate was 1.0 ml/min during the elution. Fractions containing the CesT/Map complex were pooled and dialysed against the NMR buffer for CesT without D₂O (Table 2.6).

For crystallization the CesT/Map complex was dialysed after anion exchange chromatography into gel filtration buffer (Table 2.6) and concentrated to ~10 mg/ml. The gel filtration was performed as described in 3.3.3.

6.2.4 NMR experiments

Lyophilised CesT and Tir108 were dissolved directly in NMR buffer (Table 2.4). CesT/effector samples and refolded His₇-Map84 were concentrated by ultrafiltration (see 3.3.6) to a final volume of $\sim 200 \mu\text{l}$ and 20-25 μl D₂O were added.

All NMR spectra were acquired at 303 K with the exception of the spectra used for the backbone assignment of Tir108 which were acquired at 293 K. Triple resonance experiments measured for the backbone assignments are listed in the Appendix B.1. Backbone assignments of CesT and Tir108 as well as the reassignment of the CesT backbone in CesT/effector complexes were derived from triple resonance NMR spectra as summarized in the Appendix B.2. HN, C'N and C'C_α RDCs were measured for a concentration of 12 mg/ml Pf1 in the NMR buffer (see 3.5.3). The oligomerization state was estimated as described in 3.5.4.1 and the ¹⁵N-¹H steady-state NOE of free and bound CesT were measured and analysed as described in 3.5.4.3.

A series of 2D ¹⁵N-¹H-HSQC spectra were measured to record the formation of CesT/effector complexes. To monitor the chemical shift changes of CesT upon complex formation 100 μM ²H,¹⁵N-labeled CesT was titrated with 100 μM His₇-Map84 or 300 μM Tir108. 2D ¹⁵N-¹H-TROSY-HSQC spectra were recorded at a CesT to effector ratio of 1:0.1, 1:0.2, 1:0.4 and 1:1. For CesT/Tir108, an additional 2D ¹⁵N-¹H-TROSY-HSQC spectrum was acquired for a twofold excess of Tir108. To determine the chemical shift changes of the effectors upon complex formation, CesT was added from a 1-mM stock solution to a 100- μM ¹⁵N-labeled effector sample. 2D ¹⁵N-¹H-HSQC spectra were measured at Tir108:CesT ratios of 1:0.25, 1:0.5, 1:1, 1:2, 1:4 and 1:8, at TirN:CesT ratios of 1:0.05, 1:0.1, 1:0.2, 1:0.5, 1:1 and 1:2 and at His₇-Map84:CesT ratios of 1:0.2, 1:0.4, 1:1 and 1:2.

6.3 Results and discussion

6.3.1 Experimental difficulties

The plasmid pQE30-His₆-CesT was kindly provided by the group of Gad Frankel (Imperial College London) and was initially used to prepare His-tagged CesT expressed

in *E. coli* BL21. However, the ^{15}N -HSQC and the ^{15}N -TROSY-HSQC spectrum recorded for ^{15}N -His₆-CesT suggested aggregation of the sample (Figure 6.5). NMR

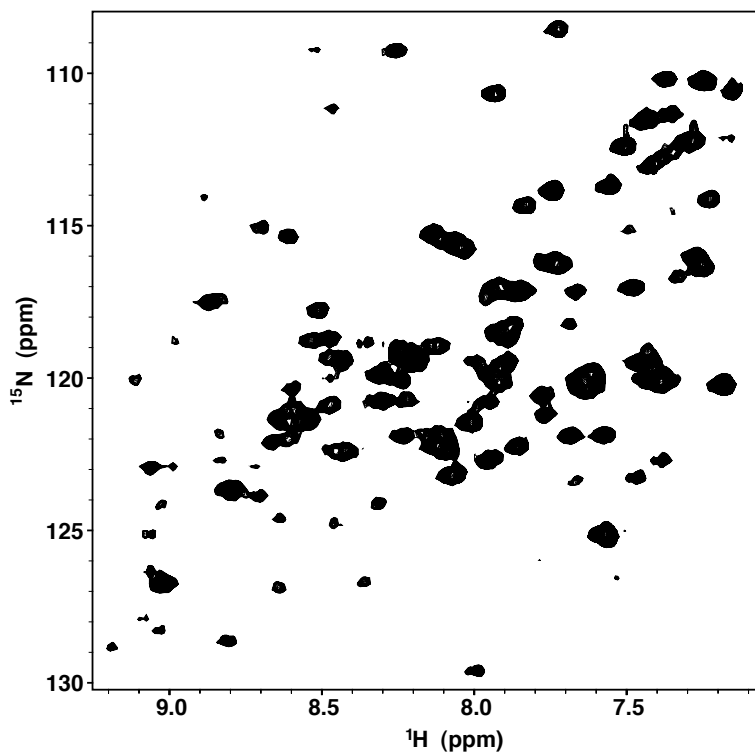


Figure 6.5: 800 MHz ^{15}N -HSQC spectrum of His₆-CesT at 303 K.

relaxation times and results from dynamic light scattering indicated a CesT trimer which is in disagreement with previous results showing a CesT dimer [154]. Furthermore, His₆-CesT was not soluble at concentrations above 300 μM and the construct did not contain a protease cleavage site for the removal of the His₆-tag. Therefore pET16bTEV-CesT was constructed to enable the preparation of CesT without a His-tag. The ^{15}N -HSQC and the ^{15}N -TROSY-HSQC spectrum for CesT without His-tag were significantly improved (Figure 6.6). Intensities were almost identical for all signals and no strong line broadening was observed. CesT without His-tag is soluble to concentrations up to 1 mM and stable for a few months if 0.5 mM EDTA is present in the NMR buffer.

6.3.2 Extent of the assignment of CesT

3D NMR experiments acquired for the backbone assignment of CesT are summarized in the Appendix Table B.2. The analysis of the triple resonance spectra (see 3.5.1.1) allowed identification and sequential assignment of 97 % of the backbone amide resonances, not including the ^{15}N of the five proline residues. Five residues could not be found in the HSQC due to overlap (Figure 6.6).

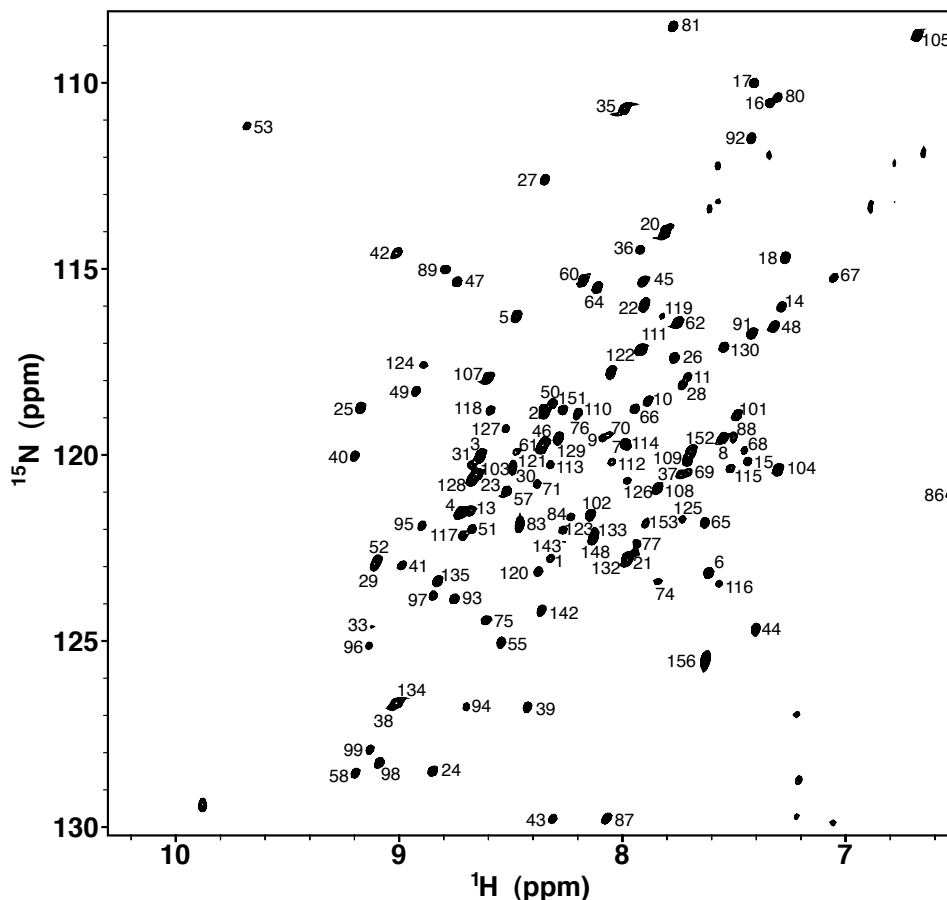


Figure 6.6: 900 MHz ^{15}N -TROSY-HSQC spectrum of CesT at 303 K. Resonance assignments are indicated with residue numbers.

Of the C' resonances 98.7 %, of the C_α resonances 98 % and of the C_β 98 % resonances have been assigned unambiguously. The chemical shift data have been deposited in the BMRB (<http://www.bmrwisc.edu>) under accession number 6451.

6.3.3 Secondary structure and dynamics of *CesT* in solution

Secondary structure elements were determined from the deviation of the C_α , C_β and C' chemical shifts from the random coil values (see 3.5.2) and are shown in Figure 6.7. Regions of secondary structure identified by the combined secondary chemical shifts

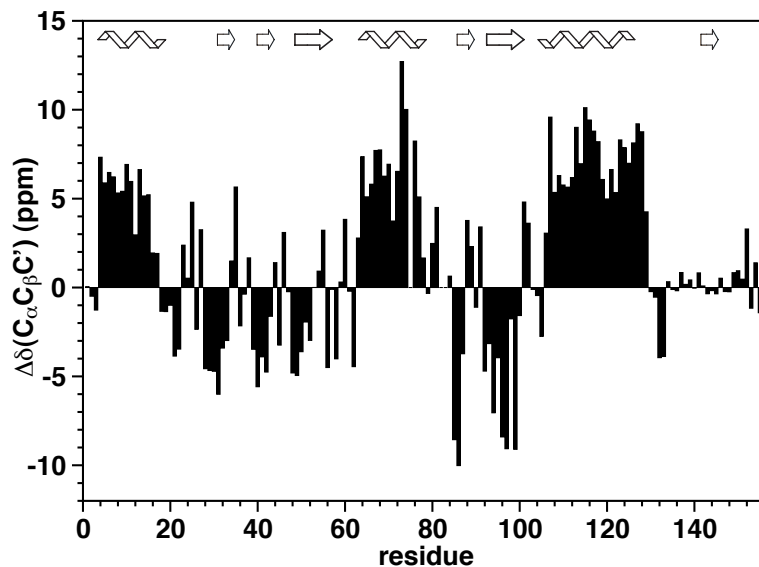


Figure 6.7: Combined secondary chemical shifts for *CesT* plotted against the residue number. Secondary structure elements of the X-ray structure are indicated.

were confirmed by the characteristic NOE patterns [7] in the 3D [^{15}N - ^1H] NOESY-HSQC. The X-ray structure reveals three α helices and six β strands: $\alpha 1$ from residue 4 to 16, $\alpha 2$ from residue 63 to 78, $\alpha 3$ from residue 106 to 128, $\beta 1$ from residue 30 to 33, $\beta 2$ from residue 38 to 42, $\beta 3$ from residue 30 to 33, $\beta 4$ from residue 83 to 87, $\beta 5$ from residue 92 to 100 and $\beta 6$ from residue 141 to 144. These regions are identical to the regions found via the secondary chemical shifts with the exception of the C-terminal part containing residues 130 to 156 where no secondary structure is indicated by the secondary chemical shifts. Therewith the additional strand $\beta 6$ of the *CesT* X-ray structure which is not found in the other X-ray structures of TTSS chaperones, is also not present for *CesT* in solution.

The amplitude of motions in the pico- to nanosecond time-scale was investigated with the steady-state heteronuclear ^{15}N - ^1H -NOE (see 3.5.4.3). Heteronuclear amide NOE values larger than 0.6 for most of the residues of *CesT* indicate rigidity of *CesT*

especially for the secondary structure regions (Figure 6.23A). Heteronuclear amide NOE values are slightly decreased for the N-terminal residues, the C-terminal region of $\alpha 2$ and in loop regions connecting secondary structure elements. In agreement with the secondary chemical shifts a highly flexible, unstructured region is indicated by the negative heteronuclear amide NOE values for the C-terminal region comprising residues 138 to 156. Thus the 20 C-terminal residues are unfolded in solution.

6.3.4 Solution structure of CesT

The main question that arises about the 2.8 Å resolution X-ray structure of CesT in comparison to the other X-ray structures of TTSS chaperones regards the relevance of the 3D domain-swap. The CesT dimer in the crystal is created by domain-swapping of the small domain comprising residues 1-34 and 139-156 and ~ 4550 Å² of the surface are buried at the swapped dimer interface. This is significantly more than the ~ 2000 Å² surface area which is buried in the CesT model structure. The overall shape of both CesT structures and their charge distribution is depicted in Figure 6.8. The shape and the charge distribution of the two structures is very different. Thus,

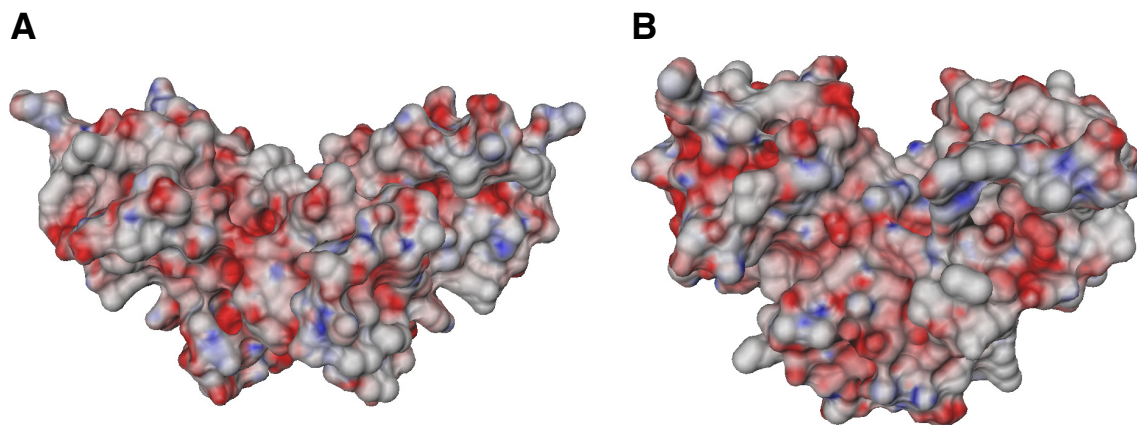


Figure 6.8: Electrostatic surface of the X-ray structure (A) and the model structure (B) of CesT with positive and negative potentials colored blue and red, respectively.

the orientation of the alignment tensor predicted using the electrostatic module of the software PALES should be very different for the two structures. Comparison of this predicted RDCs with RDCs measured in Pf1 medium can then be used to identify the structure present under solution NMR conditions. 100 backbone amide RDCs, 96 C'N

RDCs and 71 C'C $_{\alpha}$ RDCs were derived experimentally for each monomer (see 3.5.3). All backbone RDCs that could be determined accurately were used. A few RDCs had to be excluded due to peak overlap and some outlying RDCs in loop regions were excluded. RDCs for the 20 C-terminal residues could not be derived as the peaks were missing in the TROSY-based experiments. Best-fitting the observed RDCs to the crystal structure (Figure 6.9A) and to the model structure of *CesT* (Figure 6.9C) results in a correlation coefficient of 0.772 and 0.930 between experimental and back-calculated RDCs, respectively.

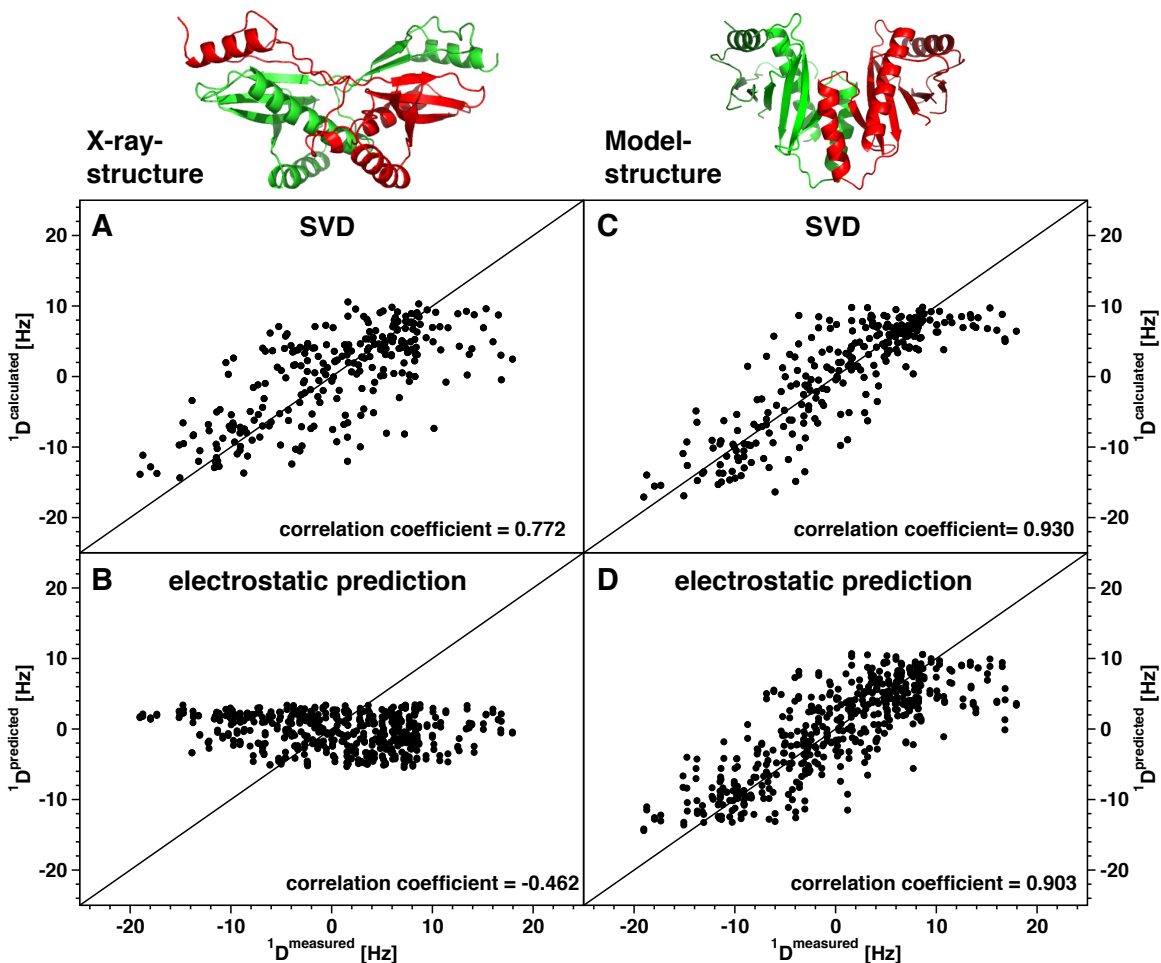


Figure 6.9: Correlation between experimental and predicted or back-calculated HN, C'N and C'C $_{\alpha}$ RDCs of *CesT*: For the X-ray structure and back-calculated RDCs (A) and charge-shape predicted RDCs (B), for the model structure and back-calculated RDCs (C) and charge-shape predicted RDCs (D). C'N and C'C $_{\alpha}$ RDCs are normalized to those of HN using the appropriate factor $\gamma_H\gamma_N r_{HN}^{-3}/\gamma_A\gamma_B r_{AB}^{-3}$ where r_{AB} is the bond distance between A and B and γ_i is the gyromagnetic ratio of i.

The back-calculated alignment tensor is characterized by $D_a = -5.4$ Hz and $R = 0.09$ for the X-ray structure and by $D_a = -7.5$ Hz and $R = 0.52$ for the model structure. The correlation coefficient from best-fitting for the X-ray structure is 0.16 less than for the model structure. These best-fit results point slightly at the non-swapped model structure as the structure which is present under solution NMR conditions. However the X-ray structure cannot be excluded based on these results.

Therefore, RDCs were predicted using the electrostatic module of PALES with the default charges placed at the positions of the ionizable side chains (see 3.5.3.3). Even though the 10 C-terminal residues of the native CesT sequence are missing in the X-ray and model structure, the net charge of the structures used for the prediction is identical to the native situation as two positive and negative charges are present among the 10 C-terminal residues. In Figure 6.9B and D, the RDCs predicted with PALES via the charge-shape model for a NaCl concentration of 0.1 M are plotted against the experimental RDCs for the two structures. With a correlation coefficient of -0.462, the RDCs predicted from the X-ray structure do not fit to the experimental values (Figure 6.9B). As opposed to the X-ray structure, the sign and size of the predicted and experimental RDCs correlate very well for the model structure with a correlation coefficient of 0.9 (Figure 6.9D).

Consequently, a non-swapped structure which is best represented by the model structure exists under solution NMR condition and the 3D domain-swap is most likely a crystallographic artefact.

6.3.5 CesT_{V116R} aggregates

The C-terminal amphipathic α -helix is the only conserved feature among TTSS chaperones. Mutation of the central residue Val116 of this α -helix of CesT to Arg was demonstrated to abrogate the interaction with Tir [31] and Map [30] raising the question whether the effects of this point mutation are a direct consequence of impaired inter- or intra-helical interactions or a indirect consequence from associated conformational changes elsewhere in CesT.

To address this question, the structure of CesT_{V116R} should be determined and com-

pared with the *CesT* structure. However, the ^{15}N -HSQC spectrum of *CesT*_{V116R} reveals a strongly aggregated protein and T_2 as well as $T_{1\rho}$ relaxation times are halved compared to *CesT*. Changing the buffer and decreasing the concentration of *CesT*_{V116R} to 100 μM did not reduce the level of aggregation. These *in vitro* results indicate a strong effect of the point mutation V116R on the stability and probably the structure of *CesT*. Therefore, it is unlikely that *CesT*_{V116R} can provide insights into the TTSS mechanism. It was therefore not used for further studies.

6.3.6 Structural characterisation of Tir and Map

As for most TTSS effectors, the CBD of Map and Tir has been shown to be located within the first ~ 100 amino-terminal residues. For Tir, it was demonstrated by Abe *et al.* [167] that residues 50-100 are essential for *CesT* binding. For Map it was demonstrated by Creasey *et al.* [30] that *CesT* interacts within the 101 amino-terminal residues and that neither residues 1-51 nor 51-101 are sufficient for the interaction. Therefore, the C-terminally truncated fragments of Tir and Map were mainly studied here.

A degradation product of Map101, coexpressed and copurified with *CesT*, was detected by SDS-PAGE (see 3.3.8). Analysis of this *CesT*/Map101 sample by ESI-MS (see 3.4.3) pointed at a mass of 9157.71 ± 0.60 for the degradation product (Figure 6.10). Further analysis by Edman degradation (see 3.3.9) revealed only two sequences with the intact N-termini of *CesT* and Map. Combining these experimental results, the degradation product of Map101 was assessed as residues 1-84 of Map (Map84). This shortened fragment was used for further studies.

All differently expressed Map constructs are insoluble in aqueous buffers. Therefore, the protein was prepared in the presence of 8 M urea via a single Ni^{2+} -column and refolded by dialysis against NMR-buffer (see 6.2.3.2). In this way, it was possible to produce His₇-Map84 at a concentration of $\sim 100 \mu\text{M}$ in the *CesT* NMR-buffer (Table 2.6) with $\sim 5\%$ impurities.

In order to estimate the secondary structure content of the N-terminal regions of Tir and Map, far-UV CD spectra of His₇-Map84 and Tir108 in NMR buffer were ac-

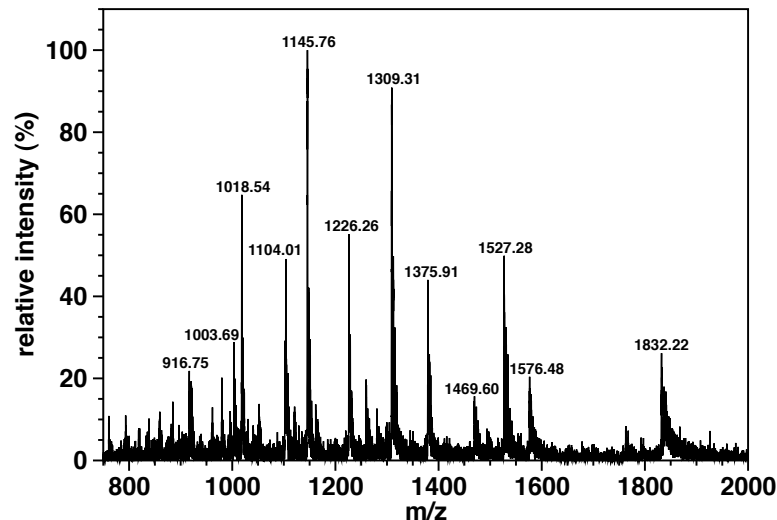


Figure 6.10: ESI-mass spectrum of coexpressed and copurified CesT/Map101 in 20 mM NH_4OAc . The signals at a m/z -ratio of 1832.22, 1527.28, 1309.31, 1145.76, 1018.54 and 916.75 correspond to a 5- to 10-fold positively charged degradation product of Map101. The less intense signals at a m/z -ratio of 1576.76, 1375.91, 1226.26, 1104.01 and 1003.69 belong to 7- to 11-fold positively charged Map101.

quired (Figure 6.11). Both proteins show far-UV CD spectra typical of an unfolded

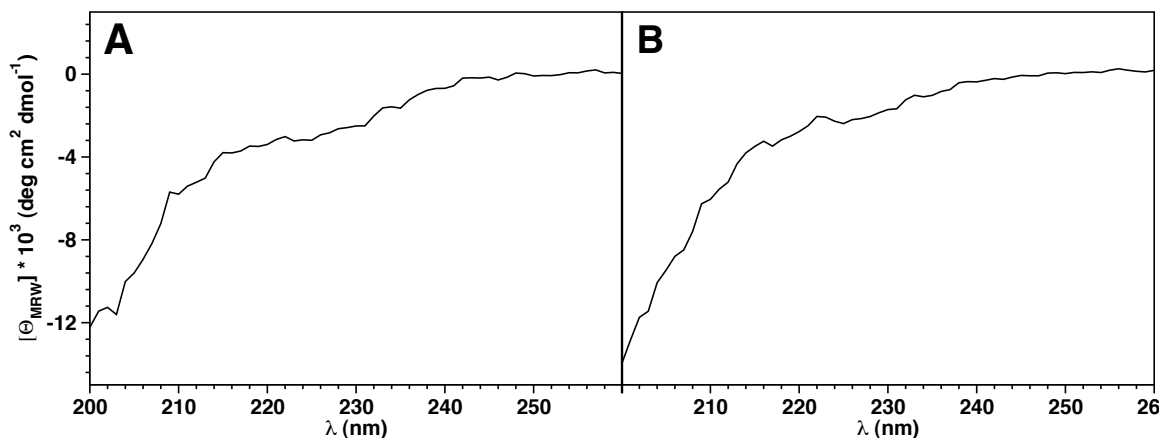


Figure 6.11: Far-UV-CD spectra in NMR buffer at 303 K of His₇-Map84 (A) and Tir108 (B).

polypeptide chain with a largely negative ellipticity around 200 nm and a negligible ellipticity at 222 nm [168]. The unfolded nature of His₇-Map84 and Tir108 is additionally evidenced by the absence of dispersion in the proton dimension of the ¹⁵N-HSQC spectra of free His₇-Map84 (Figure 6.15) and Tir108 (Figure 6.13).

6.3.7 Crystallisation of the *CesT*/Map complex

Limited proteolysis (see 3.3.10) was performed in order to detect the proteolytically stable core of Map84 which is protected by *CesT*. This approach has already been successfully applied to reveal the proteolytically stable core for the homologous chaperone/effector complexes of SptP when bound to SicP [159], of YopE when bound to SycE [162], of YopN when bound to SycN-YscB [163] and of SipA when bound to InvB [161]. However, in the case of the *CesT*/Map84 complex the chaperone was proteolysed even at low protease concentrations and short reaction times. Only a *CesT* fragment of about 15 kDa proved stable when digested with elastase. The digestion was completed at a elastase:protein ratio of 1:1000 after a reaction time of 2 h. Since Map84 was revealed as the shortest N-terminal fragment of Map which binds *CesT*, crystallisation of Map84 in complex with full length *CesT* was attempted.

Coexpressed and copurified *CesT*/Map84 complex after gel filtration (see 3.3.3) was concentrated to ~10 mg/ml for crystallisation trials using 372 different conditions from 6 crystallisation screens (see 2.5). As the C-terminal residues of *CesT* were detected as highly flexible by NMR (Figure 6.23) and *in vivo* studies of *CesT*_{E142G} demonstrated unchanged effector binding [153], crystallisation trials were also performed with a *CesT*₁₋₁₃₅/Map84 complex. Small crystals were obtained at 0.2 M NH₄OAc, 0.1 M Na₃-citrate pH 5.6 and 30 % PEG 4000 and were improved by a grid screen and increasing the complex concentration to 20 mg/ml. Crystals grown in hanging drops at the final conditions, 0.2 M NH₄OAc, 0.1 M Na₃-citrate pH 5.7 and 26 % PEG 4000, were washed and shown to contain both complex components by SDS-PAGE. Nevertheless the crystals did not diffract. Trials to improve them were not successful so far.

6.3.8 Backbone assignment and secondary structure of Tir108

Backbone and C_{β} resonances of Tir108 were assigned. The assignment is complete with the exception of Pro17 and Pro20 which are followed by a proline residue, of the amide ¹H and ¹⁵N of His14 and the carbon resonances of Met1 (Table B.10). The completeness and confirmation of the assignment was achieved with the help of a 3D

HNN experiment which establishes correlations between amide protons and amide nitrogen atoms of both neighbouring residues [169].

The secondary chemical shifts of the C_α and C' resonances, derived as described in 3.5.2, is another evidence for the unfolded state of Tir108 as all secondary chemical shift values are small (Figure 6.12).

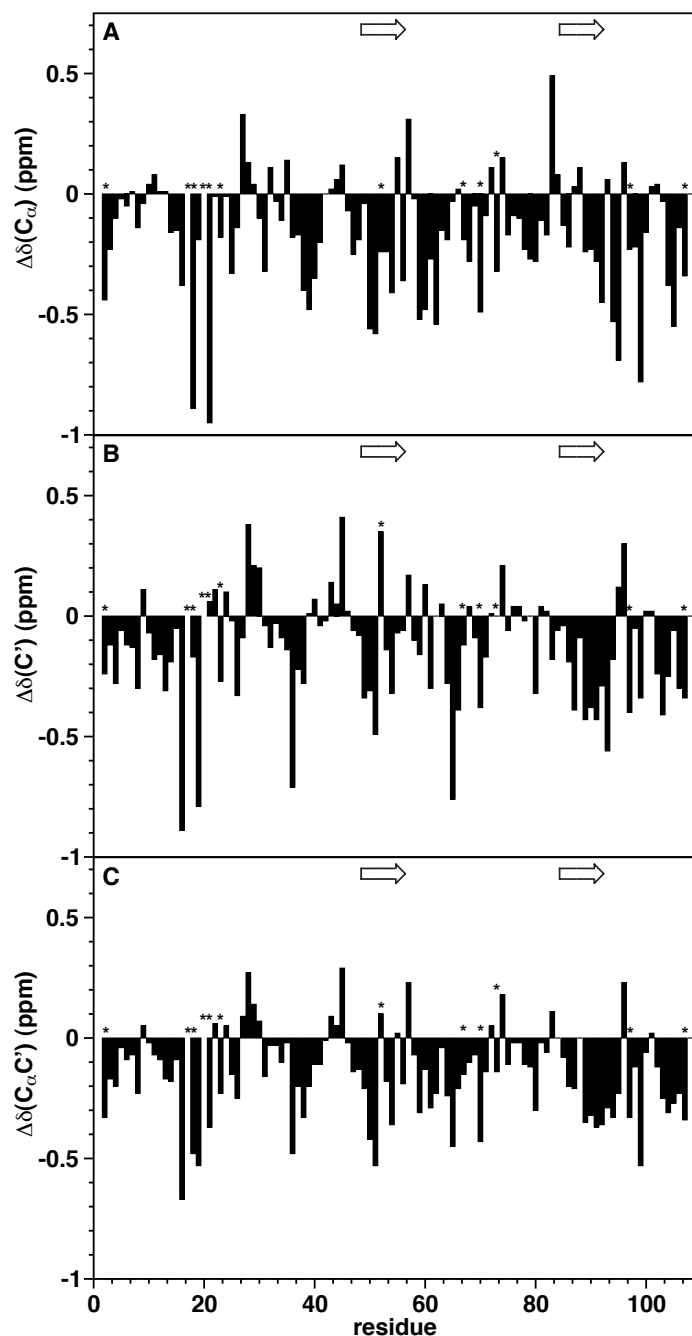


Figure 6.12: Secondary chemical shifts of Tir108 in 50 mM potassium phosphate buffer pH 6.8, 100 mM NaCl and 5 mM DTT at 293 K. Secondary chemical shifts for ΔC_α (A), $\Delta C'$ (B) and averaged for ΔC_α and $\Delta C'$ (C) plotted against the residue number. The average shifts in (C) were calculated as $[3\Delta\delta(C_\alpha) + 4\Delta\delta(C')]/7$. The two proposed β -motifs for Tir108 are indicated [161] and proline residues are labeled with an asterisk.

A rapid exchange between extended and β -structures is suggested by the negative

secondary chemical shift values for 85 out of the 104 residues. In agreement with a β -motif which is proposed as the only conserved targeting motif in type three secretion [161] and with secretion of Tir in the absence of *CesT* [29], the β -structure tendency indicates a preformed secretion signal for free Tir.

6.3.9 Effects of *CesT* binding on Tir and Map

To monitor chemical shift changes and potentially induced secondary structure formation of Tir and Map upon *CesT* binding, NMR titration experiments were performed (see 6.2.4). The start and end ^{15}N -HSQC spectra for the titration of ^{15}N -Tir108 with *CesT* are shown in Figure 6.13. Around 30 % of the peaks of Tir108 disappear

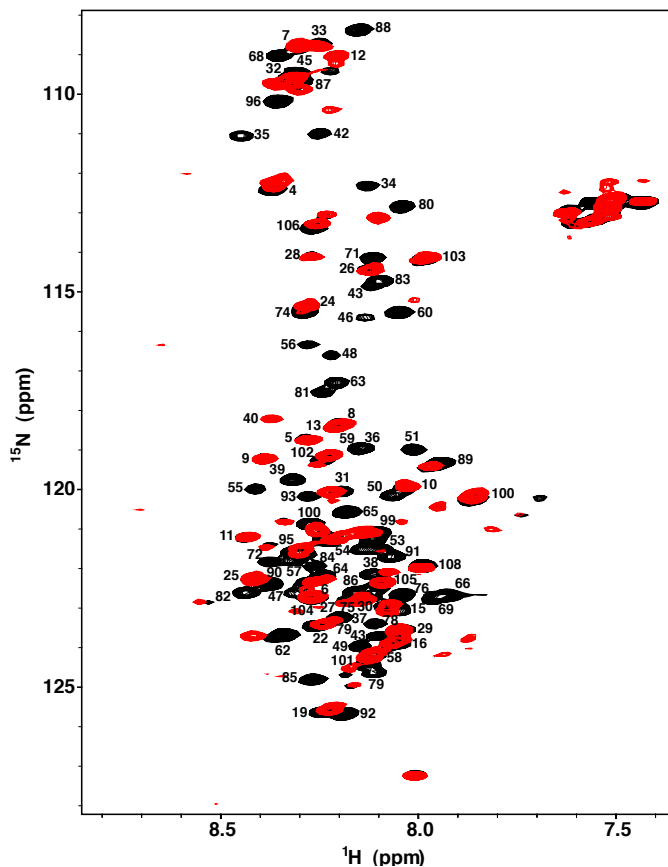


Figure 6.13: Overlay of ^{15}N -HSQC spectra of free Tir108 (black) and Tir108 bound to *CesT* (red) measured at 800 MHz and 293 K. Resonance assignments for free Tir108 are indicated with residue numbers.

upon complex formation. Changing the temperature alone or in combination with a decreased pH did not increase the number of detectable signals even with a TROSY

pulse sequence. Nevertheless, the average amide chemical shift perturbation of Tir108 upon CesT binding was calculated (see 3.5.5) and $\Delta\delta_{HN}$ values of the disappeared residues were assigned to 0.5 (Figure 6.14).

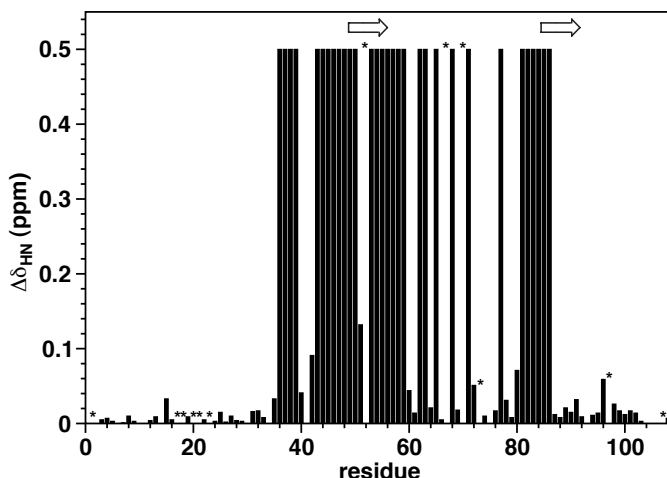


Figure 6.14: Chemical shift changes of Tir108 upon formation of the CesT/Tir108 complex. For vanished peaks $\Delta\delta_{HN}$ values were set to 0.5. Proposed β -motifs for Tir108 are indicated [161] and proline residues are labeled with an asterisk.

This way the region comprising residues 35-86 could be identified as most affected by interaction with CesT and composes most likely the CBD of Tir. In Lilic *et al.* [161], the β -motif was proposed for residues 49-56 and 85-92. The β -motif for residues 49-56 is supported by the changes of chemical shift changes of Tir108 upon CesT/Tir108 complex formation (Figure 6.14). For residues 85-92, which were proposed to form the other β -motif, strong effects could only be found for residues 85 and 86. In this region, residues 80-86 were most strongly affected by CesT-binding (Figure 6.14) and may form the second β -motif.

^{15}N -HSQC spectra of free His₇-Map84 and bound Map84 are depicted in Figure 6.15. Like in the ^{15}N -HSQC of free Tir108 and characteristic for unfolded proteins, amide proton signals of free Map84 are limited to a region between 7.75 and 8.5 ppm. However, unlike for CesT-bound Tir108 (Figures 6.13), only ~12 % of the Map84 signals are missing, four amide proton signals appear lowfield to 8.5 ppm and six signals appear highfield to 7.75 ppm (Figures 6.15). Altogether the dispersion of signals is increased by 2 ppm upon CesT/Map84 complex formation therewith suggesting ter-

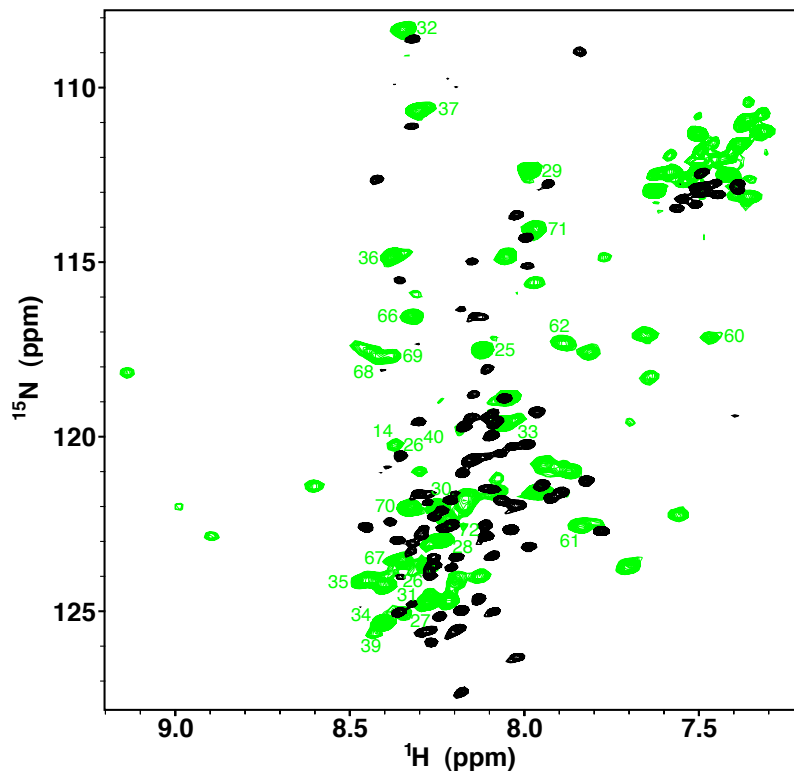


Figure 6.15: Overlay of ^{15}N -HSQC spectra of free His₇-Map84 (black) and Map84 bound to CesT (green) at 303 K. Assigned residues in the bound state are indicated by green numbers.

tiary structure formation for a part of the CBD of Map.

Map84 in complex with CesT was partially assigned. Assigned resonances are indicated in Figure 6.15 by green numbers and the assigned chemical shifts are tabulated in the Appendix Table B.11.

Broad signals of bound Map84 (Figure 6.15) and disappearing signals upon CesT binding especially for Tir108 (Figure 6.13) point at motions in the intermediate exchange regime. This disagrees with the K_d of 18.51 ± 1.97 nM for the CesT/Tir and the K_d of 16.89 ± 6.62 nM for the CesT/Map complex which were determined using an enzyme-linked immunosorbent assay (ELISA)-based approach [30]. As the full length effector proteins were used for the K_d determination, missing interactions between the C-terminal effector regions and the chaperone may explain the decreased affinity. This is suggested by signals outside the CBD of Tir which are affected upon CesT binding (Figure 6.16). However, a decreased affinity due to missing interac-

tions between the C-terminal effector regions and the chaperone is unlikely for the CesT/Map84 complex because its copurification including a gel filtration step is possible. A decreased affinity between CesT and Map84 would probably lead to their separation.

6.3.10 Effect of CesT on effector regions outside the CBD

The full N-terminal cytoplasmic part of Tir, referred to as TirN, and full length Map were prepared to unravel a potential unfolding activity of CesT on effector regions outside the CBD. As already observed for the C-terminally truncated proteins, only TirN is solubly expressed and can be concentrated to ~ 1 mM. All peaks which are present in the ^{15}N -HSQC spectrum of Tir108 overlay well with peaks of TirN. Additional signals in the ^{15}N -HSQC spectrum of TirN are highly dispersed (Figure 6.16). The signals low field to 8.5 ppm show that the C-terminal part of TirN is folded. The dispersion of TirN signals remains unchanged upon CesT binding (Figure 6.16) contradicting an unfolding activity of CesT on residues outside the CBD. This finding agrees with results from isothermal titration calorimetry which show a retained intimin binding activity in the presence of CesT [154]. However, upon CesT binding, there are at least six residues outside the CBD of Tir which disappear or show significantly decreased intensity and two residue signals which are shifted. This suggests that residues of Tir outside the CBD may contribute to the interaction with CesT.

6.3.11 Effects of Tir and Map binding on CesT

In order to explore the effector binding region on the surface of CesT, $^2\text{H}^{15}\text{N}$ -labeled CesT was titrated with Tir108 and His₇-Map84 (see 6.2.4). This approach allows for the identification of the binding surface of CesT specific for Tir and Map based on chemical shift perturbation. Start and end ^{15}N -TROSY-HSQC spectra for the CesT/Tir108 and the CesT/His₇-Map84 titration are shown in Figure 6.17. For both complexes, slow exchange is observed between free and bound CesT and almost all peak positions are changed. In order to reassign the backbone resonances of CesT in

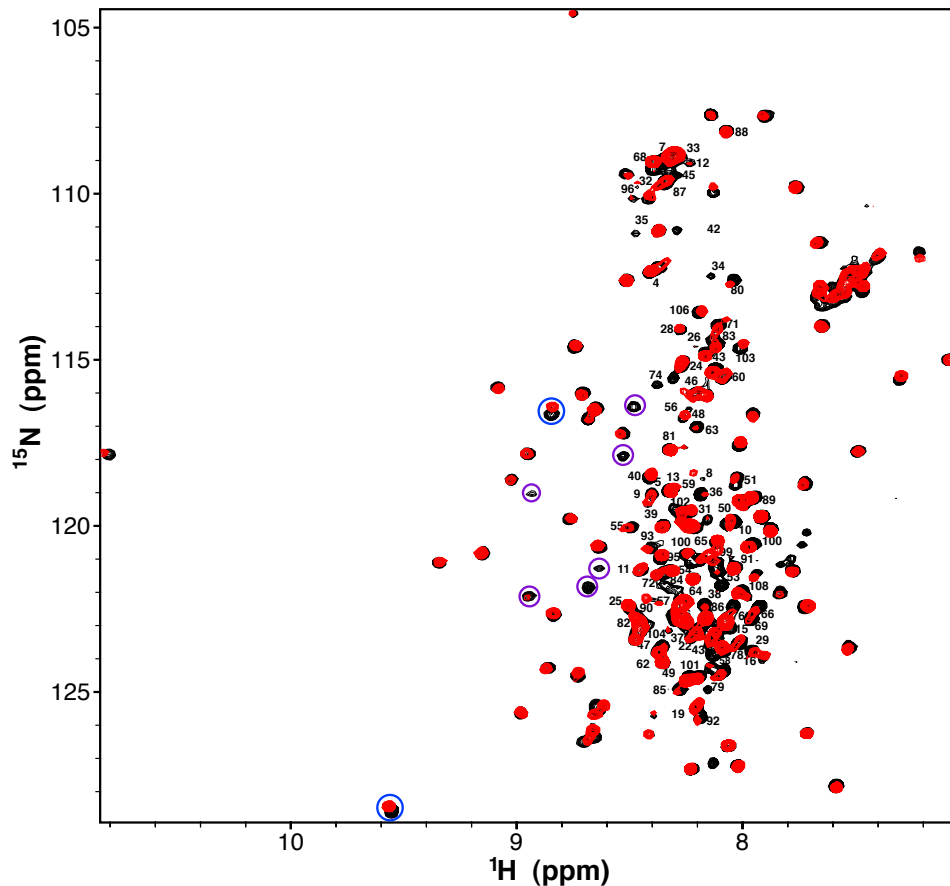


Figure 6.16: Overlay of ^{15}N -HSQC spectra of TirN and TirN bound to CesT at 303 K. Spectra were measured at 800 MHz. Peaks of Tir108 superimpose well with TirN and are labeled with the residue number. Signals which do not belong to Tir108 and are affected upon addition of CesT are indicated by circles: Violet circles mark peaks which show significantly decreased intensity or which vanished and blue circles mark shifted peaks.

complex with Tir108 as well as in complex with Map84, 3D triple resonance spectra were acquired for the two different ^2H , ^{15}N , ^{13}C -CesT/effector complex samples (Appendix Table B.5 and B.4). 93 % of the backbone amide resonances, 96 % of the C_α and C' resonances and 95 % of the C_β resonances could be reassigned for CesT in complex with Map84. To analyse the structural changes of CesT upon Map84 binding, the secondary chemical shifts were calculated and subtracted from the secondary chemical shift of free CesT (Figure 6.18). No secondary structure changes were observed for CesT upon Map binding. Furthermore, averaged amide chemical shift perturbations were determined (see 3.5.5) in order to map the binding surface

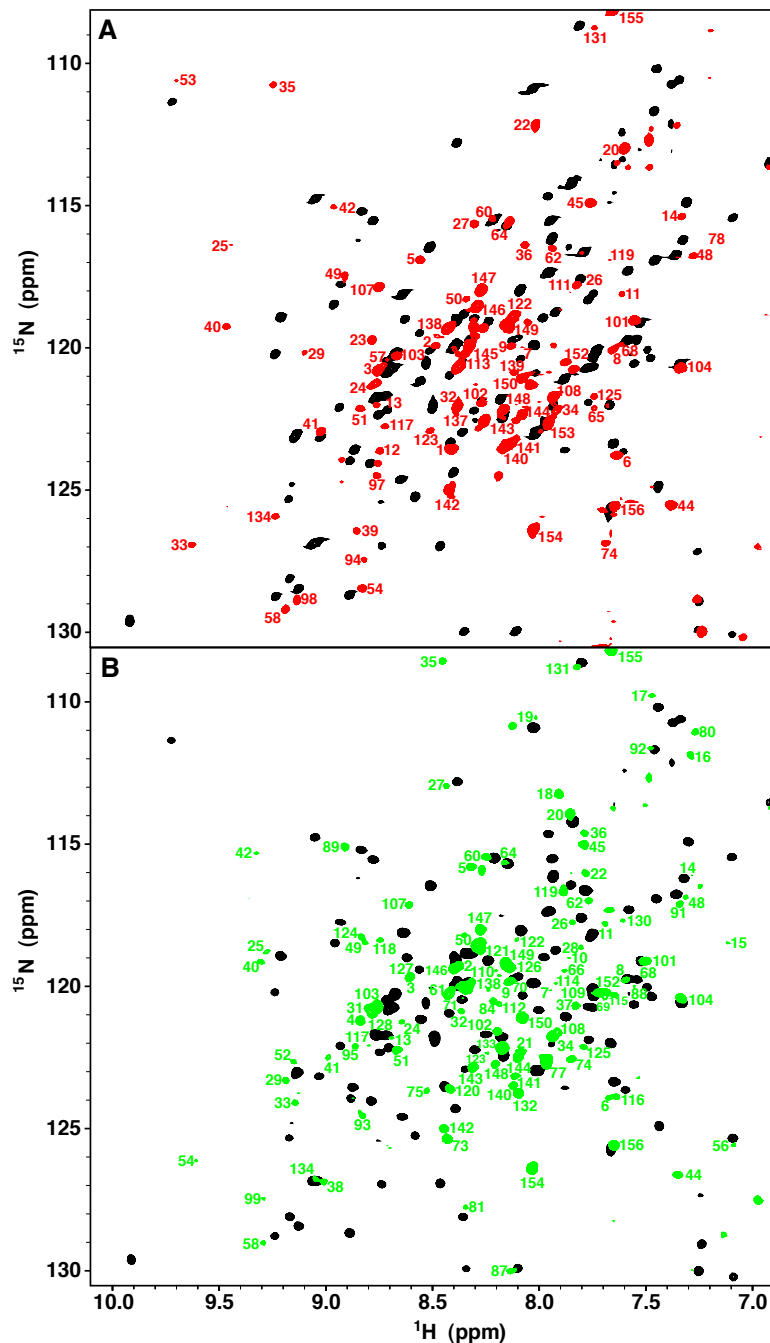


Figure 6.17: Overlay of ^{15}N -TROSY-HSQC spectra of free CesT (black) and effector-bound CesT: (A) Red peaks correspond to CesT in complex with Tir108 and (B) green peaks correspond to CesT in complex with Map84. Assigned peaks are indicated by the sequence number.

for Map84 on CesT (Figure 6.19). Strongly shifted residues in the chemical shift perturbation plot (Figure 6.19), in the structure-based sequence alignment of TTSS chaperone (Figure 6.20) and on the surface of CesT (Figure 6.21) are colored identi-

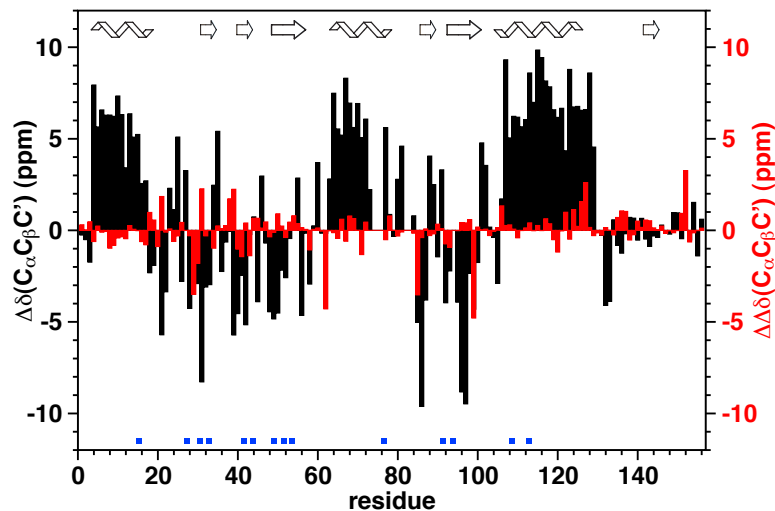


Figure 6.18: Secondary structure of *CesT* bound to Map84. Secondary chemical shifts $\Delta C_{\alpha} - \Delta C_{\beta} + \Delta C'$ of *CesT* bound to Map84 (black) and secondary chemical shift difference for $\Delta C_{\alpha} - \Delta C_{\beta} + \Delta C'$ between free and bound *CesT* (red) plotted against the residue number. Secondary structure regions of the X-ray structure are indicated. Residues proposed to form hydrophobic patches are labeled by blue squares [162].

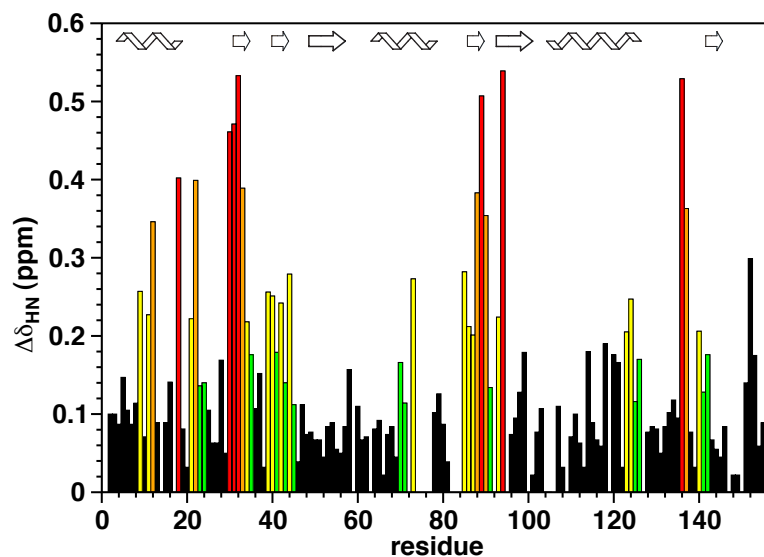


Figure 6.19: Chemical shift changes of *CesT* upon formation of the *CesT*/Map84 complex. Residues shifted >0.4 ppm, >0.3 ppm and >0.2 ppm are colored red, orange, yellow and green, respectively, and residues shifted >0.1 ppm are colored green if an adjacent residue is also shifted >0.1 ppm. Secondary structure regions of the X-ray structure are indicated.

cally with red for shifts >0.4 ppm, orange for shifts >0.3 ppm and yellow for shifts >0.2 ppm. Green indicates residues with shifts >0.1 ppm that have at least one adjacent residue shifted by >0.1 ppm. All strongly shifted peaks (>0.3 ppm) are

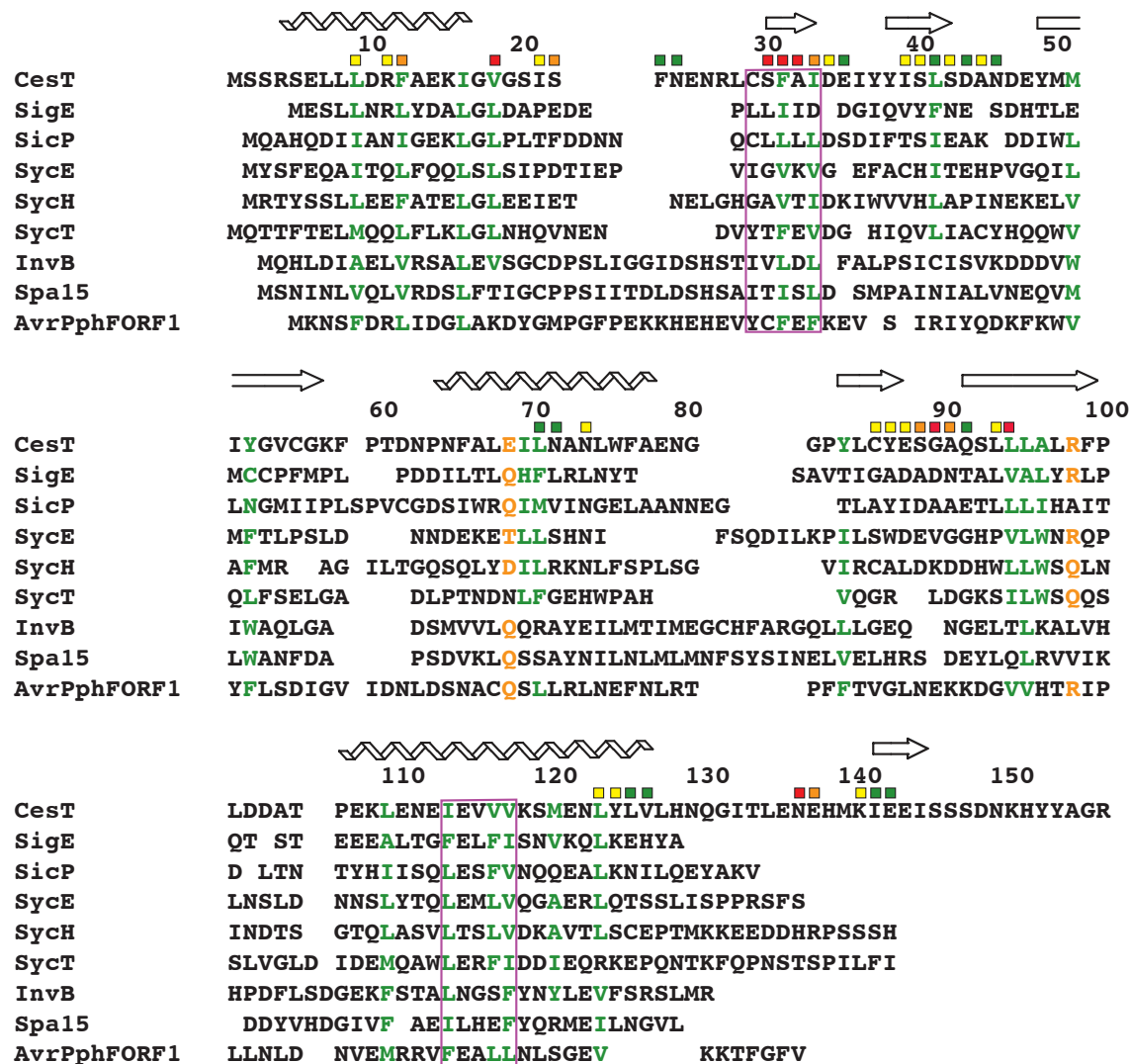


Figure 6.20: Structure-based sequence alignment of nine TTSS chaperones. Conserved residues are highlighted: green for hydrophobic and orange for polar residues. Differently colored squares above the sequence indicate residues as in Figure 6.19. The magenta boxes point at the two regions which are assumed to interact with the β -motif of the TTSS effector [161]. Secondary structure elements of CesT in the X-ray structure are displayed above the sequence.

hydrophobic, thereby indicating the hydrophobic character of the CesT/ Map84 interaction. The only exceptions are the strongly shifted residues Asn136 and Glu137. These shifts suggest a contribution of electrostatic interactions between Map84 and Glu137 of CesT to the complex formation. The ~ 0.2 ppm shifted residues 151-153 were not colored in Figures 6.19, 6.20 and 6.21 because they are not present in the X-ray structure. Altogether, residues exhibiting strong shifts correspond well to the

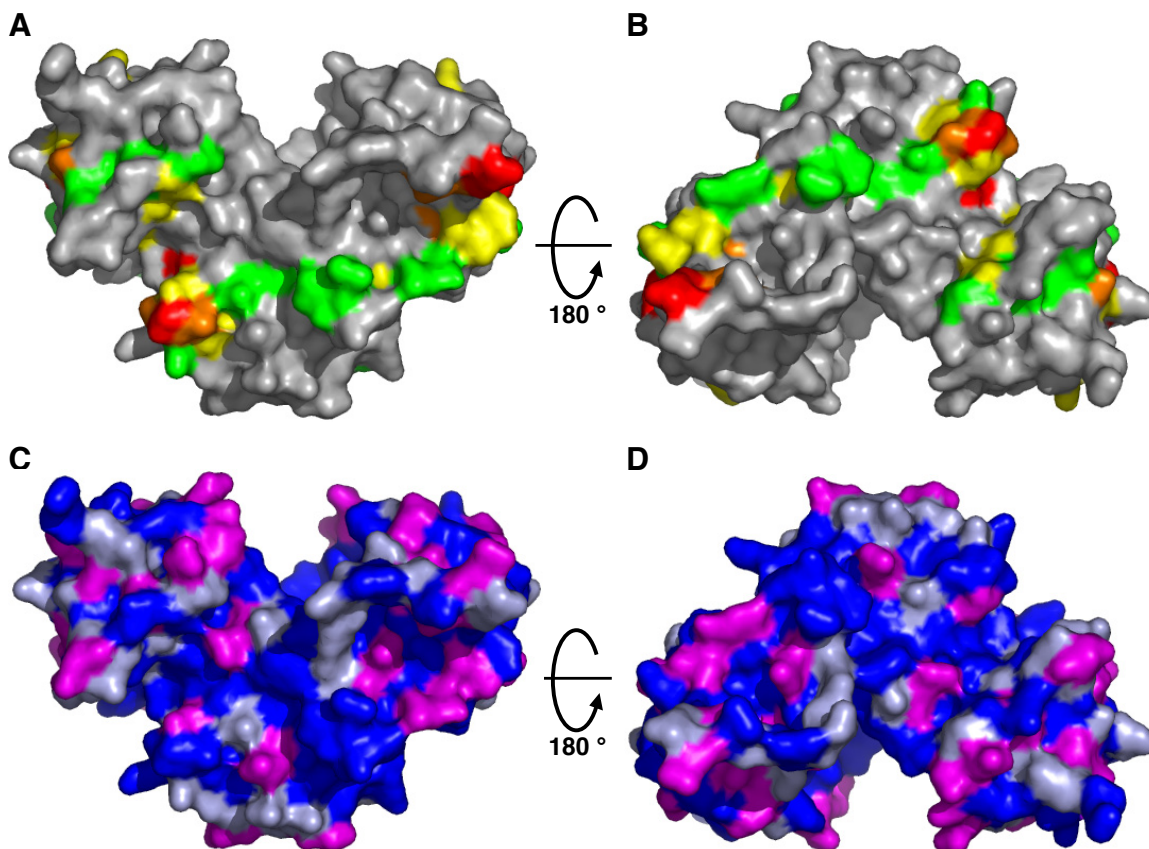


Figure 6.21: Side views of the surface representation of *CesT*. (A,B) Chemical shift changes upon Map84 binding mapped to the surface of *CesT*. Residues colored green, yellow, orange and red are identical to Figure 6.19. (C,D) Residue properties mapped to the surface of *CesT* with charged, polar and hydrophobic residues colored magenta, lightblue and blue. (A,C) are rotated by 180° around the x-axis to generate (B,D).

hydrophobic patches proposed by Birtalan *et al.* [162]. The binding mode seems to be mainly identical to the homologous chaperone/effector complexes with known structure. Large hydrophobic surfaces are buried in these homologous chaperone/effector complexes and *CesT* achieves the effector recognition in a similar way through numerous hydrophobic interactions. The mapping of the shifted residues onto the 3D surface reveals a narrow band of contact around the circumference of the chaperone. This path involves almost exclusively hydrophobic residues, revealing that the interaction is largely hydrophobic in nature. A few additional auxiliary points are also provided by polar or charged interactions and are shifted <0.3 ppm (Figure 6.21). For residues 29-33 and 113-117, which are indicated by magenta boxes in Figure 6.20, the formation of an intermolecular β -sheet with the effector has been proposed [161].

This could be experimentally verified by the chemical shift analysis for residues 29-33, which are most affected upon Map binding. However, the chemical shift perturbation for residues 113-117 is not more significant than for other regions of CesT.

For CesT in complex with Tir108, $\sim 67\%$ of the backbone amide resonances were re-assigned (Appendix Table B.9). A reason for the missing reassignments is the quality of the 3D spectra which is inferior to the 3D spectra recorded for CesT in complex with Map84. An almost complete reassignment may be achieved with a better CesT/Tir108 sample and with additional 3D spectra. Among the $\sim 67\%$ assigned backbone amide resonances, residues with the largest shifts also presented large shifts for CesT in complex with Map84 (Figure 6.22). Therefore, it is likely that the

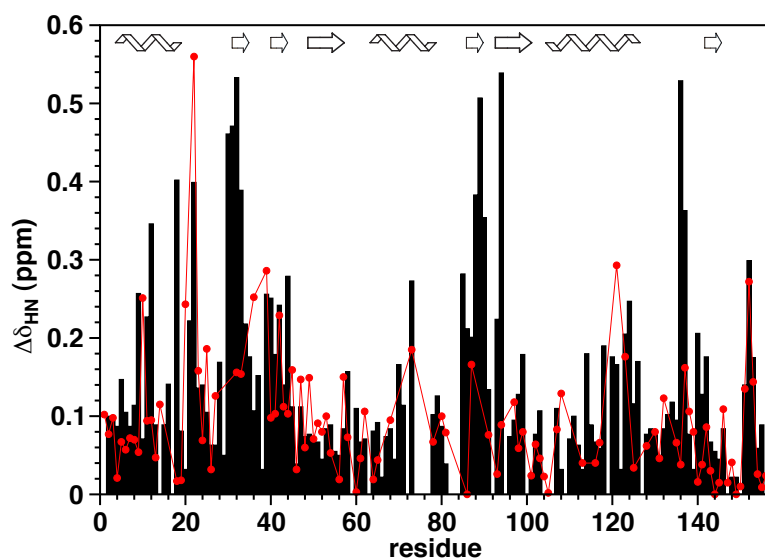


Figure 6.22: Chemical shift changes of CesT upon formation of the CesT/Tir108 complex (red circles). The black bars show the chemical shift changes of CesT upon formation of the CesT/Map84 complex for comparison. Secondary structure regions of the X-ray structure are indicated.

CesT/Map and the CesT/Tir interaction involve the same surface regions of CesT. This agrees with previous ELISA and yeast three hybrid results which demonstrate a competition between Map and Tir for CesT binding [30].

6.3.12 Properties of the CesT/Map84 complex in solution

To detect dynamic regions of the CesT/Map84 complex and differences in dynamics between free and effector-bound CesT, the amplitude of motions were determined via the steady-state heteronuclear ^{15}N - ^1H -NOE (see 3.5.4.3) measured for a uniformly ^{15}N -labeled CesT/Map84 sample (Figure 6.23). ^{15}N - ^1H -NOE values suggest overall

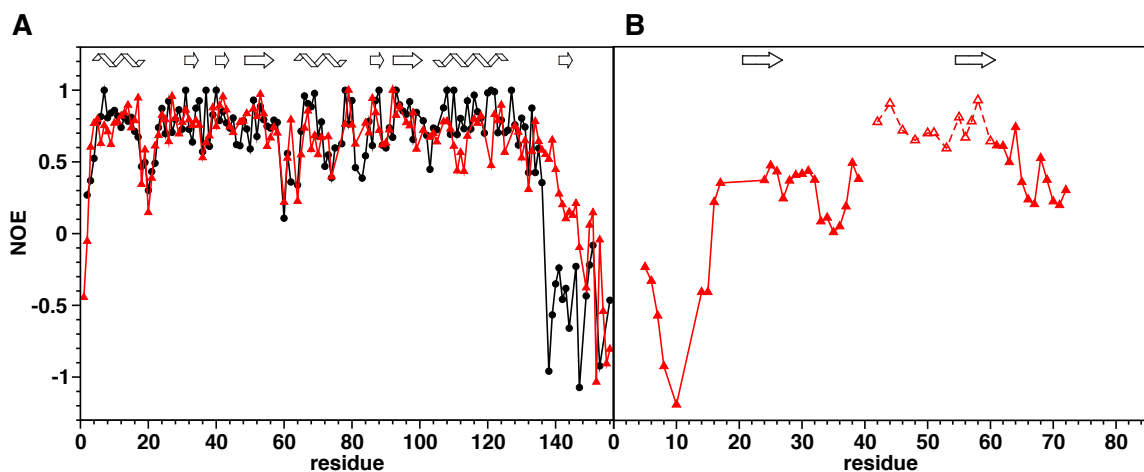


Figure 6.23: ^{15}N - ^1H -NOE values are plotted against the residue number. Values for free CesT are indicated as black circles. Values for CesT (A) and Map84 (B) in complex with each other are shown as red filled triangles and are determined for a coexpressed and copurified ^{15}N -CesT/Map84 complex sample. ^{15}N - ^1H -NOE values without assignment in (B) are indicated by empty triangles and connected by a dashed line. Secondary structure elements of the X-ray structure are indicated for CesT and proposed β -motifs for Map84 [161].

unchanged dynamics of CesT upon complex formation (Figure 6.23A). Only ^{15}N - ^1H -NOE values for residues 138-144 of CesT in complex are increased. This region includes the additional β -strand of the X-ray structure of free CesT which was not found under solution NMR conditions. However, ^{15}N - ^1H -NOE values for residues 138-144 for CesT in the complex are smaller than for the secondary structure regions and secondary chemical shifts do not indicate structure (Figure 6.18). Thus, reduced flexibility but no additional secondary structure formation is indicated for residues 138-144 of CesT in complex with Map84 suggesting a conformation for Map84-bound CesT which is between the conformation of free CesT in the crystal and in solution. A conformational change of C-terminal residues upon complex formation is also indicated by chemical shift changes ~ 0.2 ppm for residues 151-153 (Figure 6.19).

For Map84 in complex only $\sim 50\%$ of the backbone is assigned. The longest stretches with missing assignments are residues 39-58 and 10 C-terminal residues. For 12 non-assigned peaks ^{15}N - ^1H -NOE values between 0.6 and 0.8 were found. Overall ^{15}N - ^1H -NOE values indicate flexibility for most residues of Map84 in complex with CesT (Figure 6.23B). Only for residue 36 and around residue 60 ^{15}N - ^1H -NOE values >0.5 and >0.6 are found, respectively. Therefore the ^{15}N - ^1H -NOE values without assignment were assumed to belong to residues 39-58. Under this assumption residues 40-64 have ^{15}N - ^1H -NOE values above 0.6 and are thus suggested to directly interact with CesT. Due to homology to chaperone/effector complexes with known structures, β -motifs forming an intermolecular β -sheet were proposed for residues 20-27 and 54-61 of Map [161]. ^{15}N - ^1H -NOE values ~ 0.4 are found thereby pointing at restricted flexibility but not complete rigidity. Combined with the fact that the β -motif was inferred from a chaperone/effector-complex structure containing only one β -motif, one β -motif is concluded as sufficient to provide the specific recognition between CesT and its effectors. This β -motif is most likely composed of residues 54-61 of Map. For Tir this β -motif is probably formed by residues 49-56 because the other proposed β -motif is only partly affected by CesT binding (Figure 6.13). Through interactions with the second β -motif of the effector, CesT may confer a hierarchy to the secretion of its seven known effectors via the TTSS of EPEC. To confirm this, at least for one of the CesT/effector complexes, the effector has to be assigned in complex with CesT. For Map84 in complex with CesT the assignment has to be completed but ~ 15 -20 of the Map84 signals are weak or missing completely in the ^{15}N -TROSY-HSQC spectrum of Map84 after refolding the CesT/Map84 complex (Figure 6.15). But in the ^{15}N -TROSY-HSQC spectrum of coexpressed and copurified CesT/Map84 complex (Figure 6.24) there are ~ 10 sharp peaks of Map84 which are weak in the spectrum of the CesT/Map84 complex after refolding (compare Figures 6.24 and 6.15). Furthermore, there are ~ 10 additional peaks for the coexpressed CesT/Map84 complex sample (Figure 6.24) which are not found in the spectra of complex samples after refolding. This finding is in agreement with the fact that the coexpressed sample is stable for more than a year while the refolded sample shows degradation after two

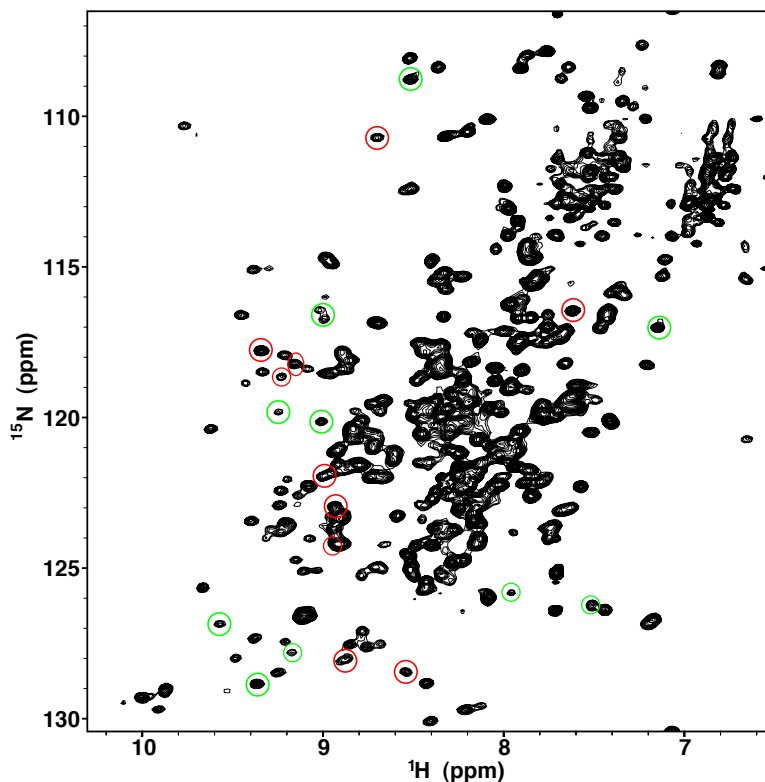


Figure 6.24: ^{15}N -TROSY-HSQC spectrum of uniformly ^{15}N -labeled CesT/Map84 complex. Red circles indicate signals from Map84 which are weak in the HSQC spectrum of the refolded complex (Figure 6.15) and green circles indicate peaks which are absent in both HSQC spectra with one of the components of the refolded complex labeled.

weaks. Thus, TROSY-based triple resonance spectra measured for a coexpressed and copurified, $^2\text{H}(75\%)$, ^{13}C , ^{15}N -labeled CesT/Map84 sample should be most suitable to complete the assignment of Map84 in complex with CesT.

As yet, $\sim 30\%$ of the expected Tir108 signals could not be detected in complex with CesT (Figure 6.13) and a stable coexpressed and copurified CesT/Tir108 complex could not be prepared. Since NMR-titration experiments with TirN and CesT suggested a role of residues outside the N-terminal 100 residues of Tir on CesT-binding (Figure 6.16), a coexpressed and copurified CesT/TirN complex should be prepared and may turn out to be best for NMR experiments and assignment of Tir in complex with CesT.

6.3.13 Conclusions

In this chapter, the role of the chaperone CesT and its chaperone/effector complexes for the TTSS mechanism were investigated using mainly NMR spectroscopy. Structural data about TTSS chaperones and their complexes from solution NMR spectroscopy are not yet available while many X-ray structures have already been solved. In the first part of this chapter, the solution structure of CesT was revealed to be the non-swapped model structure based on experimental and charge-shape predicted RDCs. Results presented in the second part of this chapter demonstrate the relevance of models and conclusions derived for homologous chaperones and chaperone/effector complexes for CesT and its complexes. Firstly, the wrapping of the effector around the chaperone dimer is indicated by a shift of almost all CesT signals in the ^{15}N -HSQC upon effector binding whereas no peaks are doubled due to a broken homodimer symmetry thus hinting at a pseudo-symmetric chaperone/effector complex in which effector residues with similar properties mediate the interaction with both monomers (Figure 6.17). Secondly, an extended, non-globular N-terminal structure of the effector in the chaperone/effector complex is consistent with the small number of signals low field to 8.5 ppm for Map84 in complex with CesT (Figure 6.15). Thirdly, the insolubility of Map in the absence of CesT agrees with the suggested chaperone function of preventing inappropriate interactions. The solubility of Tir without the chaperone agrees with experiments showing a CesT-independent secretion of Tir [29] thereby implying a pre-formed secretion signal for free Tir. Finally, the hydrophobic properties of the residues of CesT with the strongest chemical shift perturbations upon complex formation reveal the CesT/effector complex to be mediated by many hydrophobic interactions (Figure 6.19).

However, the absence of an effector core protected by the chaperone and the additional flexible ~ 20 C-terminal residues point at differences in the effector binding mode with CesT compared to the homologous chaperones and the function for the TTSS mechanism in general. The depicted results agree with and extend the current view of the role of the chaperone in the effector targeting to the TTSS. Figure 6.25

summarizes the overall extended effector targeting model, including additional chaperone/effector interactions proposed based on the results presented in this chapter.

(I) *CesT* recognizes an effector in the bacterial cytoplasm via hydrophobic surface

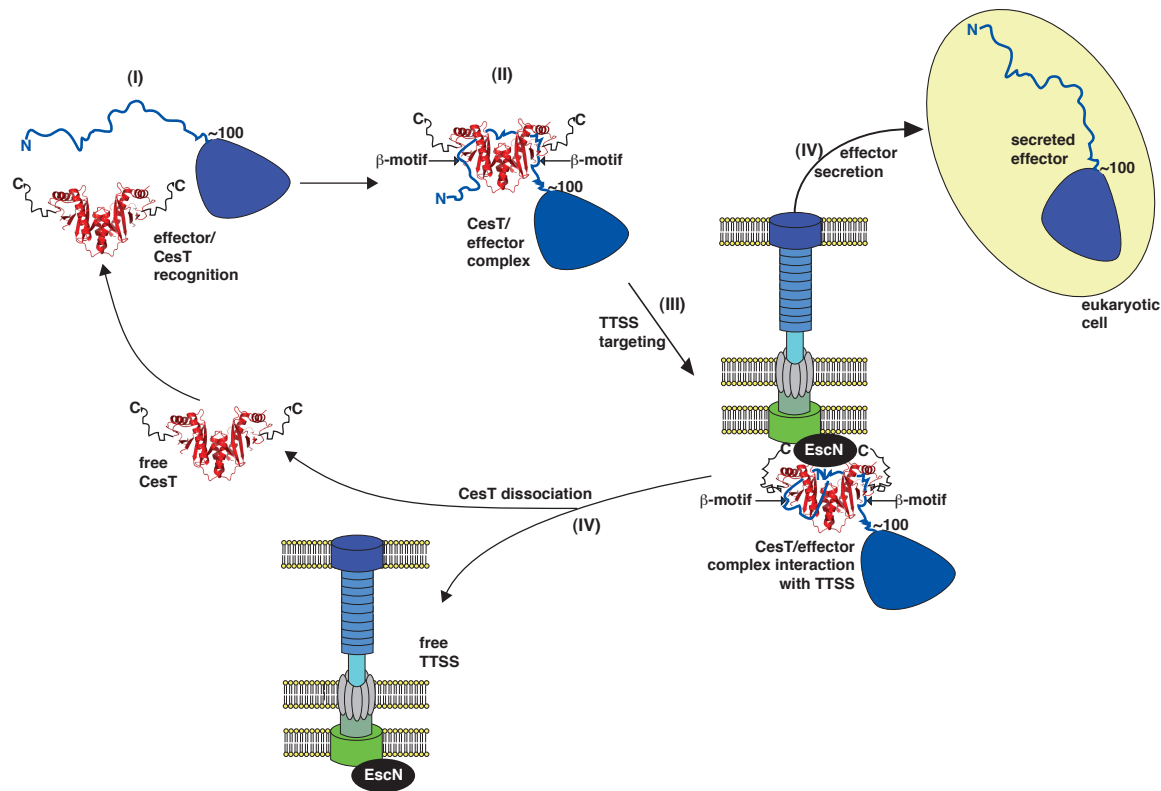


Figure 6.25: Model of effector targeting to the TTSS through *CesT* in pathogenic *E. coli*. *CesT* is shown as a red ribbon diagram with the C-terminus added as a black line and the effector in blue with a globular C-terminal and extended or unfolded ~ 100 N-terminal residues. In the bacterial cytoplasm, the effector is recognized by *CesT* (I) and the *CesT*/effector complex is formed via β -motifs (II). (III) Mediated by the unfolded C-termini of the *CesT* dimer and the unfolded N-terminal ~ 15 N-terminal residues of the effector the *CesT*/effector complex interacts with the ATPase *EscN* of the TTSS. (IV) Conformational change catalyzed by *EscN* leads to *CesT* dissociation and translocation of the effector into the eukaryotic cell.

patches. (II) Upon formation of manifold hydrophobic *CesT*/effector interactions, a part of the N-terminal ~ 100 residues wraps around the chaperone. In the complex, the most important new tertiary structure is formed between two β -motif regions of the tertiary structure lacking effector and identical regions of the *CesT* monomers. (III) The two unfolded C-terminal 20 residues of *CesT* and the ~ 15 N-terminal residues of the effector [170] direct the complex to the TTSS and mediate the association with

the ATPase EscN. (IV) Using the unfolded regions as starting points, EscN catalyses conformational changes which lead to dissociation of CesT and translocation of the unfolded effector into the eukaryotic cell.



7

Summary and future considerations

This thesis makes a contribution to the understanding of two mechanisms of bacterial pathogenicity: A new transcription regulation mechanism of the pathogen *E. faecalis* and the type three protein secretion mechanism of enteropathogenic *E. coli*. In addition a novel method for the improvement of high-resolution structure determination of symmetric oligomers by NMR is presented.

The research projects aimed mainly at the determination of unknown structures of proteins, protein-DNA and protein-protein complexes in order to unravel protein functions and mechanisms. Methods based on recent advances in NMR spectroscopy concerning sample preparation, NMR experiments as well as structure calculation have been applied. In the initial sample preparation stages, plasmids were constructed or mutations introduced to allow for the production and purification of the desired isotopically-labeled or unlabeled proteins. The second stage consisted of the application of well-established and novel NMR experiments. Finally the experimental data were analysed to generate a model for a protein-DNA complex, were utilized to develop a novel method for NMR structure determination of homodimeric proteins and were interpreted to structurally characterize a ~36 kDa homodimeric chaperone and two of its complexes. The major results of Chapter 4-6 are summarized and future experiments are proposed in the following paragraphs.

In 2002, Haas *et al.* showed that CylR2 regulates together with CylR1 the production of the highly toxic cytolysin via an autoinduction quorum sensing mechanism [104].

The results presented in Chapter 4, entitled “**Structure and DNA-binding properties of the cytolysin regulator CylR2**”, are the first structural data about this new quorum sensing mechanism. The structure of CylR2 reveals a rigid dimer containing a helix-turn-helix DNA-binding motif as part of a five helix-bundle that is extended by an antiparallel β -sheet. NMR chemical shift perturbation experiments identify surfaces involved in DNA-binding and are in agreement with a model for the CylR2/DNA complex that attributes binding specificity through a complex network of CylR2/DNA interactions. The results reveal about the new quorum sensing mechanism that CylR2 acts as a repressor of cytolysin transcription by binding to the inverted repeat 1 within the promotor region of the cytolysin operon.

Next steps towards a complete understanding of the molecular basis of cytolysin repression by the CylR1/CylR2 two-component regulatory system rely on the preparation of the putative transmembrane protein CylR1. If its production is not possible by recombinant expression and reconstitution in micelles, the recently developed cell-free expression should be attempted [171]. The proposed binding of CylR1 to CylR2 may involve the N- and C-terminal β -strands which are not found in the structurally homologous proteins SinR and 434 repressor (see 4.3.1). Once the regulation of cytolysin expression is understood, a remedy for hospital-acquired antibiotic-resistant infections caused by *E. faecalis* can be developed.

Oligomerization plays an important role in protein function. The structure determination of homooligomeric proteins by NMR spectroscopy, however, is problematic due to the difficulty of distinguishing inter- and intramonomer contacts. To attack this challenging problem, the protein CylR2 was used to develop a novel method for high-resolution structure determination of symmetric homodimers in solution which is introduced in Chapter 5, “**Using dipolar couplings and intermonomer paramagnetic relaxation enhancement to determine the structure of the CylR2 homodimer**”. Based on the structure of the monomer of CylR2, which was obtained by conventional NOE-based methods, the high-resolution structure of the CylR2 homodimer was determined by rigid-body docking. The rigid-body docking was driven

by a few backbone HN residual dipolar couplings and intermolecular long-range distances derived from relaxation enhancement observed for two paramagnetic spin label positions. The final NMR ensemble of the 15.4 kDa CylR2 homodimer deviates by 1.15 Å from the previously determined X-ray structure. In addition, it is shown that charge-shape predicted residual dipolar couplings can be used to select a structural model from an ensemble of ab initio docking solutions that is closest to the native structure. This method does not require intermolecular NOEs or assignment of side chain resonances, making it particularly useful for the structure determination of large, symmetric homooligomers.

In the next step this novel method should be applied for the solution NMR structure calculation of the 35.4 kDa homodimer CesT which is investigated in Chapter 6. Thereby it could be shown that the method is more generally valuable. Moreover, due to its size CesT is an ideal test protein for the applicability of the method en route to NMR structure determination of large oligomeric proteins up to 80 kDa.

The superordinate aim of Chapter 6, called “**Structural studies of the function of the chaperone CesT in protein secretion via the type three secretion machinery of enteropathogenic *E. coli***”, is the understanding of the type three secretion mechanism. This is a controversially discussed topic with very recent and important findings published only this or last year [146, 153, 161].

In the first half of Chapter 6, the non-swapped homology-based model structure of CesT is revealed as the solution structure solely based on the backbone assignment and backbone residual dipolar couplings. The method described in Chapter 5 can be applied for the NMR structure determination of the homodimer CesT to replace the non-swapped model structure by a solution NMR structure. The second part of the Chapter 6 presents experimental data about the complexes between CesT and its effectors Tir and Map. Hydrophobic patches on the surface of the chaperone CesT which are important for the chaperone/effector interaction were identified via backbone amide chemical shift changes and correspond to important regions of homologue chaperones. Further results are consistent with the function of CesT to keep effectors

in an unfolded, secretion competent state and in targeting its effectors to the type three secretion system.

A CesT/effector structure is required for a detailed understanding of the differences and conserved features between the CesT/effector complexes and the homologue chaperone/effector complexes. Structure determination trials made so far have been focussed on the CesT/Map complex. However, it was not possible to prepare crystals suitable for X-ray diffraction experiments. The first steps towards a solution NMR structure of the CesT/Map84 complex have already been made but more experimental data are needed to generate a model structure. Due to stability problems of the refolded CesT/Map84 sample a uniformly ^2H , ^{15}N , ^{13}C -CesT/Map84 sample should be used to acquire TROSY-based triple resonance spectra to complete the assignment of Map84 in complex with CesT and to derive secondary structure information. Intermolecular distance information will be best determined via site-directed spin-labeling for CesT/Map84 samples with one complex component ^{15}N -labeled and the spin label attached to the other complex component. These intermolecular distances could then be used together with the model structure of CesT and secondary structure information about Map84 to generate a model structure of the CesT/Map84 complex.

Regarding the translocation mechanism, it will be of particular interest to perform NMR titration experiments with the chaperone CesT, the effector Tir or Map and the ATPase EscN. Since it was shown that EscN binds CesT as well as the CesT/Tir complex the effector is probably directed towards the ATPase by CesT and unfolded by the ATPase [153, 166]. Even for the entire trimeric and more than 100 kDa large complex, ^{15}N -HSQC spectra will be accessible and will not suffer from severe signal overlap if the chaperone or the effector are ^{15}N -labeled. Thus, the interaction region and recognition between the ATPase and the chaperone could be revealed. Furthermore, very important information about changes in the fold of the effector and insights into the unfolding mechanism can be gained especially upon addition of ATP to the trimeric complex.

It is the author's hope, that the conclusions reached in this work will be a fruitful

source for further experiments finally leading to a complete understanding of these mechanisms of bacterial pathogenicity allowing for drug design to fight against the related diseases. The author also aimed for a general contribution to approaches attacking protein oligomerization and therewith its biological functions especially in homodimerization and homooligomerization.



References

- [1] T. M. Fuchs, Molecular mechanisms of bacterial pathogenicity, *Naturwissenschaften* **85**: 99–108 (1998).
- [2] J. W. Wilson, M. J. Schurr, C. L. LeBlanc, R. Ramamurthy, K. L. Buchanan, and C. A. Nickerson, Mechanisms of bacterial pathogenicity, *Postgrad Med J* **78**: 216–224 (2002).
- [3] H. Remaut and G. Waksman, Structural biology of bacterial pathogenesis, *Curr Opin Struct Biol* **14**: 161–170 (2004).
- [4] A. Abragam, The principles of nuclear magnetism, Oxford University Press, London (1961).
- [5] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton, Protein NMR spectroscopy: Principles and Practice, Academic Press Inc., San Diego (1996).
- [6] M. H. Lewitt, Spin Dynamics - Basics of Nuclear Magnetic Resonance, Wiley, Chichester (2001).
- [7] K. Wüthrich, NMR of proteins and nucleic acids, Wiley-Interscience, New York (1986).
- [8] F. Bloch, W. W. Hansen, and M. Packard, Nuclear induction., *Physical Review* **69**: 127 (1946).

- [9] E. Purcell, H. C. Torrey, and R. V. Pound, Nuclear induction., *Physical Review* **69**: 37–38 (1946).
- [10] R. R. Ernst, Nobel Lecture. Nuclear magnetic resonance Fourier transform spectroscopy, *Biosci Rep* **12**: 143–187 (1992).
- [11] K. Wüthrich, NMR studies of structure and function of biological macromolecules (Nobel lecture), *Angew Chem Int Ed Engl* **42**: 3340–3363 (2003).
- [12] A. E. Ferentz and G. Wagner, NMR spectroscopy: a multifaceted approach to macromolecular structure, *Q Rev Biophys* **33**: 29–65 (2000).
- [13] K. H. Gardner and L. E. Kay, The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins, *Annu Rev Biophys Biomol Struct* **27**: 357–406 (1998).
- [14] K. Pervushin, R. Riek, G. Wider, and K. Wüthrich, Attenuated T_2 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 USA* **94**: 12366–12371 (1997).
- [15] C. Fernández and G. Wider, TROSY in NMR studies of the structure and function of large biological macromolecules, *Curr Opin Struct Biol* **13**: 570–580 (2003).
- [16] P. Güntert, Structure calculation of biological macromolecules from NMR data, *Q Rev Biophys* **31**: 145–237 (1998).
- [17] G. M. Clore and C. D. Schwieters, Theoretical and computational advances in biomolecular NMR spectroscopy, *Curr Opin Struct Biol* **12**: 146–153 (2002).
- [18] J. J. Gray, High-resolution protein-protein docking, *Curr Opin Struct Biol* **16**: 183–193 (2006).

- [19] R. B. Russell, F. Alber, P. Aloy, F. P. Davis, D. Korkin, M. Pichaud, M. Topf, and A. Sali, A structural perspective on protein-protein interactions, *Curr Opin Struct Biol* **14**: 313–24 (2004).
- [20] 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**: 371–424 (2000).
- [21] A. Bax, G. Kontaxis, and N. Tjandra, Dipolar couplings in macromolecular structure determination, *Meth Enzymol* **339**: 127–174 (2001).
- [22] R. S. Lipsitz and N. Tjandra, Residual dipolar couplings in NMR structure analysis, *Annu Rev Biophys Biomol Struct* **33**: 387–413 (2004).
- [23] M. Blackledge, Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings, *Prog Nucl Magn Reson Spectrosc* **46**: 23–61 (2005).
- [24] D. S. Goodsell and A. J. Olson, Structural symmetry and protein function, *Annu Rev Biophys Biomol Struct* **29**: 105–153 (2000).
- [25] Y. Liu and D. Eisenberg, 3D domain swapping: as domains continue to swap, *Protein Sci* **11**: 1285–1299 (2002).
- [26] M. J. Bennett, M. P. Schlunegger, and D. Eisenberg, 3D domain swapping: a mechanism for oligomer assembly, *Protein Sci* **4**: 2455–2468 (1995).
- [27] M. J. Bennett, M. R. Sawaya, and D. Eisenberg, Deposition diseases and 3D domain swapping, *Structure* **14**: 811–824 (2006).
- [28] J. Grodberg and J. J. Dunn, ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification, *J Bacteriol* **170**: 1245–1253 (1988).
- [29] S. J. Elliott, S. W. Hutcheson, M. S. Dubois, J. L. Mellies, L. A. Wainwright, M. Batchelor, G. Frankel, S. Knutton, and J. B. Kaper, Identification of CesT,

- a chaperone for the type III secretion of Tir in enteropathogenic *Escherichia coli*, *Mol Microbiol* **33**: 1176–1189 (1999).
- [30] E. A. Creasey, R. M. Delahay, A. A. Bishop, R. K. Shaw, B. Kenny, S. Knutton, and G. Frankel, CesT is a bivalent enteropathogenic *Escherichia coli* chaperone required for translocation of both Tir and Map, *Mol Microbiol* **47**: 209–221 (2003).
- [31] R. M. Delahay, R. K. Shaw, S. J. Elliott, J. B. Kaper, S. Knutton, and G. Frankel, Functional analysis of the enteropathogenic *Escherichia coli* type III secretion system chaperone CesT identifies domains that mediate substrate interactions, *Mol Microbiol* **43**: 61–73 (2002).
- [32] E. L. Hartland, M. Batchelor, R. M. Delahay, C. Hale, S. Matthews, G. Dougan, S. Knutton, I. Connerton, and G. Frankel, Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells, *Mol Microbiol* **32**: 151–158 (1999).
- [33] R. A. Laskowski, J. A. Rullmann, M. W. MacArthur, R. Kaptein, and J. M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J Biomol NMR* **8**: 477–486 (1996).
- [34] R. M. Esnouf, Further additions to MolScript version 1.4, including reading and contouring of electron-density maps, *Acta Crystallogr D Biol Crystallogr* **55**: 938–940 (1999).
- [35] M. D. Winn, A. W. Ashton, P. J. Briggs, C. C. Ballard, and P. Patel, Ongoing developments in CCP4 for high-throughput structure determination, *Acta Crystallogr D Biol Crystallogr* **58**: 1929–1936 (2002).
- [36] E. Potterton, P. Briggs, M. Turkenburg, and E. Dodson, A graphical user interface to the CCP4 program suite, *Acta Crystallogr Sect D-Biol Crystallogr* **59**: 1131–1137 (2003).

- [37] S. R. Comeau, D. W. Gatchell, S. Vajda, and C. J. Camacho, ClusPro: An automated docking and discrimination method for the prediction of protein complexes, *Bioinformatics* **20**: 45–50 (2004).
- [38] S. R. Comeau, D. W. Gatchell, S. Vajda, and C. J. Camacho, ClusPro: a fully automated algorithm for protein-protein docking, *Nucleic Acids Res* **32**: W96–9 (2004).
- [39] S. R. Comeau and C. J. Camacho, Predicting oligomeric assemblies: N-mers a primer, *J Struct Biol* **150**: 233–44 (2005).
- [40] P. Güntert, Automated NMR structure calculation with CYANA, *Methods Mol Biol* **278**: 353–378 (2004).
- [41] A. Phillippsen, DINO: Visualizing Structural Biology, <http://www.dino3d.org> (2003).
- [42] Y. S. Jung and M. Zweckstetter, Mars – robust automatic backbone assignment of proteins, *J Biomol NMR* **30**: 11–23 (2004).
- [43] R. Koradi, M. Billeter, and K. Wüthrich, MOLMOL: a program for display and analysis of macromolecular structures, *J Mol Graph* **14**: 51–5, 29–32 (1996).
- [44] P. J. Kraulis, Molscript - a program to produce both detailed and schematic plots of protein structures, *J Appl Crystallogr* **24**: 946–950 (1991).
- [45] M. F. Sanner, A. J. Olson, and J. C. Spehner, Reduced surface: An efficient way to compute molecular surfaces, *Biopolymers* **38**: 305–320 (1996).
- [46] F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, NMR-Pipe: a multidimensional spectral processing system based on UNIX pipes, *J Biomol NMR* **6**: 277–93 (1995).
- [47] B. A. Johnson and R. A. Blevins, NMR View: A computer program for the visualization and analysis of NMR data, *J Biomol NMR* **4**: 603–614 (1994).

- [48] M. Zweckstetter and A. Bax, Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR, *J Am Chem Soc* **122**: 3791–3792 (2000).
- [49] R. A. Laskowski, M. W. Macarthur, D. S. Moss, and J. M. Thornton, Procheck - a program to check the stereochemical quality of protein structures, *J Appl Crystallogr* **26**: 283–291 (1993).
- [50] W. L. DeLano, The PyMOL molecular graphics system, <http://www.pymol.org> (2002).
- [51] E. A. Merritt and D. J. Bacon, Raster3D: Photorealistic molecular graphics, *Macromol Crystallogr, Pt B* **277**: 505–524 (1997).
- [52] G. Cornilescu, F. Delaglio, and A. Bax, Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J Biomol NMR* **13**: 289–302 (1999).
- [53] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, and G. M. Clore, The Xplor-NIH NMR molecular structure determination package, *J Magn Reson* **160**: 65–73 (2003).
- [54] K. Mullis, F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. Erlich, Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction, *Cold Spring Harb Symp Quant Biol* **51 Pt 1**: 263–273 (1986).
- [55] J. Sambrook, F. F. Fritsch, and M. Maniatis, *Molecular Cloning: A laboratory manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (1989).
- [56] U. K. Laemmli, Cleavage of structural proteins during assembly of the head of bacteriophage T4, *Nature* **227**: 680–685 (1970).
- [57] T. Okajima, T. Tanabe, and T. Yasuda, Nonurea sodium dodecyl sulfate-polyacrylamide gel electrophoresis with high-molarity buffers for the separation of proteins and peptides, *Anal Biochem* **211**: 293–300 (1993).

- [58] P. Edman, Method for determination of the amino acid sequence in peptides, *Acta Chem Scand* **4**: 283–293 (1950).
- [59] A. Fontana, P. P. de Laureto, B. Spolaore, E. Frare, P. Picotti, and M. Zambonin, Probing protein structure by limited proteolysis, *Acta Biochim Pol* **51**: 299–321 (2004).
- [60] W. L. Hubbell, A. Gross, R. Langen, and M. A. Lietzow, Recent advances in site-directed spin labeling of proteins, *Curr Opin Struct Biol* **8**: 649–656 (1998).
- [61] H. E. Mash, Y. P. Chin, L. Sigg, R. Hari, and H. Xue, Complexation of copper by zwitterionic aminosulfonic (good) buffers, *Anal Chem* **75**: 671–677 (2003).
- [62] F. X. Schmid, Spectral methods of characterizing protein conformation and conformational changes., in T. E. Creighton, editor, *Protein structure: A practical approach.*, IRL Press Oxford, Oxford (1989).
- [63] A. Bax and S. Grzesiek, Methodological advances in protein NMR, *Accounts Chem Res* **26**: 131–138 (1993).
- [64] M. Salzmann, K. Pervushin, G. Wider, H. Senn, and K. Wüthrich, TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins, *Proc Natl Acad Sci U S A* **95**: 13585–13590 (1998).
- [65] M. Salzmann, G. Wider, K. Pervushin, H. Senn, and K. Wüthrich, TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins, *J Am Chem Soc* **121**: 844–848 (1999).
- [66] A. Eletsky, A. Kienhofer, and K. Pervushin, TROSY NMR with partially deuterated proteins, *J Biomol NMR* **20**: 177–180 (2001).
- [67] T. Yamazaki, W. Lee, C. H. Arrowsmith, D. R. Muhandiram, and L. E. Kay, A suite of triple-resonance NMR experiments for the backbone assignment of N-15, C-13, H-2 labeled proteins with high-sensitivity, *J Am Chem Soc* **116**: 11655–11666 (1994).

- [68] W. Peti, C. Griesinger, and W. Bermel, Adiabatic TOCSY for C,C and H,H J-transfer, *J Biomol NMR* **18**: 199–205 (2000).
- [69] M. Zweckstetter and T. A. Holak, Robust refocusing of C-13 magnetization in multidimensional NMR experiments by adiabatic fast passage pulses, *J Biomol NMR* **15**: 331–334 (1999).
- [70] T. Yamazaki, J. D. Forman-Kay, and L. E. Kay, 2-dimensional NMR experiments for correlating $^{13}\text{C}\beta$ and $^1\text{H}\delta/\epsilon$ chemical shifts of aromatic residues in ^{13}C -labeled proteins via scalar couplings, *J Am Chem Soc* **115**: 11054–11055 (1993).
- [71] S. Schwarzingler, G. J. Kroon, T. R. Foss, J. Chung, P. E. Wright, and H. J. Dyson, Sequence-dependent correction of random coil NMR chemical shifts, *J Am Chem Soc* **123**: 2970–2978 (2001).
- [72] D. S. Wishart and B. D. Sykes, Chemical shifts as a tool for structure determination, *Methods Enzymol* **239**: 363–392 (1994).
- [73] W. J. Metzler, K. L. Constantine, M. S. Friedrichs, A. J. Bell, E. G. Ernst, T. B. Lavoie, and L. Mueller, Characterization of the three-dimensional solution structure of human profilin: ^1H , ^{13}C , and ^{15}N NMR assignments and global folding pattern, *Biochemistry* **32**: 13818–13829 (1993).
- [74] M. R. Hansen, L. Mueller, and A. Pardi, Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions, *Nat Struct Biol* **5**: 1065–1074 (1998).
- [75] M. Zweckstetter and A. Bax, Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage, *J Biomol NMR* **20**: 365–377 (2001).
- [76] G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax, Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase, *J Am Chem Soc* **120**: 6836–6837 (1998).

- [77] M. Ottiger, F. Delaglio, and A. Bax, Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra, *J Magn Reson* **131**: 373–378 (1998).
- [78] N. Tjandra and A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, *Science* **278**: 1111–1114 (1997).
- [79] V. Vijayan and M. Zweckstetter, Simultaneous measurement of protein one-bond residual dipolar couplings without increased resonance overlap, *J Magn Reson* **174**: 245–253 (2005).
- [80] J. J. Chou, F. Delaglio, and A. Bax, Measurement of one-bond ^{15}N - $^{13}\text{C}'$ dipolar couplings in medium sized proteins, *J Biomol NMR* **18**: 101–105 (2000).
- [81] C. P. Jaroniec, T. S. Ulmer, and A. Bax, Quantitative J correlation methods for the accurate measurement of $^{13}\text{C}'$ - $^{13}\text{C}\alpha$ dipolar couplings in proteins, *J Biomol NMR* **30**: 181–194 (2004).
- [82] M. Zweckstetter, G. Hummer, and A. Bax, Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases, *Biophys J* **86**: 3444–60 (2004).
- [83] V. Sklenar and A. Bax, Two-dimensional heteronuclear chemical-shift correlation of proteins at natural abundance ^{15}N and ^{13}C levels, *J Magn Reson* **71**: 379–383 (1987).
- [84] J. W. Peng and G. Wagner, Investigation of protein motions via relaxation measurements, *Methods Enzymol* **239**: 563–596 (1994).
- [85] D. J. Craik and J. A. Wilce, Studies of protein-ligand interactions by NMR, *Methods Mol Biol* **60**: 195–232 (1997).
- [86] L. Fielding, NMR methods for the determination of protein-ligand dissociation constants, *Curr Top Med Chem* **3**: 39–53 (2003).

- [87] 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**: 5355–5365 (2000).
- [88] P. A. Kosen, Spin labeling of proteins, *Methods Enzymol* **177**: 86–121 (1989).
- [89] G. M. Clore, Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization, *Proc Natl Acad Sci USA* **97**: 9021–9025 (2000).
- [90] R. M. A. Knegtel, R. Boelens, and R. Kaptein, Monte-Carlo docking of protein-DNA complexes - incorporation of DNA flexibility and experimental data, *Protein Eng* **7**: 761–767 (1994).
- [91] R. M. A. Knegtel, J. Antoon, C. Rullmann, R. Boelens, and R. Kaptein, Monty - a Monte-Carlo approach to protein-DNA recognition, *J Mol Biol* **235**: 318–324 (1994).
- [92] B. D. Jett, M. M. Huycke, and M. S. Gilmore, Virulence of *Enterococci*, *Clin Microbiol Rev* **7**: 462–78 (1994).
- [93] B. E. Murray, The life and times of the *Enterococcus*, *Clin Microbiol Rev* **3**: 46–65 (1990).
- [94] M. S. Gilmore, R. A. Segarra, and M. C. Booth, An Hlyb-type function is required for expression of the *Enterococcus faecalis* hemolysin/bacteriocin, *Infect Immun* **58**: 3914–3923 (1990).
- [95] Y. Ike, H. Hashimoto, and D. B. Clewell, Hemolysin of *Streptococcus faecalis* subspecies zymogenes contributes to virulence in mice, *Infect Immun* **45**: 528–530 (1984).

- [96] K. V. Singh, X. Qin, G. M. Weinstock, and B. E. Murray, Generation and testing of mutants of *Enterococcus faecalis* in a mouse peritonitis model, *J Infect Dis* **178**: 1416–1420 (1998).
- [97] B. D. Shepard and M. S. Gilmore, Antibiotic-resistant enterococci: the mechanisms and dynamics of drug introduction and resistance, *Microbes Infect* **4**: 215–224 (2002).
- [98] L. S. Havarstein, D. B. Diep, and I. F. Nes, A family of bacteriocin ABC transporters carry out proteolytic processing of their substrates concomitant with export, *Mol Microbiol* **16**: 229–240 (1995).
- [99] M. C. Booth, C. P. Bogie, H. G. Sahl, R. J. Siezen, K. L. Hatter, and M. S. Gilmore, Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic, *Mol Microbiol* **21**: 1175–84 (1996).
- [100] P. S. Coburn and M. S. Gilmore, The *Enterococcus faecalis* cytolysin: a novel toxin active against eukaryotic and prokaryotic cells, *Cell Microbiol* **5**: 661–669 (2003).
- [101] B. D. Jett and M. S. Gilmore, The growth-inhibitory effect of the *Enterococcus faecalis* bacteriocin encoded by pAD1 extends to the oral streptococci, *J Dent Res* **69**: 1640–5 (1990).
- [102] W. Haas and M. S. Gilmore, Molecular nature of a novel bacterial toxin: the cytolysin of *Enterococcus faecalis*, *Med Microbiol Immunol (Berl)* **187**: 183–90 (1999).
- [103] M. S. Gilmore, R. A. Segarra, M. C. Booth, C. P. Bogie, L. R. Hall, and D. B. Clewell, Genetic structure of the *Enterococcus faecalis* plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants, *J Bacteriol* **176**: 7335–44 (1994).

- [104] W. Haas, B. D. Shepard, and M. S. Gilmore, Two-component regulator of *Enterococcus faecalis* cytolysin responds to quorum-sensing autoinduction, *Nature* **415**: 84–87 (2002).
- [105] P. S. Coburn, L. E. Hancock, M. C. Booth, and M. S. Gilmore, A novel means of self-protection, unrelated to toxin activation, confers immunity to the bactericidal effects of the *Enterococcus faecalis* cytolysin, *Infect Immun* **67**: 3339–47 (1999).
- [106] L. Keller and M. G. Surette, Communication in bacteria: an ecological and evolutionary perspective, *Nat Rev Microbiol* **4**: 249–258 (2006).
- [107] J. A. Hoch, Two-component and phosphorelay signal transduction, *Curr Opin Microbiol* **3**: 165–70 (2000).
- [108] S. Rumpel, A. Razeto, C. M. Pillar, V. Vijayan, A. Taylor, K. Giller, M. S. Gilmore, S. Becker, and M. Zweckstetter, Structure and DNA-binding properties of the cytolysin regulator CylR2 from *Enterococcus faecalis*, *EMBO J* **23**: 3632–42 (2004).
- [109] F. A. Mulder, D. Schipper, R. Bott, and R. Boelens, Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilis*ins, *J Mol Biol* **292**: 111–123 (1999).
- [110] L. J. W. Shimon and S. C. Harrison, The phage 434 O(R)2/R1-69 complex at 2.5 Å resolution, *J Mol Biol* **232**: 826–838 (1993).
- [111] L. J. Beamer and C. O. Pabo, Refined 1.8 Å crystal structure of the lambda repressor-operator complex, *J Mol Biol* **227**: 177–196 (1992).
- [112] S. C. Harrison and A. K. Aggarwal, DNA recognition by proteins with the Helix-Turn-Helix motif, *Annu Rev Biochem* **59**: 933–969 (1990).
- [113] K. Pervushin, M. Billeter, G. Siegal, and K. Wüthrich, Structural role of a buried salt bridge in the 434 repressor DNA-binding domain, *J Mol Biol* **264**: 1002–1012 (1996).

- [114] A. Mondragon, S. Subbiah, S. C. Almo, M. Drottar, and S. C. Harrison, Structure of the amino-terminal domain of phage-434 repressor at 2.0 Å resolution, *J Mol Biol* **205**: 189–200 (1989).
- [115] L. Holm and C. Sander, Protein structure comparison by alignment of distance matrices, *J Mol Biol* **233**: 123–138 (1993).
- [116] R. J. Lewis, J. A. Brannigan, W. A. Offen, I. Smith, and A. J. Wilkinson, An evolutionary link between sporulation and prophage induction in the structure of a repressor:anti-repressor complex, *J Mol Biol* **283**: 907–912 (1998).
- [117] V. M. A. Ducros, R. J. Lewis, C. S. Verma, E. J. Dodson, G. Leonard, J. P. Turkenburg, G. N. Murshudov, A. J. Wilkinson, and J. A. Brannigan, Crystal structure of GerE, the ultimate transcriptional regulator of spore formation in *Bacillus subtilis*, *Journal of Molecular Biology* **306**: 759–771 (2001).
- [118] A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne, and S. C. Harrison, Recognition of a DNA operator by the repressor of phage 434 - a view at high-resolution, *Science* **242**: 899–907 (1988).
- [119] A. C. Bell and G. B. Koudelka, Operator sequence context influences amino acid-base-pair interactions in 434 repressor-operator complexes, *J Mol Biol* **234**: 542–553 (1993).
- [120] A. Simoncsits, J. Chen, P. Percipalle, S. Wang, I. Toro, and S. Pongor, Single-chain repressors containing engineered DNA-binding domains of the phage 434 repressor recognize symmetric or asymmetric DNA operators, *J Mol Biol* **267**: 118–131 (1997).
- [121] C. Wolberger, Y. Dong, M. Ptashne, and S. C. Harrison, Structure of a phage 434 Cro/DNA complex, *Nature* **335**: 789–795 (1988).
- [122] E. L. Sonnhammer, G. von Heijne, and A. Krogh, A hidden Markov model for predicting transmembrane helices in protein sequences, *Proc Int Conf Intell Syst Mol Biol* **6**: 175–182 (1998).

- [123] A. L. Breeze, Isotope-filtered NMR methods for the study of biomolecular structure and interactions, *Prog Nucl Magn Reson Spectros* **36**: 323–372 (2000).
- [124] K. J. Walters, H. Matsuo, and G. Wagner, A simple method to distinguish intermonomer nuclear overhauser effects in homodimeric proteins with C-2 symmetry, *J Am Chem Soc* **119**: 5958–5959 (1997).
- [125] F. K. Junius, S. I. O’Donoghue, M. Nilges, A. S. Weiss, and G. F. King, High resolution NMR solution structure of the leucine zipper domain of the c-Jun homodimer, *J Biol Chem* **271**: 13663–13667 (1996).
- [126] J. Voss, L. Salwinski, H. R. Kaback, and W. L. Hubbell, A method for distance determination in proteins using a designed metal ion binding site and site-directed spin labeling: evaluation with T4 lysozyme, *Proc Natl Acad Sci U S A* **92**: 12295–12299 (1995).
- [127] C. Ma and S. J. Opella, Lanthanide ions bind specifically to an added "EF-hand" and orient a membrane protein in micelles for solution NMR spectroscopy, *J Magn Reson* **146**: 381–384 (2000).
- [128] V. Gaponenko, J. W. Howarth, L. Columbus, G. Gasmi-Seabrook, J. Yuan, W. L. Hubbell, and P. R. Rosevear, Protein global fold determination using site-directed spin and isotope labeling, *Protein Sci* **9**: 302–309 (2000).
- [129] A. Leonov, B. Voigt, F. Rodriguez-Castaneda, P. Sakhaii, and C. Griesinger, Convenient synthesis of multifunctional EDTA-based chiral metal chelates substituted with an S-mesylcysteine, *Chemistry* **11**: 3342–3348 (2005).
- [130] L. W. Donaldson, N. R. Skrynnikov, W. Y. Choy, D. R. Muhandiram, B. Sarkar, J. D. Forman-Kay, and L. E. Kay, Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy, *J Am Chem Soc* **123**: 9843–9847 (2001).
- [131] J. Feeney, B. Birdsall, A. F. Bradbury, R. R. Biekofsky, and P. M. Bayley, Calmodulin tagging provides a general method of using lanthanide induced

- magnetic field orientation to observe residual dipolar couplings in proteins in solution, *J Biomol NMR* **21**: 41–48 (2001).
- [132] H. J. Dyson and P. E. Wright, Equilibrium NMR studies of unfolded and partially folded proteins, *Nat Struct Biol* **5 Suppl**: 499–503 (1998).
- [133] J. Iwahara and G. M. Clore, Detecting transient intermediates in macromolecular binding by paramagnetic NMR, *Nature* **440**: 1227–1230 (2006).
- [134] V. Gaponenko, A. S. Altieri, J. Li, and R. A. Byrd, Breaking symmetry in the structure determination of (large) symmetric protein dimers, *J Biomol NMR* **24**: 143–148 (2002).
- [135] M. Nilges, A calculation strategy for the structure determination of symmetric dimers by ^1H NMR, *Proteins* **17**: 297–309 (1993).
- [136] C. A. Bewley and G. M. Clore, Determination of the relative orientation of the two halves of the domain-swapped dimer of cyanovirin-N in solution using dipolar couplings and rigid body minimization, *J Am Chem Soc* **122**: 6009–6016 (2000).
- [137] P. H. Hirel, J. M. Schmitter, P. Dessen, G. Fayat, and S. Blanquet, Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side chain length of the penultimate amino acid, *Proc Nat Acad Sci USA* **86**: 8247–8251 (1989).
- [138] A. Dobrodumov and A. M. Gronenborn, Filtering and selection of structural models: combining docking and NMR, *Proteins* **53**: 18–32 (2003).
- [139] P. B. Card, P. J. Erbel, and K. H. Gardner, Structural basis of ARNT PAS-B dimerization: use of a common β -sheet interface for hetero- and homodimerization, *J Mol Biol* **353**: 664–77 (2005).
- [140] M. Zweckstetter, J. R. Schnell, and J. J. Chou, Determination of the packing mode of the coiled-coil domain of cGMP-dependent protein kinase I α in

- solution using charge-predicted dipolar couplings, *J Am Chem Soc* **127**: 11918–9 (2005).
- [141] J. Garmendia, G. Frankel, and V. F. Crepin, Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation, *Infect Immun* **73**: 2573–2585 (2005).
- [142] J. P. Nataro and J. B. Kaper, Diarrheagenic *Escherichia coli*, *Clin Microbiol Rev* **11**: 142–201 (1998).
- [143] P. Papatheodorou, G. Domanska, M. Oxle, J. Mathieu, O. Selchow, B. Kenny, and J. Rassow, The enteropathogenic *Escherichia coli* (EPEC) Map effector is imported into the mitochondrial matrix by the TOM/Hsp70 system and alters organelle morphology, *Cell Microbiol* **8**: 677–689 (2006).
- [144] B. Kenny and M. Jepson, Targeting of an enteropathogenic *Escherichia coli* (EPEC) effector protein to host mitochondria, *Cell Microbiol* **2**: 579–590 (2000).
- [145] B. Kenny, Mechanism of action of EPEC Type III effector molecules, *Int J Med Microbiol* **291**: 469–477 (2002).
- [146] N. M. Alto, F. Shao, C. S. Lazar, R. L. Brost, G. Chua, S. Mattoo, S. A. McMahon, P. Ghosh, T. R. Hughes, C. Boone, and J. E. Dixon, Identification of a bacterial type III effector family with G protein mimicry functions, *Cell* **124**: 133–145 (2006).
- [147] P. Wattiau, S. Woestyn, and G. R. Cornelis, Customized secretion chaperones in pathogenic bacteria, *Mol Microbiol* **20**: 255–262 (1996).
- [148] C. Parsot, C. Hamiaux, and A. L. Page, The various and varying roles of specific chaperones in type III secretion systems, *Curr Opin Microbiol* **6**: 7–14 (2003).
- [149] R. M. Macnab, The bacterial flagellum: reversible rotary propellor and type III export apparatus, *J Bacteriol* **181**: 7149–7153 (1999).

- [150] R. R. Isberg and G. Dumenil, Delivering dangerous cargoes, *Nat Struct Biol* **8**: 1006–1008 (2001).
- [151] A. L. Page and C. Parsot, Chaperones of the type III secretion pathway: jacks of all trades, *Mol Microbiol* **46**: 1–11 (2002).
- [152] J. A. Crawford and J. B. Kaper, The N-terminus of enteropathogenic *Escherichia coli* (EPEC) Tir mediates transport across bacterial and eukaryotic cell membranes, *Mol Microbiol* **46**: 855–868 (2002).
- [153] N. A. Thomas, W. Deng, J. L. Puente, E. A. Frey, C. K. Yip, N. C. Strynadka, and B. B. Finlay, CesT is a multi-effector chaperone and recruitment factor required for the efficient type III secretion of both LEE- and non-LEE-encoded effectors of enteropathogenic *Escherichia coli*, *Mol Microbiol* **57**: 1762–79 (2005).
- [154] Y. Luo, M. G. Bertero, E. A. Frey, R. A. Pfuetzner, M. R. Wenk, L. Creagh, S. L. Marcus, D. Lim, F. Sicheri, C. Kay, C. Haynes, B. B. Finlay, and N. C. J. Strynadka, Structural and biochemical characterization of the type III secretion chaperones CesT and SigE, *Nat Struct Biol* **8**: 1031–1036 (2001).
- [155] S. Birtalan and P. Ghosh, Structure of the *Yersinia* type III secretory system chaperone SycE, *Nat Struct Biol* **8**: 974–978 (2001).
- [156] J. Phan, J. E. Tropea, and D. S. Waugh, Structure of the *Yersinia pestis* type III secretion chaperone SycH in complex with a stable fragment of YscM2, *Acta Crystallogr D Biol Crystallogr* **60**: 1591–1599 (2004).
- [157] C. R. Buttner, G. R. Cornelis, D. W. Heinz, and H. H. Niemann, Crystal structure of *Yersinia enterocolitica* type III secretion chaperone SycT, *Protein Sci* **14**: 1993–2002 (2005).
- [158] M. Locher, B. Lehnert, K. Krauss, J. Heesemann, M. Groll, and G. Wilharm, Crystal structure of the *Yersinia enterocolitica* type III secretion chaperone SycT, *J Biol Chem* **280**: 31149–31155 (2005).

- [159] C. E. Stebbins and J. E. Galan, Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion, *Nature* **414**: 77–81 (2001).
- [160] A. van Eerde, C. Hamiaux, J. Perez, C. Parsot, and B. W. Dijkstra, Structure of Spa15, a type III secretion chaperone from *Shigella flexneri* with broad specificity, *EMBO Reports* **5**: 477–483 (2004).
- [161] M. Lilic, M. Vujanac, and C. E. Stebbins, A common structural motif in the binding of virulence factors to bacterial secretion chaperones, *Mol Cell* **21**: 653–664 (2006).
- [162] S. C. Birtalan, R. M. Phillips, and P. Ghosh, Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens, *Mol Cell* **9**: 971–980 (2002).
- [163] F. D. Schubot, M. W. Jackson, K. J. Penrose, S. Cherry, J. E. Tropea, G. V. Plano, and D. S. Waugh, Three-dimensional structure of a macromolecular assembly that regulates type III secretion in *Yersinia pestis*, *J Mol Biol* **346**: 1147–61 (2005).
- [164] C. E. Stebbins and J. E. Galan, Priming virulence factors for delivery into the host, *Nat Rev Mol Cell Biol* **4**: 738–743 (2003).
- [165] Y. Akeda and J. E. Galan, Chaperone release and unfolding of substrates in type III secretion, *Nature* **437**: 911–5 (2005).
- [166] A. Gauthier and B. B. Finlay, Translocated intimin receptor and its chaperone interact with ATPase of the type III secretion apparatus of enteropathogenic *Escherichia coli*, *J Bacteriol* **185**: 6747–6755 (2003).
- [167] A. Abe, M. de Grado, R. A. Pfuetzner, C. Sanchez-SanMartin, R. DeVinney, J. L. Puente, N. C. J. Strynadka, and B. B. Finlay, Enteropathogenic *Escherichia coli* translocated intimin receptor, Tir, requires a specific chaperone for stable secretion, *Mol Microbiol* **33**: 1162–1175 (1999).

- [168] V. Receveur-Brechot, J. M. Bourhis, V. N. Uversky, B. Canard, and S. Longhi, Assessing protein disorder and induced folding, *Proteins* **62**: 24–45 (2006).
- [169] S. C. Panchal, N. S. Bhavesh, and R. V. Hosur, Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and ^{15}N sequential correlations in (^{13}C , ^{15}N) labeled proteins: application to unfolded proteins, *J Biomol NMR* **20**: 135–47 (2001).
- [170] P. Ghosh, Process of protein transport by the type III secretion system, *Microbiol Mol Biol Rev* **68**: 771–95 (2004).
- [171] C. Klammt, F. Löhr, B. Schafer, W. Haase, V. Dötsch, H. Rüterjans, C. Glau-bitz, and F. Bernhard, High level cell-free expression and specific labeling of integral membrane proteins, *Eur J Biochem* **271**: 568–580 (2004).
- [172] S. B. Nabuurs, A. J. Nederveen, W. Vranken, J. F. Doreleijers, A. M. Bonvin, G. W. Vuister, G. Vriend, and C. A. Spronk, DRESS: a database of REfined solution NMR structures, *Proteins* **55**: 483–6 (2004).
- [173] J. P. Linge, M. A. Williams, C. Spronk, A. Bonvin, and M. Nilges, Refinement of protein structures in explicit solvent, *Prot Struct Funct Genet* **50**: 496–506 (2003).

A

Appendix: Sample preparation

This appendix shows SDS-gels to document the stages of the different protein purification protocols and ESI-MS spectra to verify the purified proteins.

A.1 CylR2

E. coli BL21(DE3) cells harbouring pET32a-vectors with an inserted wild type or mutant CylR2 gene were cultivated at 28 °C and the expressed proteins were purified after cell rupture via three chromatographic steps. The SDS-PAGE in Figure A.1 presents an exemplary preparation.

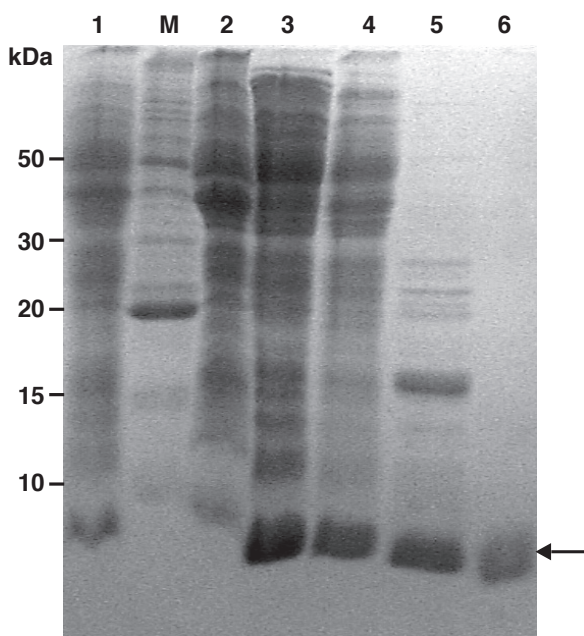


Figure A.1: Coomassie blue-stained SDS-PAGE documenting the expression and purification of CylR2 in *E. coli* BL21(DE3). The arrow indicates the CylR2 band. Lane 1, induced crude cell extract. Lane 2, pellet proteins of the sonicated cells. Lane 3, supernatant proteins of the sonicated cells. Lane 4, flow through and wash of the anion exchange chromatography. Lane 5, after cation exchange chromatography. Lane 6, purified CylR2 after RP-HPLC. Lane M, molecular weight standard.

The result of the ESI-MS analysis of the final sample after RP-HPLC is shown for CylR2_{T55C} in Figure A.2.

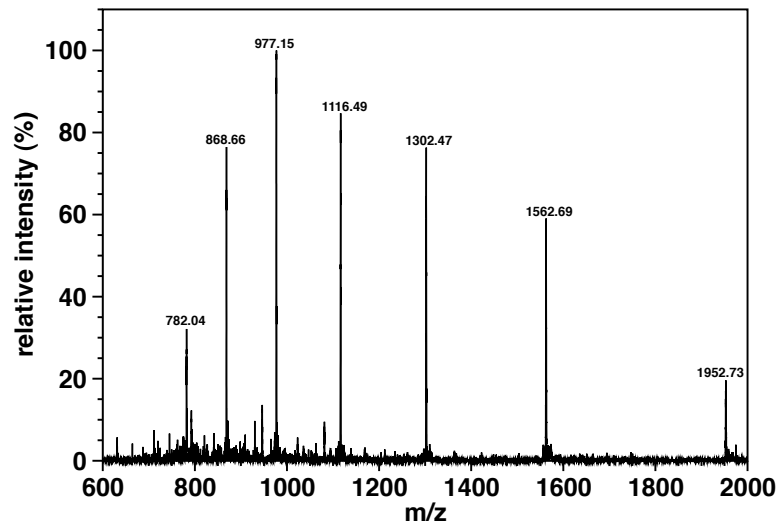


Figure A.2: ESI-MS spectrum of CylR2_{T55C} after RP-HPLC. The indicated m/z-ratios belong to 4- to 10-fold positively charged CylR2_{T55C} molecule ions.

A.2 CesT

The expression and purification of CesT is only shown together with Map84 (Figure A.8). The ESI-MS spectrum documents simultaneously the molecular weight and deuteration degree of a perdeuterated His₇-CesT sample (Figure A.3).

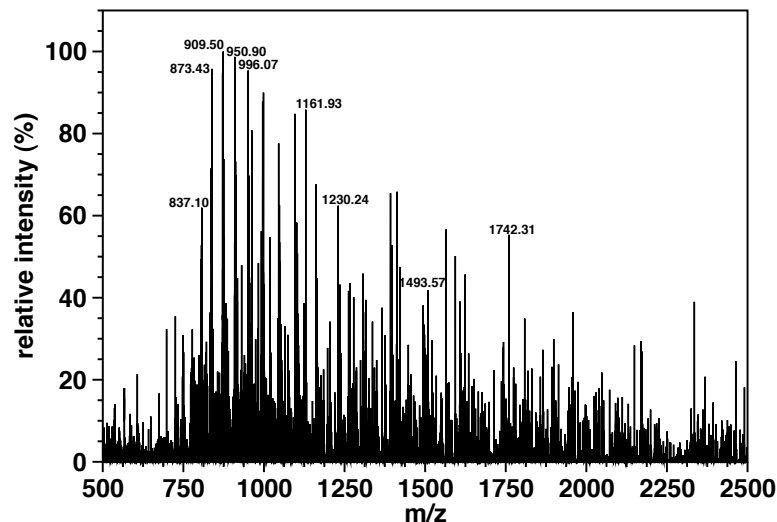


Figure A.3: ESI-MS spectrum of perdeuterated ²H-His₇-CesT in 20 mM NH₄OAc. The m/z-ratio is explicitly indicated for peaks corresponding to 99 % deuterated His₇-CesT with the number of positive charges decreasing from left to right from 25 to 12.

A.3 Map

Figure A.4 displays the expression of His₇-Map84 at 37 °C in *E. coli* BL21(DE3) and its purification.

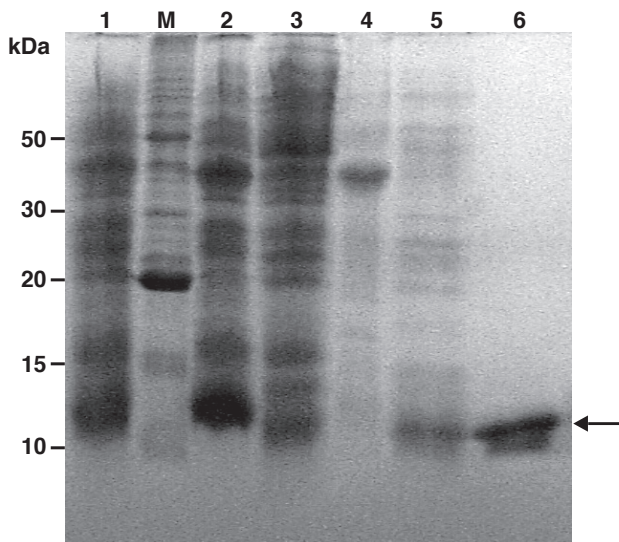


Figure A.4: Coomassie blue-stained SDS-PAGE illustrates expression and purification of His₇-Map84 in *E. coli* BL21(DE3). The arrow indicates His₇-Map84. Lane 1, induced crude cell extract. Lane 2, pellet proteins of the sonicated cells. Lane 3, supernatant proteins of the sonicated cells. Lane 4, pellet proteins after denaturing treatment of the pellet after sonication. Lane 5, supernatant proteins after denaturing treatment of the pellet after sonication. Lane 6, after Ni-NTA-agarose column. Lane M, molecular weight standard.

A.4 Tir

The expression and purification for different Tir-constructs is presented exemplarily for Tir108 in Figure A.5. ESI-MS spectra of Tir108 and TirN are shown in Figures A.6 and A.7, respectively.

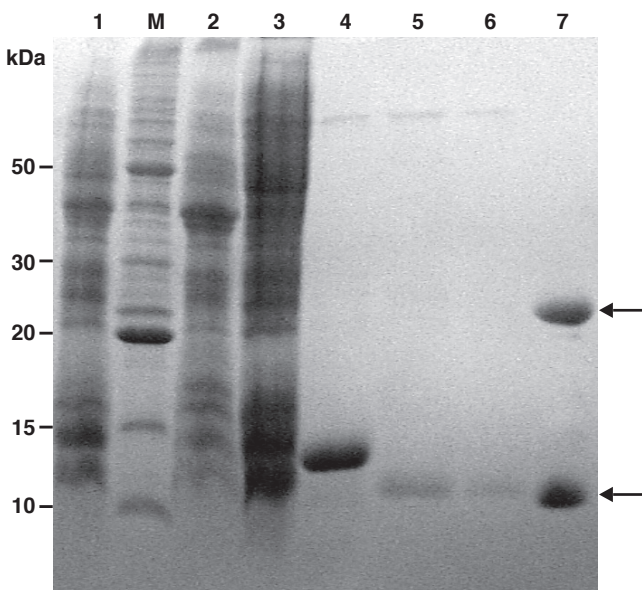


Figure A.5: Coomassie blue-stained SDS-PAGE documenting the expression and purification of Tir108 in *E. coli* BL21(DE3). The arrows indicate monomeric and dimeric Tir108. Lane 1, induced crude cell extract. Lane 2, pellet proteins of the sonicated cells. Lane 3, supernatant proteins of the sonicated cells. Lane 4, after first Ni-NTA-agarose column. Lane 5, after TEV-cleavage. Lane 6, after second Ni-NTA-agarose column. Lane 7, purified Tir108 after RP-HPLC. Lane M, molecular weight standard.

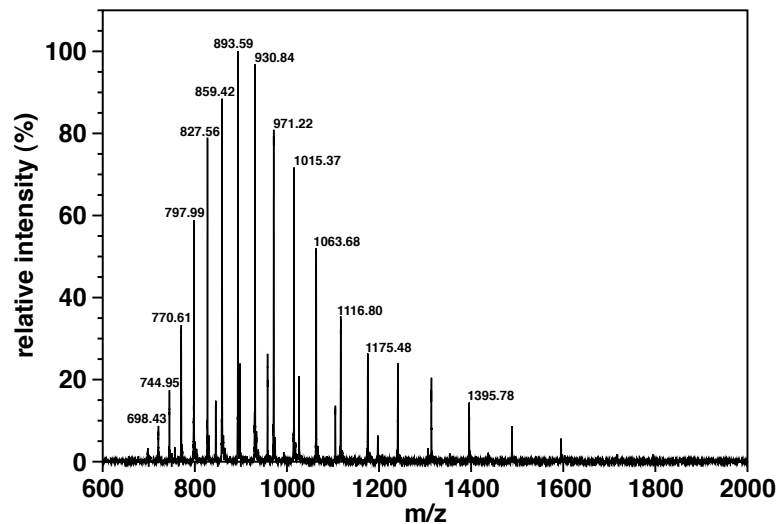


Figure A.6: ESI-MS spectrum of Tir108 after RP-HPLC. The m/z -ratio is shown the 16- and the 19- to 32-fold positively charged molecule ions of Tir108. The small peaks belong to dimeric Tir108. Two Tir108 are covalently bond via a disulfide bridge between the unique cysteine residue in each chain.

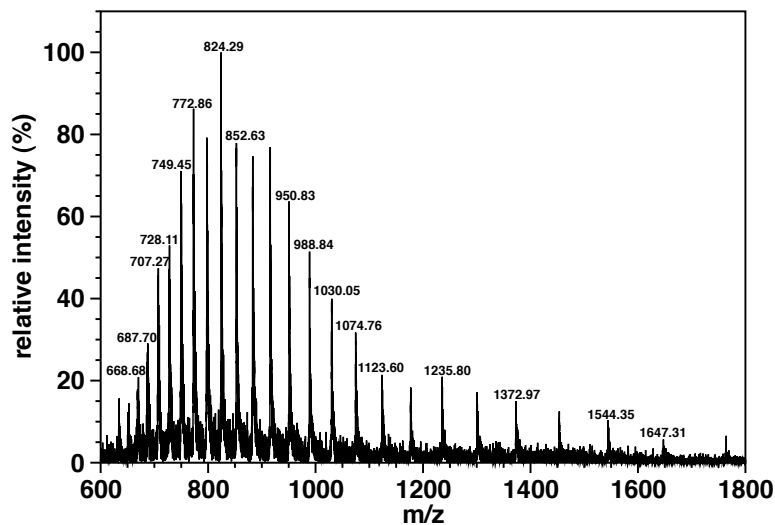


Figure A.7: ESI-MS spectrum of TirN in 20 mM NH_4OAc . The found m/z -ratios belong to TirN molecule ions with a different number of positive charges. The detected molecular is 24698.46 ± 7.87 what deviates by 0.032 % from the expected mass for TirN.

A.5 CesT/effector-complexes

The coexpression and copurification of the CesT/Map84 complex is demonstrated in Figure A.8. The ESI-MS spectra exhibit the molecular weights of coexpressed and copurified complex samples of CesT/Map84 (Figure A.9) and CesT/Tir108 (Figure A.10).

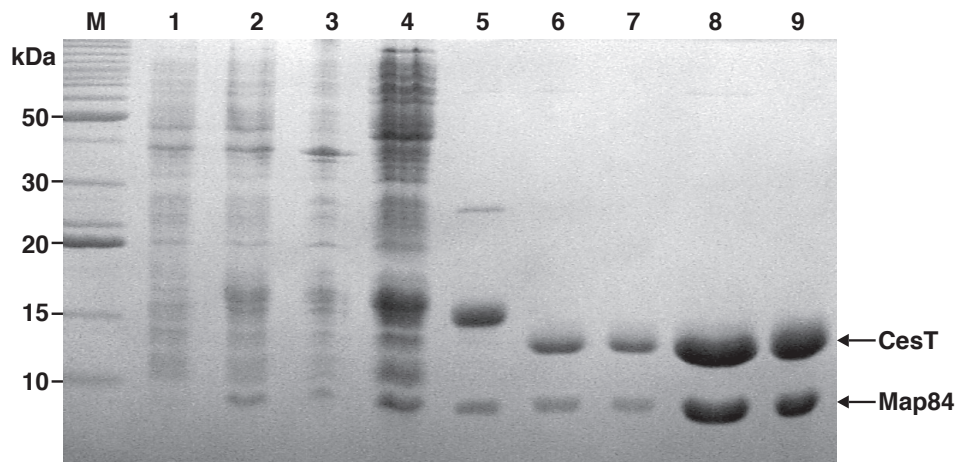


Figure A.8: Coomassie blue-stained SDS-PAGE documenting the coexpression and copurification of CesT/Map84 in *E. coli* BL21(DE3). CesT and Map84 bands are indicated. Lane M, molecular weight standard. Lane 1, uninduced crude cell extract. Lane 2, induced crude cell extract. Lane 3, pellet proteins of the sonicated cells. Lane 4, supernatant proteins of the sonicated cells. Lane 5, after first Ni-NTA-Agarose column. Lane 6, after TEV-cleavage. Lane 7, after second Ni-NTA-Agarose column. Lane 8, after anion exchange chromatography. Lane 9, purified CesT/Map84 after gel filtration.

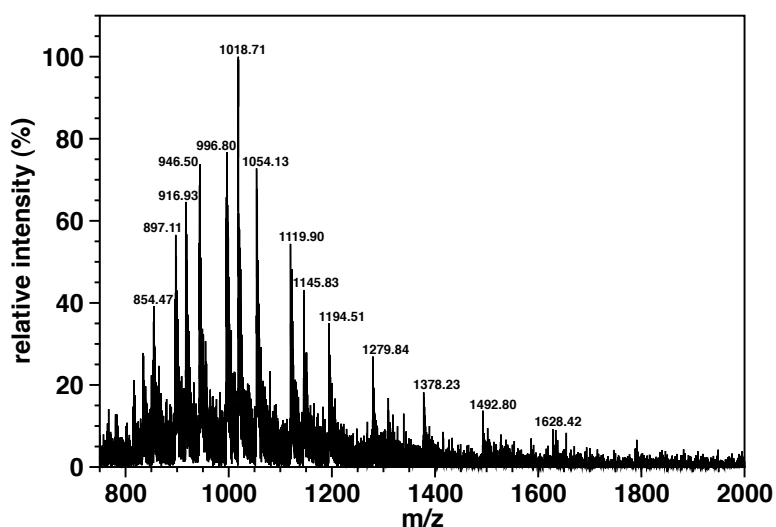


Figure A.9: ESI-MS spectrum of coexpressed and copurified CesT/Map84 in 20 mM NH_4OAc . The m/z -ratio is displayed for the 8- to 10-fold positively charged molecule ions of Map84 and for the 11- to 21-fold positively charged molecule ions of CesT.

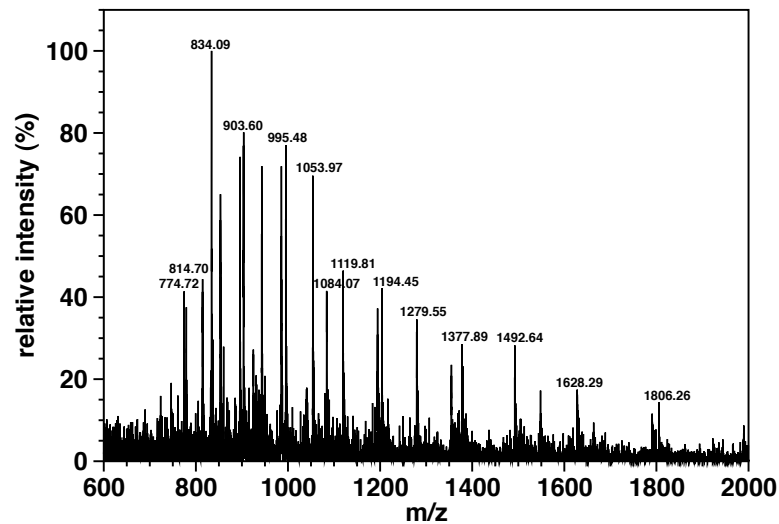


Figure A.10: ESI-MS spectrum of coexpressed and copurified CesT/Tir108 in 20 mM NH_4OAc . The m/z-ratio is indicated for the 6-, 10-, 12-, 13- and 14-fold positively charged molecule ions of Tir108 and the 11- to 18- and 22-fold positively charged molecule ions of CesT.



B

Appendix: NMR experiments and chemical shift assignments

In the following the NMR experiments measured for resonance assignments and structure calculation (B.1) as well as the obtained chemical shift assignments (B.2) which have not been deposited with the BMRB are summarized.

B.1 NMR experiments

Figure B.1 shows an overlay of the 1D NMR experiments measured with two different relaxation delays for a 10 μ M CylR2 sample at 298 K to estimate T_2 . The intensity ratio of 1.26 between the two spectra corresponds to an experimental molecular weight of \sim 17 kDa which indicates the concentration-independent dimeric state of CylR2.

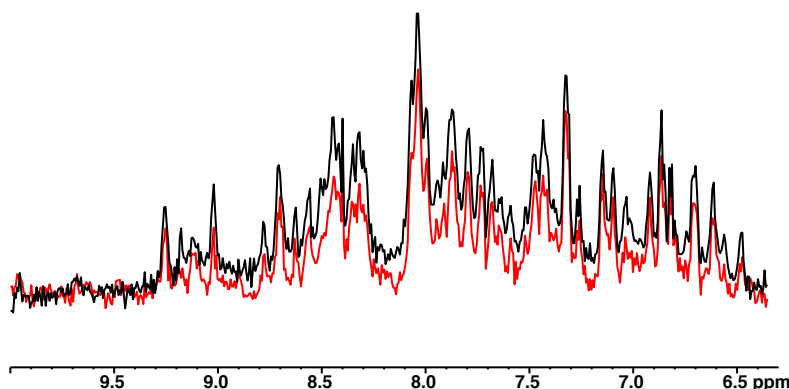


Figure B.1: Dimerization of CylR2. Overlay of the 1D NMR experiments measured for a 10 μ M CylR2 sample with a relaxation delay of 0.1 ms and 3.1 ms in black and red, respectively.

NMR spectra acquired for the assignment and structure determination of CylR2 (Table B.1), for the backbone assignment of CesT (Table B.2) and Tir108 (Table B.3), the reassignment of the CesT backbone in complex with Map84 (Table B.4) and in complex with Tir108 (Table B.5) and the accomplished backbone assignments of Map84 in complex with CesT (Table B.6) are provided in the following:

Table B.1: NMR spectra for assignment and structure determination of CylR2 measured at 298 K and with a $^{13}\text{C}/^{15}\text{N}$ -labeled CylR2 sample if not indicated differently.

	experiment	spectrometer	comment/pulse program (pp)
backbone assignment	2D ^{15}N HSQC	AVANCE 600, cryogenic probe	hsqc-ge.mz, reference experiment
	3D HNCA	AVANCE 600, cryogenic probe	hnca-ge.mz
	3D HNCO	AVANCE 600, cryogenic probe	hnco-cpd-fb.mz
	3D HN(CA)CO	AVANCE 600, cryogenic probe	Bruker pp hncacogp3d
	3D HNCACB	AVANCE 600, cryogenic probe	hncacb-cpd.mz
	3D CBCACONH	AVANCE 600, cryogenic probe	cbcaconh-J.mz
aliphatic sidechain assignment	3D HCCH-TOCSY	DRX 800	^{13}C refocusing and mixing by adiabatic pulses [69], [68]
	3D CCONH-TOCSY	AVANCE 600, cryogenic probe	
	3D HCCONH-TOCSY	AVANCE 600, cryogenic probe	
aromatic sidechain assignment	2D ^{13}C HSQC	AVANCE 900	Bruker pulse program hsqcsetgppsp
	2D $(\text{H}\beta)\text{C}\beta(\text{C}\gamma\text{C}\delta)\text{H}\delta$	AVANCE 700	hbcbcgaro.vivi.txt, [70]
	2D $(\text{H}\beta)\text{C}\beta(\text{C}\gamma\text{C}\delta\text{C}\epsilon)\text{H}\epsilon$	AVANCE 700	hbcbcgcearo.vivi.txt, [70]
distance determination	3D ^{15}N NOESY-HSQC	AVANCE 600, cryogenic probe	^{15}N -labeled, noesy3d.mz
	3D ^{13}C NOESY-HSQC	DRX 600	c13noesy_hsqc.vivi.txt
	2D ^{15}N HSQC	AVANCE 700	^{15}N -labeled, hsqc-ge.mz, PRE
	2D ^{15}N T ₂	AVANCE 700	^{15}N -labeled, T2N.sr.txt

Table B.2: NMR spectra for the backbone assignment of CesT measured at 303 K.

labeling	experiment	spectrometer	comment/pulse program (pp)
$^{13}\text{C}/^{15}\text{N}$	2D ^{15}N HSQC	AVANCE 900	hsqc-ge.mz, reference spectrum
$^2\text{H}(75\%)/^{15}\text{N}$	2D ^{15}N Trosy-HSQC	AVANCE 900	trosyfb.mz, reference spectrum
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D HNCA	AVANCE 700	hnca-ge-2H.sr.txt
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D Trosy-HN(CO)CA	AVANCE 600, cryogenic probe	trhncoca3d-QJ.cpj.txt, [81]
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D HN(CO)CACB	AVANCE 600, cryogenic probe	Bruker pp hncocacbgp2h3d, $\Delta(\text{C}_\alpha\text{C}_\beta)=3.6$ ms
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D HN(COCA)CB	AVANCE 600, cryogenic probe	Bruker pp hncocacbgp2h3d, $\Delta(\text{C}_\alpha\text{C}_\beta)=7$ ms
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D Trosy-HN(CA)CB	AVANCE 900	hncacb-tr-h2.sr, $\Delta(\text{C}_\alpha\text{C}_\beta)=7.1$ ms
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D HNCO	AVANCE 700	hnco-cpd-viv.txt
$^2\text{H}(75\%)/^{13}\text{C}/^{15}\text{N}$	3D HN(CA)CO	AVANCE 600, cryogenic probe	Bruker pp hncacogp2h3d
$^2\text{H}(75\%)/^{15}\text{N}$	3D ^{15}N NOESY-HSQC	AVANCE 900, cryogenic probe	noesy3d.sr

Table B.3: NMR spectra for the backbone assignment of Tir108 measured at 293 K with a $^{13}\text{C}/^{15}\text{N}$ -labeled sample.

experiment	spectrometer	comment/pulse program (pp)
2D ^{15}N HSQC	AVANCE 600, cryogenic probe	hsqc-ge.mz, reference experiment
3D HNCO	AVANCE 600, cryogenic probe	hnco-cpd-fb.mz
3D HN(CA)CO	AVANCE 600, cryogenic probe	Bruker pp hncacogp3d
3D HNCACB	AVANCE 600, cryogenic probe	hncacb-cpd.mz
3D CBCACONH	AVANCE 600, cryogenic probe	cbcaconh-J.mz
3D HNN	AVANCE 600, cryogenic probe	hnn.sr.txt, [169]

Table B.4: NMR spectra for the backbone assignment of CesT in complex with Map84 measured at 303 K for a $^2\text{H}(75\ \%) / ^{13}\text{C} / ^{15}\text{N}$ CesT/His₇Map84 complex sample.

experiment	spectrometer	comment/pulse program (pp)
2D ^{15}N Trosy-HSQC	AVANCE 900, cryogenic probe	trosytf3gpsi, ^{13}C decoupling added
3D Trosy-HNCA	AVANCE 900, cryogenic probe	trhnca_tr_d.rtf
3D Trosy-HN(COCA)CB	AVANCE 600, cryogenic probe	Ehncacabtrosywgrrt, $\Delta(C_\alpha C_\beta)=7$ ms
3D Trosy-HN(CA)CB	AVANCE 900, cryogenic probe	hncacb-tr-h2.sr, $\Delta(C_\alpha C_\beta)=7.1$ ms
3D Trosy-HNCO	AVANCE 900, cryogenic probe	Bruker pp trhncogp2h3d
3D Trosy-HN(CA)CO	AVANCE 600, cryogenic probe	trhncacoetgp3d, deuterium decoupling added
3D ^{15}N NOESY-Trosy-HSQC	AVANCE 900, cryogenic probe	Bruker pp noesytrtf3gp3d

Table B.5: NMR spectra for the backbone assignment of CesT in complex with Tir108 measured at 303 K with a $^2\text{H}(75\ \%) / ^{13}\text{C} / ^{15}\text{N}$ CesT/Tir108 complex sample.

experiment	spectrometer	comment/pulse program (pp)
2D ^{15}N Trosy-HSQC	AVANCE 900, cryogenic probe	trosytf3gpsi, ^{13}C decoupling added
3D Trosy-HNCA	AVANCE 900, cryogenic probe	trhnca_tr_d.rtf
3D Trosy-HN(COCA)CB	AVANCE 600, cryogenic probe	Ehncacabtrosywgrrt, $\Delta(C_\alpha C_\beta)=7$ ms
3D Trosy-HN(CA)CB	AVANCE 900, cryogenic probe	hncacb-tr-h2.sr, $\Delta(C_\alpha C_\beta)=7.1$ ms
3D Trosy-HNCO	AVANCE 600, cryogenic probe	Bruker pp trhncogp2h3d
3D Trosy-HN(CA)CO	AVANCE 600, cryogenic probe	trhncacoetgp3d, deuterium decoupling added
3D ^{15}N NOESY-Trosy-HSQC	AVANCE 900, cryogenic probe	Bruker pp noesytrtf3gp3d

Table B.6: NMR spectra for the backbone assignment of Map84 in complex with CesT measured at 303 K for a $^2\text{H}(99\ \%)$ CesT/ $^{13}\text{C} / ^{15}\text{N}$ Map84 complex sample.

experiment	spectrometer	comment/pulse program (pp)
2D ^{15}N Trosy-HSQC	AVANCE 900, cryogenic probe	trosytf3gpsi, ^{13}C decoupling added
2D ^{15}N HSQC	AVANCE 900, cryogenic probe	hsqc-ge.mz, reference experiment
3D HNCA	AVANCE 900, cryogenic probe	hnca-ge.mz
3D Trosy-HNCA	AVANCE 900, cryogenic probe	trhnca_tr_d.rtf
3D HNCACB	AVANCE 600, cryogenic probe	hncacb-cpd.mz
3D CBCACONH	AVANCE 600, cryogenic probe	cbcaconh-J.mz
3D Trosy-HNCO	AVANCE 900, cryogenic probe	tr-hnco.vivi
3D HN(CA)CO	AVANCE 600, cryogenic probe	Bruker pp hncacogp3d

B.2 Chemical shift assignments

Resonances assigned during this thesis, which are not deposited in the BMRB, are specified in the following Tables. Table B.7 gives the assigned backbone, C_β and amide side chain resonances of CylR2 in complex with its DNA. The assigned backbone and C_β resonances of CesT in complex with Map84 and in complex with Tir108 are indicated in Table B.8 and Table B.9, respectively. Assignments obtained for backbone and C_β resonances of Tir108 are stated in Table B.10. Table B.11 contains the assignments for Map84 in complex with CesT.

Table B.7: Assigned chemical shifts of CylR2 in complex with its DNA at 298 K based on 3D HNCO, CBCA(CO)NH, HNCA and 2D HSQC experiments.

	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	N $_{\delta}$	H $_{\delta}$	N $_{\epsilon}$	H $_{\epsilon}$
1	MET	181.67	55.18	33.33	-	-				
2	ILE	175.83	60.16	39.65	123.782	8.830				
3	ILE	174.29	60.58	38.11	129.925	9.426				
4	ASN	174.81	51.38	40.44	125.099	8.716	112.911	8.063/6.698		
5	ASN	175.50	51.88	39.77	122.382	9.292	112.647	6.915/6.662		
6	LEU	176.23	59.92	41.22	121.181	8.233				
7	LYS	177.43	60.51	32.50	118.245	8.881				
8	LEU	180.29	57.86	42.50	116.423	7.297				
9	ILE	178.01	63.05	36.74	119.068	8.058				
10	ARG	179.19	61.02	29.73	120.751	9.181				
11	GLU	181.61	59.88	29.04	116.461	8.607				
12	LYS	178.66	59.50	32.32	122.190	7.933				
13	LYS	175.32	55.41	31.61	117.130	7.640				
14	LYS	174.88	57.06	28.66	115.194	7.933				
15	ILE	175.45	59.45	41.43	117.674	7.284				
16	SER	175.36	58.14	65.23	124.513	9.185				
17	GLN	178.59	61.21	29.20	121.519	9.083			112.364	7.739/7.314
18	SER	177.18	62.98	-	113.589	8.417				
19	GLU	179.56	59.11	30.19	124.496	8.337				
20	LEU	177.70	57.70	42.50	120.681	8.356				
21	ALA	178.62	55.56	17.47	120.227	8.296				
22	ALA	181.83	54.87	17.90	117.791	7.739				
23	LEU	178.21	57.51	41.83	119.294	7.509				
24	LEU	175.10	54.93	43.58	115.218	7.364				
25	GLU	175.97	57.64	31.16	114.644	7.938				
26	VAL	174.32	58.55	-	110.175	7.794				
27	SER	176.82	58.09	-	120.966	9.196				
28	ARG	179.21	60.29	30.28	115.330	8.151				
29	GLN	178.66	58.74	27.43	116.400	8.563			112.311	7.552/6.914
30	THR	176.46	66.08	68.04	121.166	7.990				
31	ILE	177.24	61.31	34.90	118.166	6.707				
32	ASN	176.52	58.62	41.93	118.617	8.343				
33	GLY	175.97	46.90	-	105.072	8.710				
34	ILE	181.44	64.60	38.90	121.474	7.946				
35	GLU	178.53	60.61	29.42	121.379	8.603				
36	LYS	175.41	55.14	31.75	116.762	7.791				
37	ASN	175.29	54.82	37.01	115.132	8.096	113.177	7.583/7.100		
38	LYS	176.32	57.91	33.19	116.326	8.524				
39	TYR	173.07	55.34	36.95	111.877	7.040				
40	ASN	-	51.14	-	119.177	9.839	112.588	7.529/6.863		
41	PRO	175.99	62.69	31.9	-	-				
42	SER	172.82	57.87	65.06	125.303	7.761				
43	LEU	178.09	59.04	41.00	121.439	9.120				
44	GLN	176.87	60.13	28.18	114.805	8.819			112.640	7.449/6.927
45	LEU	178.43	57.74	40.77	119.374	7.536				
46	ALA	178.93	55.74	19.34	121.438	8.803				
47	LEU	180.58	57.81	42.54	116.257	8.481				
48	LYS	178.15	60.24	32.76	120.796	7.702				
49	ILE	176.76	67.40	38.29	120.994	8.791				
50	ALA	179.06	55.79	18.36	120.794	8.048				
51	TYR	180.39	61.73	38.51	117.457	7.980				
52	TYR	177.97	63.24	37.83	118.551	8.929				
53	LEU	176.53	54.42	39.93	115.427	8.233				
54	ASN	174.03	53.92	38.00	122.388	7.969	111.898	7.486/6.772		
55	THR	-	60.31	-	115.260	8.126				
56	PRO	178.46	62.85	32.36	-	-				
57	LEU	179.89	60.26	41.82	127.612	9.699				
58	GLU	177.01	58.87	29.78	117.102	9.980				
59	ASP	176.86	55.85	40.35	117.663	7.863				
60	ILE	174.04	64.09	39.64	120.354	8.067				
61	PHE	172.61	55.72	42.66	116.521	8.319				

continued on next page

	Res.	C'	C _α	C _β	N	H _N	N _δ	H _δ	N _ε	H _ε
62	GLN	175.09	54.20	32.34	117.231	8.517			110.852	7.535/6.624
63	TRP	174.60	57.00	31.09	125.665	9.309				
64	GLN	-	51.76	-	126.861	7.782			111.997	7.394/6.783
65	PRO	-	-	-	-	-				
66	GLU	-	-	-	-	-				

Table B.8: Assigned chemical shifts of CesT bound to Map84 at 303 K based on TROSY versions of 3D HNCO, HN(CA)CO, HNCA, HN(COCA)CB, HN(CA)CB and 2D HSQC experiments.

	Res.	C'	C _α	C _β	N	H _N		Res.	C'	C _α	C _β	N	H _N	
1	MET	54.63	32.02	175.77	-	-		79	ASN	52.59	38.80	175.33	112.724	6.781
2	SER	58.49	63.43	174.05	119.263	8.405		80	GLY	46.42	-	175.39	110.894	7.255
3	SER	57.03	64.19	175.08	120.525	8.748		81	GLY	43.75	-	104.82	105.066	8.320
4	ARG	60.00	29.99	180.03	121.088	8.829		82	PRO	-	-	-	-	-
5	SER	61.50	62.56	176.04	117.103	8.601		83	TYR	-	-	-	-	-
6	GLU	59.42	28.64	179.15	123.840	7.667		84	LEU	53.65	40.60	177.29	-	-
7	LEU	57.34	40.90	179.59	119.733	8.128		85	CYS	56.30	29.87	172.70	123.442	8.908
8	LEU	57.33	39.95	178.56	120.230	7.630		86	TYR	54.91	42.81	172.69	120.221	6.295
9	LEU	57.79	40.13	178.52	120.804	8.361		87	GLU	55.51	31.43	175.00	130.370	8.337
10	ASP	57.15	40.08	179.69	118.543	7.799		88	SER	60.04	62.65	175.58	118.570	7.006
11	ARG	58.17	29.90	179.91	119.308	7.863		89	GLY	46.84	-	175.40	116.881	9.415
12	PHE	60.73	38.98	175.63	123.752	9.125		90	ALA	50.55	18.22	177.56	123.714	9.289
13	ALA	54.57	17.63	179.99	121.686	8.814		91	GLN	56.44	25.12	175.08	116.420	7.249
14	GLU	58.27	28.91	178.87	115.951	7.277		92	SER	57.88	65.49	173.31	111.498	7.442
15	LYS	58.38	31.26	178.74	120.737	7.425		93	LEU	54.32	41.50	174.46	124.375	9.066
16	ILE	61.60	36.66	176.88	109.696	7.464		94	LEU	-	-	-	124.463	9.322
17	GLY	46.62	-	175.65	110.018	7.428		95	LEU	53.32	43.60	173.48	-	-
18	VAL	59.92	32.06	175.38	111.824	7.299		96	ALA	50.27	23.35	174.59	125.161	9.041
19	GLY	44.60	-	172.84	110.764	8.120		97	LEU	53.77	47.31	176.08	124.418	8.845
20	SER	58.19	62.62	174.53	113.825	7.850		98	ARG	55.52	30.95	175.37	128.481	9.278
21	ILE	59.68	40.90	173.97	121.506	8.190		99	PHE	53.62	40.36	177.02	126.747	9.038
22	SER	56.68	65.13	173.94	113.112	7.896		100	PRO	62.10	31.29	173.84	-	-
23	PHE	58.55	39.20	176.47	119.609	8.601		101	LEU	54.93	41.31	175.07	119.039	7.490
24	ASN	51.02	37.91	177.74	122.048	8.855		102	ASP	56.12	39.17	175.96	122.094	8.191
25	GLU	58.71	27.65	177.14	119.074	9.313		103	ASP	54.50	39.18	178.85	120.702	8.777
26	ASN	52.62	38.47	173.32	117.755	7.826		104	ALA	53.52	20.57	176.64	120.327	7.332
27	ARG	55.29	23.79	173.84	112.841	8.427		105	THR	57.75	70.05	177.47	108.720	6.706
28	LEU	53.63	45.77	177.74	116.901	7.757		106	PRO	65.14	31.20	175.26	-	-
29	CYS	58.78	28.44	174.11	123.279	9.182		107	GLU	60.16	27.94	177.47	118.302	8.749
30	SER	58.50	64.33	172.25	123.568	8.521		108	LYS	59.28	32.59	176.84	120.692	7.827
31	PHE	56.14	44.10	173.89	118.705	9.280		109	LEU	57.38	39.31	178.64	120.210	7.733
32	ALA	50.34	19.86	177.17	122.490	8.979		110	GLU	60.29	29.09	178.81	119.369	8.195
33	ILE	58.67	37.76	175.65	126.023	9.604		111	ASN	56.11	38.17	177.54	116.413	7.875
34	ASP	55.56	38.09	175.37	127.494	9.289		112	GLU	58.25	28.17	177.92	120.375	8.138
35	GLU	60.20	28.30	176.81	108.373	8.441		113	ILE	66.16	36.98	177.88	120.331	8.358
36	ILE	61.83	39.93	174.70	114.512	7.781		114	GLU	59.74	27.99	179.06	120.404	8.195
37	TYR	56.94	38.26	174.73	120.001	7.941		115	VAL	66.41	30.50	178.07	119.705	7.524
38	TYR	56.63	37.65	174.42	126.822	9.001		116	VAL	67.95	30.33	177.94	123.008	7.541
39	ILE	56.70	40.39	175.36	125.479	8.698		117	VAL	66.68	30.48	179.98	121.735	8.693
40	SER	58.12	66.38	172.43	120.600	9.546		118	LYS	58.20	30.72	177.31	118.775	8.873
41	LEU	52.55	42.23	176.33	123.978	9.138		119	SER	62.58	63.94	177.61	-	-
42	SER	54.51	65.46	175.73	115.277	9.328		120	MET	60.30	31.97	176.59	122.397	8.243
43	ASP	51.90	37.59	175.55	129.922	8.130		121	GLU	59.25	28.74	179.38	120.988	8.561
44	ALA	54.17	19.48	176.98	126.589	7.339		122	ASN	55.68	37.82	176.74	117.386	7.922
45	ASN	52.45	39.83	173.36	114.925	7.783		123	LEU	57.88	38.12	178.61	120.595	8.189
46	ASP	56.35	40.61	176.65	119.935	8.363		124	TYR	58.78	36.05	179.08	117.960	9.244
47	GLU	56.45	30.79	176.90	114.993	8.898		125	LEU	58.44	40.98	178.66	121.520	7.905
48	TYR	56.74	38.82	172.50	116.991	7.337		126	VAL	65.22	31.62	179.63	120.436	8.216
49	MET	53.95	36.92	176.33	118.246	8.827		127	LEU	58.98	-	180.13	-	-
50	MET	52.00	32.46	174.36	118.133	8.341		128	HIS	59.63	29.26	179.63	121.005	8.779
51	ILE	60.20	38.37	174.24	121.979	8.782		129	ASN	55.27	37.60	176.63	120.009	8.236

continued on next page

Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$		
52	TYR	54.74	40.04	176.03	122.569	9.143	130	GLN	55.06	28.69	175.81	117.223	7.666
53	GLY	45.45		173.50	110.542	9.692	131	GLY	45.49		174.30	108.675	7.821
54	VAL	61.89	30.14	175.24	127.955	8.784	132	ILE	60.16	38.51	174.41	123.084	8.103
55	CYS	60.93	27.63	175.44	125.168	8.626	133	THR	60.50	69.74	174.21	122.695	8.200
56	GLY	43.58		171.20	102.972	7.074	134	LEU	53.65	40.38	177.00	127.142	9.161
57	LYS	56.17	32.66	176.54	121.154	8.641	135	GLU	56.21	29.99	175.96	124.000	8.870
58	PHE	56.33	39.66	174.29	128.202	9.422	136	ASN	52.11	38.73	175.57	122.157	8.652
59	PRO	61.90	30.02	176.43	-	-	137	GLU	55.51	29.32	176.67	118.545	8.189
60	THR	62.38	68.71	174.91	115.930	8.274	138	HIS	56.25	28.94	175.02	118.509	8.254
61	ASP	53.84	39.62	175.22	120.219	8.423	139	MET	55.77	31.87	175.40	121.232	8.000
62	ASN	49.79	38.27	116.27	116.517	7.867	140	LYS	55.60	31.99	176.12	122.556	7.858
63	PRO	64.35	31.33	178.23	-	-	141	ILE	60.79	37.77	176.25	123.765	8.103
64	ASN	55.15	37.99	177.21	115.376	8.240	142	GLU	56.12	29.67	176.18	125.376	8.428
65	PHE	61.08	38.14	176.92	122.088	7.773	143	GLU	56.14	29.57	176.49	122.775	8.298
66	ALA	55.18	18.07	178.74	118.689	7.993	144	ILE	60.89	37.86	176.39	122.409	8.091
67	LEU	57.07	38.75	179.58	114.871	6.985	145	SER	57.88	63.31	174.68	119.909	8.323
68	GLU	58.82	28.27	179.65	119.875	7.584	146	SER	58.06	63.34	174.89	118.469	8.285
69	ILE	60.92	34.16	177.42	120.179	7.690	147	SER	58.31	63.27	174.34	117.908	8.267
70	LEU	57.82	42.85	180.87	118.275	8.093	148	ASP	54.14	40.49	176.00	122.070	8.160
71	ASN	57.31	40.17	178.22	119.941	8.416	149	ASN	53.10	38.06	175.36	119.109	8.147
72	ALA	54.72	18.00	180.12	-	-	150	LYS	56.26	31.64	176.41	121.013	8.071
73	ASN	55.92	-	177.27	117.721	7.604	151	HIS	55.44	29.21	174.58	119.210	8.116
74	LEU	-	-	-	-	-	152	TYR	57.61	38.15	175.18	121.672	7.930
75	TRP	-	-	-	-	-	153	TYR	57.20	38.17	175.06	122.663	7.958
76	PHE	-	-	-	-	-	154	ALA	52.07	18.47	177.78	126.360	8.026
77	ALA	55.35	18.13	179.03	-	-	155	GLY	45.00		173.09	108.141	7.656
78	GLU	57.06	28.75	176.09	115.488	7.014	156	ARG	56.89	30.49	171.00	125.509	7.638

Table B.9: Assigned chemical shifts of CesT bound to Tir108 at 303 K based on TROSY versions of 3D HNC0, HN(CA)CO, HNCA, HN(COCA)CB, HN(CA)CB and 2D HSQC experiments.

Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$		
1	MET	54.58	32.01	175.79	122.732	8.368	79	ASN	-	-	-	-	-
2	SER	58.40	63.39	174.18	118.669	8.334	80	GLY	46.08	-	175.33	110.566	7.201
3	SER	57.11	64.33	175.17	120.275	8.665	81	GLY	43.76	-	180.57	105.593	8.420
4	ARG	59.55	29.79	179.70	121.230	8.759	82	PRO	-	-	-	-	-
5	SER	61.31	-	176.08	116.874	8.556	83	TYR	-	-	-	-	-
6	GLU	59.30	-	179.17	123.730	7.622	84	LEU	-	-	-	-	-
7	LEU	57.25	40.94	179.43	119.999	8.075	85	CYS	-	-	-	-	-
8	LEU	57.38	40.41	178.04	120.083	7.657	86	TYR	55.91	-	172.18	121.261	6.571
9	LEU	57.67	-	178.61	119.929	8.129	87	GLU	55.60	-	-	130.760	8.278
10	ASP	56.70	-	-	118.120	7.612	88	SER	-	-	-	-	-
11	ARG	58.25	-	180.04	118.808	7.687	89	GLY	-	-	-	-	-
12	PHE	60.22	-	175.35	123.645	8.746	90	ALA	-	-	-	-	-
13	ALA	54.34	-	179.68	121.973	8.762	91	GLN	-	-	-	115.410	7.552
14	GLU	57.99	-	-	116.774	7.276	92	SER	-	-	-	-	-
15	LYS	-	-	-	-	-	93	LEU	54.00	-	174.20	124.546	8.759
16	ILE	-	-	-	-	-	94	LEU	52.42	-	171.08	127.453	8.817
17	GLY	-	-	-	-	-	95	LEU	-	-	-	-	-
18	VAL	-	-	-	115.417	7.328	96	ALA	-	-	-	-	-
19	GLY	44.29	-	172.11	112.215	8.017	97	LEU	53.83	-	173.63	124.763	8.924
20	SER	58.18	62.82	174.59	114.897	7.759	98	ARG	55.67	-	175.20	128.875	9.136
21	ILE	59.81	-	173.74	122.922	7.990	99	PHE	53.80	-	-	128.541	9.093
22	SER	56.88	-	174.23	116.371	8.061	100	PRO	62.13	31.25	176.07	-	-
23	PHE	58.23	-	176.67	119.719	8.783	101	LEU	54.93	-	179.01	119.028	7.554
24	ASN	50.84	-	177.73	121.424	8.799	102	ASP	56.23	-	176.70	121.833	8.189
25	GLU	58.68	-	177.49	116.378	9.441	103	ASP	54.13	-	175.40	120.819	8.594
26	ASN	52.77	-	173.14	119.250	9.466	104	ALA	53.43	20.41	177.42	120.705	7.319
27	ARG	55.24	-	-	117.802	7.816	105	THR	57.68	-	177.37	108.927	6.714
28	LEU	53.07	-	177.85	115.694	8.299	106	PRO	65.35	-	177.22	-	-
29	CYS	48.84	-	173.65	118.444	7.801	107	GLU	59.82	-	178.72	117.872	8.746

continued on next page

	Res.	C'	C _α	C _β	N	H _N		Res.	C'	C _α	C _β	N	H _N
30	SER	-	-	-	-	-	108	LYS	59.30	-	-	121.759	7.934
31	PHE	-	-	-	-	-	109	LEU	-	-	-	-	-
32	ALA	49.83	-	177.80	122.094	8.379	110	GLU	-	-	-	-	-
33	ILE	-	-	-	-	-	111	ASN	-	-	-	-	-
34	ASP	-	-	-	-	-	112	GLU	-	-	-	-	-
35	GLU	-	-	-	-	-	113	ILE	-	-	-	120.627	8.368
36	ILE	-	-	-	115.545	8.133	114	GLU	-	-	-	-	-
37	TYR	-	-	-	-	-	115	VAL	65.94	-	180.14	120.758	7.667
38	TYR	-	-	-	-	-	116	VAL	67.81	-	177.13	123.786	7.655
39	ILE	56.14	-	175.08	126.443	8.854	117	VAL	67.16	-	-	122.178	8.834
40	SER	57.89	-	172.95	120.228	9.099	118	LYS	-	-	-	-	-
41	LEU	52.16	-	176.47	122.943	9.022	119	SER	-	-	-	-	-
42	SER	54.67	-	175.83	115.049	8.963	120	MET	-	-	-	-	-
43	ASP	51.79	-	175.68	130.758	8.361	121	GLU	-	-	-	119.023	8.305
44	ALA	54.24	-	177.14	125.531	7.377	122	ASN	-	-	-	-	-
45	ASN	52.27	-	173.43	115.675	7.924	123	LEU	-	-	-	122.571	8.254
46	ASP	56.14	-	176.63	120.088	8.395	124	TYR	-	-	-	-	-
47	GLU	56.35	-	177.05	117.475	8.910	125	LEU	-	-	178.94	121.717	7.739
48	TYR	56.61	-	172.22	116.814	7.351	126	VAL	-	-	-	-	-
49	MET	53.98	37.75	176.18	118.607	8.981	127	LEU	-	-	-	-	-
50	MET	51.83	-	174.57	118.559	8.288	128	HIS	59.62	-	-	120.787	8.755
51	ILE	60.36	-	174.13	122.782	8.717	129	ASN	-	-	-	-	-
52	TYR	54.98	-	175.82	123.864	9.034	130	GLN	-	-	176.09	118.633	7.899
53	GLY	44.87	-	173.61	110.652	9.701	131	GLY	45.28	-	174.12	108.183	7.654
54	VAL	62.25	-	175.24	128.483	8.824	132	ILE	-	-	-	121.309	8.040
55	CYS	-	-	-	-	-	133	THR	-	-	-	-	-
56	GLY	-	-	-	102.898	7.117	134	LEU	-	-	-	-	-
57	LYS	-	-	176.62	121.307	8.555	135	GLU	55.77	29.74	176.09	123.950	8.924
58	PHE	56.33	-	173.82	129.203	9.187	136	ASN	52.94	38.41	175.45	119.302	8.305
59	PRO	61.82	29.98	176.53	-	-	137	GLU	56.74	29.40	176.49	122.418	8.391
60	THR	62.56	-	174.86	115.482	8.217	138	HIS	55.18	28.43	174.66	119.317	8.258
61	ASP	53.25	-	175.25	120.294	8.373	139	MET	55.28	32.14	175.82	121.150	8.095
62	ASN	49.99	-	177.88	116.664	7.798	140	LYS	55.79	32.02	176.44	123.366	8.137
63	PRO	64.29	-	-	-	-	141	ILE	60.81	37.59	176.34	123.178	8.107
64	ASN	54.85	-	177.09	115.737	8.155	142	GLU	56.13	29.44	176.31	124.978	8.414
65	PHE	60.57	-	176.59	122.134	7.736	143	GLU	56.09	29.45	176.49	122.849	8.279
66	ALA	-	-	-	-	-	144	ILE	60.85	37.77	176.42	122.300	8.074
67	LEU	-	-	-	-	-	145	SER	57.83	63.30	174.69	120.185	8.341
68	GLU	58.52	-	179.38	119.655	7.507	146	SER	58.01	63.31	174.91	119.315	8.426
69	ILE	-	-	-	-	-	147	SER	58.26	63.20	174.40	118.069	8.273
70	LEU	-	-	-	-	-	148	ASP	54.04	40.44	176.04	122.425	8.175
71	ASN	-	-	-	-	-	149	ASN	53.08	38.05	175.36	119.323	8.133
72	ALA	-	-	-	-	-	150	LYS	56.13	31.63	176.42	121.054	8.066
73	ASN	54.44	-	174.33	120.832	8.391	151	HIS	54.99	28.61	174.20	119.191	8.159
74	LEU	-	-	-	-	-	152	TYR	57.63	38.15	175.21	120.514	7.867
75	TRP	-	-	-	-	-	153	TYR	57.09	38.20	175.05	122.665	7.957
76	PHE	-	-	-	-	-	154	ALA	51.99	18.50	177.74	126.384	8.027
77	ALA	-	-	-	-	-	155	GLY	44.97	-	173.10	122.037	7.683
78	GLU	57.72	-	176.20	122.476	7.937	156	ARG	56.83	30.48	170.96	125.561	7.646

Table B.10: Assigned chemical shifts of Tir108 at 283 K based on 3D HNCO, HN(CA)CO, CBCA(CO)NH, HNCACB and 2D HSQC experiments.

	Res.	C'	C _α	C _β	N	H _N		Res.	C'	C _α	C _β	N	H _N
1	MET	-	-	-	-	-	55	ASN	175.96	53.28	38.79	8.411	119.985
2	PRO	176.84	62.89	31.91	-	-	56	SER	175.30	58.53	63.68	8.280	116.330
3	ILE	176.72	61.30	38.82	8.223	121.327	57	MET	174.74	55.93	32.87	8.315	121.799
4	GLY	176.84	45.14	-	8.373	112.411	58	ALA	176.25	52.63	19.10	8.115	124.277
5	ASN	173.72	53.10	38.88	8.285	118.749	59	ASP	177.49	54.36	41.10	8.150	119.005
6	LEU	175.39	55.54	42.06	8.282	122.423	60	SER	176.23	58.12	63.85	8.054	115.550
7	GLY	177.93	45.29	-	8.310	108.855	61	VAL	174.67	62.36	32.65	8.104	121.517
8	ASN	173.88	53.05	38.74	8.206	118.377	62	ASP	175.88	54.23	41.31	8.344	123.742

continued on next page

Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$		
9	ASN	174.93	53.20	38.72	8.392	119.244	63	SER	176.41	58.59	63.47	8.211	117.286
10	VAL	175.39	62.74	32.41	8.040	119.993	64	ARG	174.76	56.31	30.65	8.244	122.322
11	ASN	176.08	53.28	38.87	8.441	121.238	65	ASP	176.05	54.26	41.01	8.188	120.591
12	GLY	175.62	45.49	-	8.209	109.026	66	ILE	175.63	58.61	38.57	7.933	122.719
13	ASN	173.94	53.14	38.93	8.215	118.403	67	PRO	173.81	63.39	32.05	-	-
14	HIS	176.79	56.03	30.08	-	-	68	GLY	177.39	44.92	-	8.355	109.019
15	LEU	174.75	55.03	42.11	8.047	123.038	69	LEU	173.82	52.94	41.72	7.961	122.747
16	ILE	176.79	58.21	38.44	8.060	123.940	70	PRO	174.11	63.09	31.92	-	-
17	PRO	-	-	-	-	-	71	THR	177.00	61.80	69.86	8.119	114.176
18	PRO	175.05	62.55	31.93	-	-	72	ASN	174.12	51.30	38.79	8.378	121.881
19	ALA	176.22	50.20	18.04	8.248	125.646	73	PRO	174.87	63.42	31.96	-	-
20	PRO	-	-	-	-	-	74	SER	177.24	58.66	63.63	8.293	115.482
21	PRO	175.65	62.56	32.14	-	-	75	ARG	174.84	56.23	30.50	8.151	122.760
22	LEU	176.70	52.98	41.61	8.240	123.471	76	LEU	176.30	55.09	42.16	8.044	122.708
23	PRO	174.84	63.36	32.00	-	-	77	ALA	177.22	52.47	19.04	8.116	124.662
24	SER	176.94	58.51	63.55	8.289	115.438	78	ALA	177.45	52.37	19.12	8.090	122.996
25	GLN	174.84	55.87	29.42	8.421	122.360	79	ALA	177.57	52.46	19.17	8.203	123.284
26	THR	176.28	61.80	69.87	8.130	114.461	80	THR	178.04	61.60	69.90	8.044	112.838
27	ASP	174.46	54.45	41.12	8.267	122.766	81	SER	174.68	58.49	63.66	8.245	117.550
28	GLY	176.24	45.49	-	8.271	114.034	82	GLU	174.66	56.86	30.14	8.439	122.658
29	ALA	174.32	52.58	19.01	8.054	123.606	83	THR	176.85	62.37	69.61	8.098	114.740
30	ALA	177.81	52.60	18.95	8.117	122.722	84	CYS	174.57	58.62	28.03	8.267	121.945
31	ARG	177.98	56.11	30.66	8.198	120.102	85	LEU	174.50	55.23	42.12	8.270	124.801
32	GLY	176.91	45.38	-	8.326	109.630	86	LEU	177.33	55.17	42.27	8.162	122.648
33	GLY	174.66	45.21	-	8.250	108.718	87	GLY	177.78	45.34	-	8.309	109.700
34	THR	174.57	61.79	69.68	8.129	112.341	88	GLY	174.45	45.16	-	8.155	108.389
35	GLY	175.46	45.45	-	8.450	111.077	89	PHE	173.88	57.66	39.66	7.949	119.363
36	HIS	173.90	56.09	30.42	8.139	118.905	90	GLU	175.56	56.48	30.27	8.383	122.419
37	LEU	175.03	55.01	42.19	8.110	123.423	91	VAL	175.99	62.17	32.59	8.074	121.751
38	ILE	177.14	61.06	38.62	8.117	122.193	92	LEU	175.80	54.80	42.22	8.193	125.728
39	SER	176.22	57.99	63.79	8.321	119.799	93	HIS	176.95	55.61	30.14	8.279	120.230
40	SER	174.75	58.30	63.68	8.411	118.448	94	ASP	174.61	54.27	41.16	8.292	121.691
41	THR	174.97	61.83	69.82	8.117	114.827	95	LYS	176.08	56.16	32.88	8.318	121.639
42	GLY	175.29	45.23	-	8.253	111.010	96	GLY	176.81	44.62	-	8.361	110.203
43	ALA	173.94	52.63	19.20	8.108	123.757	97	PRO	177.35	63.38	32.07	-	-
44	LEU	178.04	55.48	42.13	8.240	121.229	98	LEU	177.38	55.07	42.12	8.275	120.895
45	GLY	178.19	45.35	-	8.320	109.467	99	ASP	177.33	54.41	41.09	8.108	121.150
46	SER	174.37	58.42	63.71	8.137	115.643	100	ILE	176.43	61.50	38.57	7.869	120.254
47	ARG	174.81	56.17	30.59	8.313	122.658	101	LEU	176.40	55.35	42.09	8.133	124.460
48	SER	176.35	58.26	63.70	8.222	116.696	102	ASN	177.24	53.25	38.70	8.247	119.293
49	LEU	174.48	55.22	42.17	8.149	124.030	103	THR	175.44	62.04	69.58	7.987	114.191
50	PHE	176.84	57.37	39.64	8.068	120.179	104	GLN	174.47	55.65	29.16	8.275	122.757
51	SER	175.24	55.71	63.59	8.017	119.038	105	ILE	175.66	61.00	38.70	8.109	122.517
52	PRO	177.43	63.26	32.05	-	-	106	GLY	176.32	44.36	-	8.269	113.408
53	LEU	176.80	55.16	42.07	8.117	121.285	107	PRO	178.72	63.25	32.09	-	-
54	ARG	177.44	55.96	30.74	8.142	121.538	108	SER	176.50	59.93	64.82	7.993	121.962

Table B.11: Assigned chemical shifts of Map84 in complex with CesT at 303 K based on 3D HNCO, HN(CA)CO, CBCA(CO)NH, HNCACB, HNCA and 2D HSQC experiments.

Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$	
1	MET	-	-	-	-	43	MET	-	-	-	-	
2	PHE	-	-	-	-	44	GLN	-	-	-	-	
3	SER	-	-	-	-	45	SER	-	-	-	-	
4	PRO	63.56	32.15	177.94	-	46	ASN	-	-	-	-	
5	THR	61.95	69.89	174.38	113.008	8.053	47	LEU	-	-	-	-
6	ALA	52.61	19.36	177.55	125.985	8.132	48	MET	-	-	-	-
7	MET	55.53	32.87	176.07	119.740	8.286	49	ILE	-	-	-	-
8	VAL	62.59	32.75	176.54	121.220	8.040	50	ASN	-	-	-	-
9	GLY	45.44	-	173.16	113.116	8.508	51	HIS	-	-	-	-
10	ARG	57.26	31.63	170.99	125.569	7.789	52	GLY	-	-	-	-

continued on next page

	Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$		Res.	C'	C $_{\alpha}$	C $_{\beta}$	N	H $_N$
11	ALA	-	-	-	-	-	53	LYS	-	-	-	-	-
12	LEU	-	-	-	-	-	54	LEU	-	-	-	-	-
13	ALA	52.36	19.17	178.98	-	-	55	THR	-	-	-	-	-
14	GLN	55.55	-	175.44	120.273	8.411	56	THR	-	-	-	-	-
15	ALA	52.49	19.39	177.75	125.660	8.365	57	GLN	-	-	-	-	-
16	VAL	62.25	32.92	176.25	119.431	8.151	58	LEU	-	-	-	-	-
17	THR	61.85	69.91	173.46	118.895	8.278	59	LEU	55.07	-	179.82	-	-
18	GLN	57.45	30.57	180.35	127.704	8.035	60	GLN	58.00	-	177.63	117.187	7.498
19	THR	-	-	-	-	-	61	ALA	54.50	-	-	122.607	7.880
20	LEU	-	-	-	-	-	62	VAL	64.41	-	177.38	117.385	7.9
21	ARG	-	-	-	-	-	63	ALA	53.83	-	178.50	123.767	7.741
22	PRO	63.08	32.08	176.44	-	-	64	LYS	56.53	32.72	176.81	117.125	7.6
23	ALA	52.42	19.25	177.56	124.602	8.402	65	GLN	56.32	29.48	176.29	121.041	7.915
24	VAL	62.46	32.92	176.34	120.130	8.268	66	THR	62.25	69.91	174.57	116.612	8.3
25	THR	61.86	69.87	174.51	117.568	8.161	67	ARG	55.97	31.13	176.36	123.499	8.375
26	LYS	56.54	32.92	176.19	123.753	8.317	68	SER	58.58	63.86	174.78	117.587	8.4
27	ALA	52.54	-	177.63	125.458	8.329	69	SER	58.64	63.81	174.34	117.745	8.4
28	ALA	52.74	19.19	178.16	123.028	8.277	70	ASP	54.53	41.17	176.30	122.094	8.370
29	THR	62.03	69.86	174.70	112.416	8.026	71	THR	62.28	69.59	174.49	114.142	8.014
30	GLN	55.99	29.55	175.66	-	-	72	GLN	56.22	32.91	175.65	122.305	8.266
31	ALA	52.91	19.22	178.23	124.797	8.319	73	GLN	-	-	-	-	-
32	GLY	45.47	-	174.22	108.358	8.386	74	TRP	-	-	-	-	-
33	MET	55.58	33.38	175.98	119.683	8.105	75	PHE	-	-	-	-	-
34	ALA	52.46	19.37	177.69	125.388	8.445	76	LYS	-	-	-	-	-
35	ALA	52.70	19.33	178.05	124.166	8.494	77	GLN	-	-	-	-	-
36	SER	58.98	64.10	175.22	114.813	8.414	78	GLU	-	-	-	-	-
37	GLY	45.31	-	173.99	110.657	8.343	79	GLN	-	-	-	-	-
38	MET	55.40	29.29	175.74	119.993	8.248	80	ILE	-	-	-	-	-
39	ARG	52.04	19.98	177.50	125.967	8.523	81	THR	-	-	-	-	-
40	PHE	-	-	-	-	-	82	PHE	-	-	-	-	-
41	THR	-	-	-	-	-	83	LEU	55.10	42.62	176.17	-	-
42	PRO	-	-	-	-	-	84	SER	59.94	64.88	178.56	122.015	7.808



C

Appendix: NMR restraints

In this appendix the experimentally in Pf1 medium determined RDCs (C.1) and the distance restraints (C.2) as well as the dihedral angle restraints (C.3) for the CylR2 solution structure determination are tabulated.

C.1 Experimental residual dipolar couplings

Table C.1: HN-RDCs of CylR2 measured at 298 K, 800 MHz, a Pf1 concentration of ~ 12 mg/ml and at 0.5 M NaCl.

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
2	HN	2	N	17.46	33	HN	33	N	-37.35
3	HN	3	N	17.93	34	HN	34	N	-51.42
4	HN	4	N	15.39	35	HN	35	N	-40.61
5	HN	5	N	16.20	36	HN	36	N	-20.63
6	HN	6	N	-3.24	37	HN	37	N	-29.24
7	HN	7	N	-25.41	38	HN	38	N	-33.25
8	HN	8	N	-16.49	39	HN	39	N	15.26
9	HN	9	N	2.58	43	HN	43	N	17.97
10	HN	10	N	-17.96	44	HN	44	N	19.35
11	HN	11	N	-35.67	45	HN	45	N	-2.01
12	HN	12	N	-9.13	46	HN	46	N	9.43
13	HN	13	N	3.45	47	HN	47	N	24.63
14	HN	14	N	-50.81	48	HN	48	N	12.21
15	HN	15	N	-39.17	49	HN	49	N	2.43
16	HN	16	N	-50.14	50	HN	50	N	18.90
17	HN	17	N	18.52	51	HN	51	N	20.20
18	HN	18	N	30.33	52	HN	52	N	8.15
19	HN	19	N	8.79	53	HN	53	N	9.33
20	HN	20	N	3.35	54	HN	54	N	9.76
21	HN	21	N	25.43	55	HN	55	N	19.34
22	HN	22	N	27.48	57	HN	57	N	13.65
23	HN	23	N	6.11	58	HN	58	N	13.07
24	HN	24	N	19.13	59	HN	59	N	-25.85
25	HN	25	N	16.17	60	HN	60	N	21.35
26	HN	26	N	26.91	61	HN	61	N	13.11
27	HN	27	N	2.08	62	HN	62	N	25.69
29	HN	29	N	-25.88	64	HN	64	N	18.48
31	HN	31	N	-40.69	66	HN	66	N	-15.25
32	HN	32	N	-27.32					

Table C.2: HN-RDCs of DNA-bound CylR2 measured at 298 K, 700 MHz and at a Pf1 concentration of ~ 12 mg/ml. The IPAP- ^{15}N -HSQC spectrum at 0.5 M NaCl was acquired at 800 MHz and at a Pf1 concentration of 14 mg/ml.

0 M NaCl		0.025 M NaCl		0.05 M NaCl		0.1 M NaCl		0.2 M NaCl		0.3 M NaCl		0.5 M NaCl	
Res.	D [Hz]	Res.	D [Hz]	Res.	D [Hz]	Res.	D [Hz]	Res.	D [Hz]	Res.	D [Hz]	Res.	D [Hz]
2	11.38	2	10.68	2	10.44	2	9.96	2	9.14	2	7.98	2	-8.37
3	-16.64	3	-16.50	3	-16.62	3	-14.80	3	-9.12	3	-10.06	3	8.41
4	10.94	4	8.40	4	8.86	4	6.22	4	6.92	4	5.80	4	-10.06
5	-21.68	5	-23.26	5	-19.02	5	-19.72	5	-17.36	6	0.80	5	13.37
6	-0.50	6	2.34	6	2.52	6	3.14	6	1.74	7	6.24	6	-1.66
7	11.92	7	11.06	7	9.00	7	7.66	7	7.38	8	-9.12	7	-14.19
8	-12.44	8	-12.30	8	-11.90	8	-10.82	8	-8.12	9	-0.20	8	12.00
9	-0.12	9	0.16	9	0.60	9	-0.18	9	0.26	10	9.38	10	-9.98
10	11.96	10	11.68	10	15.48	10	12.66	10	8.24	12	-10.60	11	-0.35
11	1.36	11	0.58	11	-0.38	11	-0.90	11	0.54	13	5.18	13	-3.75
12	-21.42	12	-20.70	12	-18.64	12	-17.06	12	-15.98	14	12.36	14	-19.41
13	7.52	13	10.80	13	9.92	13	8.22	13	5.78	15	-1.44	15	-4.16
14	19.14	14	14.04	14	14.66	14	10.96	14	9.50	16	9.40	16	-14.20
15	-2.06	15	-0.78	15	-1.76	15	-2.66	15	-2.52	17	-18.16	18	23.60
16	13.66	16	9.40	16	11.32	16	8.44	16	10.42	18	-18.26	20	-7.54
17	-17.88	17	-18.54	17	-16.28	17	-18.60	17	-8.74	19	0.62	21	21.02
18	-23.04	18	-29.28	18	-22.04	18	-28.64	18	-21.32	20	-3.90	22	18.70
19	4.00	19	3.44	19	-2.22	19	1.86	19	-5.16	21	-14.92	23	-7.94
20	4.60	20	1.18	20	4.98	20	-1.04	20	-0.44	22	-12.84	24	17.30
21	-22.70	21	-20.58	21	-19.22	21	-17.46	21	-18.52	24	-8.42	25	23.33
22	-18.96	22	-18.94	22	-15.34	22	-17.40	22	-14.66	25	-14.16	26	8.76
23	11.54	23	15.52	23	13.26	23	10.08	23	11.40	26	-3.78	29	-23.22
24	-10.32	24	-10.62	24	-16.42	24	-12.52	24	-10.88	29	3.60	31	-23.12
25	-26.56	25	-22.72	25	-17.94	25	-20.08	25	-15.38	30	12.38	32	-17.60
26	-6.48	26	-5.90	26	-5.14	26	-4.10	26	-1.90	32	8.96	34	-14.06
29	19.46	29	10.26	29	7.42	29	14.42	29	6.18	33	12.14	35	-14.61
30	21.56	30	18.30	30	14.74	30	11.82	30	10.24	34	12.10	36	-27.32
31	-1.48	31	-2.90	31	-2.44	31	-5.24	31	-1.00	35	9.42	37	-25.07
32	17.54	32	15.84	32	17.70	32	11.30	32	14.14	36	13.50	38	-5.36
33	22.94	33	22.66	33	21.32	33	17.70	33	13.24	37	21.32	39	-4.08
34	11.98	34	16.66	34	15.64	34	8.44	34	8.02	38	6.44	40	7.17
35	5.70	35	8.10	35	10.10	35	6.82	35	7.96	39	5.22	42	1.97
36	20.44	36	20.58	36	18.00	36	15.54	36	12.76	40	-0.96	43	15.03
37	30.08	37	25.12	37	27.84	37	21.52	37	17.82	43	-1.76	44	-15.25
38	-1.52	38	-1.80	38	-0.02	38	1.14	38	1.42	44	7.14	45	-10.20
39	6.30	39	5.74	39	5.78	39	2.00	39	5.58	45	0.28	46	11.39
40	-0.66	40	-0.52	40	-2.50	40	1.28	40	-0.64	47	1.26	47	-1.08
43	-11.26	43	-1.98	43	-5.92	43	-4.46	43	-2.54	48	13.12	49	0.21
44	11.74	44	8.48	44	9.06	44	7.16	44	8.90	49	1.04	50	4.99
45	7.68	45	6.96	45	6.84	45	3.96	45	3.88	50	-7.34	51	-8.33
46	-8.12	46	-8.42	46	-6.00	46	-6.68	46	-6.26	51	7.72	52	-15.24
47	7.58	47	4.46	47	6.04	47	7.12	47	6.74	52	7.74	54	-2.10
48	19.32	48	15.92	48	15.72	48	16.22	48	13.64	54	7.14	57	0.97
49	2.86	49	2.84	49	2.86	49	7.28	49	2.34	55	15.90	59	-25.78
50	-9.00	50	-9.98	50	-10.82	50	-7.62	50	-6.62	57	-3.72	61	-5.11
51	13.08	51	12.66	51	12.06	51	10.60	51	9.96	59	15.90	62	12.29
52	17.94	52	14.52	52	13.00	52	12.78	52	12.68	60	-13.54	63	-10.05
54	15.10	54	14.40	54	14.88	54	13.94	54	11.94	61	-1.98	64	-9.02
55	22.38	55	21.08	55	20.22	55	19.50	55	17.70	62	-11.24	66	-9.63
57	-5.62	57	-6.34	57	-3.40	57	-3.18	57	-5.20	63	5.16		
59	21.04	59	20.04	59	19.06	59	13.98	59	14.90	64	9.66		
60	-16.80	60	-15.94	60	-19.08	60	-19.30	60	-13.70	66	7.38		
61	-0.06	61	1.06	61	0.98	61	-1.26	61	-1.58				
62	-17.44	62	-15.88	62	-12.08	62	-12.10	62	-14.08				
63	8.70	63	9.14	63	7.32	63	6.80	63	7.54				
64	13.64	64	13.22	64	13.22	64	12.48	64	10.68				
66	11.86	66	11.16	66	9.82	66	9.18	66	8.16				

Table C.3: HN-RDCs of CesT measured at 303 K, 700 MHz and at a Pf1 concentration of ~ 10 mg/ml.

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
5	HN	5	N	6.71	67	HN	67	N	8.47
6	HN	6	N	2.48	68	HN	68	N	0.55
7	HN	7	N	-6.59	70	HN	70	N	13.37
8	HN	8	N	-3.88	71	HN	71	N	4.83
9	HN	9	N	-0.43	72	HN	72	N	6.68
10	HN	10	N	4.02	74	HN	74	N	7.91
11	HN	11	N	-2.68	75	HN	75	N	4.36
13	HN	13	N	4.85	77	HN	77	N	8.67
14	HN	14	N	-0.81	78	HN	78	N	5.03
15	HN	15	N	-5.87	80	HN	80	N	0.02
16	HN	16	N	7.69	86	HN	86	N	-6.87
17	HN	17	N	2.99	87	HN	87	N	-2.71
18	HN	18	N	-8.44	89	HN	89	N	1.52
20	HN	20	N	4.80	92	HN	92	N	7.17
23	HN	23	N	6.07	93	HN	93	N	-12.25
24	HN	24	N	1.56	94	HN	94	N	-10.17
26	HN	26	N	-3.36	95	HN	95	N	-13.22
27	HN	27	N	-9.64	96	HN	96	N	-3.00
28	HN	28	N	-5.08	97	HN	97	N	-9.81
29	HN	29	N	-11.37	98	HN	98	N	-8.36
30	HN	30	N	-2.95	99	HN	99	N	5.93
31	HN	31	N	-11.41	101	HN	101	N	-2.81
33	HN	33	N	1.26	102	HN	102	N	-7.33
34	HN	34	N	6.90	103	HN	103	N	6.16
36	HN	36	N	9.99	104	HN	104	N	8.35
37	HN	37	N	-2.67	105	HN	105	N	1.03
38	HN	38	N	-1.69	107	HN	107	N	3.49
39	HN	39	N	-18.74	108	HN	108	N	5.94
40	HN	40	N	-11.45	109	HN	109	N	2.19
41	HN	41	N	-11.65	110	HN	110	N	6.53
42	HN	42	N	-3.18	111	HN	111	N	7.70
43	HN	43	N	-1.18	112	HN	112	N	5.84
44	HN	44	N	5.01	113	HN	113	N	2.84
45	HN	45	N	2.79	114	HN	114	N	12.51
46	HN	46	N	-4.29	115	HN	115	N	3.65
47	HN	47	N	6.20	116	HN	116	N	3.26
48	HN	48	N	10.05	117	HN	117	N	7.43
49	HN	49	N	5.93	118	HN	118	N	6.30
51	HN	51	N	-9.74	119	HN	119	N	5.39
52	HN	52	N	-9.10	120	HN	120	N	7.97
53	HN	53	N	-17.98	121	HN	121	N	14.14
54	HN	54	N	-11.09	122	HN	122	N	11.54
55	HN	55	N	-8.66	123	HN	123	N	7.83
56	HN	56	N	1.94	124	HN	124	N	5.26
57	HN	57	N	-8.10	125	HN	125	N	7.36
58	HN	58	N	12.66	126	HN	126	N	8.46
60	HN	60	N	-11.47	127	HN	127	N	4.67
61	HN	61	N	-6.21	128	HN	128	N	6.57
65	HN	65	N	4.10	133	HN	133	N	8.38
66	HN	66	N	6.53	134	HN	134	N	6.17

Table C.4: NC'-RDCs of CesT measured at 303 K, 700 MHz and at a Pf1 concentration of ~ 10 mg/ml.

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
4	N	3	C'	-1.06	69	N	68	C'	-0.22
5	N	4	C'	0.34	71	N	70	C'	1.00

continued on next page

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
7	N	6	C'	0.25	73	N	72	C'	1.79
8	N	7	C'	-1.47	74	N	73	C'	-0.32
9	N	8	C'	0.89	77	N	76	C'	-1.57
10	N	9	C'	0.17	79	N	78	C'	-0.42
11	N	10	C'	-0.20	81	N	80	C'	0.46
13	N	12	C'	0.30	86	N	85	C'	1.71
14	N	13	C'	-0.39	87	N	86	C'	-1.23
15	N	14	C'	-1.47	88	N	87	C'	0.82
16	N	15	C'	1.64	89	N	88	C'	-0.39
17	N	16	C'	-0.34	90	N	89	C'	0.74
19	N	18	C'	0.17	91	N	90	C'	-2.03
20	N	19	C'	0.28	92	N	91	C'	0.17
21	N	20	C'	-0.70	93	N	92	C'	0.87
22	N	21	C'	0.78	95	N	94	C'	0.14
23	N	22	C'	-0.50	96	N	95	C'	-0.97
25	N	24	C'	1.21	97	N	96	C'	0.03
26	N	25	C'	1.08	98	N	97	C'	-0.72
28	N	27	C'	-0.49	99	N	98	C'	0.08
30	N	29	C'	1.15	101	N	100	C'	0.80
33	N	32	C'	-0.93	102	N	101	C'	-1.21
38	N	37	C'	0.03	103	N	102	C'	0.05
39	N	38	C'	0.87	104	N	103	C'	0.57
40	N	39	C'	-0.78	105	N	104	C'	0.18
41	N	40	C'	-0.26	107	N	106	C'	0.40
42	N	41	C'	-1.40	108	N	107	C'	-0.15
43	N	42	C'	0.55	109	N	108	C'	-0.22
45	N	44	C'	0.32	110	N	109	C'	-0.09
46	N	45	C'	0.86	111	N	110	C'	-0.46
47	N	46	C'	-0.73	112	N	111	C'	0.91
48	N	47	C'	0.54	113	N	112	C'	-1.85
49	N	48	C'	0.21	114	N	113	C'	0.94
50	N	49	C'	0.43	115	N	114	C'	-1.30
51	N	50	C'	1.46	116	N	115	C'	-0.06
52	N	51	C'	-0.14	117	N	116	C'	-1.05
54	N	53	C'	-0.84	118	N	117	C'	-0.14
55	N	54	C'	-0.82	119	N	118	C'	0.75
56	N	55	C'	0.88	120	N	119	C'	-0.38
57	N	56	C'	0.17	121	N	120	C'	-0.06
60	N	59	C'	0.39	122	N	121	C'	-1.15
61	N	60	C'	0.16	123	N	122	C'	1.16
62	N	61	C'	0.39	124	N	123	C'	-1.61
64	N	63	C'	0.78	125	N	124	C'	-0.18
65	N	64	C'	-1.19	126	N	125	C'	0.13
66	N	65	C'	1.49	128	N	127	C'	-1.58
67	N	66	C'	-0.09	133	N	132	C'	-0.33
68	N	67	C'	-0.97	134	N	133	C'	0.36

Table C.5: $C\alpha C'$ -RDCs of CesT measured at 303 K, 600 MHz and at a Pf1 concentration of ~ 10 mg/ml.

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
4	$C\alpha$	4	C'	0.92	57	$C\alpha$	57	C'	0.09
6	$C\alpha$	6	C'	1.30	59	$C\alpha$	59	C'	-1.61
7	$C\alpha$	7	C'	1.28	60	$C\alpha$	60	C'	1.07
8	$C\alpha$	8	C'	-0.93	61	$C\alpha$	61	C'	-0.35
9	$C\alpha$	9	C'	-1.86	64	$C\alpha$	64	C'	0.89
10	$C\alpha$	10	C'	1.47	65	$C\alpha$	65	C'	1.28
12	$C\alpha$	12	C'	-1.73	66	$C\alpha$	66	C'	0.51
13	$C\alpha$	13	C'	-0.79	68	$C\alpha$	68	C'	1.15
14	$C\alpha$	14	C'	0.96	76	$C\alpha$	76	C'	-1.38
15	$C\alpha$	15	C'	0.63	80	$C\alpha$	80	C'	1.09
16	$C\alpha$	16	C'	-1.82	86	$C\alpha$	86	C'	1.27

continued on next page

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
18	C α	18	C'	0.73	87	C α	87	C'	2.94
20	C α	20	C'	-0.56	90	C α	90	C'	0.91
22	C α	22	C'	0.03	91	C α	91	C'	0.30
25	C α	25	C'	-1.07	92	C α	92	C'	0.24
26	C α	26	C'	2.99	93	C α	93	C'	1.48
29	C α	29	C'	0.02	96	C α	96	C'	2.67
30	C α	30	C'	-0.45	97	C α	97	C'	-0.09
32	C α	32	C'	2.38	100	C α	100	C'	-0.92
36	C α	36	C'	-1.66	103	C α	103	C'	-0.43
39	C α	39	C'	3.18	104	C α	104	C'	-1.27
40	C α	40	C'	0.67	107	C α	107	C'	-1.78
41	C α	41	C'	-0.30	108	C α	108	C'	1.68
42	C α	42	C'	-1.09	109	C α	109	C'	0.95
44	C α	44	C'	-0.42	110	C α	110	C'	0.73
45	C α	45	C'	-0.72	112	C α	112	C'	-1.22
46	C α	46	C'	1.46	114	C α	114	C'	-2.69
47	C α	47	C'	0.75	115	C α	115	C'	1.06
48	C α	48	C'	0.56	116	C α	116	C'	-0.89
49	C α	49	C'	1.24	117	C α	117	C'	-0.81
50	C α	50	C'	-2.46	121	C α	121	C'	-0.31
51	C α	51	C'	1.40	123	C α	123	C'	0.06
52	C α	52	C'	0.03	124	C α	124	C'	1.35
53	C α	53	C'	1.53	126	C α	126	C'	1.83
55	C α	55	C'	1.13	130	C α	130	C'	-0.96
56	C α	56	C'	-0.67					

Table C.6: HN-RDCs of Tir108 measured at 293 K, 700 MHz and at a Pfl concentration of ~ 10 mg/ml.

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
3	HN	3	N	0.14	57	HN	57	N	3.41
4	HN	4	N	1.06	58	HN	58	N	2.50
6	HN	6	N	1.07	60	HN	60	N	0.21
7	HN	7	N	-3.34	61	HN	61	N	-0.52
8	HN	8	N	4.25	62	HN	62	N	1.21
9	HN	9	N	1.20	63	HN	63	N	0.56
10	HN	10	N	3.55	64	HN	64	N	-0.22
11	HN	11	N	-0.93	65	HN	65	N	0.57
12	HN	12	N	0.50	66	HN	66	N	1.34
15	HN	15	N	1.33	68	HN	68	N	1.42
16	HN	16	N	4.97	69	HN	69	N	0.78
19	HN	19	N	3.05	71	HN	71	N	2.21
22	HN	22	N	1.77	72	HN	72	N	0.62
24	HN	24	N	0.58	74	HN	74	N	1.02
25	HN	25	N	1.08	75	HN	75	N	-0.20
26	HN	26	N	-0.81	76	HN	76	N	0.15
27	HN	27	N	0.06	77	HN	77	N	0.43
28	HN	28	N	5.86	78	HN	78	N	1.76
29	HN	29	N	1.07	79	HN	79	N	2.68
30	HN	30	N	-2.58	80	HN	80	N	4.09
31	HN	31	N	2.14	81	HN	81	N	4.83
32	HN	32	N	2.28	82	HN	82	N	0.22
33	HN	33	N	0.70	86	HN	86	N	5.84
34	HN	34	N	-0.83	87	HN	87	N	0.71
35	HN	35	N	1.63	88	HN	88	N	4.12
36	HN	36	N	3.65	89	HN	89	N	4.64
37	HN	37	N	-0.05	90	HN	90	N	2.14
38	HN	38	N	0.93	91	HN	91	N	3.98
39	HN	39	N	4.26	92	HN	92	N	2.62
41	HN	41	N	3.48	93	HN	93	N	-0.60
42	HN	42	N	2.84	94	HN	94	N	5.00
43	HN	43	N	2.97	95	HN	95	N	0.43

continued on next page

Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]	Residue 1	Atom 1	Residue 2	Atom 2	D [Hz]
44	HN	44	N	-0.39	96	HN	96	N	-0.50
45	HN	45	N	-2.75	98	HN	98	N	-0.14
46	HN	46	N	-0.15	99	HN	99	N	-0.83
47	HN	47	N	6.87	100	HN	100	N	1.07
48	HN	48	N	0.64	101	HN	101	N	0.78
49	HN	49	N	-0.25	102	HN	102	N	-1.75
50	HN	50	N	0.49	103	HN	103	N	2.70
51	HN	51	N	-0.52	104	HN	104	N	-0.71
53	HN	53	N	-0.91	105	HN	105	N	0.35
54	HN	54	N	2.33	106	HN	106	N	0.36
56	HN	56	N	3.41	108	HN	108	N	0.64

C.2 Distance restraints for CylR2

Automatically assigned intramonomer proton distances are listed in Table C.7 and intermonomer distances from PRE and NOE can be found in Table C.8 and C.9, respectively.

Table C.7: Intramolecular NOE distances of CylR2 which were automatically assigned using Cyana 2.0.

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
43	HN	43	HD2*	4.0	33	HA2	36	HB2	4.0
36	HN	36	HG1	4.0	47	HB1	47	HD1*	4.0
43	HD2*	63	HH2	4.0	47	HB1	48	HB1	4.0
36	HA	36	HG1	4.0	60	HB	61	HB2	4.0
36	HA	36	HG2	4.0	10	HA	15	HB	4.0
43	HA	43	HD2*	4.0	6	HB1	60	HG2*	4.0
43	HB2	43	HD2*	4.0	53	HB2	60	HD1*	4.0
34	HA	38	HN	4.0	29	HA	32	HB1	2.0
14	HN	14	HG2	4.0	30	HN	32	HB1	4.0
14	HN	14	HG1	4.0	57	HB2	58	HG2	4.0
13	HN	13	HG2	4.0	58	HA	58	HG2	4.0
8	HN	8	HD2*	4.0	9	HB	55	HG2*	4.0
34	HA	39	HN	4.0	62	HG1	63	HN	4.0
5	HA	8	HD2*	4.0	17	HN	17	HG1	4.0
33	HA2	34	HA	4.0	49	HG2*	49	HG11	3.0
34	HA	39	HB1	4.0	36	HB1	36	HD2	2.0
34	HA	34	HG2*	4.0	41	HB1	42	HN	4.0
38	HN	38	HG2	4.0	36	HB2	38	HN	3.0
38	HN	38	HG1	4.0	62	HN	62	HB2	4.0
8	HD2*	12	HN	4.0	63	HN	63	HB1	3.0
38	HA	38	HG2	4.0	36	HN	36	HB2	2.0
38	HA	38	HG1	4.0	45	HG	48	HD2	4.0
13	HA	14	HG2	4.0	21	HN	31	HG12	4.0
8	HB1	8	HD2*	3.0	53	HN	53	HD2*	4.0
23	HB1	23	HD1*	4.0	43	HD1*	57	HD2*	4.0
42	HB2	43	HN	4.0	29	HB2	30	HN	4.0
42	HB1	43	HN	4.0	31	HN	31	HG12	4.0
8	HD1*	10	HN	4.0	23	HN	23	HG	4.0
8	HD1*	11	HN	4.0	12	HN	13	HG2	4.0
8	HD1*	9	HN	4.0	5	HB1	60	HG2*	4.0
8	HN	8	HD1*	4.0	9	HG2*	60	HG2*	4.0
8	HA	8	HD1*	4.0	2	HD1*	63	HA	2.0
8	HD1*	12	HE2	4.0	15	HD1*	16	HN	4.0
8	HD1*	12	HE1	4.0	3	HD1*	64	HN	4.0
5	HB2	8	HD1*	4.0	51	HN	52	HD2	4.0
8	HD1*	11	HB1	4.0	34	HN	39	HD1	4.0
8	HB2	8	HD1*	4.0	34	HN	39	HD2	4.0
45	HB2	45	HD2*	3.0	38	HN	39	HD1	4.0
5	HN	60	HA	4.0	39	HD1	40	HN	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
24	HD2*	49	HN	4.0	39	HD2	40	HN	4.0
45	HD2*	46	HN	4.0	49	HA	52	HE2	4.0
6	HD2*	9	HN	4.0	58	HG2	63	HZ2	4.0
45	HN	45	HD2*	4.0	3	HG12	4	HN	4.0
24	HN	24	HD2*	4.0	2	HA	3	HA	4.0
24	HA	24	HD2*	3.0	3	HA	4	HA	4.0
6	HA	60	HA	4.0	24	HG	49	HA	4.0
6	HA	6	HD2*	3.0	51	HE1	52	HA	4.0
24	HD2*	49	HA	4.0	50	HB*	51	HD1	4.0
5	HB1	60	HA	4.0	24	HD2*	51	HD1	4.0
5	HB2	60	HA	4.0	60	HN	61	HD2	4.0
24	HD2*	48	HE1	4.0	51	HD1	52	HD2	4.0
24	HD2*	48	HE2	4.0	30	HB	31	HN	4.0
24	HD2*	49	HG12	3.0	30	HB	31	HB	4.0
6	HD2*	9	HB	3.0	30	HN	30	HG2*	4.0
55	HG2*	60	HA	4.0	24	HN	26	HG1*	4.0
60	HA	60	HD1*	4.0	20	HB2	20	HG	2.0
24	HD2*	48	HG1	3.0	48	HG1	49	HG12	4.0
24	HD2*	48	HG2	4.0	60	HG11	61	HN	4.0
53	HN	53	HG	4.0	2	HG12	3	HN	4.0
48	HG1	49	HA	4.0	59	HN	60	HG11	4.0
47	HG	48	HN	4.0	50	HA	60	HG11	4.0
53	HA	53	HG	3.0	56	HD2	60	HG11	4.0
44	HA	47	HG	4.0	57	HA	60	HG11	4.0
43	HN	43	HD1*	4.0	2	HG12	63	HB2	4.0
9	HA	12	HN	4.0	43	HD2*	57	HD1*	4.0
52	HA	52	HD1	4.0	48	HA	52	HD2	4.0
43	HA	43	HD1*	3.0	47	HB1	48	HE1	4.0
43	HB2	43	HD1*	4.0	48	HE1	49	HN	4.0
23	HD1*	52	HA	4.0	50	HB*	51	HB2	4.0
23	HD2*	52	HA	4.0	50	HB*	51	HB1	4.0
9	HA	9	HG11	2.0	56	HB2	58	HG1	3.0
20	HN	20	HG	4.0	43	HD2*	61	HB2	4.0
20	HA	20	HG	4.0	63	HB2	64	HN	4.0
23	HB1	23	HD2*	4.0	63	HB1	64	HN	4.0
20	HG	21	HN	4.0	63	HB2	64	HB2	4.0
50	HB*	51	HA	4.0	2	HG11	63	HB2	4.0
31	HA	31	HG2*	3.0	63	HB1	63	HZ3	4.0
26	HN	26	HG1*	4.0	62	HN	63	HB1	4.0
51	HA	51	HD2	4.0	57	HB2	63	HZ2	4.0
26	HA	26	HG1*	4.0	52	HN	52	HE2	4.0
24	HB2	26	HG1*	4.0	63	HN	63	HE3	4.0
24	HB1	26	HG1*	4.0	33	HA1	39	HD1	4.0
10	HA	12	HN	4.0	11	HB1	11	HG2	3.0
17	HA	20	HN	4.0	40	HA	41	HB1	4.0
3	HA	4	HN	3.0	41	HD2	42	HN	4.0
3	HA	4	HB2	4.0	41	HD1	42	HN	4.0
48	HA	52	HN	4.0	61	HB2	63	HH2	4.0
48	HA	51	HN	4.0	48	HN	48	HG2	4.0
55	HG2*	60	HN	4.0	2	HD1*	63	HN	4.0
48	HA	52	HE2	4.0	2	HA	64	HN	4.0
4	HD21	35	HA	4.0	2	HA	63	HA	3.0
48	HA	51	HD1	4.0	1	HA	2	HG12	4.0
55	HA	55	HG2*	4.0	2	HG2*	63	HA	4.0
55	HG2*	59	HA	4.0	3	HA	4	HB1	4.0
48	HA	51	HB1	4.0	3	HB	4	HA	4.0
55	HG2*	59	HB1	4.0	2	HN	3	HG11	4.0
55	HG2*	60	HG11	4.0	2	HA	3	HG11	4.0
48	HA	48	HB1	2.0	3	HD1*	63	HA	4.0
57	HA	57	HG	4.0	4	HA	62	HN	4.0
57	HA	60	HD1*	4.0	4	HA	4	HD21	4.0
55	HG2*	60	HD1*	4.0	4	HA	60	HA	4.0
55	HG2*	60	HG2*	4.0	41	HB2	61	HE1	4.0
48	HA	48	HG2	4.0	63	HA	63	HE3	4.0
7	HN	35	HA	4.0	2	HG11	63	HA	4.0
21	HN	28	HA	4.0	63	HA	64	HA	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
28	HA	31	HN	4.0	2	HA	63	HB1	4.0
2	HA	2	HG11	4.0	62	HA	63	HB1	4.0
43	HN	46	HB*	4.0	2	HA	63	HB2	4.0
29	HA	32	HN	4.0	63	HA	64	HB2	4.0
45	HN	46	HB*	4.0	3	HN	64	HG2	4.0
46	HB*	61	HE2	4.0	3	HN	64	HG1	4.0
46	HB*	61	HD2	4.0	43	HA	46	HN	4.0
11	HA	11	HG2	4.0	43	HG	44	HN	4.0
46	HB*	57	HD1*	4.0	31	HG12	32	HA	4.0
45	HB1	46	HB*	4.0	13	HG2	14	HN	4.0
46	HB*	60	HB	4.0	13	HG1	14	HN	4.0
45	HB2	46	HB*	4.0	10	HA	13	HG1	4.0
43	HD1*	46	HB*	3.0	5	HA	8	HD1*	4.0
43	HD2*	46	HB*	4.0	8	HD1*	9	HA	4.0
29	HA	31	HN	4.0	43	HB1	43	HD1*	4.0
29	HA	30	HA	4.0	9	HA	11	HN	4.0
19	HA	22	HN	4.0	30	HA	31	HA	4.0
49	HD1*	61	HE2	4.0	31	HA	34	HA	4.0
49	HA	49	HD1*	4.0	21	HB*	31	HA	4.0
49	HB	49	HD1*	4.0	10	HA	13	HN	4.0
45	HB1	49	HD1*	3.0	35	HN	35	HA	3.0
19	HA	22	HB*	4.0	15	HA	15	HD1*	4.0
49	HD1*	60	HG2*	4.0	31	HG2*	32	HB2	3.0
43	HA	46	HB*	4.0	47	HB1	48	HN	4.0
5	HN	60	HG2*	4.0	47	HB2	48	HN	4.0
21	HN	26	HG2*	4.0	39	HN	39	HB1	3.0
60	HG2*	61	HN	4.0	39	HN	39	HB2	3.0
60	HN	60	HG2*	4.0	56	HB1	59	HN	4.0
24	HN	26	HG2*	4.0	3	HN	3	HB	4.0
60	HG2*	61	HE2	4.0	2	HA	3	HN	4.0
60	HG2*	61	HD2	4.0	57	HN	57	HB2	4.0
26	HG2*	31	HN	4.0	57	HN	58	HN	4.0
4	HA	60	HG2*	4.0	3	HN	64	HN	4.0
60	HG2*	61	HA	4.0	41	HA	42	HN	3.0
26	HA	26	HG2*	4.0	41	HB2	42	HN	3.0
26	HG2*	30	HA	4.0	2	HG2*	63	HN	4.0
25	HA	26	HG2*	4.0	3	HB	63	HN	4.0
46	HA	60	HG2*	4.0	18	HN	19	HN	4.0
6	HA	60	HG2*	3.0	16	HA	19	HN	4.0
60	HA	60	HG2*	4.0	15	HG2*	19	HN	4.0
5	HB2	60	HG2*	4.0	3	HN	4	HN	4.0
6	HD2*	60	HG2*	3.0	4	HN	5	HN	4.0
60	HG2*	60	HG12	4.0	4	HN	4	HD22	4.0
50	HB*	57	HN	4.0	2	HN	2	HG12	4.0
49	HN	50	HB*	4.0	34	HN	36	HN	4.0
34	HG2*	35	HN	4.0	11	HN	12	HN	3.0
31	HG2*	32	HN	4.0	34	HN	35	HN	4.0
34	HG2*	37	HN	4.0	31	HA	34	HN	4.0
9	HN	9	HG2*	4.0	8	HD1*	12	HN	4.0
17	HE21	31	HG2*	4.0	9	HG2*	12	HN	4.0
9	HG2*	13	HN	4.0	12	HN	12	HA	3.0
17	HE22	31	HG2*	4.0	34	HN	34	HG2*	4.0
34	HG2*	61	HE1	4.0	12	HN	13	HN	2.0
31	HN	31	HG2*	4.0	5	HN	6	HN	4.0
34	HG2*	61	HZ	4.0	4	HA	5	HN	3.0
4	HD21	34	HG2*	4.0	4	HB1	5	HN	4.0
34	HG2*	40	HA	4.0	43	HN	43	HB2	4.0
50	HB*	56	HA	4.0	43	HN	43	HB1	4.0
31	HG2*	32	HA	4.0	6	HN	7	HN	4.0
9	HA	9	HG2*	2.0	35	HN	37	HN	4.0
17	HA	31	HG2*	3.0	35	HN	35	HG2	4.0
28	HA	31	HG2*	3.0	48	HN	48	HB1	4.0
48	HA	50	HB*	4.0	6	HN	6	HD1*	4.0
31	HG2*	32	HB1	4.0	48	HN	48	HG1	4.0
8	HA	11	HG2	4.0	48	HA	50	HN	4.0
34	HG2*	41	HB1	3.0	49	HB	50	HN	3.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
9	HG2*	9	HG12	2.0	6	HN	6	HB1	2.0
9	HG2*	55	HG2*	3.0	45	HG	46	HN	4.0
50	HB*	60	HD1*	3.0	46	HN	47	HB1	4.0
31	HG2*	31	HD1*	3.0	46	HN	46	HB*	4.0
34	HG2*	60	HG2*	4.0	50	HN	50	HB*	3.0
8	HA	11	HN	4.0	6	HN	9	HD1*	4.0
3	HG2*	62	HN	4.0	16	HA	17	HN	4.0
21	HN	22	HB*	4.0	49	HN	49	HD1*	4.0
20	HN	22	HB*	4.0	48	HN	49	HN	4.0
22	HB*	25	HN	4.0	45	HA	49	HN	4.0
22	HN	22	HB*	3.0	49	HN	49	HB	4.0
47	HA	47	HG	2.0	49	HN	49	HG12	4.0
8	HA	8	HD2*	2.0	10	HN	10	HB2	4.0
45	HA	45	HD2*	4.0	9	HB	10	HN	4.0
15	HG2*	22	HB*	4.0	10	HN	10	HB1	4.0
3	HN	3	HG2*	4.0	49	HN	49	HG11	4.0
3	HG2*	5	HN	4.0	6	HD2*	10	HN	4.0
3	HG2*	4	HN	4.0	9	HG2*	10	HN	4.0
3	HG2*	5	HD21	4.0	46	HB*	49	HN	4.0
3	HG2*	5	HD22	4.0	30	HN	31	HN	4.0
3	HG2*	4	HA	4.0	30	HN	30	HB	4.0
2	HA	3	HG2*	4.0	21	HN	22	HN	4.0
3	HA	3	HG2*	4.0	21	HN	31	HG2*	4.0
20	HA	20	HD2*	4.0	20	HN	21	HN	2.0
2	HG2*	3	HN	3.0	20	HN	20	HB2	2.0
2	HN	2	HG2*	4.0	20	HN	20	HB1	4.0
2	HG2*	62	HN	4.0	60	HN	61	HN	4.0
32	HA	34	HN	4.0	60	HN	60	HG12	3.0
20	HN	21	HB*	4.0	44	HB2	45	HN	4.0
15	HG2*	20	HN	4.0	44	HB1	45	HN	4.0
2	HG2*	64	HN	4.0	45	HN	45	HB1	4.0
14	HA	15	HN	3.0	45	HN	45	HG	4.0
15	HN	15	HG2*	4.0	45	HN	45	HD1*	4.0
2	HG2*	61	HD1	4.0	8	HB1	9	HN	2.0
2	HA	2	HG2*	4.0	31	HN	31	HG11	4.0
15	HA	15	HG2*	4.0	9	HN	10	HN	4.0
22	HA	25	HA	4.0	9	HN	9	HG12	4.0
15	HG2*	19	HA	4.0	30	HG2*	31	HN	4.0
2	HG2*	61	HB1	4.0	9	HN	9	HD1*	3.0
15	HG2*	19	HB2	3.0	52	HN	52	HD2	4.0
15	HG2*	19	HB1	3.0	51	HN	52	HN	3.0
3	HN	63	HA	4.0	51	HD1	52	HN	4.0
18	HN	21	HB*	4.0	51	HB2	52	HN	4.0
21	HN	21	HB*	3.0	31	HN	32	HN	4.0
63	HA	64	HN	3.0	32	HN	32	HB1	2.0
21	HB*	22	HN	2.0	7	HN	8	HN	4.0
21	HB*	31	HN	4.0	6	HG	7	HN	4.0
63	HA	63	HD1	4.0	6	HB2	7	HN	4.0
37	HA	38	HA	4.0	59	HN	60	HN	4.0
21	HB*	31	HB	3.0	22	HN	23	HN	4.0
21	HB*	26	HG2*	3.0	56	HA	59	HN	4.0
5	HN	61	HA	4.0	59	HN	59	HB2	4.0
36	HA	38	HN	4.0	59	HN	59	HB1	4.0
50	HA	55	HN	4.0	55	HG2*	59	HN	4.0
61	HA	61	HD1	4.0	58	HN	59	HN	4.0
61	HA	61	HD2	4.0	51	HN	51	HB1	2.0
46	HA	61	HE2	4.0	58	HN	58	HG2	3.0
46	HA	49	HB	4.0	58	HN	58	HG1	4.0
36	HA	36	HD1	4.0	57	HB2	58	HN	4.0
2	HG2*	61	HA	4.0	50	HB*	51	HN	4.0
21	HA	26	HN	4.0	51	HN	51	HB2	3.0
21	HA	24	HN	4.0	38	HB2	39	HN	4.0
50	HA	55	HB	4.0	57	HD2*	58	HN	3.0
13	HA	13	HG2	3.0	61	HA	62	HN	3.0
21	HA	26	HG2*	3.0	61	HB1	62	HN	4.0
1	HA	2	HA	4.0	13	HN	14	HN	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
13	HA	14	HG1	4.0	13	HN	13	HG1	4.0
9	HG2*	13	HA	4.0	8	HN	9	HN	4.0
60	HN	60	HD1*	4.0	11	HA	13	HN	4.0
60	HD1*	61	HE2	4.0	13	HN	13	HB2	4.0
50	HA	60	HD1*	3.0	8	HN	8	HB1	3.0
34	HD1*	41	HB2	4.0	8	HN	8	HB2	3.0
60	HG2*	60	HD1*	4.0	61	HN	61	HD2	4.0
34	HD1*	35	HN	4.0	11	HN	11	HG1	2.0
51	HN	51	HD1	4.0	60	HB	61	HN	4.0
34	HD1*	39	HD2	4.0	47	HN	47	HB1	4.0
34	HD1*	39	HA	4.0	10	HN	11	HN	4.0
51	HD1	52	HA	4.0	47	HN	48	HN	4.0
34	HA	34	HD1*	4.0	61	HN	61	HB2	3.0
6	HA	60	HD1*	4.0	47	HN	47	HB2	3.0
31	HA	34	HD1*	4.0	47	HN	47	HD2*	4.0
34	HD1*	41	HB1	3.0	46	HB*	47	HN	4.0
23	HG	52	HD1	4.0	47	HN	47	HD1*	4.0
24	HD2*	52	HD2	4.0	36	HN	37	HN	4.0
48	HG1	52	HD2	4.0	36	HN	36	HD2	4.0
54	HA	55	HN	3.0	47	HN	57	HG	4.0
54	HA	54	HD21	4.0	36	HN	36	HG2	2.0
39	HA	39	HD1	4.0	43	HN	44	HN	3.0
33	HA2	39	HD1	4.0	22	HN	25	HN	4.0
2	HB	61	HD1	4.0	44	HN	46	HB*	4.0
41	HB2	61	HD1	4.0	44	HN	45	HN	4.0
39	HD2	45	HG	4.0	34	HB	36	HN	4.0
30	HG2*	39	HD2	4.0	43	HD2*	44	HN	4.0
49	HN	49	HG2*	4.0	52	HN	53	HN	4.0
49	HG2*	50	HN	4.0	53	HN	53	HB2	3.0
13	HN	15	HD1*	4.0	37	HN	38	HN	4.0
49	HG2*	61	HE2	4.0	38	HN	39	HN	3.0
49	HG2*	50	HA	4.0	22	HA	25	HN	4.0
10	HA	15	HD1*	4.0	55	HN	55	HG2*	4.0
13	HB2	15	HD1*	4.0	37	HN	37	HB2	4.0
2	HD1*	3	HN	4.0	26	HG2*	28	HN	4.0
2	HD1*	62	HN	4.0	26	HG1*	27	HN	4.0
2	HD1*	61	HD1	4.0	17	HN	18	HN	3.0
2	HA	2	HD1*	4.0	16	HA	18	HN	4.0
2	HD1*	3	HA	4.0	18	HN	18	HB2	3.0
2	HD1*	63	HB1	4.0	18	HN	18	HB1	3.0
3	HN	3	HD1*	4.0	5	HD21	9	HD1*	4.0
2	HN	3	HD1*	4.0	21	HB*	26	HN	4.0
3	HD1*	4	HN	4.0	26	HN	26	HG2*	4.0
3	HD1*	64	HE21	4.0	25	HB2	26	HN	4.0
3	HD1*	64	HE22	4.0	33	HN	34	HN	4.0
2	HA	3	HD1*	4.0	49	HN	50	HN	4.0
3	HA	3	HD1*	4.0	48	HB2	49	HN	4.0
3	HD1*	64	HG2	4.0	47	HB1	49	HN	4.0
3	HD1*	64	HG1	4.0	15	HN	15	HD1*	4.0
5	HA	8	HN	4.0	15	HN	15	HB	3.0
5	HA	60	HA	4.0	50	HB*	52	HN	4.0
5	HA	7	HN	4.0	35	HN	36	HN	4.0
4	HA	61	HD1	4.0	37	HA	38	HN	3.0
56	HD2	59	HA	4.0	5	HB1	5	HD21	3.0
56	HD2	59	HB1	4.0	1	HA	1	HG*	4.0
5	HA	8	HB1	4.0	1	HB*	2	HN	4.0
55	HG2*	56	HD2	4.0	1	HG*	64	HB2	4.0
55	HG2*	56	HD1	4.0	1	HG*	64	HB1	4.0
2	HG2*	4	HA	4.0	2	HA	64	HG*	4.0
9	HD1*	10	HN	4.0	2	HD1*	65	HG*	4.0
9	HD1*	11	HN	4.0	2	HD1*	65	HD*	4.0
9	HD1*	59	HA	4.0	3	HB	64	HE2*	4.0
6	HA	9	HD1*	4.0	3	HG2*	4	HB*	4.0
9	HA	9	HD1*	4.0	3	HG2*	62	HE2*	4.0
9	HD1*	60	HA	4.0	3	HD1*	62	HG*	4.0
9	HD1*	59	HB1	4.0	3	HD1*	62	HE2*	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
8	HD2*	9	HD1*	4.0	3	HD1*	64	HG*	3.0
9	HD1*	60	HG2*	4.0	3	HD1*	64	HE2*	4.0
37	HA	40	HA	4.0	4	HB*	4	HD21	2.0
31	HD1*	33	HN	4.0	4	HB*	4	HD22	3.0
31	HD1*	34	HN	4.0	4	HB*	5	HN	4.0
31	HN	31	HD1*	4.0	4	HB*	6	HN	4.0
31	HD1*	32	HA	4.0	4	HB*	34	HG2*	4.0
31	HA	31	HD1*	3.0	4	HB*	61	HE1	4.0
31	HB	31	HD1*	4.0	4	HB*	62	HN	4.0
20	HB2	31	HD1*	4.0	6	HD1*	10	HB*	4.0
20	HG	31	HD1*	4.0	6	HD1*	35	HG*	3.0
33	HA2	35	HN	4.0	7	HN	7	HB*	3.0
33	HA1	39	HB2	4.0	7	HN	7	HD*	4.0
33	HA1	36	HB2	4.0	7	HN	7	HE*	4.0
33	HA2	36	HG1	4.0	7	HA	7	HG*	4.0
61	HB1	61	HD1	4.0	7	HA	7	HD*	4.0
47	HB2	47	HD1*	4.0	7	HB*	8	HN	4.0
8	HB2	9	HN	4.0	7	HG*	7	HD*	2.0
5	HA	8	HB2	4.0	7	HG*	7	HE*	3.0
5	HB1	8	HB2	4.0	7	HD*	7	HE*	2.0
48	HN	48	HE1	4.0	7	HD*	8	HN	4.0
48	HE2	52	HE2	4.0	7	HD*	8	HA	4.0
57	HB2	58	HG1	4.0	7	HE*	8	HN	4.0
23	HB2	23	HD1*	4.0	8	HD1*	12	HD*	3.0
23	HB2	23	HD2*	4.0	8	HD1*	12	HE*	4.0
4	HB2	5	HN	4.0	9	HA	12	HB*	3.0
4	HN	4	HB2	4.0	9	HA	12	HD*	4.0
4	HN	4	HB1	4.0	9	HG2*	13	HE*	4.0
20	HB2	21	HN	3.0	10	HA	10	HG*	3.0
20	HB1	21	HN	4.0	10	HB*	10	HD*	3.0
6	HB1	6	HD1*	3.0	10	HB*	11	HN	4.0
20	HB1	31	HG2*	4.0	10	HG*	20	HG	4.0
45	HB1	46	HN	3.0	10	HD*	20	HB1	4.0
6	HN	6	HB2	4.0	10	HD*	20	HG	4.0
43	HB1	43	HD2*	4.0	10	HD*	20	HD*	3.0
6	HB2	60	HG2*	4.0	12	HN	12	HB*	3.0
34	HN	39	HB2	4.0	12	HN	12	HG*	4.0
59	HB1	60	HG11	4.0	12	HN	12	HD*	4.0
55	HG2*	59	HB2	4.0	12	HA	12	HB*	2.0
39	HB1	40	HN	4.0	12	HA	12	HG*	2.0
39	HB2	40	HN	4.0	12	HA	12	HD*	4.0
34	HN	39	HB1	4.0	12	HB*	12	HD*	3.0
39	HB1	39	HD2	3.0	12	HB*	13	HN	4.0
39	HB2	39	HD1	3.0	12	HG*	12	HE*	3.0
5	HB2	5	HD22	4.0	12	HG*	13	HN	4.0
5	HB1	5	HD22	4.0	13	HN	13	HD*	4.0
39	HA	39	HB1	3.0	13	HA	13	HD*	3.0
34	HG11	39	HB1	4.0	13	HA	13	HE*	4.0
56	HB2	59	HB2	4.0	13	HA	14	HG*	4.0
5	HB1	8	HB1	4.0	13	HB1	13	HE*	4.0
5	HB2	8	HB1	4.0	13	HB1	14	HG*	4.0
5	HB1	8	HD1*	4.0	13	HG2	13	HE*	3.0
30	HG2*	39	HB1	4.0	14	HN	14	HG*	4.0
60	HN	60	HB	3.0	14	HN	14	HD*	4.0
60	HB	61	HD2	4.0	14	HA	14	HG*	3.0
6	HB2	60	HB	4.0	14	HA	14	HD*	4.0
2	HN	2	HB	4.0	14	HG*	14	HD*	2.0
31	HN	32	HB2	4.0	15	HN	15	HG1*	4.0
31	HN	32	HB1	4.0	15	HA	15	HG1*	4.0
32	HA	32	HB2	3.0	15	HA	16	HB*	4.0
29	HA	32	HB2	3.0	15	HA	19	HB*	4.0
3	HB	4	HN	4.0	15	HB	19	HB*	4.0
3	HB	62	HN	4.0	15	HG2*	19	HB*	3.0
2	HA	3	HB	4.0	15	HG2*	19	HG*	4.0
3	HB	62	HA	4.0	15	HG1*	19	HB*	4.0
58	HA	58	HG1	4.0	16	HN	16	HB*	3.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
51	HB1	52	HN	4.0	16	HN	19	HB*	3.0
34	HB	35	HN	4.0	16	HB*	17	HN	4.0
34	HN	34	HB	4.0	16	HB*	17	HA	4.0
31	HA	34	HB	4.0	16	HB*	17	HB*	4.0
52	HN	52	HB2	3.0	16	HB*	17	HG*	4.0
52	HN	52	HB1	4.0	16	HB*	18	HN	4.0
11	HN	11	HG2	2.0	16	HB*	19	HN	4.0
11	HG2	12	HN	4.0	16	HB*	19	HB*	4.0
52	HB2	52	HD2	3.0	17	HN	17	HB*	4.0
10	HA	11	HG2	4.0	17	HA	17	HG*	4.0
23	HG	52	HB1	4.0	17	HB*	18	HB*	4.0
24	HD2*	52	HB2	4.0	17	HG*	17	HE2*	3.0
35	HN	35	HG1	4.0	17	HG*	31	HG2*	4.0
51	HE1	52	HD1	4.0	17	HG*	32	HA	4.0
51	HE1	52	HD2	4.0	17	HE2*	31	HG2*	4.0
39	HA	39	HE1	4.0	17	HE2*	32	HA	4.0
51	HB1	51	HE1	4.0	18	HN	18	HB*	2.0
48	HB2	52	HE2	3.0	18	HN	28	HD*	4.0
30	HG2*	39	HE2	3.0	18	HB*	19	HN	4.0
24	HD2*	52	HE2	3.0	18	HB*	20	HN	4.0
37	HN	37	HB1	4.0	18	HB*	21	HN	4.0
9	HN	9	HB	4.0	18	HB*	28	HB*	4.0
6	HA	9	HB	4.0	18	HB*	28	HD*	2.0
31	HB	32	HN	4.0	19	HN	19	HB*	3.0
26	HB	27	HN	4.0	19	HA	19	HG*	2.0
31	HN	31	HB	3.0	19	HB*	20	HN	3.0
20	HB1	31	HB	4.0	19	HG*	22	HB*	4.0
26	HB	30	HG2*	4.0	20	HN	20	HD*	4.0
55	HN	55	HB	4.0	20	HN	23	HD*	4.0
55	HB	56	HD2	3.0	20	HA	20	HD*	3.0
55	HB	60	HD1*	4.0	20	HD*	31	HD1*	2.0
44	HN	44	HG2	4.0	21	HA	24	HB*	4.0
44	HN	44	HG1	4.0	21	HB*	28	HD*	4.0
44	HG2	45	HN	4.0	22	HB*	23	HD*	4.0
44	HG1	45	HN	4.0	23	HA	23	HD*	3.0
44	HA	44	HG2	4.0	23	HB*	23	HD*	2.0
44	HA	44	HG1	4.0	23	HB*	52	HE1	4.0
17	HN	17	HG2	4.0	23	HB*	52	HE2	4.0
62	HG2	63	HN	4.0	23	HD*	52	HA	4.0
38	HB1	39	HN	4.0	23	HD*	52	HB2	4.0
64	HA	64	HG2	4.0	23	HD*	52	HE1	4.0
64	HA	64	HG1	4.0	23	HD*	53	HN	4.0
3	HG12	64	HG2	4.0	24	HB*	24	HG	2.0
3	HG12	64	HG1	4.0	24	HB*	24	HD1*	3.0
48	HB1	49	HN	4.0	24	HB*	24	HD2*	2.0
45	HA	48	HB2	4.0	24	HB*	26	HG1*	4.0
48	HB2	48	HE1	4.0	24	HB*	26	HG2*	2.0
48	HB2	48	HE2	4.0	24	HD1*	45	HD*	4.0
24	HD2*	48	HB1	4.0	25	HN	25	HG*	2.0
24	HD2*	48	HB2	4.0	27	HB*	28	HN	4.0
48	HB2	48	HG1	2.0	28	HA	28	HB*	2.0
7	HN	7	HB2	4.0	28	HA	28	HG*	4.0
7	HN	7	HB1	4.0	28	HA	28	HD*	4.0
7	HB2	8	HN	4.0	28	HB*	28	HD*	3.0
7	HB1	8	HN	4.0	28	HB*	29	HN	4.0
62	HN	62	HB1	4.0	29	HN	29	HB*	2.0
36	HN	36	HB1	4.0	29	HN	29	HG*	4.0
56	HB2	59	HN	4.0	29	HA	29	HG*	3.0
41	HB1	61	HE1	4.0	29	HB*	30	HN	3.0
49	HA	49	HG12	4.0	29	HB*	32	HB1	4.0
49	HG2*	49	HG12	2.0	31	HD1*	35	HG*	4.0
63	HN	63	HB2	4.0	34	HG2*	35	HG*	4.0
64	HN	64	HB2	3.0	34	HD1*	41	HG*	3.0
63	HB1	63	HE3	3.0	35	HA	35	HG*	2.0
3	HD1*	64	HB1	4.0	35	HB*	36	HN	3.0
16	HN	19	HB2	4.0	36	HN	36	HE*	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
16	HN	19	HB1	4.0	36	HA	36	HE*	4.0
19	HB2	20	HN	3.0	36	HB2	36	HE*	4.0
19	HB1	20	HN	3.0	36	HB1	36	HE*	3.0
34	HG11	35	HN	4.0	36	HD1	36	HE*	2.0
11	HN	11	HB2	4.0	37	HN	37	HB*	3.0
11	HN	11	HB1	4.0	37	HB*	37	HD21	2.0
11	HB2	12	HN	2.0	37	HB*	38	HA	4.0
34	HN	34	HG11	4.0	38	HN	38	HB*	2.0
24	HG	25	HN	4.0	38	HN	38	HG*	4.0
34	HG11	39	HD2	4.0	38	HB*	39	HN	4.0
11	HA	11	HB1	2.0	38	HD*	38	HE*	2.0
36	HA	36	HD2	3.0	39	HA	40	HB*	4.0
36	HB2	36	HD2	4.0	39	HE2	45	HD*	4.0
6	HD1*	10	HB1	4.0	40	HB*	40	HD2*	3.0
9	HN	9	HG11	4.0	40	HB*	41	HA	4.0
26	HB	30	HA	4.0	41	HG*	42	HN	4.0
30	HA	30	HG2*	3.0	41	HG*	61	HD1	4.0
2	HN	2	HG11	4.0	41	HG*	61	HZ	4.0
29	HB1	30	HN	4.0	42	HB*	43	HN	4.0
24	HA	24	HD1*	4.0	42	HB*	43	HB*	4.0
24	HD1*	49	HG12	4.0	43	HB*	43	HD1*	3.0
3	HG2*	3	HG12	3.0	43	HB*	43	HD2*	3.0
57	HN	57	HG	4.0	43	HB*	44	HN	4.0
3	HN	3	HG11	4.0	44	HN	44	HB*	2.0
3	HN	3	HG12	4.0	44	HN	44	HG*	4.0
3	HA	3	HG11	3.0	44	HA	44	HG*	3.0
6	HD1*	35	HA	4.0	44	HB*	45	HN	4.0
6	HD1*	10	HB2	4.0	44	HG*	45	HN	4.0
55	HG2*	60	HG12	4.0	45	HN	45	HD*	3.0
49	HA	49	HG2*	3.0	45	HA	45	HD*	3.0
49	HA	49	HG11	4.0	45	HB2	45	HD*	2.0
47	HD1*	57	HG	2.0	45	HD*	46	HN	4.0
6	HD1*	60	HG2*	4.0	45	HD*	48	HG2	4.0
31	HG11	32	HN	4.0	45	HD*	48	HG1	4.0
25	HN	25	HB1	4.0	45	HD*	48	HD2	4.0
28	HA	28	HG2	4.0	45	HD*	48	HE2	4.0
31	HA	31	HG11	4.0	45	HD*	48	HE1	4.0
28	HA	28	HG1	4.0	45	HD*	49	HG12	4.0
47	HD1*	48	HN	4.0	49	HA	53	HD*	4.0
20	HA	20	HD1*	4.0	53	HD*	60	HG2*	4.0
47	HA	47	HD1*	4.0	53	HD*	60	HG12	4.0
24	HA	25	HB2	4.0	54	HN	54	HB*	4.0
44	HA	47	HD1*	4.0	54	HB*	54	HD21	3.0
43	HD2*	47	HD1*	4.0	55	HA	56	HG*	4.0
30	HB	32	HN	4.0	56	HG*	59	HA	4.0
45	HD1*	46	HN	4.0	56	HG*	59	HB2	4.0
53	HN	53	HD1*	4.0	58	HB*	63	HH2	4.0
57	HD2*	63	HH2	4.0	61	HA	62	HB*	4.0
47	HA	47	HD2*	4.0	62	HN	62	HB*	3.0
57	HA	57	HD2*	4.0	62	HN	62	HG*	4.0
45	HA	45	HD1*	4.0	62	HA	62	HG*	4.0
57	HB2	57	HD2*	3.0	62	HB*	62	HG*	2.0
45	HB2	45	HD1*	3.0	62	HB*	62	HE2*	4.0
57	HN	57	HD2*	3.0	62	HG*	62	HE2*	3.0
24	HN	24	HD1*	4.0	63	HA	65	HD*	4.0
24	HD1*	49	HD1*	4.0	63	HB2	65	HB*	4.0
43	HD2*	57	HG	4.0	63	HB2	65	HG*	4.0
31	HG2*	31	HG12	3.0	63	HB1	64	HG*	4.0
31	HB	31	HG12	2.0	63	HB1	65	HG*	4.0
55	HB	60	HG11	4.0	64	HN	64	HG*	4.0
60	HN	60	HG11	3.0	64	HN	65	HG*	4.0
33	HA2	39	HB2	4.0	64	HA	64	HG*	4.0

Table C.8: Intermolecular distances of CylR2 from PRE

Residue 1	Atom 1	Residue 2	Atom 2	distance [Å]	Residue 1	Atom 1	Residue 2	Atom 2	distance [Å]
40	NZ	68	HN	30.00	55	NZ	68	HN	22.71
40	NZ	69	HN	30.00	55	NZ	69	HN	24.59
40	NZ	70	HN	28.60	55	NZ	70	HN	24.30
40	NZ	71	HN	30.00	55	NZ	71	HN	24.15
40	NZ	72	HN	30.00	55	NZ	72	HN	23.20
40	NZ	73	HN	30.00	55	NZ	73	HN	27.76
40	NZ	74	HN	29.89	55	NZ	74	HN	30.00
40	NZ	75	HN	24.60	55	NZ	75	HN	30.00
40	NZ	76	HN	30.00	55	NZ	76	HN	27.39
40	NZ	77	HN	30.00	55	NZ	78	HN	25.77
40	NZ	79	HN	30.00	55	NZ	79	HN	30.00
40	NZ	80	HN	30.00	55	NZ	80	HN	30.00
40	NZ	81	HN	30.00	55	NZ	81	HN	30.00
40	NZ	82	HN	29.09	55	NZ	82	HN	30.00
40	NZ	83	HN	27.07	55	NZ	83	HN	25.31
40	NZ	84	HN	30.00	55	NZ	84	HN	30.00
40	NZ	85	HN	30.00	55	NZ	85	HN	23.71
40	NZ	86	HN	30.00	55	NZ	86	HN	30.00
40	NZ	87	HN	30.00	55	NZ	87	HN	25.92
40	NZ	89	HN	20.75	55	NZ	88	HN	30.00
40	NZ	90	HN	20.97	55	NZ	89	HN	28.36
40	NZ	92	HN	24.56	55	NZ	90	HN	30.00
40	NZ	93	HN	20.89	55	NZ	92	HN	28.58
40	NZ	95	HN	27.59	55	NZ	95	HN	30.00
40	NZ	96	HN	21.53	55	NZ	96	HN	30.00
40	NZ	97	HN	24.16	55	NZ	97	HN	30.00
40	NZ	98	HN	30.00	55	NZ	98	HN	30.00
40	NZ	99	HN	30.00	55	NZ	99	HN	30.00
40	NZ	101	HN	30.00	55	NZ	102	HN	30.00
40	NZ	102	HN	25.26	55	NZ	104	HN	29.12
40	NZ	104	HN	30.00	55	NZ	105	HN	25.06
40	NZ	105	HN	21.61	55	NZ	109	HN	23.78
40	NZ	109	HN	15.81	55	NZ	110	HN	22.77
40	NZ	110	HN	7.00	55	NZ	111	HN	22.91
40	NZ	111	HN	15.79	55	NZ	112	HN	20.68
40	NZ	112	HN	20.20	55	NZ	113	HN	23.90
40	NZ	113	HN	7.00	55	NZ	114	HN	20.80
40	NZ	114	HN	18.10	55	NZ	115	HN	23.89
40	NZ	115	HN	17.85	55	NZ	116	HN	22.71
40	NZ	116	HN	18.44	55	NZ	117	HN	23.52
40	NZ	117	HN	15.84	55	NZ	118	HN	21.25
40	NZ	118	HN	19.13	55	NZ	119	HN	23.87
40	NZ	119	HN	18.82	55	NZ	120	HN	25.20
40	NZ	120	HN	20.71	55	NZ	124	HN	17.74
40	NZ	123	HN	21.56	55	NZ	125	HN	21.29
40	NZ	124	HN	23.01	55	NZ	126	HN	17.33
40	NZ	125	HN	20.73	55	NZ	127	HN	23.01
40	NZ	126	HN	19.34	55	NZ	128	HN	18.66
40	NZ	127	HN	18.49	55	NZ	129	HN	15.61
40	NZ	128	HN	21.43	55	NZ	130	HN	19.09
40	NZ	129	HN	23.48	55	NZ	132	HN	13.61
40	NZ	130	HN	22.74	2	HN	121	NZ	22.71
2	HN	106	NZ	30.00	3	HN	121	NZ	24.59
3	HN	106	NZ	30.00	4	HN	121	NZ	24.30
4	HN	106	NZ	28.60	5	HN	121	NZ	24.15
5	HN	106	NZ	30.00	6	HN	121	NZ	23.20
6	HN	106	NZ	30.00	7	HN	121	NZ	27.76
7	HN	106	NZ	30.00	8	HN	121	NZ	30.00
8	HN	106	NZ	29.89	9	HN	121	NZ	30.00
9	HN	106	NZ	24.60	10	HN	121	NZ	27.39
10	HN	106	NZ	30.00	12	HN	121	NZ	25.77
11	HN	106	NZ	30.00	13	HN	121	NZ	30.00

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
13	HN	106	NZ	30.00	14	HN	121	NZ	30.00
14	HN	106	NZ	30.00	15	HN	121	NZ	30.00
15	HN	106	NZ	30.00	16	HN	121	NZ	30.00
16	HN	106	NZ	29.09	17	HN	121	NZ	25.31
17	HN	106	NZ	27.07	18	HN	121	NZ	30.00
18	HN	106	NZ	30.00	19	HN	121	NZ	23.71
19	HN	106	NZ	30.00	20	HN	121	NZ	30.00
20	HN	106	NZ	30.00	21	HN	121	NZ	25.92
21	HN	106	NZ	30.00	22	HN	121	NZ	30.00
23	HN	106	NZ	20.75	23	HN	121	NZ	28.36
24	HN	106	NZ	20.97	24	HN	121	NZ	30.00
26	HN	106	NZ	24.56	26	HN	121	NZ	28.58
27	HN	106	NZ	20.89	29	HN	121	NZ	30.00
29	HN	106	NZ	27.59	30	HN	121	NZ	30.00
30	HN	106	NZ	21.53	31	HN	121	NZ	30.00
31	HN	106	NZ	24.16	32	HN	121	NZ	30.00
32	HN	106	NZ	30.00	33	HN	121	NZ	30.00
33	HN	106	NZ	30.00	36	HN	121	NZ	30.00
35	HN	106	NZ	30.00	38	HN	121	NZ	29.12
36	HN	106	NZ	25.26	39	HN	121	NZ	25.06
38	HN	106	NZ	30.00	43	HN	121	NZ	23.78
39	HN	106	NZ	21.61	44	HN	121	NZ	22.77
43	HN	106	NZ	15.81	45	HN	121	NZ	22.91
44	HN	106	NZ	7.00	46	HN	121	NZ	20.68
45	HN	106	NZ	15.79	47	HN	121	NZ	23.90
46	HN	106	NZ	20.20	48	HN	121	NZ	20.80
47	HN	106	NZ	7.00	49	HN	121	NZ	23.89
48	HN	106	NZ	18.10	50	HN	121	NZ	22.71
49	HN	106	NZ	17.85	51	HN	121	NZ	23.52
50	HN	106	NZ	18.44	52	HN	121	NZ	21.25
51	HN	106	NZ	15.84	53	HN	121	NZ	23.87
52	HN	106	NZ	19.13	54	HN	121	NZ	25.20
53	HN	106	NZ	18.82	58	HN	121	NZ	17.74
54	HN	106	NZ	20.71	59	HN	121	NZ	21.29
57	HN	106	NZ	21.56	60	HN	121	NZ	17.33
58	HN	106	NZ	23.01	61	HN	121	NZ	23.01
59	HN	106	NZ	20.73	62	HN	121	NZ	18.66
60	HN	106	NZ	19.34	63	HN	121	NZ	15.61
61	HN	106	NZ	18.49	64	HN	121	NZ	19.09
62	HN	106	NZ	21.43	66	HN	121	NZ	13.61
63	HN	106	NZ	23.48					
64	HN	106	NZ	22.74					

NZ is the nitrogen of the MTSL ring.

Table C.9: Intermolecular NOE distances of CylR2 which were manually assigned.

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
47	HG	127	HB*	4.0	114	HA	63	HE3	4.0
113	HG	61	HB*	4.0	48	HA	129	HE3	4.0
47	HG	124	HG2	4.0	114	HE*	63	HH2	4.0
113	HG	58	HG2	4.0	48	HE*	129	HH2	4.0
47	HG	124	HG1	4.0	114	HG2	63	HH2	4.0
113	HG	58	HG1	4.0	48	HG2	129	HH2	4.0
44	HA	68	HG2*	4.0	117	HD1	63	HE3	4.0
110	HA	2	HG2*	4.0	51	HD1	129	HE3	4.0
41	HG2	109	HB1	4.0	117	HD1	63	HZ2	4.0
107	HG2	43	HB1	4.0	51	HD1	129	HZ2	4.0
47	HD1*	129	HB1	4.0	117	HB*	63	HZ2	4.0
113	HD1*	63	HB1	4.0	51	HB*	129	HZ2	4.0
48	HG2	129	HZ2	4.0	110	HA	63	HD1	4.0
114	HG2	63	HZ2	4.0	44	HA	129	HD1	4.0
47	HA	129	HB1	4.0	113	HD2*	63	HB2	4.0
113	HA	63	HB1	4.0	47	HD2*	129	HB2	4.0

continued on next page

Residue 1	H1	Residue 2	H2	distance [Å]	Residue 1	H1	Residue 2	H2	distance [Å]
48	HA	129	HE3	4.0	113	HB2	63	HB2	4.0
114	HA	63	HE3	4.0	47	HB2	129	HB2	4.0
47	HB1	129	HB1	4.0	113	HB1	63	HB2	4.0
113	HB1	63	HB1	4.0	47	HB1	129	HB2	4.0
44	HG*	129	HB*	4.0	113	HD2*	63	HB1	4.0
110	HG*	63	HB*	4.0	47	HD2*	129	HB1	4.0
114	HE*	63	HZ2	4.0	113	HB*	63	HB1	4.0
48	HE*	129	HZ2	4.0	47	HB*	129	HB1	4.0

C.3 Dihedral angle restraints for CylR2

Table C.10: Dihedral angle restraints for CylR2 derived from the chemical shifts with TALOS.

Residue	PHI	lower angle	upper angle	Residue	PSI	lower angle	upper angle
2	-140.0	-100.0	2	127.3	156.5		
3	-140.0	-99.5	3	112.0	129.2		
4	-140.0	-99.6	4	124.0	144.7		
6	-70.0	-40.0	6	-60.0	-20.0		
7	-66.9	-54.7	7	-50.6	-40.4		
8	-67.9	-56.3	8	-45.2	-35.9		
9	-76.2	-58.2	9	-51.4	-38.7		
10	-66.4	-56.2	10	-45.7	-33.8		
11	-68.5	-61.5	11	-44.4	-34.6		
12	-67.5	-59.1	12	-37.2	-19.8		
17	-63.3	-57.2	17	-45.0	-30.5		
18	-67.5	-59.2	18	-48.3	-28.7		
19	-71.0	-59.6	19	-49.3	-38.1		
20	-72.3	-59.4	20	-43.3	-33.1		
21	-70.1	-58.9	21	-49.1	-33.6		
22	-70.7	-60.7	22	-45.1	-36.7		
23	-72.3	-61.3	23	-38.8	-21.4		
24	-104.4	-85.2	24	-4.6	14.6		
28	-60.9	-52.7	28	-45.9	-28.9		
29	-73.4	-59.8	29	-49.9	-29.1		
30	-71.6	-56.1	30	-50.8	-28.0		
31	-72.8	-60.0	31	-44.5	-31.5		
32	-69.5	-52.3	32	-49.4	-38.2		
33	-66.0	-60.4	33	-45.4	-32.6		
34	-71.6	-59.5	34	-47.4	-36.6		
35	-71.4	-60.0	35	-35.5	-18.8		
43	-70.0	-40.0	43	-60.0	-20.0		
44	-62.6	-56.6	44	-49.9	-36.3		
45	-69.3	-57.9	45	-48.3	-33.6		
46	-71.9	-60.6	46	-48.3	-40.3		
47	-70.3	-61.7	47	-45.1	-34.9		
48	-68.4	-60.5	48	-47.9	-35.5		
49	-69.1	-61.9	49	-44.4	-37.9		
50	-63.9	-54.9	50	-45.8	-33.7		
51	-70.0	-58.8	51	-51.3	-35.1		
52	-72.8	-57.9	52	-41.7	-29.3		
57	-66.9	-54.9	57	-40.6	-21.3		
58	-71.2	-59.6	58	-44.0	-28.4		
59	-70.0	-30.0	59	-60.0	-20.0		
60	-70.0	-30.0	60	-60.0	-20.0		
61	-142.9	-112.9	61	142.3	162.7		
62	-146.4	-105.6	62	126.0	147.8		
63	-123.2	-90.2	63	118.4	135.1		



D

Appendix: Xplor input files

Xplor input files used for water refinement, rigid-body docking and simulated annealing in chapter 5 are specified in the following.

D.1 Water refinement

The protocol used for water refinement is given below. It is slightly different from the one in literature [172], [173].

```
#!/usr/bin/env python

import string,os,sys,profile,cPickle,math,socket,time,shutil,fnmatch,glob,random,pdb_file
import nmv_dct,nmv_dsc,nmv_xplor

# MAIN SCRIPT
# =====

# SET THE GENERAL CONFIGURATION FILE AND INPUT PARAMETER FILE
configfile           = os.path.join(os.getcwd(),'nmr_waterrefine.conf')
parameterfile       = os.path.join(os.getcwd(),'nmr_waterrefine.pars')

# READ THE GENERAL CONFIGURATION FILE
print 'Reading CONFIGURATION file %s...\n'%configfile
config               = nmv_dct.read(configfile)

# READ THE REFINEMENT PARAMETER FILE
print 'Reading PARAMETER file %s...\n'%parameterfile
refinepars           = nmv_dct.read(parameterfile)

# NOTE THAT PARAMETERS NEEDED FOR THE XPLOR RUN ARE READ IN USING NMV_XPLOR.PY

# SET PATHNAMES
path                 = os.path.realpath(refinepars['PROJECTDIR'])
tmppath              = os.path.realpath(refinepars['TMPDIR'])
scriptspath          = os.path.realpath(refinepars['SCRIPTSDIR'])
tablepath            = os.path.realpath(refinepars['TABLEDIR'])

# SET INITIAL PARAMETERS FOR THE XPLOR RUN
ninputstructures     = int(refinepars['NINPUTSTRUCTURES'])

# SET BASE FILENAMES
analyzed_pdb_base    = 'analyzed'
refined_pdb_base     = 'refined'
analyzed_subdirname  = 'analyzed'
refined_subdirname   = 'refined'
```

```

# SET SOFTWARE AND PARAMETERS FOR THE VALIDATION ROUTINES
profit                = config['PROFIT']
procheck_run         = config['PROCHECK_RUN']
prochecknmr_run     = config['PROCHECKNMR_RUN']
procheck_dir        = os.path.realpath(config['PROCHECK_DIR'])
whatif_run          = config['WHATIF_RUN']
zonefile            = os.path.realpath(refinepars['ZONEFILE'])
energysort          = refinepars['ENERGYSORT']
selection           = refinepars['SELECTION']

# SET FURTHER OPTIONS
run_cluster         = refinepars['RUN_CLUSTER']
queu_cluster       = refinepars['QUEUE_CLUSTER']
run_checks         = refinepars['RUN_CHECKS']

#####
# ANALYSIS INPUT STRUCTURES #
#####

# CHECK WHETHER ANALYSIS WAS DONE
# SET THE OUTPUTCOORDINATESPATH FOR THE ANALYSIS OF THE INPUT STRUCTURES
outputcoordinatespath = os.path.join(path,analyzed_subdirname)

filename = os.path.join(outputcoordinatespath,'summary_%s.txt'%analyzed_pdb_base)
if os.path.exists(filename):
    print 'Analysis of the input structures is finished'
else:
    # SET TEMPORARY FILES
    checkscriptfilename = os.path.join(outputcoordinatespath,'checkscr.tmp')
    procheckpath        = os.path.join(outputcoordinatespath,'procheck')
    whatcheckpath       = os.path.join(outputcoordinatespath,'whatcheck')

    # RE-CREATE ANALYSIS DIRECTORY/CREATE ANALYSIS DIRECTORY IF NOT PRESENT:
    # REMOVE ALL PREVIOUS TEMPORARY FILES FOR CLEAN START
    if os.path.exists(outputcoordinatespath):
        print 'Cleaning directory %s'%outputcoordinatespath
        nmv_dsc.removedir(outputcoordinatespath)
        os.mkdir(outputcoordinatespath)
    else:
        print 'Creating directory %s'%outputcoordinatespath
        os.mkdir(outputcoordinatespath)

    # ANALYZE THE INPUT STRUCTURES IN THE XPLOR FORCEFIELD USED FOR THE WATERREFINEMENT
    # INCLUDES VIOLATION AND ENERGY ANALYSIS
    # THE NMV_XPLOR MODULE CREATES THE SCRIPTS NEEDED TO CALCULATE STRUCTURES ON A CLUSTER
    # USING THE PARAMETER AND CONFIG FILES
    # AND GENERATES A 'jobs' SUBDIRECTORY(=RUNPATH) IN WHICH THE XPLOR RUN IS PERFORMED
    nmv_xplor.analyze(parameterfile,configfile)

    # CHECK EVERY 10 SECS WHETHER ALL STRUCTURES ARE ANALYZED (NEEDED WHEN RUNNING ON A CLUSTER)
    # todo: check acceptance criteria and loop until we have enough accepted structures
    done=0
    while not done:
        time.sleep(10)
        if len(glob.glob(os.path.join(outputcoordinatespath,analyzed_pdb_base+'*.pdb')))==ninputstructures:
            done=1

    # RUN THE OTHER CHECKS
    # CREATE INPUT SCRIPTS FOR THE CHECKING ROUTINES
    print 'Preparing for protein structure checks of the analyzed input structures'
    checkscript=open(checkscriptfilename,'w')
    checkscript.write('#!/bin/tcsh\n')

    # SET ENVIRONMENT VARIABLES
    checkscript.write('setenv prodir %s\n'%procheck_dir)

    # WRITE A XPLOR FILE-LIST CONTAINING THE INFORMATION ABOUT THE ANALYZED INPUT STRUCTURES

```

```

# FILES ARE SORTED ON RESTRAINT ENERGY (SET ENERGYSORT), THE FUNCTION USES THE PDB-HEADER INFORMATION
checkscript.write('%s -writexplorfilelist %s %s %s\n'%(os.path.join(scriptspath, 'nmv_nmrcheck.py'), \
    outputcoordinatespath, \
    analyzed_pdb_base, \
    energysort))

# CALCULATE CORNILESCU Q-FACTORS (IF DEFINED IN THE PARAMETERFILE), READ THE LOGFILE CREATED BY XPLOR
for line in open(parameterfile, 'r').readlines():
    if line[:9] in ['SANI_FILE', 'DIPO_FILE'] and len(string.split(line))==3:
        calccornilQ = 1
        break
    else: calccornilQ = 0
if calccornilQ:
    checkscript.write('%s -calccornilQ %s %s \n'%(os.path.join(scriptspath, 'nmv_nmrcheck.py'), \
        outputcoordinatespath, \
        analyzed_pdb_base))

# AUTOMATICALLY CHECK THE ANALYZED INPUT STRUCTURES WITH PROCHECK AND WHATIF
# PROFIT IS USED FOR RMSD CALCULATIONS
if run_checks=='y':
    checkscript.write('%s -checkall %s %s %s %s %s %s %s %s %s\n'%(os.path.join(scriptspath, 'nmv_nmrcheck.py'), \
        profit, \
        procheck_run, \
        prochecknmr_run, \
        procheck_dir, \
        whatif_run, \
        outputcoordinatespath, \
        analyzed_pdb_base, \
        zonefile, \
        selection, \
        tmpspath))
elif run_checks=='n':
    print 'Skipping WHATIF, PROCHECK protein structure checks, and PROFIT rmsd calculations'

# CREATE A SUMMARY FILE THAT CONTAINS ALL INFORMATION
checkscript.write('%s -summary %s %s \n'%(os.path.join(scriptspath, 'nmv_nmrcheck.py'), \
    outputcoordinatespath, \
    analyzed_pdb_base))

# CLOSE THE SCRIPTFILE
checkscript.close()

# SUBMIT THE SCRIPT EITHER TO THE CLUSTER OR ON A SINGLE PROCESSOR MACHINE
os.system('/bin/chmod +x %s'%checkscriptfilename)
print 'Starting structure checks of the analyzed input structures'
if run_cluster=='y':
    os.system('%s %s &'%(queu_cluster, checkscriptfilename))
else:
    os.system('%s'%checkscriptfilename)

# CHECK EVERY 10 SECS WHETHER THE CHECKING ROUTINES ARE FINISHED (NEEDED WHEN RUNNING ON A CLUSTER)
done=0
while not done:
    time.sleep(10)
    if os.path.exists(os.path.join(outputcoordinatespath, 'summary_%s.txt'%analyzed_pdb_base)):
        done=1

# COMPRESS JOBS, PROCHECK AND WHATCHECK SUBDIRECTORIES
print 'Compressing files...'
current_location=os.getcwd()
os.chdir(outputcoordinatespath)
os.system('tar cf - jobs | gzip > jobs.tgz')
if run_checks=='y':
    os.system('tar cf - procheck | gzip > procheck.tgz')
    os.system('tar cf - whatcheck | gzip > whatcheck.tgz')
os.chdir(current_location)

# REMOVE THE SCRIPTFILES AND PROCHECK AND WHATCHECK DIRECTORIES

```

```

# IN THE FINISHED CYCLE
for file in glob.glob(os.path.join(outputcoordinatespath,'*.tmp*')):
    os.remove(file)
nmv_dsc.removedir(os.path.join(outputcoordinatespath,'jobs'))
if run_checks=='y':
    nmv_dsc.removedir(procheckpath)
    nmv_dsc.removedir(whatcheckpath)
print 'Analysis of the input structures is finished'

#####
# REFINEMENT OF INPUT STRUCTURES #
#####

# SET THE OUTPUTCOORDINATESPATH FOR THE ACTUAL REFINEMENT OF THE INPUT STRUCTURES
outputcoordinatespath = os.path.join(path,refined_subdirname)

# CHECK WHETHER ANALYSIS WAS DONE
filename = os.path.join(outputcoordinatespath,'summary_%s.txt'%refined_pdb_base)
if os.path.exists(filename):
    print 'Analysis of refined structures is finished'
else:
    # SET TEMPORARY FILES
    checkscriptfilename = os.path.join(outputcoordinatespath,'checkscr.tmp')
    procheckpath = os.path.join(outputcoordinatespath,'procheck')
    whatcheckpath = os.path.join(outputcoordinatespath,'whatcheck')

    # CREATE REFINED STRUCTURES DIRECTORY IF NOT PRESENT:
    if not os.path.exists(outputcoordinatespath):
        print 'Creating directory %s'%outputcoordinatespath
        os.mkdir(outputcoordinatespath)

    # DO THE ACTUAL WATERREFINEMENT
    nmv_xplor.waterrefine(parameterfile,configfile)

    # CHECK EVERY 10 SECS WHETHER ALL STRUCTURES ARE CALCULATED (NEEDED WHEN RUNNING ON A CLUSTER)
    # check acceptance criteria and loop until we have enough accepted structures
    done=0
    while not done:
        time.sleep(10)
        if len(glob.glob(os.path.join(outputcoordinatespath,refined_pdb_base+'*.pdb*')))==ninputstructures:
            done=1

    # RUN THE OTHER CHECKS
    # CREATE INPUT SCRIPTS FOR THE CHECKING ROUTINES
    print 'Preparing for protein structure checks of the refined input structures'
    checkscript=open(checkscriptfilename,'w')
    checkscript.write('#!/bin/tcsh\n')

    # SET ENVIRONMENT VARIABLES
    checkscript.write('setenv prodir %s\n'%procheck_dir)

    # WRITE A XPLOR FILE-LIST CONTAINING THE INFORMATION ABOUT THE ANALYZED INPUT STRUCTURES
    # FILES ARE SORTED ON RESTRAINT ENERGY (SET ENERGYSORT), THE FUNCTION USES THE PDB-HEADER INFORMATION
    checkscript.write('%s -writexplorfilelist %s %s %s\n'%(os.path.join(scriptspath,'nmv_nmrcheck.py'),\
        outputcoordinatespath,\
        refined_pdb_base,\
        energysort))

    # CALCULATE CORNILESCU Q-FACTORS (IF DEFINED IN THE PARAMETERFILE), READ THE LOGFILE CREATED BY XPLOR
    for line in open(parameterfile,'r').readlines():
        if line[:9] in ['SANI_FILE','DIPO_FILE'] and len(string.split(line))==3:
            calccornilQ = 1
            break
        else: calccornilQ = 0
    if calccornilQ:
        checkscript.write('%s -calccornilQ %s %s %s\n'%(os.path.join(scriptspath,'nmv_nmrcheck.py'),\
            outputcoordinatespath,\

```



```

        refined_pdb_base))

# AUTOMATICALLY CHECK THE INPUT STRUCTURES WITH PROCHECK AND WHATIF
# PROFIT IS USED FOR RMSD CALCULATIONS
if run_checks=='y':
    checkscript.write('%s -checkall %s %s %s %s %s %s %s %s %s %s\n'%(os.path.join(scriptspath,'nmv_nmrcheck.py'),\
        profit,\
        procheck_run,\
        prochecknmr_run,\
        procheck_dir,\
        whatif_run,\
        outputcoordinatespath,\
        refined_pdb_base,\
        zonefile,\
        selection,\
        tmpspath))
elif run_checks=='n':
    print 'Skipping WHATIF, PROCHECK protein structure checks, and PROFIT rmsd calculations'

# CREATE A SUMMARY FILE THAT CONTAINS ALL INFORMATION
checkscript.write('%s -summary %s %s\n'%(os.path.join(scriptspath,'nmv_nmrcheck.py'),\
    outputcoordinatespath,\
    refined_pdb_base))
checkscript.close()

# SUBMIT THE SCRIPT EITHER TO THE CLUSTER OR ON A SINGLE PROCESSOR MACHINE
os.system('/bin/chmod +x %s'%checkscriptfilename)
print 'Starting structure checks of the refined input structures'
if run_cluster=='y':
    os.system('%s %s &'%(queu_cluster,checkscriptfilename))
else:
    os.system('%s'%checkscriptfilename)

# CHECK EVERY 10 SECS WHETHER THE CHECKING ROUTINES ARE FINISHED (NEEDED WHEN RUNNING ON A CLUSTER)
done=0
while not done:
    time.sleep(10)
    if os.path.exists(os.path.join(outputcoordinatespath,'summary_%s.txt'%refined_pdb_base)):
        done=1

# COMPRESS JOBS, PROCHECK AND WHATCHECK SUBDIRECTORIES
print 'Compressing files...'
current_location=os.getcwd()
os.chdir(outputcoordinatespath)
os.system('tar cf - jobs | gzip > jobs.tgz')
if run_checks=='y':
    os.system('tar cf - procheck | gzip > procheck.tgz')
    os.system('tar cf - whatcheck | gzip > whatcheck.tgz')
os.chdir(current_location)

# REMOVE THE SCRIPTFILES AND PROCHECK AND WHATCHECK DIRECTORIES
# IN THE FINISHED CYCLE
for file in glob.glob(os.path.join(outputcoordinatespath,'*.tmp*')):
    os.remove(file)
nmv_dsc.removedir(os.path.join(outputcoordinatespath,'jobs'))
if run_checks=='y':
    nmv_dsc.removedir(procheckpath)
    nmv_dsc.removedir(whatcheckpath)
print 'Refinement and analysis of the input structures is finished'

```

An example parameter file as used for the refinement of monomeric CylR2:

```
## *****
## *                               *
## *           R E F I N E . P A R S           *
## *                               *
## *****
## This parameter file contains all necessary information for a refinement
## in explicit solvent using xplor

## NOTES FOR USAGE:
## use absolute pathnames for the directories
##
## Multiple noe, dihedral rdc etc restraint files can be used.
## these files are separated using OBLIGATORY indices, which are defined as the string trailing the first ".":
## eg: NOE_FILE.1 has the index '1'
##     SANI_FILE.JHNC has the index 'JHNC'
## All parameters that are needed for each restraint file are indexed accordingly.
## Indices can be any string or number with up to 4 characters, longer indices will be truncated in XPLOR
## The index will refer to the classname in the xplor run. The pdb headers will
## use the indices for reference.
##
## In case a parameter should not be used uncomment (#) the line, delete the line
## or leave the part behind the equals sign blank. In the latter case a warning will be printed
##
## IMPORTANT: do not mix SANI and DIPO restraints in a single run!! Use either one of them

#####
## FILES AND DIRECTORIES:#
#####

## Set a temporary directory (default='/tmp'):
## This is needed since some programs cannot deal with long pathnames
## These programs are run in the TMPDIR when necessary
## use names up to 20 chars for the TMPDIR

TMPDIR      = /tmp

## Directory where the project is run:

PROJECTDIR  = /home/mpg4/MBPC/030/srumpel/cyana/waterrefine_org/run33

## Location of structure file:

PSFDIR      = /home/mpg4/MBPC/030/srumpel/cyana/waterrefine_org/run33/psf

## Location of restraint tables:

TABLEDIR    = /home/mpg4/MBPC/030/srumpel/cyana/waterrefine_org/run33/tables

## Location of input coordinate files:

COORDINDIR  = /home/mpg4/MBPC/030/srumpel/cyana/waterrefine_org/run33/input

## Location of python scripts:

SCRIPTSDIR  = /home/mpg4/MBPC/030/srumpel/cyana/waterrefine_org/run33/scripts

## Total number of inputstructures:

NINPUTSTRUCTURES = 20

## Base filename for input structures (pdb_base_xxx.pdb):

PDB_BASE    = 33rb

## Filename of the XPLOR psf file:
```

```
PSF_FILE      = 33.psf

## Filename of the pdb file describing the different axis systems of
## the media used for orienting the molecules:
## All axis systems should be described in a single pdb file
## Note that the nomenclature in the axis pdb file should correspond to the ARIA nomenclature

PDB_AXIS      = axis.pdb

## Filenames of the pdb files describing the different axis systems of
## the media used for orienting the molecules:
## All axis systems should be described in a single psf file
## Note that the nomenclature in the axis psf file should correspond to the ARIA nomenclature

PSF_AXIS      = axis_500.psf

## Filenames and parameters of the XPLOR NOE tables:
## NOE_AVE      = Type of averaging used for the NOE restraints
## NOE_ACCEPT   = NOE restraints acceptance criterium (Angstrom)

NOE_FILE.1    = 33_noe.tbl
NOE_AVE.1     = sum
NOE_ACCEPT.1  = 0.5

#NOE_FILE.hbon = hbond.tbl
#NOE_AVE.hbon  = sum
#NOE_ACCEPT.hbon = 0.5

## Filenames and parameters of the XPLOR CDIH tables:
## CDIH_ACCEPT: Dihedral angle restraints acceptance criterium (degrees)

CDIH_FILE.1   = 33_dihe.tbl
CDIH_ACCEPT.1 = 5.0

DIPO_FILE.NHf = hn.tbl
DIPO_COEFA.NHf = 0.0
DIPO_COEFB.NHf = -22.573
DIPO_COEFC.NHf = 0.217
DIPO_SCALE.NHf = 1
DIPO_FORCE.NHf = 0.1
DIPO_POTEN.NHf = harmonic
DIPO_AVE.NHf   = sum
DIPO_TYPE.NHf  = fixd
DIPO_SIGN.NHf  = on
DIPO_ACCEPT.NHf = 0.3

#####
## REFINEMENT PARAMETERS:##
#####

## Number of steps for the heating stage:
## TIME FOR HEATING STAGE: 0.015*HEAT_STEPS ps
## 5 CYCLES, TIMESTEP 0.003 ps
## Normally HEAT_STEPS = 100

HEAT_STEPS    = 120

## Number of steps for the heating stage:
## TIME FOR HEATING STAGE: 0.004*HOT_STEPS ps
## 1 CYCLE, TIMESTEP 0.004 ps
## Normally HOT_STEPS = 2000

HOT_STEPS     = 2000

## Number of steps for the heating stage:
## TIME FOR HEATING STAGE: 0.080*HEAT_STEPS ps
## 20 CYCLES, TIMESTEP 0.004 ps
```

```

## Normally COOL_STEPS = 200

COOL_STEPS   = 200

## Random seed number used for velocity generation in XPLOR:

SEED         = 12397

## Zonefile for RMSD calculations in PROFIT:

ZONEFILE     = /home/mpg1/MBPC/yjung/progs/waterrefine/tables/zones

## Atom selection for RMSD calculations in PROFIT (All= separate calculations of Heavy, CA and backbone RMSD)
## Options: All, Heavy, Ca and BB

SELECTION    = All

## Type of Energy Sorting of pdb files:
## Options: restraints, noe, total

ENERGYSORT   = restraints

#####
## FURTHER OPTIONS:#
#####

## Run the refinement procedure on a cluster (y/n):

RUN_CLUSTER  = n

## Que command for run on a cluster:

QUEUE_CLUSTER = /usr/local/pbs/bin/qsub -l nodes=1:ppn=1

## Run structure checks (WHATIF, PROCHECK and PROFIT) (y/n)
## If set to 'n', none of the WHATIF, PROCHECK and PROFIT checks and RMSD calculations are performed

RUN_CHECKS   = n

```

For other refinements the filenames and paths as well as the dipolar coupling coefficients were changed.

D.2 Rigid body docking

The Xplor input file for rigid-body docking described by Clore [89] was slightly changed. The used input file is indicated.

```

REMARKS very slow cooling with increase of vdw
evaluate ($seed=546372920)

set seed $seed end
{* read in the PSF file and initial structures          *}

param @parallhdg_CMS_H.pro @TOPPAR:par_axis_3.pro
end

structure @CylMTSL_complex.psf @axis_new.psf end

coor @axis_xyz_3.pdb
coor @meanQexplor.pdb
evaluate ($count =0)
while ($count < 7) loop file
    evaluate ($count = $count + 1)
evaluate ($file = "B_" + encode($count)+ ".pdb")

```

```

coor @$file

coor copy end

constraints inter (resid 1:66) (resid 67:132) end
{* resid 1:66 is protein cylr2A and resid 67:132 is cylr2B *}

vector do (refx=x) (all)
vector do (refy=y) (all)
vector do (refz=z) (all)

{* set the weights for the experimental energy terms          *}

evaluate ($knoe = 0.01)  {* noes                               *}
evaluate ($asym = 0.1)  {* slope of NOE potential            *}

noe
  reset
  nres = 5000
  class all
  @N40T55_5Accorr.tbl
end

noe
  ceiling 1000
  averaging all sum
  potential all square  {* soft                               *}
  scale all $knoe
  sqconstant all 1.0
  sqexponent all 2

end

evaluate ($ksani = 0.001)

sani
  nres=4000
  class hnBw
  force $ksani
  potential harmonic
  coeff 0.0 -24.0 0.13
  @hnB.tbl

  class hnBf
  force 0.0  {* These are explicitly set to zero and not *}
             {* modified, in order to use them for a      *}
             {* free-R factor.                             *}

  potential harmonic
  coeff 0.0 -24.0 0.13
  @hnB.tbl

  class hnAw
  force $ksani
  potential harmonic
  coeff 0.0 -24.0 0.13
  @hnA.tbl

  class hnAf
  force 0.0  {* Set to zero for free-R factor.          *}
  potential harmonic
  coeff 0.0 -20.4 0.13
  @hnA.tbl

end

sani print threshold=0.0 class hnBw end
evaluate ($rms_sani_hnBw=$result)
evaluate ($R_sani_hnBw=$result*100/19.13)

```

```

    evaluate ($viol_sani_hnBw=$violations)

sani print threshold=0.0 class hnBf end
    evaluate ($rms_sani_hnBf=$result)
    evaluate ($R_sani_hnBf=$result*100/19.13)
    evaluate ($viol_sani_hnBf=$violations)

sani print threshold=0.0 class hnAw end
    evaluate ($rms_sani_hnAw=$result)
    evaluate ($R_sani_hnAw=$result*100/19.13)
    evaluate ($viol_sani_hnAw=$violations)

sani print threshold=0.0 class hnAf end
    evaluate ($rms_sani_hnAf=$result)
    evaluate ($R_sani_hnAf=$result*100/19.13)
    evaluate ($viol_sani_hnAf=$violations)

evaluate ($rcon = 0.003)

parameters
  nbonds
    atom
    nbxmod 4      {* Set at 4, because internal (torsion-angle dynamics are *}
                  {* used. *}
    wmin = 0.01  {* warning off *}
    cutnb = 4.5  {* nonbonded cutoff *}
    tolerance 0.5
    repel= 0.9   {* scale factor for vdW radii = 1 ( L-J radii) *}
    rexp = 2     {* exponents in (r^irex - R0^irex)^rexp *}
    irex = 2
    rcon=$rcon   {* actually set the vdW weight *}
  end
end

vector do (vx=maxwell(5.)) (all)
vector do (vy=maxwell(5.)) (all)
vector do (vz=maxwell(5.)) (all)

constraints fix (resid 500 and name 00) end
{* Fix dipolar couplings axis. *}

!evaluate ($cool_steps = 24000)
evaluate ($init_t = 1500.01)

vector do (mass = 30.0) (not resid 500)
    {* all atom masses are set the same so that *}
    {* motion of each atom is equally likely in *}
    {* the dynamics. Note that the dynamics is *}
    {* not a realistic model of the motion of *}
    {* the system; it is a method of minimizing *}
    {* a highly complicated function. *}
vector do (mass = 100.0) (resid 500)
vector do (fbeta = 10.0) (all)  {* coupling to heat bath *}

evaluate ($count =0)
while ($count < 7)  {* Only want one structure to start the refinement *}
  loop structure
  evaluate ($count = $count + 1)

  vector do (x=xcomp) (all)
  vector do (y=ycomp) (all)
  vector do (z=zcomp) (all)

  evaluate ($ini_rad = 0.9)      evaluate ($fin_rad = 0.75)
  evaluate ($ini_con= 0.004)    evaluate ($fin_con= 1.0)
  evaluate ($ini_ang = 1.0)     evaluate ($fin_ang = 1.0)
  evaluate ($ini_imp = 1.0)     evaluate ($fin_imp = 1.0)

```

```

evaluate ($ini_noe = 0.01)      evaluate ($fin_noe = 30.0)
evaluate ($knoe = $ini_noe)    {* force constant for NOEs *}

evaluate ($ini_sani = 0.001)    evaluate ($fin_sani = 0.1)
evaluate ($ksani = $ini_sani)  {* force constant for dipolar couplings. *}

{* We are only using the "working" dipolar couplings in the refinement. The *}
{* "free" ones are used to generate a free-R-like factor.                    *}
sani class hnBw force $ksani end
sani class hnAw force $ksani end

flags
  exclude *
  include vdw noe sani
  {* Use van der Waals, intermolecular NOEs and dipolar couplings to orient *}
  {* the rigid proteins with respect to each other.                          *}
end

parameters
  nbonds
    atom
    nbxmod 4      {* Set at 4, because internal dynamics are                *}
                  {* used for refinement.                                  *}
    wmin = 0.01  {* warning off                                           *}
    cutnb = 4.5  {* nonbonded cutoff                                       *}
    tolerance 0.5
    repel= 0.8   {* scale factor for vdW radii = 1 ( L-J radii)           *}
    rexp = 2     {* exponents in (r^irex - RO^irex)^rexp                  *}
    irex = 2
    rcon = 4     {* actually set the vdW weight                            *}
  end
end

dynamics internal
  reset
  itype=powell
  stepsize = 0.1
  nstep=1000
  depred=1
  fix = (resid 1:66)
  group = (resid 67:132)
  group = (resid 500)  {* Treat each protein and the axis as rigid bodies *}
  etol = 0.0000001
  gtol = 0.0000001
  nprint= 1
end

evaluate ($final_t = 500)      { K }
evaluate ($tempstep = 25)     { K }

evaluate ($ncycle = ($init_t-$final_t)/$tempstep)

evaluate ($bath = $init_t)
evaluate ($k_vdw = $ini_con)
evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
evaluate ($radius= $ini_rad)
evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))
evaluate ($k_ang = $ini_ang)
evaluate ($ang_fac = ($fin_ang/$ini_ang)^(1/$ncycle))
evaluate ($k_imp = $ini_imp)
evaluate ($imp_fac = ($fin_imp/$ini_imp)^(1/$ncycle))
evaluate ($noe_fac = ($fin_noe/$ini_noe)^(1/$ncycle))
evaluate ($knoe = $ini_noe)

evaluate ($sani_fac = ($fin_sani/$ini_sani)^(1/$ncycle))
evaluate ($ksani = $ini_sani)

```

```

flags
  exclude *
  include vdw noe sani
  {* Use only van der Waals repulsion, intermolecular NOE data, and dipolar *}
  {* couplings to orient the two rigid proteins with respect to one another *}
end

vector do (vx=maxwell(5.)) (all)
vector do (vy=maxwell(5.)) (all)
vector do (vz=maxwell(5.)) (all)

evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
  evaluate ($i_cool=$i_cool+1)

  evaluate ($bath = $bath - $tempstep)
  evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
  evaluate ($radius=max($fin_rad,$radius*$radfact))
  evaluate ($k_ang = $k_ang*$ang_fac)
  evaluate ($k_imp = $k_imp*$imp_fac)
  evaluate ($knoe = $knoe*$noe_fac)

  evaluate ($ksani = $ksani*$sani_fac)

  parameter
    nbonds
    cutnb=4.5 rcon=$k_vdw nbxmod=4 repel=$radius
  end
end
noe scale all $knoe end

sani class hnBw force $ksani end
sani class hnAw force $ksani end

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end
end loop cool

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end

flags exclude * include noe sani end

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end

flags exclude * include noe sani vdw end

noe scale all 60.0 end
parameter nbonds rcon 3 repel 0.75 end end

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end

print threshold=0.5 noe
evaluate ($rms_noe=$result)

```



```

    evaluate ($violations_noe=$violations)
print thres=0.05 bonds
    evaluate ($rms_bonds=$result)
print thres=5. angles
    evaluate ($rms_angles=$result)
print thres=5. impropers
    evaluate ($rms_impropers=$result)

sani print threshold=0.0 class hnBw end
evaluate ($rms_sani_hnBw=$result)
evaluate ($R_sani_hnBw=$result*100/19.13)
evaluate ($viol_sani_hnBw=$violations)

sani print threshold=0.0 class hnBf end
evaluate ($rms_sani_hnBf=$result)
evaluate ($R_sani_hnBf=$result*100/19.13)
evaluate ($viol_sani_hnBf=$violations)

sani print threshold=0.0 class hnAw end
evaluate ($rms_sani_hnAw=$result)
evaluate ($R_sani_hnAw=$result*100/19.13)
evaluate ($viol_sani_hnAw=$violations)

sani print threshold=0.0 class hnAf end
evaluate ($rms_sani_hnAf=$result)
evaluate ($R_sani_hnAf=$result*100/19.13)
evaluate ($viol_sani_hnAf=$violations)

remarks =====
remarks repel=0.75
remarks      overall vdw,noe, sani
remarks energies: $ener $vdw, $noe, $sani,
remarks =====
remarks      bonds,angles,impropers,noe
remarks bonds etc: $rms_bonds,$rms_angles,$rms_impropers, $rms_noe
remarks =====
remarks      noe
remarks violations : $violations_noe
remarks =====
remarks sani hnBw hnAw hnBf hnAf Da=-14.9 R=0.20
remarks Rms sani: $rms_sani_hnBw $rms_sani_hnAw $rms_sani_hnBf $rms_sani_hnAf
remarks R factor sani: $R_sani_hnBw $R_sani_hnAw $R_sani_hnBf $R_sani_hnAf
remarks viol sani: $viol_sani_hnBw $viol_sani_hnAw $viol_sani_hnBf $viol_sani_hnAf
remarks =====

evaluate ($file2 = "rigid_cool2x_" + encode($count) + ".pdb")
write coor output=$file2 end

flags exclude * include noe sani end
      {* Use just intermolecular NOEs and dipolar couplings for this step. *}
noe scale all 60.0 end

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end

flags exclude * include noe sani vdw end

noe scale all 60.0 end
parameter nbonds rcon 3 repel 0.78 end end

dynamics internal
  itype=powell
  nstep=1000
  depred=1
end

```

```

print threshold=0.5 noe
  evaluate ($rms_noe=$result)
  evaluate ($violations_noe=$violations)
print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
print thres=5. angles
  evaluate ($rms_angles=$result)
print thres=5. impropers
  evaluate ($rms_impropers=$result)

sani print threshold=0.0 class hnBw end
  evaluate ($rms_sani_hnBw=$result)
  evaluate ($R_sani_hnBw=$result*100/19.13)
  evaluate ($viol_sani_hnBw=$violations)

sani print threshold=0.0 class hnBf end
  evaluate ($rms_sani_hnBf=$result)
  evaluate ($R_sani_hnBf=$result*100/19.13)
  evaluate ($viol_sani_hnBf=$violations)

sani print threshold=0.0 class hnAw end
  evaluate ($rms_sani_hnAw=$result)
  evaluate ($R_sani_hnAw=$result*100/19.13)
  evaluate ($viol_sani_hnAw=$violations)

sani print threshold=0.0 class hnAf end
  evaluate ($rms_sani_hnAf=$result)
  evaluate ($R_sani_hnAf=$result*100/19.13)
  evaluate ($viol_sani_hnAf=$violations)

remarks =====
remarks repel=0.78
remarks      overall vdw,noe, sani
remarks energies: $ener $vdw, $noe, $sani,
remarks =====
remarks      bonds,angles,impropers,noe
remarks bonds etc: $rms_bonds,$rms_angles,$rms_impropers, $rms_noe
remarks =====
remarks      noe
remarks violations :  $violations_noe
remarks =====
remarks  sani hnBw hnAw hnBf hnAf  Da=-24.0 R=0.13
remarks Rms sani: $rms_sani_hnBw $rms_sani_hnAw $rms_sani_hnBf $rms_sani_hnAf
remarks R factor sani: $R_sani_hnBw $R_sani_hnAw $R_sani_hnBf $R_sani_hnAf
remarks viol sani: $viol_sani_hnBw $viol_sani_hnAw $viol_sani_hnBf $viol_sani_hnAf
remarks =====

evaluate ($file3 = "2nd_rigid_cool2x_" + encode($count) + ".pdb")
write coor output=$file3 end

end loop structure
end loop file
stop

```

D.3 Simulated annealing

After the rigid-body docking protocol, a standard simulated annealing protocol based on a Powell minimization in cartesian coordinate space was applied and is given below.

```
REMARKS Semi-rigid fit of two known structures using NOEs and dipolar couplings
REMARKS very slow cooling with increase of vdw
evaluate ($seed=778666339)

set seed $seed end

param @TOPPAR:parallhdg_new.pro @TOPPAR:par_axis_3.pro end

{* read in the PSF file and initial structure *}

structure @../cylr2B_complex_test.psf @../axis_new.psf end

coor @../ABexplor.pdb

collapse
  assign (resid 1:66 or resid 67:132 ) 50.0 17.50
                                     {* force, radius*}
  {* The entire sequence of both proteins is constrained here. *}
  {* Basically, we expect the complex to be roughly globular. *}
  scale 1.0
end

{* set the weights for the experimental and a priori energy terms *}

evaluate ($knoe = 1.0)      {* noes                               *}
evaluate ($asym = 0.1)     {* slope of NOE potential           *}
evaluate ($kcdi = 10.0)    {* torsion angle force constant     *}
evaluate ($krama = 1.0)    {* database-derived potential force constant. *}

{* Read experimental restraints *}
set echo on message on end

noe
  reset
  nrestraints = 10000          {* allocate space for NOEs *}
  class all @../WT_HN2A.tbl
                             @../33_noeAB.tbl
end
set echo on message on end

evaluate ($knoe = 30.0)
noe
  ceiling 1000
  averaging all sum
  potential all square
  scale all $knoe
  sqconstant all 1.0
  sqexponent all 2
end

evaluate ($ksani = 0.01)

sani
  nres=4000
  class hnBw
  force $ksani
  potential harmonic
  coeff 0.0 -24.7 0.07
  @../hnB.tbl

  class hnBf
  force 0.0
```

```

potential harmonic
coeff 0.0 -24.7 0.07
@../hnB.tbl

class hnAw
force $ksani
potential harmonic
coeff 0.0 -24.7 0.07
@../hnA.tbl

class hnAf
force 0.0
potential harmonic
coeff 0.0 -24.7 0.07
@../hnA.tbl

end

evaluate ($rcon = 0.003)

parameters
nbonds
  atom
  nbxmod 4      /* This is set to 4 (bonds), because we are using a */
                /* internal dynamics which uses torsion angles for */
                /* dynamics steps.                               */
  wmin = 0.01  /* warning off                                         */
  cutnb = 4.5  /* nonbonded cutoff                                             */
  tolerance 0.5
  repel= 0.9   /* scale factor for vdW radii = 1 ( L-J radii)                 */
  rexp = 2     /* exponents in (r^irex - R0^irex)^rexp                         */
  irex = 2
  rcon=$rcon   /* actually set the vdW weight                                  */
  end
end

restraints dihed
  scale $kcdi
  nass = 5000
  @../33_diheAB.tbl
end
set message on echo on end

rama
nres=10000
@GAUSSIANS:shortrange_gaussians.tbl
@GAUSSIANS:new_shortrange_force.tbl
end
set message on echo on end
@GAUSSIANS:newshortrange_setup.tbl /* Using the standard rama database- */
/* derived tables.                  */

evaluate ($init_t = 3000.01)
evaluate ($tol = $init_t/1000)

vector do (mass = 100.0) (all) /* all atom masses are set the same so that */
/* motion of each atom is equally likely in */
/* the dynamics. Note that the dynamics is */
/* not a realistic model of the motion of */
/* the system; it is a method of minimizing */
/* a highly complicated function.         */
vector do (fbeta = 10.0) (all) /* coupling to heat bath */

coor copy end

/* Generate Structures 1 -> 100 */

evaluate ($count =15)

```

```

while ($count < 100) loop structure
  evaluate ($count = $count + 1)

  {====>}                                     {*Filename(s) for embedded coordinates.*}

  vector do (x=xcomp) (all)
  vector do (y=ycomp) (all)
  vector do (z=zcomp) (all)

  flags
    exclude *
    include bonds angles impr vdw noe cdih rama sani coll
    {* turn on experimental data-based energy terms for dynamics and a priori *}
    {* restraints, such as the rama torsion correlation term and the radius of*}
    {* gyration collapse term.                                         *}
  end

  evaluate ($ini_rad = 0.9)          evaluate ($fin_rad = 0.80)
  evaluate ($ini_con= 0.004)        evaluate ($fin_con= 4.0)
  evaluate ($ini_ang = 0.4)          evaluate ($fin_ang = 1.0)
  evaluate ($ini_imp = 0.1)          evaluate ($fin_imp = 1.0)

  evaluate ($ini_noe = 2.0)          evaluate ($fin_noe = 30.0)
  evaluate ($knoe = $ini_noe)        {* NOE force constant          *}
  evaluate ($ini_rama = 0.002)        evaluate ($fin_rama = 1.0)
  evaluate ($krama = $ini_rama)       {* database potential scaling  *}
  evaluate ($kcdi = 10.0)             {* torsion angle force constant *}
  evaluate ($ini_sani = 0.01)         evaluate ($fin_sani = 0.13)
  evaluate ($ksani = $ini_sani)

  sani class hnBw force $ksani end
  sani class hnAw force $ksani end

  noe
    averaging all sum
    potential all square
    scale all $knoe
    sqconstant all 1.0
    sqexponent all 2
  end

  restraints dihed
    scale $kcdi
  end

  rama
    scale $krama
  end

  evaluate ($rcon = 1.0)

  parameters
    nbonds
      atom
      nbxmod 4          {* This is set to 4 (bonds), because the internal *}
                       {* dynamics refines using torsion dynamics   *}
      wmin = 0.01      {* warning off                               *}
      cutnb = 100      {* nonbonded cutoff                               *}
      tolerance 45
      repel= 1.2       {* scale factor for vdW radii = 1 ( L-J radii)   *}
                       {* This is so large, because it is only applied to CA *}
                       {* atoms. (see constraints interaction statement) *}
      rexp = 2         {* exponents in (rirex - R0irex)rexp          *}
      irex = 2
      rcon=$rcon       {* actually set the vdW weight                               *}
    end
  end
end

```

```

constraints
  interaction (not name ca) (all)
  weights * 1 angl 0.4 impr 0.1 vdw 0 elec 0 end
  interaction (name ca) (name ca)
  weights * 1 angl 0.4 impr 0.1 vdw 1.0 end
  { * Due to this set-up, van der Waals forces are turned off except for * }
    { * CA atoms during the high temperature steps of refinement. This * }
    { * allows atoms to pass through one another in the high temperature * }
    { * phase of annealing. This increases the conformational space that * }
    { * can be sampled in the high temperature phase. However, the choice * }
    { * of dynamics type (in this case torsional) will also affect the * }
    { * conformational space accessible. * }
end

dynamics internal
  reset
  group (resid 500 )
  hinge rotate (resid 500)
    { * rigid regions on cylr2A * }
  fix (resid 1:66 and
    ((name ca or name c or name n or name o or name cb or
    name hn or name ha or name ha#) or
    not (resid 1:3 or resid 39:44 or resid 46:48 or resid 50:52 or resid 56:59
    or resid 61:63 or resid 64:66 )))

  { * rigid regions on cylr2B * }
  group (resid 67:132 and
    ((name ca or name c or name n or name o or name cb or
    name hn or name ha or name ha#) or
    not not (resid 67:69 or resid 105:110 or resid 112:114 or resid 116:118 or
    resid 122:125 or resid 127:129 or resid 130:132 )))

  { * Keeping the rigid part of the structures, while adjusting the fit in * }
  { * the flexible regions for cylr2B. * }

  { * group together rigid ring atoms * }

  set message off echo off end
  evaluate ( $res = 1 )
  while ($res le 385) loop group
    group (resid $res and resname PHE and
      (name CG or name CD1 or name CD2 or name CE1 or name CE2 or name
      CZ))
    group (resid $res and resname HIS and
      (name CG or name ND1 or name CD2 or name CE1 or name NE2))
    group (resid $res and resname TYR and
      (name CG or name CD1 or name CD2 or name CE1 or
      name CE2 or name CZ))
    group (resid $res and resname TRP and
      (name CG or name CD1 or name CD2 or name NE1 or
      name CE2 or name CE3 or name CZ2 or name CZ3 or
      name CH2))
    evaluate ( $res = $res + 1 )
  end loop group
  set message on echo on end

  auto torsion
  maxc 10000
end

{ * initial high-temperature equilibration * }

vector do (vx = maxwell($init_t)) (all)
vector do (vy = maxwell($init_t)) (all)
vector do (vz = maxwell($init_t)) (all)

dynamics internal

```

```

itype=pc6
etol=$tol
tbath=$init_t
response= 5
nstep=4000
endtime = 10
timestep= 0.002
nprint = 50
end

parameters
  nbonds
    atom
    nbxmod 4      { * This is set to 4 (bonds), because we are using a * }
                  { * database potential (delphic) that refines 3-bond * }
                  { * interactions (torsion angles) * }
    wmin = 0.01  { * warning off * }
    cutnb = 4.5  { * nonbonded cutoff * }
    tolerance 0.5
    repel= 0.9   { * scale factor for vdW radii = 1 ( L-J radii) * }
                  { * This is reset because all atoms are subject to * }
                  { * van der Waals forces. * }
    rexp = 2     { * exponents in (rirex - R0irex)rexp * }
    irex = 2
    rcon =1.0    { * actually set the vdW weight * }
  end
end

evaluate ($kcdi = 200)
restraints dihed

  scale $kcdi

end

evaluate ($final_t = 100)    { K }
evaluate ($tempstep = 25)   { K }

evaluate ($ncycle = ($init_t-$final_t)/$tempstep)
evaluate ($nstep = int(5000*4.0/$ncycle))
evaluate ($endtime= $nstep*0.002)

evaluate ($bath = $init_t)
evaluate ($tol=$bath/1000)
evaluate ($k_vdw = $ini_con)
evaluate ($k_vdwfact = ($fin_con/$ini_con)^(1/$ncycle))
evaluate ($radius= $ini_rad)
evaluate ($radfact = ($fin_rad/$ini_rad)^(1/$ncycle))
evaluate ($k_ang = $ini_ang)
evaluate ($ang_fac = ($fin_ang/$ini_ang)^(1/$ncycle))
evaluate ($k_imp = $ini_imp)
evaluate ($imp_fac = ($fin_imp/$ini_imp)^(1/$ncycle))
evaluate ($noe_fac = ($fin_noe/$ini_noe)^(1/$ncycle))
evaluate ($knoe = $ini_noe)
evaluate ($rama_fac = ($fin_rama/$ini_rama)^(1/$ncycle))
evaluate ($krama = $ini_rama)
evaluate ($sani_fac = ($fin_sani/$ini_sani)^(1/$ncycle))
evaluate ($ksani = $ini_sani)

vector do (vx = maxwell($bath)) (all)
vector do (vy = maxwell($bath)) (all)
vector do (vz = maxwell($bath)) (all)

evaluate ($i_cool = 0)
while ($i_cool < $ncycle) loop cool
  evaluate ($i_cool=$i_cool+1)

  evaluate ($bath = $bath - $tempstep)

```

```

evaluate ($tol = $bath/1000)
evaluate ($k_vdw=min($fin_con,$k_vdw*$k_vdwfact))
evaluate ($radius=max($fin_rad,$radius*$radfact))
evaluate ($k_ang = $k_ang*$ang_fac)
evaluate ($k_imp = $k_imp*$imp_fac)
evaluate ($knoe = $knoe*$noe_fac)
evaluate ($krama = $krama*$rama_fac)
evaluate ($ksani = $ksani*$sani_fac)

constraints interaction (all) (all) weights
* 1 angles $k_ang improper $k_imp
end end
parameter nbonds
cutnb=4.5 rcon=$k_vdw
  nbxmod=4      {* This is set to 4 (bonds), because we are using a *}
                {* database potential (delphic) that refines 3-bond *}
                {* interactions(torsion angles).                    *}
  repel=$radius
end      end
noe scale all $knoe end

sani class hnBw force $ksani end
sani class hnAw force $ksani end

rama scale $krama end

dynamics internal
etol=$tol
tbath=$bath
nstep=$nstep
  response=5
  endtime=$endtime
  nprint=25
end
end loop cool

{* Reset dynamics for Powell minimizations. *}
dynamics internal
reset
cloop=False

{* dipolar axis *}
group (resid 500 )

  {* rigid regions on cylr2A *}
  fix (resid 1:66 and
  ((name ca or name c or name n or name o or name cb or
  name hn or name ha or name ha#) or
  not (resid 1:3 or resid 39:44 or resid 46:48 or resid 50:52 or resid 56:59
  or resid 61:63 or resid 64:66 )))

  {* rigid regions on cylr2B *}
  group (resid 67:132 and
  ((name ca or name c or name n or name o or name cb or
  name hn or name ha or name ha#) or
  not not (resid 67:69 or resid 105:110 or resid 112:114 or resid 116:118 or
  resid 122:125 or resid 127:129 or resid 130:132 )))

maxe 10000
itype=powell
nstep=20000
maxcalls=20000
nprint=25
etol=1e-7
gtol=0.001
depred=0.0001
end

```



```

flags exclude * include noe sani end
{* Refinement with only intermolecular NOEs and dipolar couplings. *}

noe scale all 60.0 end
sani class hnBw force 2.0 end
sani class hnAw force 2.0 end

constraints inter (resid 1:66) (resid 67:132) end

dynamics internal
  reset
  itype="powell"
  stepsize = 0.1
  nstep=20000
  maxcalls=20000
  depred=0.0001
  fix = (resid 1:66)
  group = (resid 67:132)
  group = (resid 500)
  {* Each protein is treated as a rigid body for this *}
  {* Powell minimization step. *}
  etol = 0.000001
  gtol = 0.0000001
  nprint= 50
end

flags exclude * include noe sani vdw coll end
{* refinement with van der Waals repulsion and collapse term, as well as *}
{* NOEs and dipolar couplings. Exclude * removes all previously-set terms*}

noe scale all 60.0 end
parameter nbonds rcon 4 repel 0.80 end end

dynamics internal
  nstep=20000
  maxcalls=20000
  depred=0.0001
  nprint=50
end

print thres=0.05 bonds
  evaluate ($rms_bonds=$result)
print thres=5. angles
  evaluate ($rms_angles=$result)
print thres=5. impropers
  evaluate ($rms_impropers=$result)
print threshold=0.5 noe
  evaluate ($rms_noe=$result)
  evaluate ($violations_noe=$violations)
print threshold=5. cdih
  evaluate ($rms_cdih=$result)
  evaluate ($violations_cdih=$violations)
sani print threshold=2.0 class hnBw end
  evaluate ($rms_sani_hnBw=$result)
  evaluate ($R_sani_hnBw=$result*100/19.13)
  evaluate ($viol_sani_hnBw=$violations)
sani print threshold=2.0 class hnBf end
  evaluate ($rms_sani_hnBf=$result)
  evaluate ($R_sani_hnBf=$result*100/19.13)
  evaluate ($viol_sani_hnBf=$violations)
sani print threshold=2.0 class hnAw end
  evaluate ($rms_sani_hnAw=$result)
  evaluate ($R_sani_hnAw=$result*100/19.13)
  evaluate ($viol_sani_hnAw=$violations)
sani print threshold=2.0 class hnAf end
  evaluate ($rms_sani_hnAf=$result)
  evaluate ($R_sani_hnAf=$result*100/19.13)
  evaluate ($viol_sani_hnAf=$violations)

```

```
remarks =====
remarks overall,bonds,angles,improper,vdw,cdih,noe, rama,sani, coll
remarks energies: $ener, $bond, $angl, $impr, $vdw, $cdih, $noe, $rama, $sani, $coll
remarks =====
remarks          bonds,angles,impropers,cdih,noe
remarks bonds etc: $rms_bonds,$rms_angles,$rms_impropers,$rms_cdiH,$rms_noe
remarks =====
remarks violations : $violations_cdiH $violations_noe
remarks sani hnBw hnAw hnBf hnAf Da=-24.7 R=0.07
remarks Rms sani: $rms_sani_hnBw $rms_sani_hnAw $rms_sani_hnBf $rms_sani_hnAf
remarks R factor sani: $R_sani_hnBw $R_sani_hnAw $R_sani_hnBf $R_sani_hnAf
remarks viol sani: $viol_sani_hnBw $viol_sani_hnAw $viol_sani_hnBf $viol_sani_hnAf
remarks =====
{====>}          {*Name(s) of the family of final structures.*}
evaluate ($file2 = "sa_coolWT2A2nd_dihe_intra_sc_" + encode($count) + ".sa")
write coor select (resid 1:132) output=$file2 end

end loop structure
end loop file
stop
```



Curriculum Vitae

Personal data

Name	Sigrun Rumpel
Date of birth	November 23 rd , 1978
Place of birth	Göttingen, Germany
Citizenship	German

Education and studies

1985-1991	Primary and grammar School, Göttingen, Germany
07/1998	General qualification for university entrance
11/1998	Beginning of biochemistry studies at the University of Bayreuth, Germany
10/2000	Intermediate diploma in biochemistry at the University of Bayreuth, Germany
2000-2002	Advanced studies of biochemistry at the University of Bayreuth, Germany
04/2002	Final examination in biochemistry at the University of Bayreuth, Germany
05/2002- 11/2002	Diploma thesis with the title "Expression and NMR-spectroscopic examination of a PTH(1-34)-analogon and the N-terminal domain of the PTH-receptor" at the chair of biopolymers from Prof. Dr. Paul Rösch
01/2003- 09/2006	PhD. thesis with the title "Protein NMR studies of two systems involved in bacterial pathogenicity " in the department of NMR-based structural biology of Prof. Dr. Christian Griesinger at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany

Other experience

1996-2002	Tennis trainer for children and adults with trainer certificate since 1997
-----------	--

Lebenslauf

Persönliche Daten

Name	Sigrun Rumpel
Geburtstag	23.11.1978
Geburtsort	Göttingen
Staatsangehörigkeit	deutsch

Schulbildung

1985-1989	Grundschule Herberhausen der Stadt Göttingen
1989-1991	Orientierungsstufe Lutherschule in Göttingen
1991-1998	Theodor-Heuss-Gymnasium in Göttingen
07/1998	Allgemeine Hochschulreife

Studium

11/1998	Beginn des Diplomstudienganges Biochemie an der Universität Bayreuth
10/2000	Vordiplom Biochemie an der Universität Bayreuth
2000-2002	Hauptstudium Biochemie an der Universität Bayreuth
04/2002	Diplomprüfung
05/2002- 11/2002	Anfertigung der Diplomarbeit mit dem Titel „Expression und NMR-spektroskopische Untersuchung eines PTH(1-34)-Analogons und der N-terminalen Domäne des PTH-Rezeptors“ am Lehrstuhl für Biopolymere von Prof. Dr. Paul Rösch
01/2003- 09/2006	Anfertigung der vorliegenden Arbeit mit dem Titel „Untersuchungen mittels Protein NMR an zwei Systemen mit Einfluss auf bakterielle Pathogenität“ in der Abteilung NMR-basierte Strukturbiologie von Prof. Dr. Christian Griesinger am Max Planck Institut für Biophysikalische Chemie in Göttingen

Sonstige Tätigkeiten

1996-2002	Tennistrainerin für Kinder und Erwachsene mit Trainerschein seit 1997
-----------	---