

NMR-based Structural and Dynamical Studies on Non-Native Variants of Hen Egg White Lysozyme

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim
Fachbereich Biochemie, Chemie und Pharmazie
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
CHRISTIAN SCHLÖRB
aus Schotten

Frankfurt am Main
2007



(D30)

Vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Harald Schwalbe

Gutachter: Prof. Dr. Harald Schwalbe und Prof. Dr. Volker Dötsch

Datum der Disputation: 14. November 2007

Die schönste Erfahrung, die ein Mensch machen kann, ist die Verzauberung! Sie ist die Quelle aller wahren Kunst und Wissenschaft. Wem dieses Gefühl fremd ist, wer nicht mehr staunend inne halten kann und in Ehrfurcht verharren kann, der ist eigentlich schon tot, dessen Augen sind blind!

- *Albert Einstein* -

Table of Contents

List of Figures	ix
List of Tables	xiii
List of Abbreviations	xv
Preface	xvii
Abstract	xix
1. Introduction	1
1.1. Protein folding and non-native states of proteins	1
1.1.1. Overview	1
1.1.2. The investigation of protein folding and non-native states of proteins	8
1.2. Background: Nuclear magnetic resonance spectroscopy	10
1.2.1. Overview	10
1.2.2. NMR parameters utilized in this thesis	12
1.3. Hen egg white lysozyme	27
1.4. Motivation and aims	33
2. Materials and Methods	35
2.1. Sample preparation	35
2.1.1. Overview	35
2.1.2. Expression and purification	36
2.1.3. S-Methylation of cysteine residues	38
2.1.4. Site-directed mutagenesis	38
2.1.5. Biochemical analysis	39
2.1.6. NMR samples	39
2.2. Backbone resonance assignment	40
2.2.1. Overview	40

Table of Contents

2.2.2. Acquisition of two- and three-dimensional spectra and sequential backbone assignment	40
2.3. Assignment of tryptophan side chains	46
2.3.1. General assignment strategy	46
2.3.2. The HN(CACB)CG experiment	48
2.3.3. The HN(CD)CG experiment	49
2.4. Photo-CIDNP experiments	53
2.5. Protein Dynamics	56
2.5.1. Heteronuclear relaxation rates	56
2.5.2. Relaxation dispersion	57
2.6. Residual dipolar couplings	58
2.6.1. Experimental determination of RDCs	58
2.6.2. Simulation of RDCs	59
2.7. Residual structure in organic solvents	61
3. Results and Discussion	63
3.1. Sample preparation	63
3.2. Backbone resonance assignment	66
3.3. Chemical shift analysis	71
3.4. Assignment of tryptophan side chains	74
3.5. Photo-CIDNP experiments	77
3.5.1. Two-dimensional ^{15}N - ^1H photo-CIDNP	77
3.5.2. Two-dimensional ^{13}C - ^1H photo-CIDNP	81
3.6. Protein dynamics by heteronuclear relaxation	83
3.7. Residual dipolar couplings	90
3.8. Residual structure in organic solvents	96
3.9. Conclusions and outlook	104
A. Appendix	107
A.1. List of buffers, media and primers	107
A.1.1. M9 minimal media	107
A.1.2. Protein purification buffers	108
A.1.3. Primer for site-directed mutagenesis	108
A.2. NMR acquisition and processing parameters	109
A.3. Resonance assignments	111
A.3.1. HEWL-S ^{Me}	111
A.3.2. <i>all-Ala</i> -HEWL	115

A.3.3. W62G- and W108G- <i>all-Ala</i> -HEWL	119
A.4. Heteronuclear relaxation rates	123
A.5. Residual dipolar couplings	129
A.6. Pulse programs	133
A.6.1. The 2D ^{15}N - ^1H photo-CIDNP pulse program	133
A.6.2. The 2D ^{13}C - ^1H photo-CIDNP pulse program	136
A.6.3. The HN(CACB)CG pulse program	139
A.6.4. The HN(CD)CG pulse program	144
B. References	149
Zusammenfassung	171
Danksagung	177
Curriculum Vitae	179
List of Publications	181

List of Figures

1.1.	Schematic representation of the folding funnel	2
1.2.	Overview of protein folding and misfolding in cells	3
1.3.	^1H - ^{15}N -HSQC spectra of refolded HEWL and <i>all-Ala</i> -HEWL	13
1.4.	Photo-CIDNP of proteins	15
1.5.	Photo-CIDNP cycle	16
1.6.	Orientation of a bond vector relative to the magnetic field B_0 and the molecular frame	18
1.7.	Schematic representation of a bond vector in the principal axis system	19
1.8.	Distribution of RDC values for the possible orientations	20
1.9.	Time-scales of molecular motions and NMR-techniques	23
1.10.	Dependence of R_1 , R_2 and the heteronuclear NOE on the overall rotational correlation time	24
1.11.	Dependence of ^{15}N R_2 heteronuclear relaxation rates on the correlation times for global and internal motions and the order parameter from a model-free approach	26
1.12.	NMR structure of native hen egg white lysozyme	28
1.13.	Distribution of hydrophobicity in HEWL	29
1.14.	^{15}N R_2 relaxation rates in HEWL- S^{Me} and single point mutants of HEWL- S^{Me}	30
1.15.	Conformations with the amino acid sequence of HEWL randomly chosen to represent distinct radii of hydration (R_h)	32
2.1.	Backbone resonances assignment experiments for (non-native) proteins (I)	43
2.2.	Backbone resonances assignment experiments for (non-native) proteins (II)	44
2.3.	Backbone resonances assignment experiments for (non-native) proteins (III)	45
2.4.	Assignment strategy for the tryptophan side chain resonances	46

List of Figures

2.5. Pulse sequence of the HN(CACB)CG experiment	51
2.6. Pulse sequence of the HN(CD)CG experiment	52
2.7. Pulse sequence of the 2D photo-CIDNP experiment	53
2.8. Laser set-up for photo-CIDNP experiments	54
2.9. Riboflavin 5'-mononucleotide (FMN)	55
3.1. SDS polyacrylamide gel after <i>inclusion body</i> washing and solubilization	64
3.2. SDS polyacrylamide gel after ion exchange chromatography	64
3.3. ^1H - ^{15}N -HSQC spectra of refolded HEWL and <i>all-Ala</i> -HEWL	65
3.4. Sequential backbone resonance assignment using the HNCACB experiment	67
3.5. Sequential backbone resonance assignment using the HNN experiment	68
3.6. Annotated ^1H - ^{15}N -HSQC spectrum of <i>all-Ala</i> -HEWL	70
3.7. Secondary chemical shift plots for $^1\text{H}_\alpha$ and $^{13}\text{C}_\alpha$ of <i>all-Ala</i> -HEWL . . .	72
3.9. Results of the assignment strategy for the tryptophan side chain indole resonances	75
3.10. ^1H - ^{15}N -HSQC and 2D photo-CIDNP spectra of <i>all-Ala</i> -HEWL and W62G- <i>all-Ala</i> -HEWL	78
3.11. CIDNP/HSQC normalized intensity ratios of <i>all-Ala</i> -HEWL and W62G- <i>all-Ala</i> -HEWL	79
3.12. ^1H - ^{13}C -HSQC and ^{13}C - ^1H 2D photo-CIDNP spectra of <i>all-Ala</i> -HEWL .	81
3.13. Nomenclature of tryptophan side chains	82
3.14. Heteronuclear relaxation rates for <i>all-Ala</i> -HEWL	84
3.15. Comparison of R_2 relaxation rates of <i>all-Ala</i> -HEWL and HEWL-S ^{Me} . .	85
3.16. Comparison of R_2 relaxation rates of <i>all-Ala</i> -HEWL and W108G- <i>all-Ala</i> -HEWL	87
3.17. ^{15}N single-quantum relaxation rates as a function of CPMG frequency for various residues of <i>all-Ala</i> -HEWL	89
3.18. Quadrupolar splitting in HDO in a stretched polyacrylamide gel . . .	90
3.19. Occurrence of distinct NH-RDC-values in <i>all-Ala</i> -HEWL aligned in a 7 % polyacrylamide gel	91
3.20. NH-RDCs over the residues of <i>all-Ala</i> -HEWL and W62G	92
3.21. NH-RDCs in native HEWL	93
3.22. Simulated NH-RDCs from <i>flexible-meccano</i>	94
3.23. Circular dichroism spectra of HEWL-S ^{Me} and native HEWL	97
3.24. Circular dichroism spectra of HEWL-S ^{Me} in trifluoroethanol	99

3.25. Circular dichroism spectra of HEWL-S ^{Me} in ethanol	100
3.26. Circular dichroism spectra of HEWL-S ^{Me} in ethanol after different times of incubation	101
3.27. Circular dichroism spectra of HEWL-S ^{Me} in ethanol at pH 2 and 10 . . .	102
3.28. ¹ H- ¹⁵ N-HSQC spectrum of <i>all-Ala</i> -HEWL	102
3.29. ¹ H- ¹⁵ N-HSQC spectra of <i>all-Ala</i> -HEWL in 30 % and 60 % ethanol . . .	103
3.30. ¹ H- ¹⁵ N-HSQC spectra of <i>all-Ala</i> -HEWL in 90 % ethanol and in 50 % TFE	103
B.1. Ergebnis und Strategie der Zuordnung der Tryptophan-Seitenketten- Resonanzen in <i>all-Ala</i> -HEWL	172
B.2. Normalisierte CIDNP/HSQC Intensitätsverhältnisse von <i>all-Ala</i> -HEWL und W62G- <i>all-Ala</i> -HEWL	173
B.3. Vergleich der R ₂ Relaxationsraten von <i>all-Ala</i> -HEWL und W108G- <i>all-</i> <i>Ala</i> -HEWL	174
B.4. NH-RDCs gezeigt über die Sequenz von <i>all-Ala</i> -HEWL und W62G- <i>all-Ala</i> -HEWL	175

List of Tables

2.1. Approximate chemical shifts in tryptophan residues of <i>all-Ala</i> -HEWL	47
3.1. Chemical shifts of tryptophan side chain nuclei of <i>all-Ala</i> -HEWL	74
3.2. Heteronuclear ^{15}N transverse relaxation rates in the tryptophan side chains of <i>all-Ala</i> -HEWL and W62G- <i>all-Ala</i> -HEWL	88
3.3. Fractional helicity for various samples of HEWL	98
A.1. M9 minimal media	107
A.2. Inclusion bodies solubilization buffers	108
A.3. Ion exchange buffers	108
A.4. Primers for site-directed mutagenesis	108
A.5. NMR acquisition and processing parameters for the assignment of HEWL-S ^{Me}	109
A.6. NMR acquisition and processing parameters for the resonance assignment of <i>all-Ala</i> -HEWL	110
A.7. Resonance assignment for HEWL-S ^{Me}	111
A.8. Resonance assignment for <i>all-Ala</i> -HEWL	115
A.9. Resonance assignment for W62G- <i>all-Ala</i> -HEWL and W108G- <i>all-Ala</i> -HEWL	119
A.10. Relaxation rates for <i>all-Ala</i> -HEWL	123
A.11. ^{15}N transverse relaxation rates for W108G- <i>all-Ala</i> -HEWL	127
A.12. Residual dipolar couplings for <i>all-Ala</i> -HEWL and W62G- <i>all-Ala</i> -HEWL	129

List of Abbreviations

1D, 2D, 3D	one dimensional, two dimensional, three dimensional
AFM	atomic force microscopy
ANS	1-anilino-8-naphthalene sulfonate
APS	ammonium persulfate
BMRB	Biological Magnetic Resonance Data Bank
CARA	computer aided resonance assignment
CD	circular dichroism
CIDNP	chemically induced dynamic nuclear polarization
CM	carboxymethyl
CPMG	Carr-Purcell-Meiboom-Gill
CSA	chemical shift anisotropy
CSI	chemical shift index
CTAB	cetyl trimethyl ammonium bromide
DHPC	dihexanoylphosphatidylcholine
DIPSI	composite pulse decoupling in the presence of scalar interactions
DMPC	dimyristoylphosphatidylcholine
DNA	desoxy-ribonucleic acid
DSC	differential scanning calorimetry
DSS	2,2-dimethyl-2-silapentane-5-sulfonic acid
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylene diamine tetracetic acid
EM	electron microscopy
EtOH	ethanol
FMN	riboflavin 5'-mononucleotide
FPLC	fast protein liquid chromatography
FRET	fluorescence resonance energy transfer
FTIR	Fourier transform infrared
GARP	globally optimized alternating phase rectangular pulse

List of Abbreviations

GdnHCl	guanidine hydrochloride
hetNOE	heteronuclear nuclear Overhauser effect
HEWL	hen egg white lysozyme
HMQC	heteronuclear multiple-quantum coherence
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence
INEPT	insensitive nuclei enhanced by polarization transfer
IPAP	in-phase/anti-phase
IPTG	isopropyl β -D-1-thiogalactopyranoside
IUPAC	International Union of Pure and Applied Chemistry
LB	Luria broth
MS	mass spectrometry
MWCO	molecular weight cut-off
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
OD ₆₀₀	optical density at 600 nm
PAG	polyacrylamide gel
PFG	pulse field gradient
PG-SLED	pulse gradient-simulated echo longitudinal encode-decode
pH	negative decadic logarithm of the H ⁺ ion activity
pI	isoelectric point
ppm	parts per million
RDC	residual dipolar coupling
rpm	rounds per minute
SAXS	small-angle X-ray scattering
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis
TEMED	N,N,N',N'-tetramethylethylenediamine
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
TOCSY	total correlation spectroscopy
Tris	tris-(hydroxymethyl)-aminomethane
UV	ultra-violet
w/v	weight per volume
XEASY	ETH automated spectroscopy for X window systems

Preface

This thesis is based on the work conducted in the group of Prof. Dr. Harald Schwalbe at the *Institute for Organic Chemistry and Chemical Biology* and the *Center for Biomolecular Magnetic Resonance* at the *Johann Wolfgang Goethe-University* in Frankfurt am Main in the period between May 2003 and May 2007. An *abstract* of this thesis can be found directly following this preface.

The thesis describes experiments for the investigation of structural and dynamical properties of non-native states of hen egg white lysozyme and several variants and mutants thereof and discusses enhanced and newly developed methods as well as the novel insights from the obtained results in the context of preexisting data. It is divided in three major chapters:

The introductory *Chapter 1* (starting at page 1) gives an overview of protein folding and the nature of non-native states of proteins. This section is followed by a short discussion of methods used for the elucidation of these states and a more detailed overview of NMR parameters suitable for the study of proteins in general and non-native states of proteins in particular. The chapter is concluded by a survey of the current knowledge on hen egg white lysozyme, its folding and non-native states and a short outline of the aims of the work this thesis is based on.

Chapter 2 (starting at page 35) describes the materials, methods and experiments used of the work that is summarized in this thesis.

Chapter 3 (starting at page 63) includes the results and findings and their discussion in the general context.

In addition, an *appendix* (starting on page 107) lists the more extended tables of experimental conditions and results that otherwise would harm the readability of the previous chapters. The cited *references* can be found starting from page 149. A *summary* in German can be found near the end of the thesis (page 171). A section with the *acknowledgements* (page 177) and a short *curriculum vitae* of the author (page 179) conclude this document.

Abstract

The formation and maintenance of a defined three-dimensional structure is a prerequisite for most proteins in order to fulfill their function in the native context. However, there are proteins, which are intrinsically unstructured and thus natively unfolded. In addition, the misfolding and aggregation of many proteins can lead to severe diseases. The investigation of non-native states of proteins significantly contributes to the understanding of protein folding and misfolding. Nuclear magnetic resonance (NMR) spectroscopy is the only known technique that can provide information on structure and dynamics of non-native states of proteins at atomic resolution.

Unfolded and non-native states of proteins have to be treated as ensembles of rapidly interconverting conformers and their observed properties are ensemble and time averaged. In this thesis, hen egg white lysozyme (HEWL) and mutants thereof have been investigated by NMR spectroscopy. The reduction of its four disulfide bridges and the successive methylation of the cysteine residues renders HEWL permanently non-native ('HEWL-S^{Me}'). Alternatively, the exchange of the eight cysteines for alanines results in very similar states ('*all-Ala*-HEWL'). Under these conditions, HEWL-S^{Me} and *all-Ala*-HEWL do not resemble random coil conformations, but exhibit residual secondary and tertiary structure. The presence of hydrophobic clusters and long-range interactions around the proteins six tryptophan residues and the modulation of these properties by single-point mutants has been observed.

For the NMR spectroscopic investigation, HEWL has been isotopically labelled in *E. coli* by expression into inclusion bodies. After purification, the $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{N}}$, $^{13}\text{C}_{\alpha}$, $^{13}\text{C}_{\beta}$, $^{13}\text{C}'$, $^1\text{H}_{\alpha}$ and $^1\text{H}_{\beta}$ resonances of HEWL-S^{Me} and *all-Ala*-HEWL have been assigned almost completely using three-dimensional NMR experiments. The analysis of secondary chemical shifts revealed regions in the proteins sequence — particularly around the six tryptophan residues — with significantly populated α -helix like conformations. In order to further elucidate the influence of the tryptophan side chains, a set of two new pulse sequences has been developed that allowed for the successful assignment of the $^{13}\text{C}_{\gamma}$, $^{15}\text{N}_{\epsilon}$ and $^1\text{H}^{\text{N}}_{\epsilon}$ resonances in these side chains.

Abstract

This knowledge was eventually exploited in the interpretation of two-dimensional ^{15}N - ^1H photo-CIDNP spectra, which revealed a differential solvent accessibility of the tryptophan residues in *all-Ala*-HEWL but not in the single point mutant W62G-*all-Ala*-HEWL. In addition, heteronuclear R_2 relaxation rates have been determined for the indole $^{15}\text{N}_\epsilon$ nuclei of *all-Ala*-HEWL and W62G. While in the wild-type like *all-Ala*-HEWL, the rates are different among the six tryptophan residues, in W62G they are more uniform. Together with relaxation data from the amide backbone, these results indicate the significant destabilization of the hydrophobic clusters in the absence of W62. In contrast, in the W108G mutant the profile of the R_2 relaxation rates was not found to be significantly altered. No evidence was found by $R_{1\rho}$ relaxation rates and relaxation dispersion measurements for conformational exchange on slower (micro- to millisecond) timescales.

Residual dipolar couplings have been determined for non-native HEWL in order to retrieve structural information of these states. The differences of the W62G and the wild-type like non-native HEWL is also picked up in NH-RDCs of these proteins aligned in polyacrylamide gels. Significant positive RDCs are observed in the regions of the hydrophobic clusters in *all-Ala*-HEWL, but to a much lesser degree in W62G. So far, all attempts to simulate RDCs from generated non-native ensembles failed even when including long-range contacts or specific ϕ, ψ backbone angle propensities. However, the measured RDCs can be used to cross-validate structural ensembles of non-native HEWL generated by molecular dynamics simulations that are based on restraints from the other experimental data, such as the differential solvent accessibilities from the photo-CIDNP experiments and the data on the hydrophobic clustering gained from the combined mutational and relaxation studies.

Finally, non-native HEWL has been investigated for the first time using two-dimensional NMR in organic solvents, which are able to induce secondary structures and ultimately lead to amyloid formation. Under these conditions severe line broadening was observed, which was attributed to exchange between different — mostly α -helical — conformations.

In summary, in this thesis methods have been developed, optimized and successfully applied for the structural and dynamical characterization of non-native states of proteins and the effect of single-point mutants on the properties of such ensembles has been investigated. Data has been gained that can considerably contribute to the further elucidation of the nature of non-native states of HEWL by molecular dynamics simulations.

1. Introduction

1.1. Protein folding and non-native states of proteins

1.1.1. Overview

Proteins are the main mediators of biochemical reactions and biomechanical activities of living cells and organisms. Their functions are very diverse and encompass all kinds of activities ranging from synthesis and degradation of other bio(macro)-molecules and signal transduction to motion and regulation of biological processes. They consist of amino acids and are produced by formation of peptide bonds following a mRNA template in a process called *translation* at the ribosomes — organelles consisting of RNA and proteins. In many cases, proteins are modified during synthesis or post-translationally in order to gain their functional structure or localization. The process of the formation of a more or less defined three-dimensional structure from the amino acid sequence of the polypeptide chain is called *protein folding* (Anfinsen, 1973). Protein structure can be described at different levels: The sequence of amino acids is designated *primary structure*, the formation of repetitive structural elements stabilized by hydrogen-bonds such as α -helices and β -sheets is called *secondary structure*, while the overall shape of a protein is referred to as the *tertiary structure* and is stabilized for example by disulfide bonds between cysteine residues or by hydrophobic interactions. The organization of proteins to oligomeric entities occurs at the level of the *quaternary structure*. Thus, protein folding, together with the *transcription* of DNA into mRNA and the translation of mRNA into a polypeptide, is the key process in the transformation of genetic information into biological activity.

However, in living cells protein folding usually takes place within seconds after initiation of the translation and in many cases — especially for smaller proteins — this process is even faster. After many decades of research dedicated to this problem, the process of protein folding is still under investigation. In particular, the very short time in which proteins are folded after translation is remarkable. Thus, protein folding cannot be explained by a simple random search through the space of allowed conformations of the peptide backbone after or during synthesis at the ribosome, for

1. Introduction

such an approach would take more than 100 billion years even for a small protein of 100 residues (Levinthal, 1969).

Instead, protein folding is a highly cooperative process in which domains and sub-domains of a protein can fold parallelly and partially folded intermediate states are populated. During protein folding, secondary structure is formed locally and contacts between residues that are far apart in the sequence occur in addition. The main tertiary interactions are the formation of disulfide bridges and hydrophobic clusters. The formation of tertiary contacts instead of the rapid formation of secondary structure elements is assumed to be the rate-limiting step in protein folding in theoretical folding models (Yon, 2002).

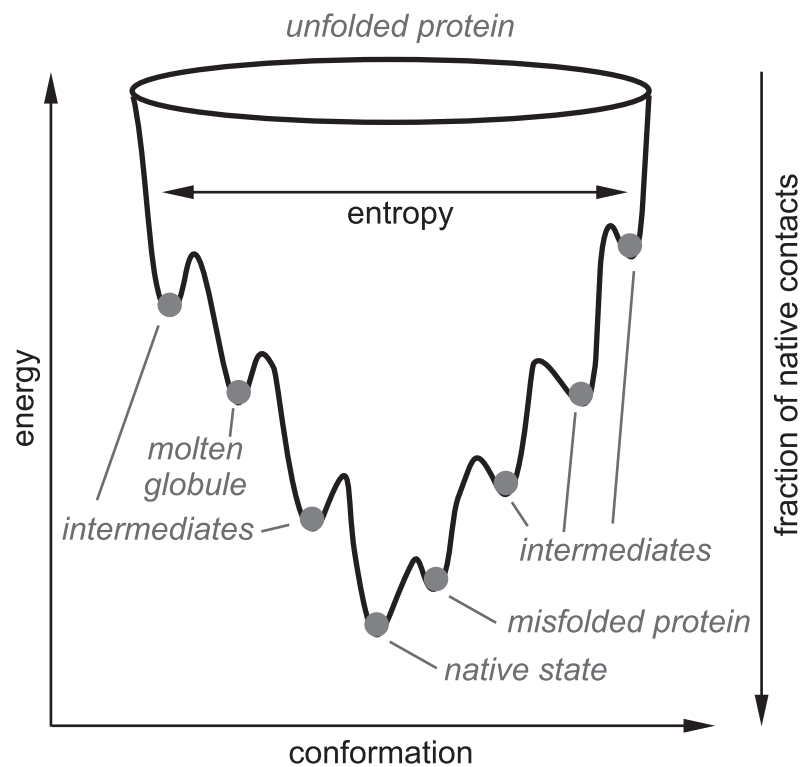


Figure 1.1.: Schematic representation of the folding funnel. The width of the funnel is a measure for the entropy, while its depth is a measure of the energy (modified after Yon, 2002).

Protein folding can be described using energy landscapes and folding funnel models defined by the amino acid sequence in order to treat the thermodynamic and kinetic properties of the transition of an ensemble of non-native states into the dominant native state (Dobson, 2003). During such a transition, parallel pathways can be

1.1. Protein folding and non-native states of proteins

followed. In a folding funnel representation (as in figure 1.1 on the facing page) the number of possible conformations and the entropy of the protein decrease from the completely unfolded states at the top of the funnel to the completely folded native state at the bottom. The steepness of the funnel is a measure for the rate of folding. During folding partially folded or misfolded proteins can get trapped in local minima and finally aggregate.

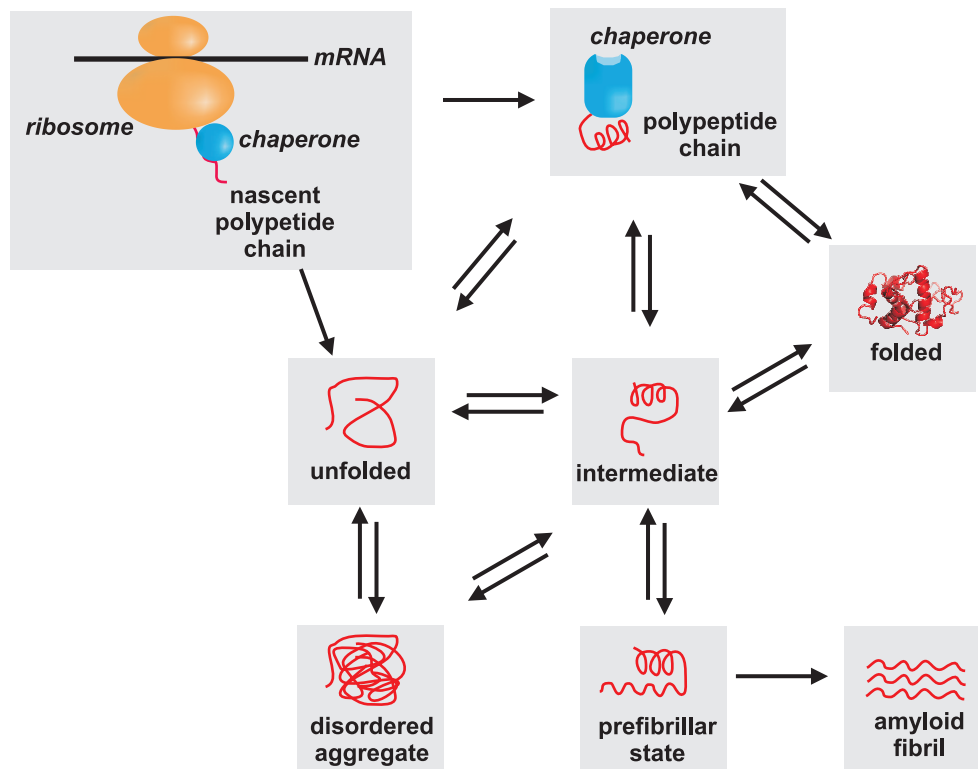


Figure 1.2.: Overview of protein folding and misfolding in living cells. Peptide chains are depicted in red, the ribosome in orange and molecular chaperones in blue. The folded protein is symbolized by a conformer of the liquid NMR structural ensemble of native hen egg white lysozyme (Schwalbe *et al.*, 2001, PDB: 1E8L).

An overview of different states a protein can assume during folding in a living cell is shown in figure 1.2. Already during synthesis at the ribosomes, the nascent polypeptide chains can be bound by auxiliary proteins that assist with the correct folding. This is necessary to prevent proteins from aggregation or other unwanted interactions and to facilitate the process of folding. Such helper proteins are called *molecular chaperones*, a term first used for the protein nucleoplasmin which facilitates the assembly of nucleosomes from histones and DNA (Laskey *et al.*, 1978). Hartl and

1. Introduction

coworkers define a molecular chaperone as ‘a protein which transiently binds to and stabilizes an unstable conformer of another protein, and through regulated binding and release, facilitates its correct fate *in vivo*: be it folding following *de novo* synthesis, transit across a membrane, or stress-induced denaturation, oligomeric assembly, interaction with other cellular components, switching between active and inactive conformations, intracellular transport, or proteolytic degradation, either singly or with the help of co-factors’ (Agashe and Hartl, 2000; Hendrick and Hartl, 1993). In addition to molecular chaperones, other proteins such as *protein disulfide isomerases* (Freedman *et al.*, 2002) and *peptidyl prolyl cis/trans isomerases* (Shaw, 2002) assist in protein folding by catalyzing the formation of disulfide bridges and the isomerization of cis/trans prolyl-peptidyl bonds.

A massive drawback for protein folding is the high molecular crowding and the resulting excluded volume effects in living cells (300 g/L of protein and other macromolecules (Ellis, 2001a,b)), so that nascent, partially folded and misfolded proteins can easily aggregate due to interactions of their exposed hydrophobic surfaces. By protecting exposed hydrophobic surfaces of proteins, molecular chaperones inhibit aggregation. Protein *aggregation* itself is defined as the ‘association of two or more polypeptide chains to form nonfunctional structures’ (Ellis, 2006), whereas protein *misfolding* involves the formation of conformations that cannot proceed to the native conformation on a biological relevant timescale. A protein conformation is designated to be *native* when the protein is existent in its biological functional and relevant form. However, there are proteins which are *intrinsically unstructured* and thus *natively unfolded* (Wright and Dyson, 1999). Such proteins or protein domains are mostly characterized by a low compactness, the absence of globularity, a low secondary structure content and a high flexibility. They usually have a low content of hydrophobic amino acid residues but a high net charge and a low complexity, e.g. manifested in repetitive sequences (Uversky, 2002). Other members of the class of natively unfolded proteins are more compact and resemble a *molten globule* state. Molten globules are folding intermediates occurring on the (re-)folding pathway of some but not all proteins. They are relatively compact states, comprise a high amount of secondary structure, but a lack well-defined tertiary structure (Creighton, 1997; Ptitsyn, 1995). Their NMR spectra usually contain very broad peaks due to conformational exchange on the intermediate timescale.

As of May 2007, more than 460 proteins have been reported to be completely or partially natively unfolded (Database of Protein Disorder (DisProt) (Vucetic *et al.*, 2005)). Many of them function as regulators in key cellular processes and fold upon

1.1. Protein folding and non-native states of proteins

ligand binding. One benefit of proteins being natively unstructured is their high flexibility and the ability to bind a high number of binding partners. However, the existence of such proteins requires the reconsideration of a fundamental paradigm in biochemistry, the *structure-function relationship*. It describes the correlation of a particular function of a protein to a defined structure and has already been postulated more than a century ago (Fischer, 1894).

A large number of severe and widespread human diseases is related to protein misfolding and aggregation. These range from neurodegenerative disorders such as *Alzheimer's disease*, *Parkinson's disease* and *spongiform encephalopathies* to *amyloidoses* and *Type II diabetes*. Most of these diseases are sporadic, many are hereditary and a few are even transmissible (Chiti and Dobson, 2006). The failure of a protein in folding correctly or remaining in the correctly folded state can lead to a reduction of the quantity of the respective protein in the cell and thereby keeping it from playing its normal role. However, the vast majority of protein misfolding diseases is associated with the conversion of soluble peptides or proteins into insoluble extra- or intracellular fibrillar aggregates. These highly ordered structures are designated *amyloid fibrils*. The proteins capable of forming amyloid fibrils are very diverse in structure, size and function, some are natively unfolded, such as the amyloid β peptide responsible for Alzheimer's disease or the N-terminal domain of the *prion* proteins, that can cause spongiform encephalopathies. Nevertheless, the amyloid fibrils are structurally similar, having a diameter of 5 to 13 nm. They are straight, rigid and unbranched and usually comprise β strands perpendicular to and backbone hydrogen bonds parallel to the fibril axis (Rochet and Lansbury, 2000). Amyloid fibrillogenesis in most cases requires the partial unfolding of the fibril-forming proteins. The amyloid fibrils are formed by a *nucleated growth* mechanism, in which initially nuclei are formed from aggregation-prone species of the protein. This lag phase is followed by an exponential growth phase.

Hence, the characterization of non-native states of proteins (both on and off the folding pathway) is adding to the understanding of protein folding and misfolding and thus to the understanding of folding-related diseases. Non-native states of proteins can be produced *in vitro* either by *denaturation* of the native protein or directly by (bio-)synthesis of non-native variants of the protein. Denaturation can be achieved for example by a heat shock or maintaining a low pH. The addition of high concentrations of denaturing chaotropic agents, such as urea or guanidine hydrochloride (GdnHCl), yields unfolded states by destabilization of hydrogen bonds. Detergents, such as sodium dodecyl sulfate (SDS), and disulfide reducing reactants,

1. Introduction

such as 2-mercaptoethanol or dithiothreitol (DTT), are widely used as denaturants. Organic solvents, particularly alcohols, can also act as denaturants, but unlike other chemical denaturants they do not induce random coil structure but rather stabilize well-ordered local structures such as α -helices, though the induced secondary structures are not necessarily native-like (Hamada and Goto, 2005). The denaturation of proteins often is reversible, allowing for example the determination of melting curves.

A completely unstructured polypeptide chain would resemble a *random coil*, i.e. a statistical ensemble of rapidly interconverting conformers with a random distribution of (sterically allowed) bond angle distributions and no non-random interactions (McCarney *et al.*, 2005; Smith *et al.*, 1996b). However, even under highly denaturing conditions, many proteins still exhibit residual local structure and/or long-range contacts (Mittag and Forman-Kay, 2007), which can be native-like or non native-like. For the first time, residual structure in a protein under strongly denaturing conditions was reported for the *434-repressor* protein, for which native-like hydrophobic contacts in the presence of 7 M urea have been reported (Neri *et al.*, 1992). To date, many examples are known of proteins with residual secondary structure in their denatured or unfolded state, among them the acyl-CoA binding protein (ACBP) (Fieber *et al.*, 2004), α -synuclein (Bussell and Eliezer, 2001), apomyoglobin (Yao *et al.*, 2001), barnase (Arcus *et al.*, 1995), the immunoglobulin binding domain of streptococcal protein G (Frank *et al.*, 1995; Sari *et al.*, 2000), the SH3 domain of drk (Blanco *et al.*, 1998) and hen egg white lysozyme (Schwalbe *et al.*, 1997; Wirmer *et al.*, 2004). Examples for which long-range interactions or tertiary structure have been reported under denaturing conditions include staphylococcal nuclease (Gillespie and Shortle, 1997a,b; Shortle and Ackerman, 2001), apomyoglobin (Lietzow *et al.*, 2002), α -synuclein (Bertoncini *et al.*, 2005), ACBP (Fieber *et al.*, 2004; Lindorff-Larsen *et al.*, 2004), mutants of the Fyn SH3 domain (Neudecker *et al.*, 2006), protein L (Yi *et al.*, 2000), human and bovine α -lactalbumin (Wirmer *et al.*, 2006a) and hen egg white lysozyme (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004). However, other proteins such as ubiquitin show no significant residual structure in their denatured state and therefore resemble a random coil (Wirmer *et al.*, 2006b), although recent studies suggest the significant population of conformations with native-like β -hairpins even under highly denaturing conditions (Meier *et al.*, 2007). The unfolded state of a protein is generally determined by its primary sequence. Local secondary structure and long-range interactions can act as nucleation sites for the folding towards the native conformation.

1.1. Protein folding and non-native states of proteins

Ensembles of *random coil* conformers can be calculated by Monte Carlo procedures using ϕ, ψ distributions extracted from high resolution structures from protein databases (Fiebig *et al.*, 1996; Schwalbe *et al.*, 1997). Even highly denatured proteins have to be treated as heteropolymer chains with residue-specific and local sequence-dependent properties (Tran and Pappu, 2006).

1. Introduction

1.1.2. The investigation of protein folding and non-native states of proteins

The properties of unfolded or non-native states of proteins as well as the process of protein folding can be investigated using a wide variety of different biophysical techniques (Buchner and Kiefhaber, 2005; Dobson, 2004; Plaxco and Dobson, 1996). The secondary structure of proteins can be probed by far-ultraviolet (UV) (180-250 nm) circular dichroism (CD) spectroscopy, which is based on the differential absorption of circularly polarized light by secondary structure elements (Kelly *et al.*, 2005). Similarly, Fourier transform infrared (FTIR) spectroscopy can provide information on the backbone conformation and thus the secondary structure. In contrast, near-UV (250-350 nm) CD can report on certain aspects of the tertiary structure of proteins, such as the environment of aromatic side chains and the existence of disulfide bridges and therefore is particularly valuable to monitor protein folding. The same is true for fluorescence and ultraviolet absorbance spectroscopy, either of the intrinsic fluorophores or chromophores of a protein or with the use of external probes such as fluorescence quenchers or external dyes (e.g. 1-anilino-8-naphthalene sulfonate (ANS)). These agents are suitable for the investigation of extended hydrophobic patches and clefts and the exposure of aromatic side chains to the solvent. Fluorescence resonance energy transfer (FRET) can be used to monitor distances in a protein during folding processes or to screen for protein variants with increased thermodynamic stability (Philipps *et al.*, 2003). Small angle X-ray scattering (SAXS) allows for the determination of the radius of gyration of a protein state of interest and thus yields very useful data e.g. of misfolded species. Deuterium pulse labelling techniques combined with mass spectrometry (MS) can report on the formation and the persistence of hydrogen bonds in folding intermediates and thus provide information on the secondary structure.

For the investigation of protein aggregates and amyloid structures, multiple techniques are available: Different dyes (including Congo red and thioflavin) can be used to probe and stain amyloid fibers, while X-ray fiber diffraction, solid-state NMR, atomic force microscopy (AFM) and electron microscopy (EM) report on the structure of such entities at different levels.

However, most of the insights on non-aggregated structures from the methods described above, can also be obtained using liquid-state nuclear magnetic resonance (NMR) spectroscopical techniques, which in addition allow for atomic resolution. The potency of solution NMR spectroscopy for the investigation of non-native states of proteins and protein folding and the corresponding NMR parameters will be dis-

1.1. Protein folding and non-native states of proteins

cussed in the following sections.

Time-resolved data on protein (re-)folding can be obtained by the combination of spectroscopic methods with stopped-flow techniques. Another possibility — especially when combined with NMR — is the investigation of protein folding induced by laser-triggered ion release (Kühn and Schwalbe, 2000; Wirmer *et al.*, 2001).

1. Introduction

1.2. Background: Nuclear magnetic resonance spectroscopy

1.2.1. Overview

Nuclear magnetic resonance (NMR) spectroscopy is a very powerful tool to investigate structures, dynamics and interactions of biomacromolecules. Its main advantages include the possibility to investigate at atomic resolution and to study the molecules of interest in solution. Challenges of this method include the need for relatively high concentrations (usually in the mM range), a drawback that in part has been overcome by the development of cryogenic probes and high field magnets, which enable experiments at the medium to high μM concentration range. A limitation of liquid-state NMR spectroscopy still is the size of the investigated molecules, which not only leads to significant resolution problems due to overlapping signals, but also to severe broadening of the signals. The larger a molecule is, the slower it tumbles in solution. This in turn leads to an increase of relaxation processes and thereby to line broadening and resolutional and experimental efficiency problems. In addition, intermediate chemical exchange can also lead to severe line broadening.

The prerequisite for the investigation of biomacromolecules at atomic resolution is the assignment of resonances to particular nuclei, which itself is based on the labeling of the molecules of interest with NMR active isotopes. Modern assignment experiments are based on the incorporation of ^{13}C and ^{15}N nuclei into the molecules and the selective transfer of magnetization between ^1H , ^{13}C and ^{15}N nuclei. This will be explained in detail in section 2.2 on page 40.

A large number of NMR parameters can be used for the characterization of proteins and other biomacromolecules. Those used for the investigation of non-native lysozyme in this thesis will be discussed in more detail in section 1.2.2 on page 12. It should be emphasized, that the NMR parameters observed for non-native proteins are averaged over time and the ensemble of interconverting conformers.

Chemical shifts are not only a measure for the position of distinct peaks in the spectra, but also can provide information on secondary structures (see page 12).

Scalar coupling constants, particularly vicinal (^3J) coupling constants, can report on protein conformations since they depend on the torsion angle between the connected spins. The relationship between coupling constants and dihedral angles is given by so-called *Karplus* relations (Karplus, 1963). Karplus relations exist for a large number of coupling constants reporting on the backbone torsion angles (ϕ , ψ) in proteins. Since the Karplus relations are not single-valued, but give up to four possible solutions for every constant, multiple experiments have to be combined in

1.2. Background: Nuclear magnetic resonance spectroscopy

order to obtain the correct main chain torsion angles. The analysis of coupling constants in the case of non-native states of proteins is performed by comparison with predictions from a random coil model (Fiebig *et al.*, 1996; Schwalbe *et al.*, 1997; Smith *et al.*, 1996a,b) or with experimental values from random coil peptides (Plaxco *et al.*, 1997).

For folded proteins, the homonuclear *nuclear Overhauser effect* (NOE) yields through-space distance restraints up to 5-6 Å that are routinely used for structure calculations and are also used for sequential assignments. In non-native proteins NOEs are rarely observed because in the ensemble of conformers on average two protons are not close enough (Dyson and Wright, 2004).

Long-distance information can instead be gained from *paramagnetic spin labels* covalently attached to cysteine, histidine or lysine residues or the N-terminus of the protein. *Paramagnetic resonance enhancement* (PRE) leads to broadened amide proton peaks in ^1H - ^{15}N -HSQC spectra of residues in up to 15 Å distance to the spin label and has already been applied to non-native proteins (Argirevic, 2006; Gillespie and Shortle, 1997a,b). Nitroxide containing spin-labels or paramagnetic ions such as Mn^{2+} or Gd^{3+} are most frequently used in such approaches.

NMR *diffusion* measurements can be used to determine the *hydrodynamic radii* (R_h) of spherical proteins or population averaged ensembles of non-native states. Pulse field gradient (PFG) NMR experiments are performed for this purpose with a PG-SLED (pulse gradient-simulated echo longitudinal encode-decode) sequence with varying strengths of a bipolar gradient pulse. Determination of the hydrodynamic radius for the protein of interest is achieved by comparison with an internal standard, e.g. dioxane.

Amide proton *hydrogen exchange* methods can be used to identify protons that are involved in hydrogen bonds or are buried in the protein core, because these protons are protected against exchange with the solvent. Hydrogen exchange therefore can provide valuable information on protein folding intermediates, unfolded states and the folding kinetics, when combined with pulse labeling techniques. Since the exchange of hydrogen for deuterium (and vice versa) is strongly pH dependent, pH jumps can be used for the quenching of the exchange reaction.

An introduction to secondary chemical shifts, heteronuclear relaxation, photo-CIDNP and residual dipolar couplings (RDCs) and the application of these parameters for the characterization of non-native proteins is given in the next section.

1. Introduction

1.2.2. NMR parameters utilized in this thesis

Secondary chemical shifts

After the assignment of the backbone resonances of a given peptide chain, chemical shift values are available and can readily be used to extract residue-specific structural information assuming their correct referencing. Since chemical shifts depend on the exact chemical environment of the nucleus of interest, they do not only reflect a sequence-dependency but are also influenced by the presence of secondary structure (Spera and Bax, 1991; Wishart *et al.*, 1991) and tertiary contacts. Secondary chemical shifts ($\Delta\delta$) are the difference of the measured and correctly referenced chemical shifts (δ_{exp}) and random coil chemical shifts (δ_{rc}) from the literature:

$$\Delta\delta = \delta_{exp} - \delta_{rc} \quad (1.1)$$

An NMR spectrum of an ideal random coil polypeptide would resemble a spectrum of the mixture of the amino acids comprising this peptide (McDonald and Phillips, 1969). Random coil chemical shifts have been collected from random coil regions of proteins (Wishart and Sykes, 1994a; Wishart *et al.*, 1992) and unstructured peptides (Schwarzinger *et al.*, 2000). Furthermore, effects of neighboring residues on the random coil chemical shifts of distinct residue types have been extracted from suitable peptides (Schwarzinger *et al.*, 2001; Wishart *et al.*, 1995a). In general, spectra of non-native states of proteins exhibit a considerable smaller signal dispersion than the spectra of folded proteins. Figure 1.3 on the facing page illustrates this at the example of the ^1H - ^{15}N -HSQC spectra of refolded, native HEWL and non-native *all-Ala*-HEWL. There is more dispersion in the $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{H}}$ and $^{13}\text{C}'$ resonances in unstructured proteins than in the $^1\text{H}_{\alpha}$, $^1\text{H}_{\beta}$, $^{13}\text{C}_{\alpha}$ and $^{13}\text{C}_{\beta}$ resonances, reflecting the sensitivity of the former nuclei to the nature of the neighboring amino acids in the primary sequence of the protein (Yao *et al.*, 2001).

In α -helical structures the $^{13}\text{C}_{\alpha}$ and $^{13}\text{C}'$ chemical shifts are shifted downfield, while they are shifted upfield in β -sheets. For $^{13}\text{C}_{\beta}$ and $^1\text{H}_{\alpha}$ the situation is inverse (Wishart and Sykes, 1994b; Zhang *et al.*, 2003).

In so-called *chemical shift index* (CSI) plots, chemical shift deviations from the random coil chemical shifts are normalized to 1, 0 and -1 depending on their direction and extent and then plotted against the sequence of the protein (Wishart and Sykes, 1994a; Wishart *et al.*, 1992). This provides a quick and reliable method to identify and illustrate stretches of secondary structural elements in proteins.

1.2. Background: Nuclear magnetic resonance spectroscopy

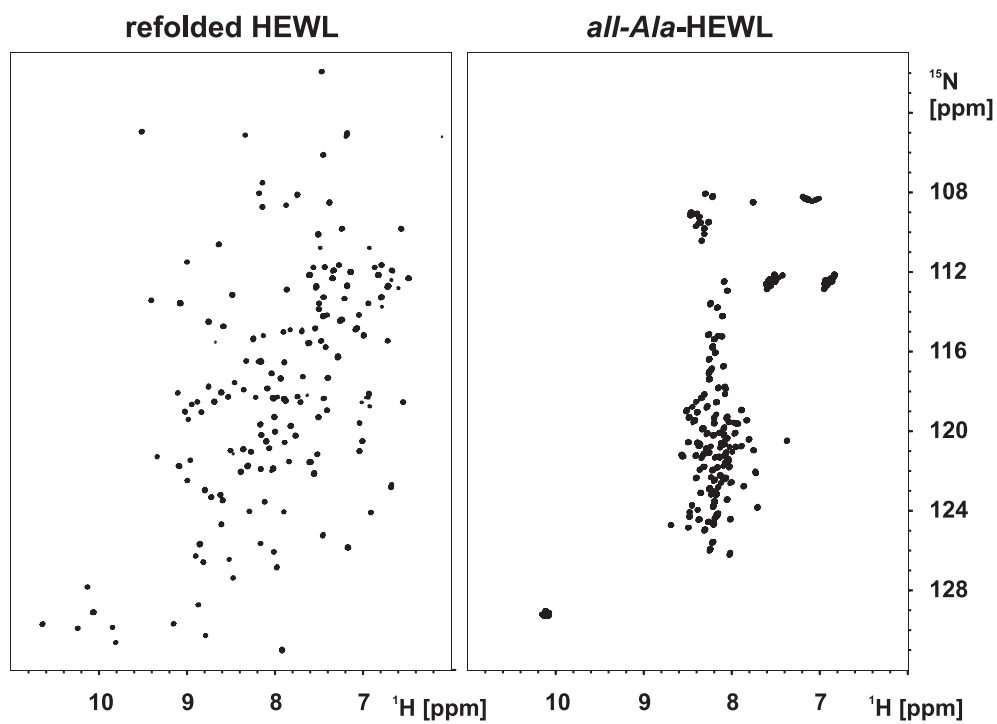


Figure 1.3.: Differences in chemical shift dispersion: ^1H - ^{15}N -HSQC spectra of refolded HEWL at pH 3.8 (left) and *all-Ala*-HEWL at pH 2.0 (right).

1. Introduction

Photo-CIDNP

CIDNP is an effect known in NMR spectroscopy already for 40 years and historically but misleadingly named after the term *Chemically induced dynamic nuclear polarization* (Bargon, 2006; Bargon *et al.*, 1967; Lawler, 1967; Ward and Lawler, 1967). Certain lines in NMR spectra are selectively enhanced by temporary interactions of radical pairs and can be detected as absorptive or emissive NMR signals. The IUPAC defines CIDNP as a ‘Non-Boltzmann nuclear spin state distribution produced in thermal or photochemical reactions, usually from colligation and diffusion, or disproportionation of radical pairs, and detected by NMR spectroscopy by enhanced absorption or emission signals’ (Nic *et al.*, 2006). Since the discovery of this phenomenon and its elucidation (Closs, 1969; Kaptein and Oosterhoff, 1969), photo-CIDNP methods have particularly been developed further for the investigation of biomacromolecules (Hore and Broadhurst, 1993; Kaptein *et al.*, 1978; Mok and Hore, 2004). In such experiments, the radical is temporarily generated by irradiation with light, usually a laser beam is used. A major advantage of CIDNP methods versus conventional NMR spectroscopy is its signal enhancement and thus the improved sensitivity.

Photo-CIDNP NMR methods have two major applications for the investigation of protein folding and unfolded proteins: Firstly, they can be used to probe solvent accessibilities of aromatic residues (Broadhurst *et al.*, 1991; Kaptein *et al.*, 1978; Lyon *et al.*, 1999) and secondly they can be used in combination with time-resolved NMR spectroscopy to follow the kinetics of protein folding (Dobson and Hore, 1998; van Nuland *et al.*, 1998; Wirmer *et al.*, 2001).

Photo-CIDNP NMR experiments rely on the temporary interaction of a photoexcited dye — primarily a flavin such as riboflavin 5'-mononucleotide (FMN) — with an aromatic amino acid side chain (tryptophan, tyrosine and histidine) or to a much lesser degree with methionine. The photoexcitation of the dye is usually achieved by a short laser pulse. The laser beam is coupled into the NMR-tube with optical fibers and triggered by the NMR pulse program. Applying the photo-CIDNP method to a protein results in the signal enhancement and in a better resolution in the aromatic region of the NMR spectrum. Intensities of photo-CIDNP signals depend on the polarization efficiency (Trp > Tyr \gg His \gg Met) and on the solvent accessibility of the residues. In ^1H photo-CIDNP emissive signals are observed for tyrosine residues, whereas the signals of tryptophan and histidine are absorptive. Figure 1.4 on the next page illustrates the photo-CIDNP effect on the example of the aromatic region of the spectra of native α -lactalbumin: Comparable signal-to-noise ratios re-

1.2. Background: Nuclear magnetic resonance spectroscopy

quire much more scans (in this case 128 vs. 1) for conventional one-dimensional NMR spectra than for the photo-CIDNP spectrum. Photo-CIDNP-spectra are usually recorded as differences between light (photo-excited) and dark spectra. During a photo-CIDNP experiment a photo cycle as shown in figure 1.5 on the following page is passed through. Upon light irradiation by the laser beam, the photosensitizer (e.g. FMN) is excited to a triplet state and a triplet radical pair is formed after an electron (or a hydrogen in the case of tyrosine) is transferred from the amino acid side chain to the photosensitizer. This triplet radical pair can either evolve to a singlet radical pair and eventually by an electron back transfer to the ground state of the two reactants ("recombination products"), or it can separate to form individual free radicals and finally after an encounter of the two species and recombination the ground states ("escape products"). The chemical reactivities of the recombination and escape pathways depend on the nuclear spin state. This results eventually in a non-Boltzmann distribution of the spins in the aromatic side chain and therefore a signal enhancement is observed (Hore and Broadhurst, 1993; Mok and Hore, 2004).

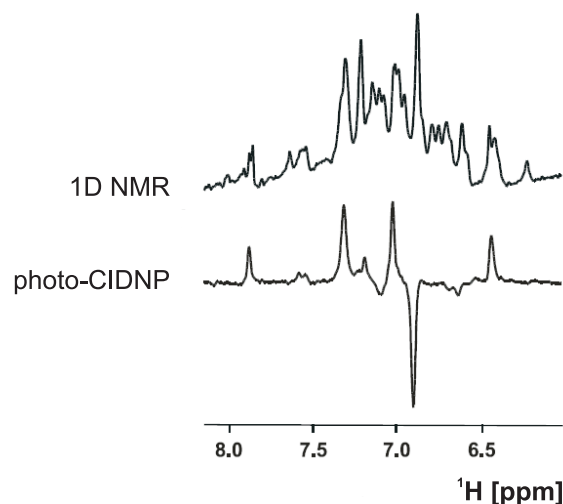


Figure 1.4.: The photo-CIDNP effect: Comparison of a standard one-dimensional NMR spectrum (128 scans) with a photo-CIDNP spectrum (1 scan) of folded α -lactalbumin. Spectra taken from Wirmer (2005).

In solution NMR spectroscopy photo-CIDNP is usually performed in ^1H spectra, although higher signal enhancements have been reported for ^{15}N nuclei (Lyon *et al.*, 1999). An important advantage of photo-CIDNP NMR for unfolded states of proteins is the better resolution due to the fact that only few residues are polarized. It is therefore often possible to assign peaks in unlabeled proteins to the respective

1. Introduction

residues and furthermore follow the changes in accessibility in different unfolded states (Broadhurst *et al.*, 1991).

Heteronuclear two-dimensional photo-CIDNP experiments have been introduced to overcome the resolution limitations of one-dimensional NMR (Scheek *et al.*, 1984, 1985). A $^{15}\text{N}, ^1\text{H}$ heteronuclear photo-CIDNP experiment has been reported and applied to native and urea-denatured HEWL (Lyon *et al.*, 1999). This experiment has the advantage, that only ^{15}N -bound protons are detected and a much larger signal enhancement is observed due to the differences in the hyperfine coupling and in the Boltzmann polarization of ^{15}N .

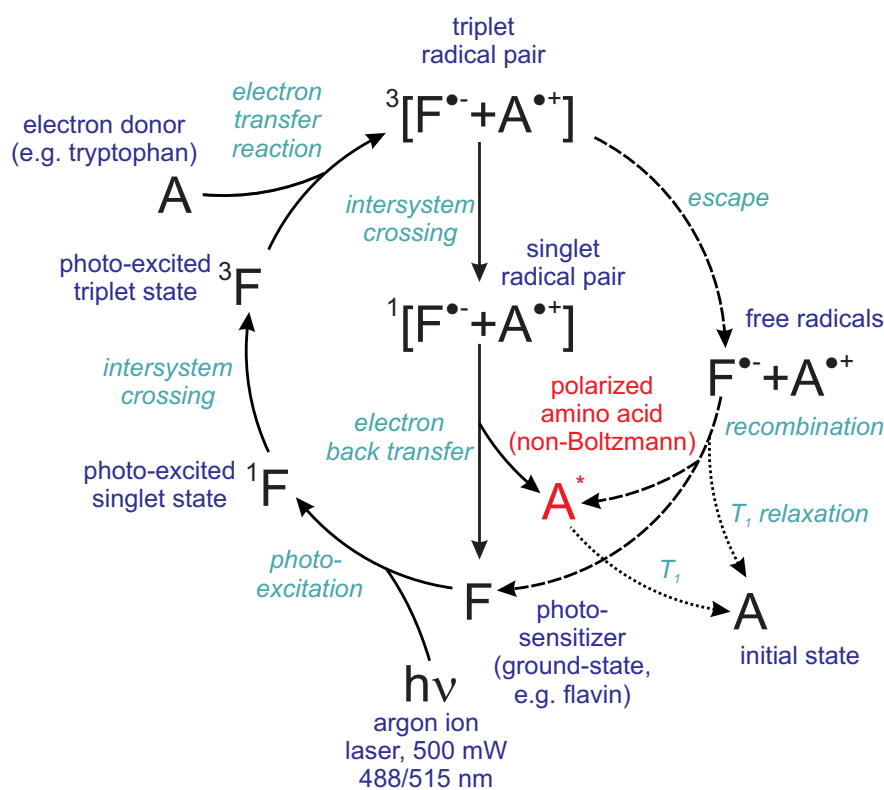


Figure 1.5.: The photo-CIDNP cycle. F is the photosensitizer (e.g. FMN) and A is the solvent-accessible amino acid side chain.

Residual dipolar couplings

Residual dipolar couplings (RDCs) are very valuable parameters in structure determination by modern solution NMR spectroscopy (Lipsitz and Tjandra, 2004; Prestegard *et al.*, 2000). They yield relative orientations of bond vectors, which can be put into calculations as long-distance restraints. Dipolar couplings itself are based on the direct interaction of two proximate magnetic dipoles through space depending on their orientation with respect to the external magnetic field B_0 . However, in solution dipolar couplings are averaged out because molecules tumble isotropically and all orientations are equally probable. If, for any reason, one orientation is preferred, residual dipolar couplings can be detected. The size of the RDCs is tunable by the degree of anisotropy, i.e. to which extent the molecules are aligned in the medium. Usually the residual dipolar couplings between two directly bonded nuclei (e.g. ^1H - ^{15}N) are determined. An alignment of a factor of 10^{-3} to 10^{-4} as compared to an ordered solid is desired in order to obtain RDC values in the order of 25 Hz (as opposed to 25 kHz in a crystal) for the amide ^1H - ^{15}N bond in protein backbones (Prestegard *et al.*, 2004).

The value of the dipolar coupling D_{ij} for a pair (i,j) of spin $\frac{1}{2}$ nuclei in a magnetic field is dependent on their effective distance r_{ij} and the angle θ of their internuclear vector (e.g. their bond) to the magnetic field (see figure 1.6 on the next page). Equation (1.2) describes this dependency, where $\gamma_{i,j}$ are the gyromagnetic ratios, μ_0 is the permittivity of space and h is Planck's constant. The averaging over the fast molecular motion in the laboratory frame is denoted by the brackets around the angular term, which is only zero for a completely isotropic situation.

$$D_{ij} = -\frac{\gamma_i\gamma_j\mu_0h}{8\pi^3r_{ij}^3} \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \quad (1.2)$$

The averaging can also be expressed by the introduction of a molecular frame in which the mean orientation of the whole molecule is given with respect to the magnetic field (angles ζ_X , ζ_Y and ζ_Z in figure 1.6). The orientation of the internuclear vector of interest is then given relative to the molecular frame (angles $\tilde{\zeta}_X$, $\tilde{\zeta}_Y$ and $\tilde{\zeta}_Z$). Therefore $\cos\theta$ in equation (1.2) has to be replaced by equation (1.3):

$$\cos\theta_{ij} = \begin{pmatrix} \cos\zeta_x \\ \cos\zeta_y \\ \cos\zeta_z \end{pmatrix} \begin{pmatrix} \cos\tilde{\zeta}_x \\ \cos\tilde{\zeta}_y \\ \cos\tilde{\zeta}_z \end{pmatrix} = \cos\zeta_x\cos\tilde{\zeta}_x + \cos\zeta_x\cos\tilde{\zeta}_z + \cos\zeta_z\cos\tilde{\zeta}_z \quad (1.3)$$

1. Introduction

This yields equation (1.4) for the left term of equation (1.2), assuming the molecule to be rigid:

$$\left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle = \frac{3}{2} \langle (\cos\zeta_x \cos\zeta_x + \cos\zeta_x \cos\zeta_z + \cos\zeta_z \cos\zeta_z)^2 \rangle - \frac{1}{2} \quad (1.4)$$

With $C_i = \cos\theta_i$ and $c_i = \cos\zeta_i$ this can be transformed to equation (1.5):

$$\begin{aligned} \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle = & \frac{3}{2} [\langle c_x^2 \rangle C_x^2 + \langle c_y^2 \rangle C_y^2 + \langle c_z^2 \rangle C_z^2 \\ & + 2\langle c_x c_y \rangle C_x C_y + 2\langle c_x c_z \rangle C_x C_z + 2\langle c_y c_z \rangle C_y C_z] - \frac{1}{2} \end{aligned} \quad (1.5)$$

Here, the averaging only occurs for the orientation relative to B_0 , since the orientation of the internuclear vector with respect to the molecular frame was assumed to be rigid. The preferential orientational averaging can then be described in terms of a symmetric 3x3 order matrix A (i.e. the *alignment tensor*) (Saupe, 1968), whose elements A_{kl} are:

$$A_{kl} = \frac{3}{2} \langle \cos\zeta_k \cos\zeta_l \rangle - \frac{1}{2} \delta_{kl} \quad (1.6)$$

δ_{kl} in equation (1.6) is the *Kronecker Delta* and the trace of the matrix is zero.

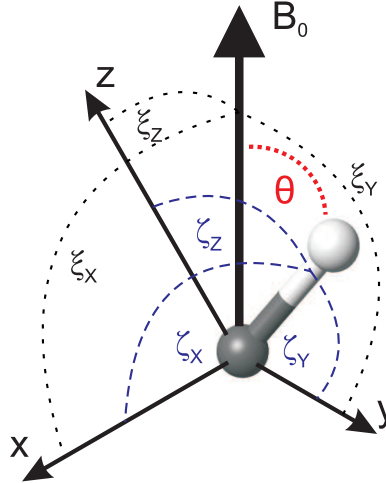


Figure 1.6.: Orientation of a bond vector relative to the B_0 field (angle θ) and the molecular frame (angles ζ_X , ζ_Y and ζ_Z). The orientation of B_0 relative to the molecular frame is given by the angles ζ_X , ζ_Y and ζ_Z . Modified after Blackledge (2005) and Mathieu (2007).

Individual RDCs can then be described by equation (1.7) on the facing page:

1.2. Background: Nuclear magnetic resonance spectroscopy

$$D_{ij} = -\frac{\gamma_i \gamma_j \mu_0 \hbar}{8\pi^3 r_{ij,eq}^3} \sum_{kl} A_{kl} \cos \zeta_k \cos \zeta_l \quad (1.7)$$

$r_{ij,eq}^3$ has been implemented in this equation to account for the averaging of this term and in order to gain a static equation. Diagonalization of the so-called *Saupe matrix* gives a description that is more convenient to handle and introduces a new molecular frame, the *principal axis system* (PAS, see figure 1.7) with the new diagonal elements A_{xx} , A_{yy} and A_{zz} . The off-diagonal elements in this matrix are zero and the orientation of the PAS with respect to the coordinate frame can be described by a three-dimensional *Euler rotation* $R(\alpha, \beta, \gamma)$.

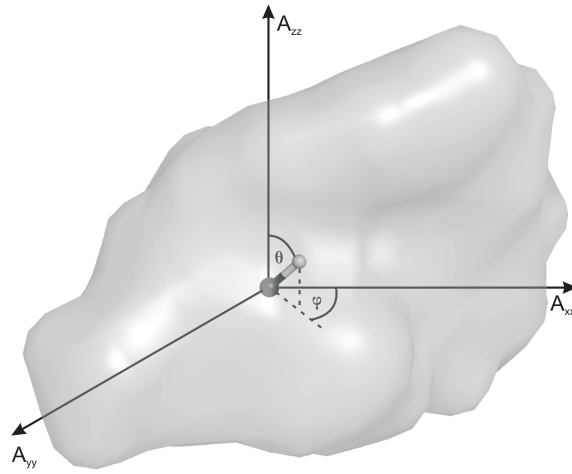


Figure 1.7.: Schematic representation of a bond vector in the principal axis system. The polar angles ϕ and θ are used to define its orientation in the principal axis system. Modified after Blackledge (2005) and Mathieu (2007).

Therefore, the measured RDC can be described using the polar angles θ and ϕ (figure 1.7) by equation (1.8):

$$D_{ij}(\theta, \phi) = -\frac{\gamma_i \gamma_j \mu_0 \hbar}{8\pi^3 r_{ij,eff}^3} [A_{zz} \cos^2 \theta + A_{xx} \sin^2 \theta \cos^2 \phi + A_{yy} \sin^2 \theta \sin^2 \phi] \quad (1.8)$$

The absolute value of A_{zz} by convention is the largest component, while A_{xx} is the smallest. In this case the axial component is $A_a = A_{zz}/2$ and the rhombic component is $A_r = (1/3)(A_{xx} - A_{yy})$:

1. Introduction

$$D_{ij}(\theta, \phi) = -\frac{\gamma_i \gamma_j \mu_0 h}{16\pi^3 r_{ij,eff}^3} \left[A_a (3\cos^2\theta - 1) + \frac{3}{2} A_r \sin^2\theta \cos 2\phi \right] \quad (1.9)$$

The maximum value for a residual dipolar coupling will be reached for a orientation of the internuclear vector along the z-axis of the alignment tensor and more generally the five parameters A_a , A_r , α , β and γ determine the alignment tensor. Therefore, five or more independent sets of RDCs are needed to determine these parameters or alternatively a singular value decomposition can be performed (Losonczi *et al.*, 1999).

However, a measured RDC still will not yield an unambiguous orientation of the internuclear vector (see figure 1.8). To solve this problem, one can either measure sets of RDCs for different internuclear vectors, whose relative orientations to each other are known or one can measure in different independent alignment media. A combination of both approaches is possible by measuring RDCs for different pairs of nuclei in only two independent alignment media (Al-Hashimi *et al.*, 2000).

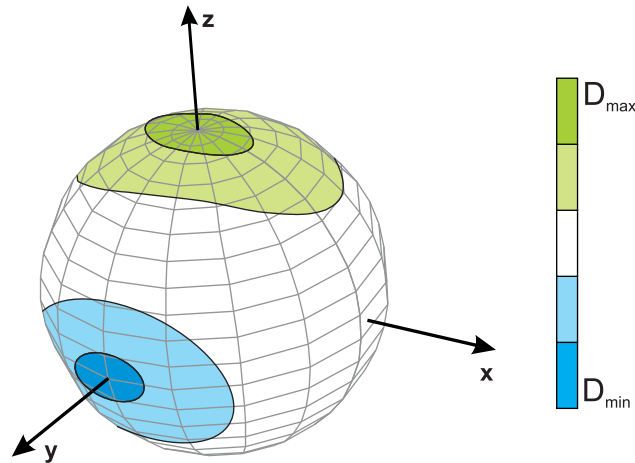


Figure 1.8.: Distribution of RDC values for the possible orientations. The RDC value is color coded, a maximal RDC is observed when the orientation is along the z-axis, the minimal value at an orientation along y. Modified after Blackledge (2005) and Mathieu (2007).

The situation in unfolded proteins is different: While for a folded protein the sign and value of a RDC is determined by the orientation of each bond vector with respect to the alignment tensor of the whole molecule (or molecular domain), unfolded proteins have to be treated as chains of statistical segments. Each segment has its own alignment tensor and in addition, the RDCs are dependent on the presence of tran-

1.2. Background: Nuclear magnetic resonance spectroscopy

sient secondary structure. In such polypeptide chains, conformational averaging and internal motions result in very small residual dipolar couplings. In ideal *random flight* chains smooth bell-like distributions have to be expected (Louhivuori *et al.*, 2003; Obolensky *et al.*, 2007) that can be explained by extended or polyproline II conformations (Mohana-Borges *et al.*, 2004).

Significant RDC have been measured in proteins under denaturing conditions, which have been attributed to the persistence of native-like structure in the cases of the $\Delta 131\Delta$ fragment of staphylococcal nuclease in 8 M urea (Gebel *et al.*, 2006; Shortle and Ackerman, 2001) and denatured eglin C (Ohnishi *et al.*, 2004) and to local conformational propensities in the cases of apomyoglobin (Mohana-Borges *et al.*, 2004) and the thermally unfolded B1 domain of protein G (Ding *et al.*, 2004). Long-range interactions have been shown to significantly modulate the RDCs in the natively unfolded α -synuclein (Bernadó *et al.*, 2005a) and acid-denatured acyl-CoA binding protein (ACBP) (Fieber *et al.*, 2004).

It has been shown that RDCs can be predicted in some cases for unfolded proteins from ensembles of unfolded structures generated using statistical ϕ, ψ propensities taken from loop regions of databases (Bernadó *et al.*, 2005b; Jha *et al.*, 2005). This method has been refined by including information on the bulkiness of amino acid side chains (Cho *et al.*, 2007) and long-range contacts (Bernadó *et al.*, 2005a). For the natively unfolded tau protein, the presence of β and β -turn structural motifs is reflected in the measured residual dipolar couplings (Mukrasch *et al.*, 2007a,b). Long-range H^N-H^N -RDCs have been used in urea-denatured ubiquitin to probe the significant population of native-like β -hairpin conformations under these conditions (Meier *et al.*, 2007).

Partial alignment of samples to gain an anisotropic distribution of orientations of the molecules of interest can be achieved by a variety of different methods. In early times, high magnetic fields have been used to align molecules exploiting the anisotropies in their magnetic susceptibilities (Tolman *et al.*, 1995). Similarly, paramagnetic lanthanide ions have been incorporated in some proteins in order to align them along the magnetic field (Bertini *et al.*, 2000; Ma and Opella, 2000). Lanthanide-binding tags (Martin *et al.*, 2007; Wöhnert *et al.*, 2003) can be attached for that purpose to proteins that do not bind lanthanides specifically. However, special *alignment media* have been developed in which biomacromolecules can be aligned. The first such medium contained so-called *bicelles*, i.e. a mixture of different phosphatidylcholine derivatives that form liquid crystalline particles, which align in the magnetic field and thereby lead to a net anisotropic orientation of the sample molecules by steric

1. Introduction

obstruction (Losonczi and Prestegard, 1998; Tjandra and Bax, 1997). In a similar way, solutions containing filamentous bacteriophages have been used (Clore *et al.*, 1998; Hansen *et al.*, 1998), though the alignment of the sample molecules in this case is based on electrostatic interactions with the highly charged coat protein of the phages.

A very different method to induce partial alignment in an NMR sample is the use of compressed or stretched polymeric gels. The alignment in such strained gels is independent of the direction of the magnetic field and the orientation of the sample molecules is based on the anisotropic environment. Highly cross-linked polyacrylamide gels (Sass *et al.*, 2000; Tycko *et al.*, 2000) are most commonly used. Compressed gels with a smaller diameter than the diameter of an NMR tube are prepared and then compressed until they fill the whole diameter of the tube (Tycko *et al.*, 2000). In contrast, stretched polyacrylamide gels are casted with a larger diameter and then pressed into the NMR tube by suitable devices. Thereby, strands of the polymer with preferred direction are produced. Strained polyacrylamide gels are particularly well suited for the use with unfolded proteins that could otherwise bind to hydrophobic surfaces of most of the other available media. In addition, polyacrylamide gels can be used at low pH, a condition that is often a requirement for the investigation of unfolded proteins.

Variations and combinations of the described methods and other techniques for the alignment of samples are discussed in more detail in Prestegard *et al.* (2004).

Protein dynamics by heteronuclear relaxation

Many biological processes, e.g. ligand binding, protein folding or enzymatic catalysis, require conformational changes in proteins. Therefore, proteins must not be completely rigid but contain more or less flexible regions. Figure 1.9 illustrates the timescales of different motions in proteins and of NMR parameters suitable for their investigation.

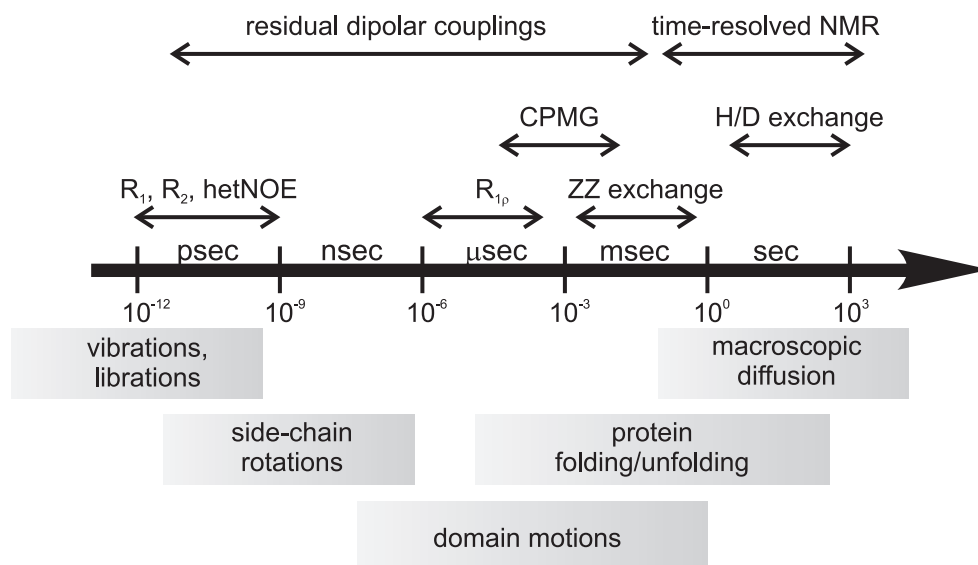


Figure 1.9.: Time-scales of molecular motions and NMR-techniques. Modified after Boehr *et al.* (2006).

Conformational dynamics on timescales faster than the overall rotational correlation time τ_c of the molecule can be accessed by spin relaxation methods. For a spherical molecule τ_c is proportional to the volume of the molecule and the viscosity of the medium. The amide ^{15}N spin in proteins is the most frequently used nucleus to probe backbone dynamics. The heteronuclear ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and the heteronuclear ^1H - ^{15}N -NOE (hetNOE) are sensitive to motions on the subnanosecond timescale and to slow conformational exchange in the microsecond timescale. These NMR parameters are not only influenced by local fluctuations (e.g. of the NH bond vector) but are also dependent on the overall rotational correlation time τ_c of the molecule. This dependence is shown for an isotropically tumbling rigid molecule in figure 1.10 on the following page.

1. Introduction

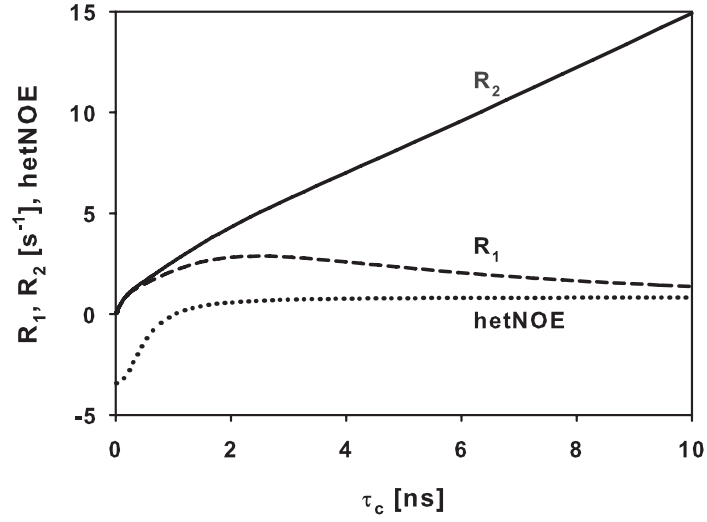


Figure 1.10.: Dependence of R_1 , R_2 and the heteronuclear NOE on the overall rotational correlation time τ_c of the molecule in the absence of internal motions and isotropic tumbling. Modified after Wirmer *et al.* (2005).

The R_1 and R_2 heteronuclear relaxation rates and the hetNOE of an amide ^{15}N nucleus with its bound proton are dominated by the dipolar interaction and the *chemical shift anisotropy* (CSA) (Wagner, 1993). The exact formulae in terms of spectral density functions are given in equations (1.10) to (1.12) (Farrow *et al.*, 1994):

$$R_1 = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2J(\omega_N) \quad (1.10)$$

$$R_2 = \frac{d^2}{2}[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{6}[4J(0) + 3J(\omega_N)] \quad (1.11)$$

$$\text{hetNOE} = 1 + \frac{d^2}{R_1} \frac{\gamma_H}{\gamma_N} [6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)] \quad (1.12)$$

The constants c^2 (for the CSA contribution) and d^2 (for the dipolar interaction) are defined in equations (1.13) and (1.14), respectively:

$$c^2 = \frac{2}{15} \gamma_N^2 H_0^2 (\sigma_{\parallel} - \sigma_{\perp})^2 \quad (1.13)$$

1.2. Background: Nuclear magnetic resonance spectroscopy

$$d^2 = 0.1 \frac{\gamma_H^2 \gamma_N^2 h^2}{4\pi^2} \left\langle \frac{1}{r_{NH}^3} \right\rangle^2 \quad (1.14)$$

γ_H and γ_N are the gyromagnetic ratios of ^1H and ^{15}N , ω_H and ω_N are the Larmor frequencies, r_{NH} is the ^1H - ^{15}N internuclear distance, h is Planck's constant, H_0 is the magnetic field strength, and $\sigma_{\parallel} - \sigma_{\perp}$ gives the parallel and perpendicular components of the chemical shift tensor. To account for conformational exchange contributions (R_{ex}) in addition to the dipolar interactions ($T_{2,DD}$) and the CSA contributions ($T_{2,CSA}$), equation (1.11) for the transverse relaxation (R_2) has to be expanded:

$$R_2 = \frac{1}{T_{2,DD}} + \frac{1}{T_{2,CSA}} + R_{ex} \quad (1.15)$$

Usually, for a protein the complete set of R_1 and R_2 relaxation rates and hetNOE data are recorded. The analysis of these relaxation rates for folded proteins in terms of backbone dynamics is generally done using the *model-free* formalism (Lipari and Szabo, 1982a,b). This approach assumes that the overall tumbling of the protein (τ_c) and the internal motions of the backbone amides (τ_e) occur on different timescales. Therefore, at a known τ_c value, internal motions can be analyzed independently. The order parameter S^2 is a measure for the spatial restriction of the internal motion, it can assume values from 0 to 1, with 1 being the completely rigid case. Figure 1.11 on the next page illustrates the dependence of ^{15}N heteronuclear R_2 relaxation rates on the internal motion (τ_e , S^2) and the global correlation time τ_c . The higher the internal motions the smaller are the measured relaxation rates. However, for unfolded proteins the assumption of the separability of global and internal motions is no longer valid, and thus the Lipari-Szabo model-free approach cannot be applied.

An alternative to the measurement of transverse relaxation rates (R_2) is the determination of the rotating frame spin-lattice relaxation rates $R_{1\rho}$ (Dayie and Wagner, 1994; Deverell *et al.*, 1970; Szyperski *et al.*, 1993). $R_{1\rho}$ is in principal the same as R_2 , but deviations can be observed in the presence of conformational exchange, since the exchange contributions are included in the R_2 rates but not in $R_{1\rho}$. In general, slow (micro- to millisecond) dynamics can be investigated not only by $R_{1\rho}$ experiments, but also by CPMG relaxation dispersion methods (Palmer *et al.*, 2001). Both techniques rely on the dispersion in the transverse relaxation rates as a function of the applied radio frequency field strength and exchange rates can be extracted from fits of model functions to the relaxation dispersion profile (Palmer, 2004). These methods are not restricted to ^{15}N amide relaxation, but have also been successfully applied to $^{13}\text{C}_{\alpha}$ nuclei (Lundström and Akke, 2005).

1. Introduction

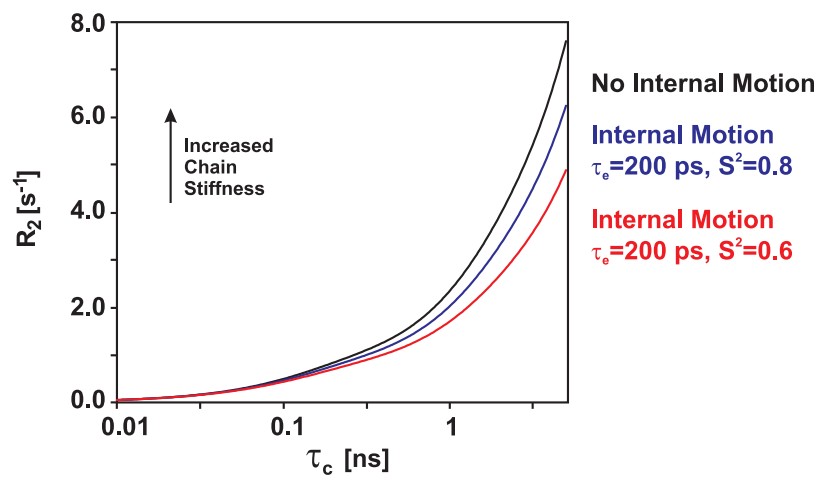


Figure 1.11.: Dependence of ^{15}N R_2 heteronuclear relaxation rates on the correlation times for global (τ_c) and internal (τ_e) motions and the order parameter (S^2) from a model-free approach.

1.3. Hen egg white lysozyme

Hen egg white lysozyme (HEWL) is one of the best studied proteins, its structure and folding has been investigated for many decades. Lysozymes (systematically classified as EC.3.2.1.17) are ubiquitous in many species and play an important role in the defense against pathogenic bacteria. They hydrolyze β -1,4-glycosidic bonds between N-acetyl muramic acid and N-acetylglucosamine which occur in peptidoglycans found in the cell walls of certain microorganisms, especially Gram-positive bacteria (Stryer, 1998). The term *lysozyme* for this class of enzymes has been formed by Alexander Fleming in 1922, when discovering these proteins and their functions (Fleming, 1922). Human lysozyme occurs mainly in mucosal secretions and despite its beneficial role, several single-point mutations have been found to be associated with hereditary forms of systemic amyloidosis (Merlini and Bellotti, 2005; Pepys *et al.*, 1993).

The lysozyme from hen egg white is a monomeric 129-residue protein with a molecular mass of approximately 14.5 kDa (Canfield, 1963) and its native structure has been first determined by X-ray crystallography in 1965 (Blake *et al.*, 1965). It consists of two domains: the α -domain (residues 1-35 and 85-129) comprises four α -helices and a 3_{10} -helix, the β -domain (residues 36-84) contains a triple-stranded anti-parallel β -sheet, a 3_{10} -helix and an extended loop region. In the native state, its eight cysteine residues form four disulfide bridges. Figure 1.12 on the following page shows one member of the high resolution NMR solution structure ensemble (Schwalbe *et al.*, 2001), the two domains are colored differently.

The kinetics of the refolding of HEWL from denaturing conditions with its four disulfide bridges intact has been investigated by a variety of techniques, including hydrogen-exchange labeling NMR and MS and stopped-flow CD, fluorescence and absorbance spectroscopy. Under the *in vitro* refolding conditions, a hydrophobic collapse that is accompanied by the formation of native-like secondary structure is observed within a few milliseconds (Matagne and Dobson, 1998). After the initial hydrophobic collapse, the folding kinetics are heterogeneous, intermediate states are populated and multiple slow and fast folding pathways exist. Cis/trans-isomerization of one or both prolines in HEWL can slow down the folding for a significant portion of the population. During refolding of HEWL, intermediates are observed for parts of the population that show a structured α -domain and the folding of the β -domain seems to be the rate limiting step (Matagne *et al.*, 1998).

Oxidative refolding from the disulfide reduced form of the protein is more com-

1. Introduction

plicated as the correct disulfide bridges have to be formed by the end and thus the refolding of reduced HEWL takes much longer than for the oxidized form. Intermediate states with one, two and three disulfide bridges are being formed and significantly populated (Fischer, 1996).

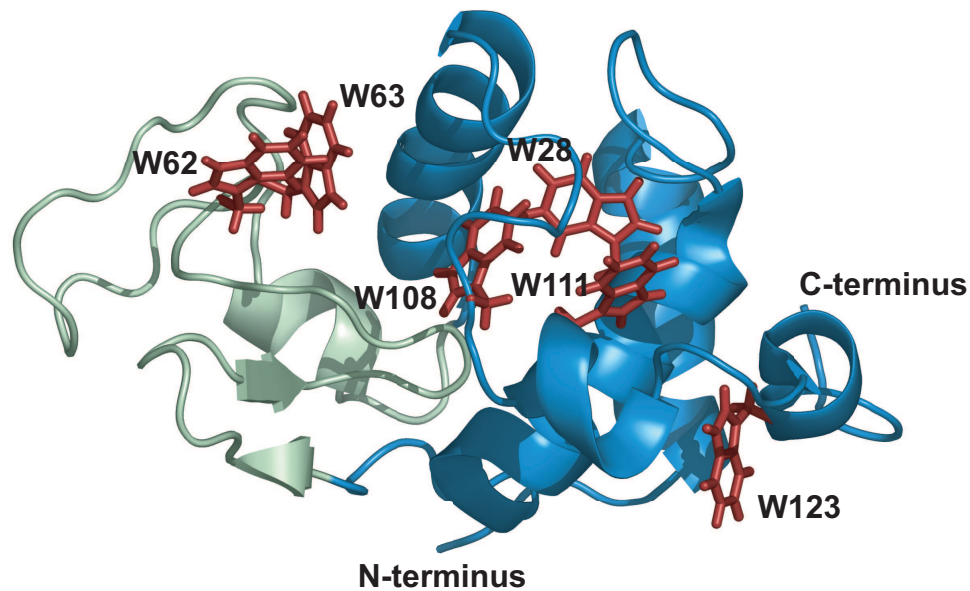


Figure 1.12.: One member of the high resolution NMR solution structure ensemble of native hen egg white lysozyme (PDB: 1E8L) (Schwalbe *et al.*, 2001). The tryptophan residues are highlighted in red, the α domain is colored in blue, the β domain is in light green. This image has been produced using PyMOL (DeLano Scientific, San Carlos, CA, USA).

HEWL was one of the first proteins to be investigated with NMR spectroscopy (Cohen and Jardetzky, 1968; Meadows *et al.*, 1967; Sternlicht and Wilson, 1967). However, it was not until more than 20 years later that an assignment of the backbone and first structural data was available (Redfield and Dobson, 1988; Smith *et al.*, 1991). The first solution NMR structure of hen egg white lysozyme was published in 1993 (Smith *et al.*, 1993). The assignment of backbone amide and side chain ^{15}N resonances in native HEWL allowed for the investigation of the dynamics by heteronuclear relaxation measurements. The relaxation-derived order parameters, which are a measure of the flexibility of the backbone and the side chains, have been found to agree very well with the structure of the protein (Buck *et al.*, 1995a).

Not only native HEWL and the process of its refolding have been studied, but also the nature of non-native states under various denaturing conditions has been a subject to research for many years. The structural and dynamical properties of oxidized

(i.e. disulfide intact) and reduced, S-methylated HEWL in 8 M urea have been studied by NMR already in 1997 using short- and medium-range NOEs and $^3J(\text{H}^{\text{N}}, \text{H}^{\alpha})$ coupling constants to probe residual local structure (Schwalbe *et al.*, 1997). In the same work longitudinal (R_1), transverse (R_2) and rotating-frame ($R_{1\rho}$) relaxation rates and the heteronuclear ^1H - ^{15}N -NOE have been used to investigate the dynamics of the backbone under these conditions. The experimental data for both states have been compared to predicted values for a random coil (Fiebig *et al.*, 1996) and found to agree well for most of the chain, although significant deviations especially in the relaxation rates have been observed for parts of the sequence. Differences between the unbranched reduced lysozyme chain and the cross-linked, disulfide-intact peptide chain have been observed as expected in the flexibility of the regions around the disulfide bridges (Collins *et al.*, 2005; Schwalbe *et al.*, 1997). The deviations from the random coil behavior have primarily been attributed to long-range interactions and hydrophobic clustering (Klein-Seetharaman *et al.*, 2002). Furthermore, these effects have also been observed in the reduced and S-methylated HEWL (HEWL- S^{Me}) in water at pH 2. Hydrophobic clusters and long-range interactions in non-native states have also been reported for a variety of other proteins (Neri *et al.*, 1992; Ropson and Frieden, 1992; Saab-Rincón *et al.*, 1996), suggesting the role of these hydrophobic clusters as nucleation sites in protein folding.

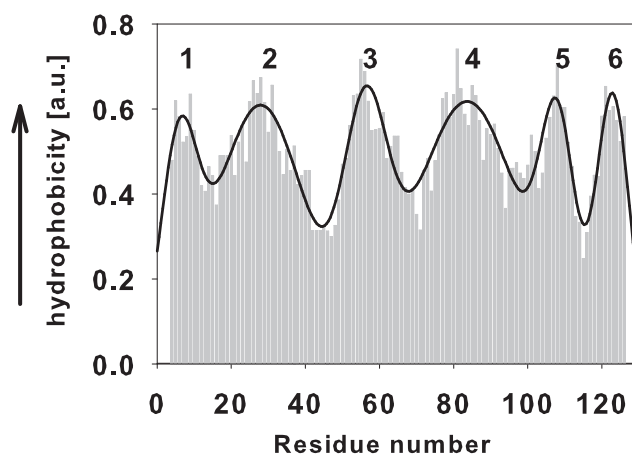


Figure 1.13.: Distribution of hydrophobicity in HEWL. Normalized hydrophobicity values are calculated according to Abraham and Leo (1987). A Gaussian least-squares fitting of hydrophobic clusters is shown as a black curve, the clusters are designated 1-6 (data taken from Wirmer *et al.*, 2004).

1. Introduction

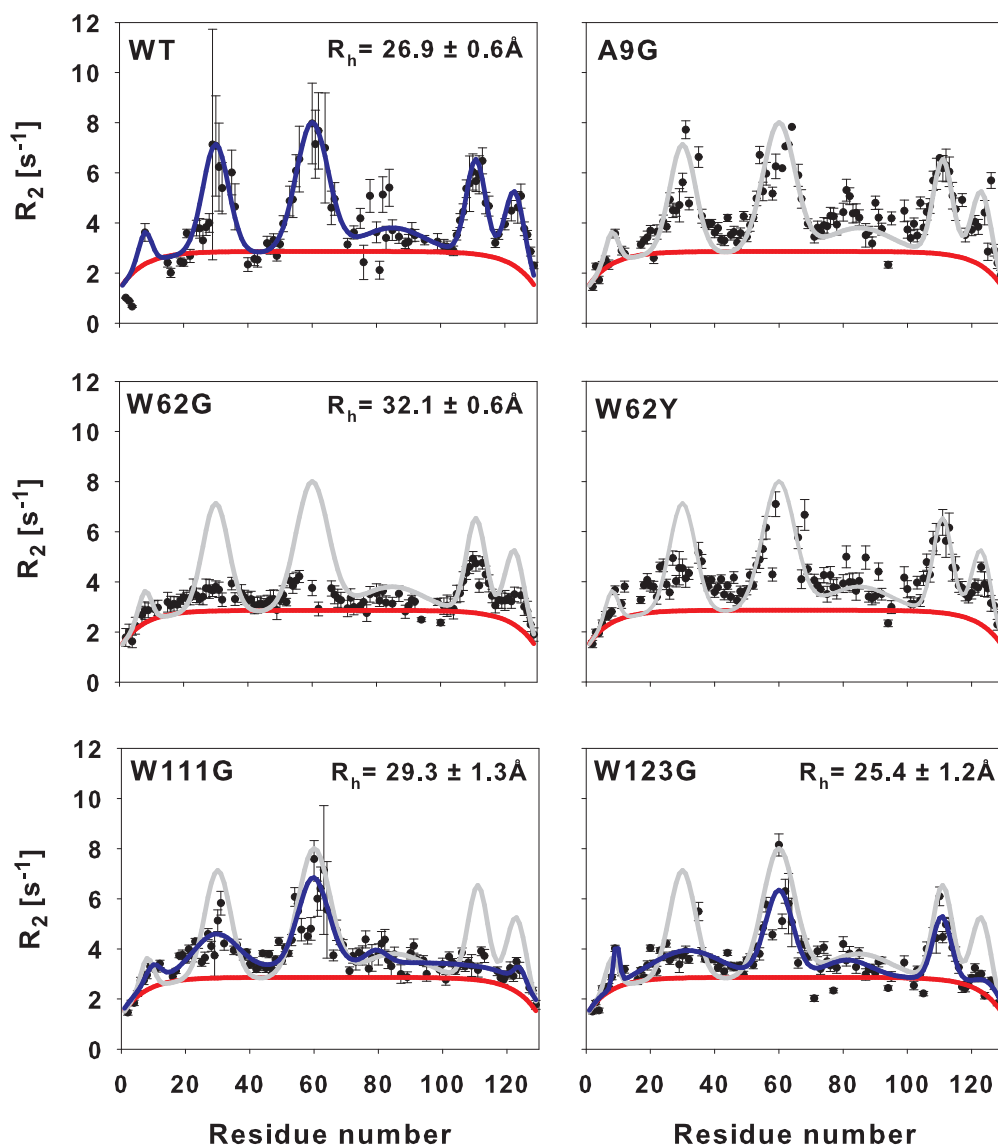


Figure 1.14.: ^{15}N R_2 relaxation rates in HEWL- S^{Me} and the A9G, W62G, W62Y, W111G, and W123G single point mutants of HEWL- S^{Me} . The baselines from the segmental motion model (see equation (3.1) on page 86) are given in red, the Gaussian fits are indicated by blue curves. The grey curves in the plots for the single point mutants indicate the relaxation rates for the wild-type. Diffusion derived radii of hydration (R_h) are given for HEWL- S^{Me} and some of the mutants (all data taken from Wirmer *et al.*, 2004).

1.3. Hen egg white lysozyme

Six hydrophobic clusters have been identified in HEWL-S^{Me} from elevated R₂ relaxation rates (compare upper left section of figure 1.14 on the facing page): around residues A9-A11 (cluster 1), around W28 (cluster 2), around W62/W63 (cluster 3), region around A82 (cluster 4), around W108/W111 (cluster 5) and around W123 (cluster 6) (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004). Interestingly, the elevated relaxation rates coincide with the calculated hydrophobicity (Abraham and Leo, 1987) along the amino acid sequence of HEWL very nicely (see figure 1.13 on page 29). These findings underlined the importance of hydrophobic residues and tryptophans in particular and led to the introduction of a single point mutation at position 62 to replace this tryptophan with a glycine. This mutation resulted in the nearly complete disappearance of the elevated relaxation rates in all of the clusters, suggesting long-range interactions of W62 with the other hydrophobic clusters and the absence of such clusters in the W62G mutant of HEWL-S^{Me}. The situation is somewhat different in the oxidized form of HEWL and W62G-HEWL in 8 M urea, where the intact disulfide bridges lead to additional restraints in flexibility and hence to higher relaxation rates in these regions. The role of W62 in the mediation of the hydrophobic clustering and the long-range contacts in the denatured states is remarkable, given its high mobility and its position at the surface in the native protein (Blake *et al.*, 1965; Buck *et al.*, 1995b; Schwalbe *et al.*, 2001). Taken all data together, native-like as well as non native-like hydrophobic contacts seem to exist in non-native states of HEWL.

A number of conservative and non-conservative single-point mutants of HEWL-S^{Me} has been prepared in order to further investigate the importance of hydrophobic residues for the residual tertiary structure in non-native states of this protein (Wirmer *et al.*, 2004). The replacement of W62 by a tyrosine leads to a very similar relaxation profile as for HEWL-S^{Me}, while A9G only attenuates the first cluster and W111G and W123G significantly modulate all clusters. W111G weakens cluster 5, in which it is located and clusters 2 and 6, while for the central cluster 3 only minor effects are observed. In W123G cluster 6 disappears and cluster 2, 3 and 5 are weakened (figure 1.14 on the facing page).

The modulation of the hydrophobic clustering in non-native lysozyme by the single-point mutants is also reflected in the compactness of these mutants. This has been shown by the determination of the radii of hydration (R_h) by diffusion NMR (Wirmer *et al.*, 2004): While native HEWL has an R_h of 20.5 Å, HEWL-S^{Me} in water has an R_h of 26.9 Å. Even more extended are the W62G and W111G mutants with radii of hydration of 32.1 Å and 29.3 Å. The W123G mutant of HEWL-S^{Me} has a

1. Introduction

comparable compactness as the wild-type HEWL-S^{Me}. The more the hydrophobic clusters are weakened, the more extended are the non-native states of HEWL. Figure 1.15 illustrates the effect of compactness on the basis of randomly chosen conformations with the R_h values corresponding to the ones for the different mutants of HEWL-S^{Me}.

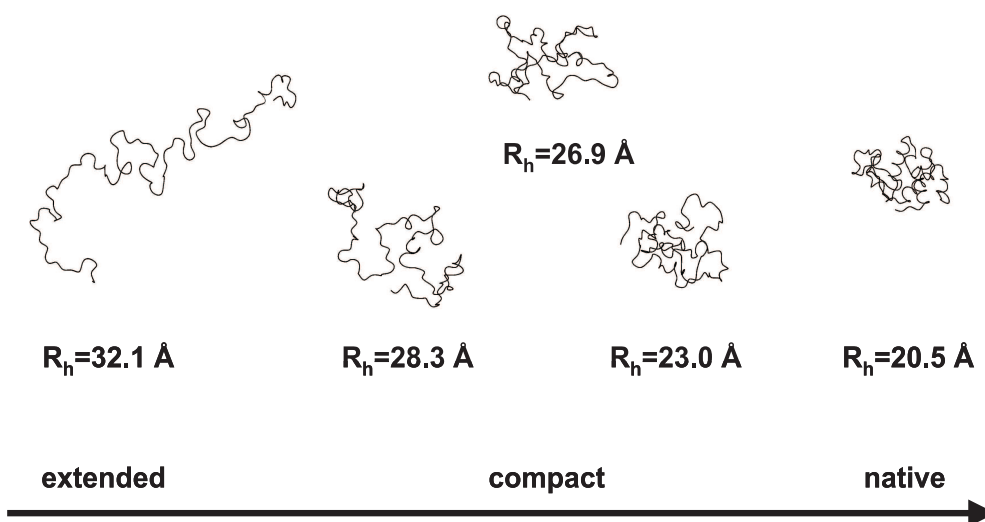


Figure 1.15.: Conformations with the amino acid sequence of HEWL randomly chosen to represent distinct radii of hydration (R_h) (modified after Wirmer *et al.*, 2004).

1.4. Motivation and aims

As outlined in the preceding sections, the investigation of non-native states of proteins can contribute to the understanding of protein folding and misfolding, especially to the early events on the way of structure formation. Hen egg white lysozyme has been a model protein in this context for many years. Its non-native states exhibit significant residual secondary and tertiary structure, and thus cannot be interpreted as random coils. The exact nature of the non-native states of hen egg white lysozyme can be modulated by single-point mutants.

Hitherto, the existing data on the nature of non-native states of HEWL was derived from either the overall properties of the non-native ensemble of conformers or residue-specific for the peptide backbone. Since the side chain type determines the hydrophobic clustering and especially the tryptophan residues play the key role in the long-range interactions, a method is desired to look at these residues at the side chain level. In the context of NMR, the prerequisite in achieving this goal is the assignment of resonances in these side chains. This can be addressed by the development and application of novel pulse sequences.

So far, mainly the dynamical properties of the non-native states of HEWL have been investigated at atomic resolution, while the non-local structural information is rare, especially regarding the environment of the tryptophan residues. Consequently, the second major aim of this thesis is the determination of structural NMR parameters, such as residual dipolar couplings (RDCs) that can add to the understanding of the situation in non-native lysozyme and mutants thereof.

In addition, the dynamical properties of a lysozyme variant with all cysteines replaced by alanines are to be investigated at different time-scales by relaxation measurements to pick up fast and possible slower motions.

2. Materials and Methods

2.1. Sample preparation

2.1.1. Overview

The synthetic genes encoding for hen egg white lysozyme (HEWL) with an N-terminal methionine (designated *wild-type* HEWL for simplification throughout this text) (Schlörb, 2003) and its respective *all-Ala* mutant with all cysteine residues replaced by alanines (Ackermann, 2003), were obtained from Entelechon, Regensburg, Germany in vector pCR4-TOPO, and subsequently cloned into the pET11a expression vector. The N-terminal methionine, which is not present in the wild-type hen protein, has been introduced for expression in *E. coli*. This modification has been shown to have no effect on the structure and dynamics of the non-native states of this protein (Schlörb *et al.*, 2005). Traditionally, hen egg white lysozyme has been rendered non-native by maintaining a pH value of 2.0, by reduction of its four disulfide bridges and the successive S-methylation of the eight cysteine residues (Schlörb, 2003). In order to circumvent the drawbacks of this method, such as unwanted side-reactions, incomplete methylation and losses during purification, an alternative approach has been introduced, in which all cysteines in hen egg white have been replaced by alanines (designated *all-Ala*-HEWL) (Ackermann, 2003). This variant of the protein has very similar properties at pH 2.0 as compared to the S-methylated protein and therefore is a suitable replacement in the studies of the non-native states of hen egg white lysozyme.

All chemicals used in this work were at least of analytical grade and ordered at Sigma-Aldrich/Fluka (St. Louis, MO, USA) or Carl Roth GmbH (Karlsruhe, Germany) unless indicated otherwise.

2. Materials and Methods

2.1.2. Expression and purification

Transformation of BL21(DE3) cells

Aliquots of competent *E. coli* BL21(DE3) cells (Stratagene, La Jolla, CA, USA) have been transformed with the pET11a (Novagen, Merck, Darmstadt, Germany) based vectors containing the gene coding for hen egg white lysozyme and its mutants, respectively. The cells have been thawed on ice and 1.0 μ L plasmid DNA (approx. 200 ng/mL) has been added to 50 μ L of the cell suspension. After a 45 min incubation period on ice, the reaction mix has been heat-shocked for 45 sec at 42°C and cooled for 2 min on ice. Following the addition of 250 μ L SOC medium (Invitrogen, Karlsruhe, Germany), the cells were allowed to grow for 40 min at 37°C and 170 rpm in a shaker. The cell suspension was then plated on ampicillin-containing (100 μ g/mL) LB agar plates and incubated over night at 37°C. The pET11a based vectors contain a gene for ampicillin resistance and a IPTG inducible T7 promoter for protein expression.

Expression of isotope-enriched proteins

Hen egg white lysozyme and its variants have been expressed heterologously in *E. coli* to allow for the uniform isotope enrichment of either nitrogen nuclei alone or nitrogen and carbon nuclei simultaneously with NMR active isotopes (^{15}N and ^{13}C). The ^{15}N , ^{13}C labeling is necessary for the sequential assignment of resonances using NMR experiments depending on the correlation of ^{15}N and ^{13}C nuclei, while ^{15}N -only labeled proteins have been used for all other NMR experiments in this work. Uniform ^{15}N and ^{13}C isotope labeling was assured by providing ^{15}N ammonium chloride and $^{13}\text{C}_6$ D-glucose (both obtained from Cambridge Isotope Laboratories, Andover, MA, USA) in the growth medium as the sole source for nitrogen and carbon, respectively. M9 medium (Sambrook and Russel, 2001) was used as a minimal medium for isotope enrichment during protein expression.

A 4 mL LB-culture (containing 100 μ g/mL ampicillin) was allowed to grow for 3-4 hours at 37°C and 170 rpm in a shaker after inoculation with a single colony from a plate. This pre-pre-culture was then transferred into a 50 mL M9-pre-culture (see table A.1 on page 107 for the composition of M9 minimal medium). The pre-culture was incubated over night at 37°C and 170 rpm in a shaker to a OD_{600} of about 2.0. Cells were harvested by centrifugation at 5,500 g for 10 min (Heraeus Biofuge primo R) and resuspended in M9 medium. The centrifugation/resuspension step was repeated twice for washing. After washing, a 2 L M9-main-culture was inoculated

2.1. Sample preparation

from the pre-culture to a initial OD₆₀₀ of about 0.1. Cells were allowed to grow to a OD₆₀₀ of 0.8 at 37°C and 170 rpm in a shaker and then protein expression was induced by the addition of 1 mg/L IPTG. After 3 hours of protein expression, the cells were harvested by centrifugation for 30 min at 6,000 g in a centrifuge (Beckman Coulter Avanti J-20XP). The cell pellet was stored at -80°C until further processing.

Solubilization of inclusion bodies

Recombinant hen egg white lysozyme and all of its mutants and variants have been expressed into inclusion bodies in *E. coli*. The inclusion bodies are formed since the protein cannot be folded correctly in the *E. coli* cells. In the case of *wild-type* HEWL, the redox conditions inside the cells do not allow the formation of the four disulfide bridges (Makrides, 1996); for *all-Ala*-HEWL the formation of the native conformation is completely impossible as there are no cysteine residues present.

For the solubilization of the inclusion bodies, the cells were resuspended after thawing in 40 mL (per initial L M9 medium) *sonication buffer* (containing 50 mM Tris, 25 % sucrose, and 1 mM EDTA, pH 7.5; all buffers for the solubilization of the inclusion bodies are summarized in table A.2 on page 108), 10 units DNase I (1 µL; New England Biolabs, Ipswich, MA, USA) and 1.5 mL 100 mM MgCl₂ were added and the suspension was sonicated four times for 1 min at 40 % power to disrupt the cells. The cell extract was washed once by centrifugation (10,000 g, 30 min), resuspended in *sonication buffer* and spun down again at the same conditions. The pellet obtained was resuspended in *washing buffer* (containing 20 mM Tris, 1 % Triton X-100, and 1 mM EDTA, pH 7.5) and centrifuged for 30 min at 10,000 g. This additional washing step removes the cell membrane fragments. To finally solubilize the inclusion bodies, the pellet was resuspended in *denaturation buffer* (containing 20 mM Tris, 50 mM NaCl, 8 M urea, 5 mM EDTA, and 0.1 M DTT, pH 7.5) and spun down for 10 min at 10,000 g. The supernatant contained the solubilized inclusion bodies.

Protein purification

The solubilized inclusion bodies were further purified by ion exchange chromatography, the buffers used for the purification are listed in table A.3 on page 108. The isoelectric point (pI) calculated from the amino acid sequence (Gasteiger *et al.*, 2005) is between 9.3 and 10.5 for all variants of HEWL used in this work, therefore a cation exchange column (CM sepharose, 8 ml column volume, column dimensions: 1.0 cm x 10 cm x 8.0 cm) was chosen for purification at a buffer pH of 7.5. The column was

2. Materials and Methods

equilibrated with 10 column volumes of buffer A (50 mM Tris, 50 mM NaCl, 4 M urea, 1 mM EDTA, and 5 mM 2-mercaptoethanol) before the solubilized inclusion bodies were loaded. The protein was eluted from the column at a flow of 0.7 mL/min using a linear salt gradient from buffer A to buffer B (50 mM Tris, 300 mM NaCl, 4 M urea, 1 mM EDTA, and 5 mM 2-mercaptoethanol). A Bio-Rad *BioLogic LP* (Bio-Rad, Hercules, CA, USA) FPLC system equipped with a UV light detector (280 nm) and a fraction collector (fraction size 3.5 mL) was used for the chromatography. Suitable fractions containing the desired protein (judged from SDS-PAGE, 15 % gels) were pooled. For *all-Ala*-HEWL and its single point mutants the pooled fractions were dialyzed against 7 L deionized water pH 2 for 4-5 times in a dialyzing tube (3,500 Da MWCO, Spectrum Laboratories, Rancho Dominguez, CA, USA). The solution was then freeze-dried, resuspended in approx. 5 mL water pH 2 and further purified using reversed-phase HPLC with a linear water-to-acetonitrile (containing 0.1 % TFA) gradient on an RP4-column.

For the case of *wild-type* HEWL see section 2.1.3.

2.1.3. S-Methylation of cysteine residues

The pooled fractions of *wild-type* HEWL were concentrated to a final concentration of about 1 mg/mL using centrifugal filter devices (Vivaspin 15, 5,000 Da MWCO, Sartorius, Göttingen, Germany). The cysteine groups of the protein were methylated as described before (Heinrikson, 1971; Wirmer, 2005) using methyl-p-nitrobenzenesulfonate in buffer containing 8 M urea and 25 % (v/v) acetonitrile. Reduced S-methylated HEWL (designated HEWL-S^{Me}) was dialyzed against deionized water pH 2, purified by reversed-phase HPLC using a linear water-to-acetonitrile gradient, and freeze-dried as described for *all-Ala*-HEWL.

2.1.4. Site-directed mutagenesis

The genes coding for W62G-*all-Ala*-HEWL and W108G-*all-Ala*-HEWL have been generated by site-directed single point mutation of the *all-Ala*-HEWL gene in the pET11a vector using the *QuikChange*® *Site-Directed Mutagenesis Kit* (Stratagene, La Jolla, CA, USA) and appropriate forward and backward primers (see table A.4 on page 108).

2.1.5. Biochemical analysis

A 15 % SDS polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970) has been conducted for various samples during protein expression, inclusion body solubilization and protein purification to monitor these processes and check for protein purity. The gels have been stained with *Coomassie Brilliant Blue*.

To check the success of the PCR for the site-directed mutagenesis, 0.8 % agarose gels have been casted, run for the mutated vectors and stained with ethidium bromide (Sambrook and Russel, 2001).

2.1.6. NMR samples

The NMR samples of HEWL-S^{Me}, *all-Ala*-HEWL and the mutants thereof have been prepared by dissolving the lyophilized protein in deionized water pH 2 containing 10 % D₂O. Insoluble fractions of the proteins have been pelleted down by centrifugation and thereby removed from the samples. The concentrations of the samples were determined by UV light absorption at 280 nm using the extinction coefficients calculated from the amino acid sequence ($\epsilon = 37,980 \text{ M}^{-1} \text{ cm}^{-1}$ and $\epsilon = 37,470 \text{ M}^{-1} \text{ cm}^{-1}$ for HEWL-S^{Me} and *all-Ala*-HEWL, respectively) (Gasteiger *et al.*, 2005). Protein concentrations in the NMR samples have been adjusted to approx. 0.5 - 0.7 mM unless otherwise stated. Norell 507-HP standard NMR tubes (Norell Inc., Landisville, NJ, USA) and *Shigemi*-type NMR tubes (Shigemi Inc., Allison Park, PA, USA) have been used unless otherwise stated.

2.2. Backbone resonance assignment

2.2.1. Overview

The assignment of the backbone resonances is the prerequisite for the further characterization of structure, interactions and dynamics of any given protein at atomic resolution by NMR spectroscopy. The assignment strategy involves the labeling of the protein with the NMR-observable ^{15}N and ^{13}C isotopes. Therefore, the non-native lysozyme variants investigated in this thesis have been labeled uniformly with these isotopes as described in section 2.1.2 on page 36. The labeling of nitrogen and carbon nuclei together with the abundant ^1H nuclei allows for the intra- and inter-residual correlation of $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{H}}$, $^{13}\text{C}_{\alpha}$, $^1\text{H}_{\alpha}$ and $^{13}\text{C}'$ nuclei through scalar couplings along bonds in a series of three-dimensional NMR experiments (Sattler *et al.*, 1999), thus enabling the sequential mapping of resonances to the protein's amino acid sequence. In addition, some of these assignment experiments can also provide correlations to the side chain ^1H and ^{13}C resonances.

2.2.2. Acquisition of two- and three-dimensional spectra and sequential backbone assignment

A different set of three-dimensional NMR experiments has been conducted for HEWL- S^{Me} and *all-Ala*-HEWL, respectively, as different combinations of NMR experiments can in fact yield comparable results in the process of assigning the protein backbone. Figure 2.1, figure 2.2 and figure 2.3 on pages 43 to 45 summarize the three-dimensional experiments used for the assignment of HEWL- S^{Me} and *all-Ala*-HEWL and illustrate the magnetization transfers and assignment paths. However, the core experiment in both cases was the HNCACB experiment (Muhandiram and Kay, 1994; Wittekind and Mueller, 1993), which correlates the $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei with the $^{13}\text{C}_{\alpha,i}$, $^{13}\text{C}_{\beta,i}$, $^{13}\text{C}_{\alpha,i-1}$ and $^{13}\text{C}_{\beta,i-1}$ nuclei. In principle, HNCACB spectra can yield complete sequential $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{H}}$, $^{13}\text{C}_{\alpha}$ and $^{13}\text{C}_{\beta}$ assignments. Unfortunately, in non-native proteins the low dispersion of NMR signals leads to significant overlapping of resonances and assignment ambiguities. Therefore, additional experiments not only provide the resonance frequencies of other nuclei, but also give important information to unambiguously assign the peaks to specific amino acid residues. The CBCA(CO)NH experiment (Grzesiek and Bax, 1992a, 1993) correlates the $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei only with the $^{13}\text{C}_{\alpha,i-1}$ and $^{13}\text{C}_{\beta,i-1}$ nuclei and thereby helps to distinguish *i* and *i-1* signals from each other. A prerequisite of the sequential assignment using

2.2. Backbone resonance assignment

three-dimensional spectra is the picking of backbone amide $^1\text{H}^{\text{N}}_i$, $^{15}\text{N}^{\text{H}}_i$ correlation peaks in highly resolved two-dimensional HSQC (Bodenhausen and Ruben, 1980) spectra. All HSQC spectra recorded were of the fast-HSQC-type (FHSQC) to minimize losses due to chemical exchange of amide protons with the solvent water (Mori *et al.*, 1995).

The chemical shift patterns in the spectra in many cases can serve as good indicators for the identification of the amino acid type especially in non-native proteins, where these shifts closely resemble random coil chemical shifts available from the literature (e.g. Schwarzinger *et al.*, 2000). Alanine, glycine, threonine and serine residues are particularly easy to identify because of their unique peak patterns and chemical shifts and thus suitable starting points for the sequential assignment. In addition, previous assignments of similar states of HEWL under similar conditions have been available (Grimshaw, 1999; Schlörb, 2003), which also contributed to the initial assignments of HEWL-S^{Me}, despite their incomplete and partially inaccurate data (Grimshaw, 1999).

For HEWL-S^{Me}, in many cases the amino acid type of a specific spin system has been determined using the distinct peak patterns in (H)CC(CO)NH and H(CC)-(CO)NH (both Montelione *et al.*, 1992) spectra. These experiments correlate $^{13}\text{C}_{i-1}$ or $^1\text{H}_{i-1}$ nuclei of the side chains by TOCSY-type transfers (Bax and Davis, 1985; Shaka *et al.*, 1988) and via scalar couplings with the backbone $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei. The $^1\text{H}_{\alpha,i-1}$ and $^1\text{H}_{\beta,i-1}$ peaks have been picked from HBHA(CO)NH (Grzesiek and Bax, 1993) spectra, where these nuclei are correlated with the backbone $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei. The backbone carbonyl ^{13}C resonances have been picked in a HNCO (Grzesiek and Bax, 1992b; Kay *et al.*, 1994; Schleucher *et al.*, 1993) spectrum, which correlates the $^{13}\text{C}'_{i-1}$ with the $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei.

For *all-Ala*-HEWL, an approach with three lines of sequential correlations has been used in order to speed-up the assignment process: In addition to the HNCACB and CBCA(CO)NH experiments, a HN(CA)CO (Clubb *et al.*, 1992) spectrum has been recorded, which correlates the $^{13}\text{C}'_{i-1}$ and $^{13}\text{C}'_i$ with the $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei. The third correlation stems from a HNN (Bhavesh *et al.*, 2001; Panchal *et al.*, 2001) experiment, which correlates $^{15}\text{N}^{\text{H}}_{i-1}$ and $^{15}\text{N}^{\text{H}}_{i+1}$ with the $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei. Additionally, HBHA(CO)NH and (H)CC(CO)NH experiments have been conducted and $^{13}\text{C}_{\alpha,i-1}$ and $^{13}\text{C}_{\alpha,i}$ correlations with $^1\text{H}^{\text{N}}_i$ and $^{15}\text{N}^{\text{H}}_i$ nuclei have been extracted from a HNCA (Grzesiek and Bax, 1992b; Kay *et al.*, 1994) spectrum to further support the HNCACB data.

However, the assignment strategy applied in the case of *all-Ala*-HEWL is favor-

2. Materials and Methods

able over the set of experiments used for the assignment of HEWL-S^{Me} as it provides three independent lines of sequential connectivities with the same amount of measurement time (compare section 3.2 on page 66).

All assignment experiments conducted (except for the HNN experiment) are standard Bruker implementations of the published pulse sequences released with XWIN-NMR 3.5 using *watrgate* (Piotto *et al.*, 1992; Sklenar *et al.*, 1994) water suppression. All spectra have been processed using either XWIN-NMR 3.5 or TOPSPIN 1.3 software (Bruker Biospin, Karlsruhe, Germany), the acquisition parameters are listed in tables A.5 and A.6. The sequential assignments of HEWL-S^{Me} and *all-Ala*-HEWL have been carried out using the XEASY (version 1.5; Bartels *et al.*, 1995) and CARA (version 1.5.4; Keller, 2004) software, respectively. The ¹H chemical shifts have been referenced directly to the external standard 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), while ¹⁵N and ¹³C chemical shifts have been referenced indirectly to the resulting ¹H carrier frequency (Cavanagh *et al.*, 1996; Wishart *et al.*, 1995b).

2.2. Backbone resonance assignment

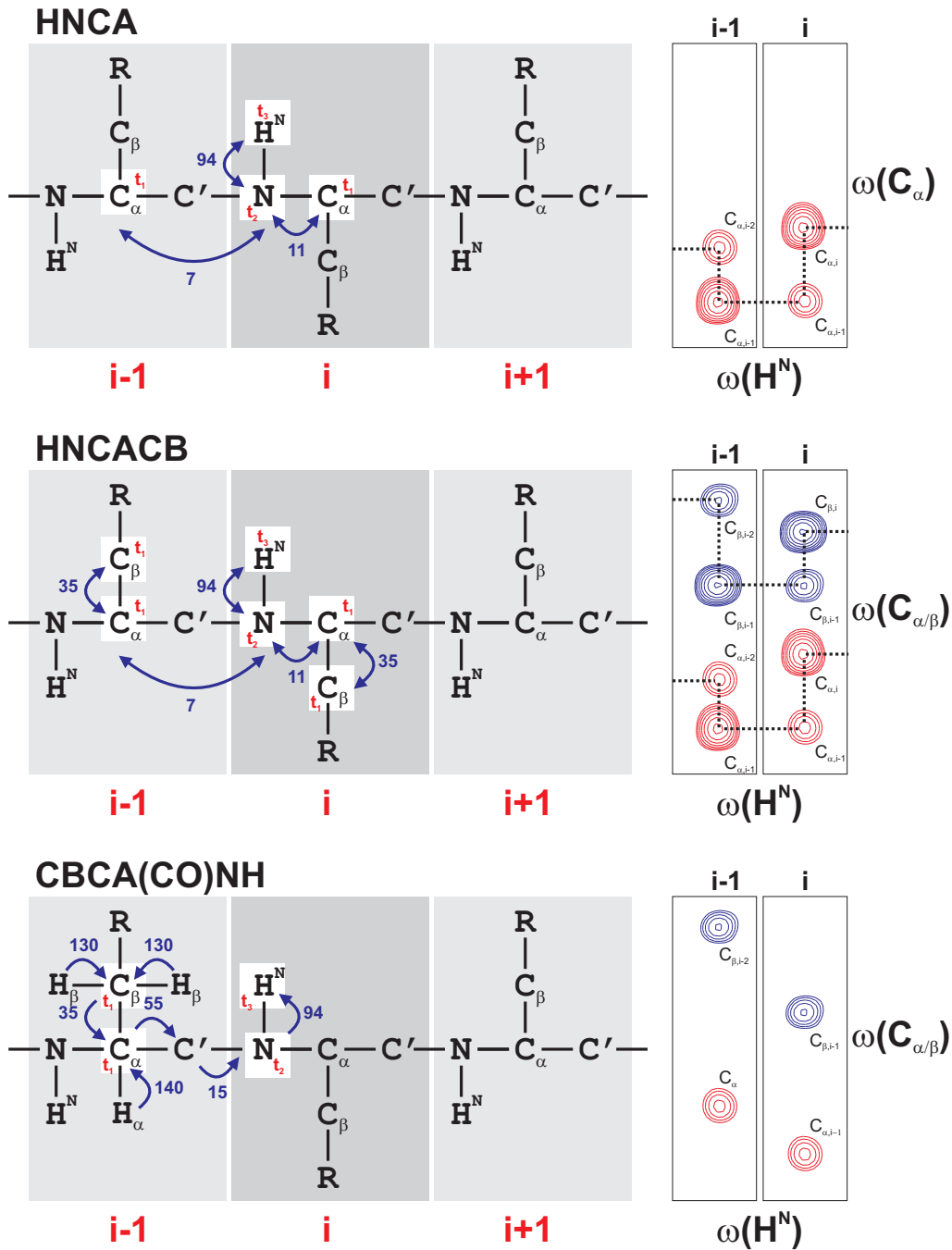


Figure 2.1.: Relevant backbone resonance assignment experiments for (non-native) proteins (I). Resonance transfer by scalar coupling is indicated by blue arrows. Coupling constants (in Hz) for the transfers are given in blue. Detected nuclei are indicated by white boxes and the symbols for the detection periods are given (t_1 , t_2 and t_3). The resonance patterns for relevant planes in the experiments are indicated on the left for the adjacent residues i and $i-1$ (modified after Duchardt, 2005).

2. Materials and Methods

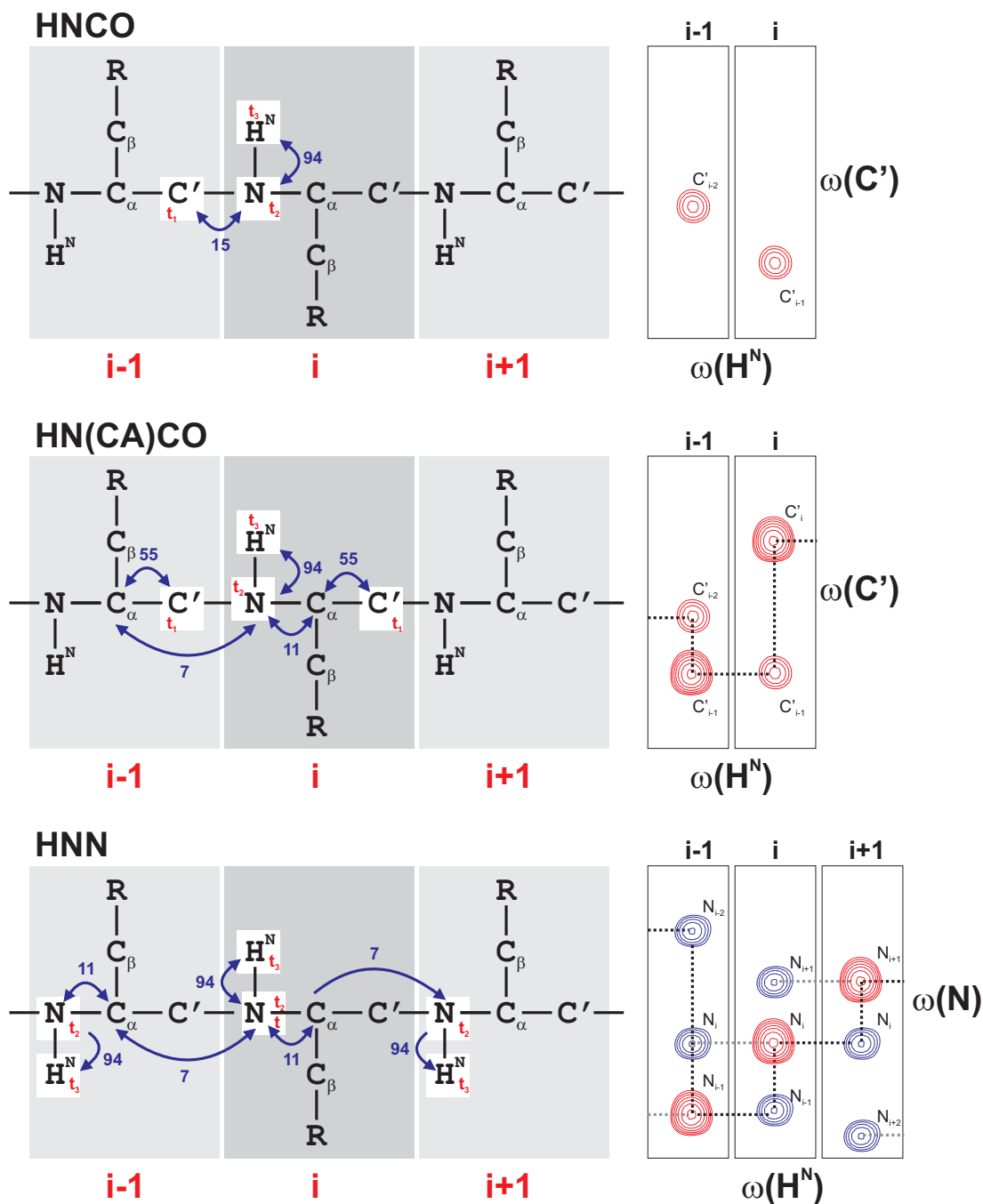


Figure 2.2.: Relevant backbone resonance assignment experiments for (non-native) proteins (II). Resonance transfer by scalar coupling is indicated by blue arrows. Coupling constants (in Hz) for the transfers are given in blue. Detected nuclei are indicated by white boxes and the symbols for the detection periods are given (t_1 , t_2 and t_3). The resonance patterns for relevant planes in the experiments are indicated on the left for the adjacent residues i , $i+1$ and $i-1$.

2.2. Backbone resonance assignment

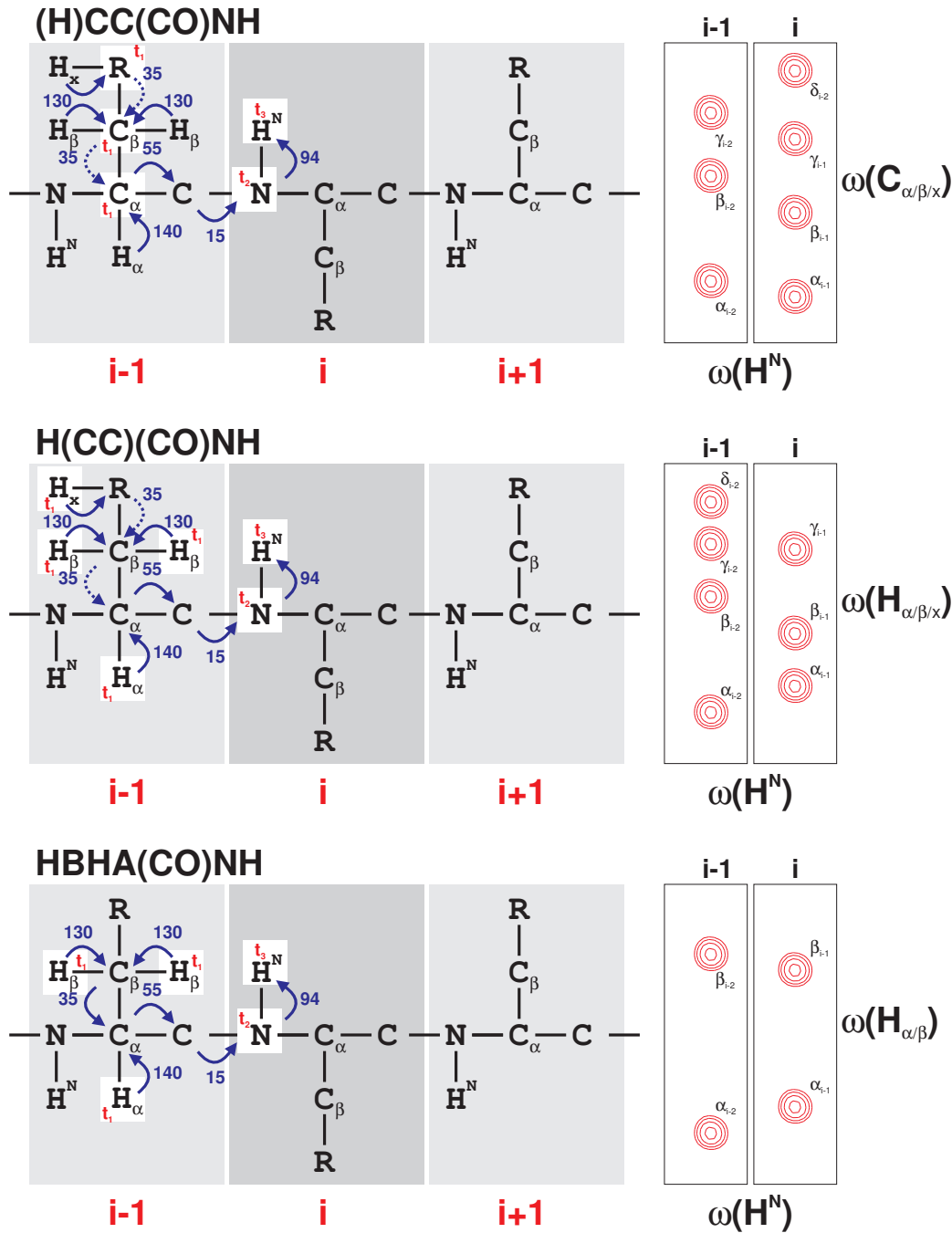


Figure 2.3.: Relevant backbone resonance assignment experiments for (non-native) proteins (III). Resonance transfer by scalar coupling is indicated by solid blue arrows, transfers through TOCSY mixing sequences are indicated by broken blue arrows. Coupling constants (in Hz) for the transfers are given in blue. Detected nuclei are indicated by white boxes and the symbols for the detection periods are given (t_1 , t_2 and t_3). The resonance patterns for relevant planes in the experiments are indicated on the left for the adjacent residues i and $i-1$.

2.3. Assignment of tryptophan side chains

2.3.1. General assignment strategy

As tryptophan side chains play key roles in the formation of hydrophobic clusters and the mediation of long-range contacts in non-native states of proteins (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004), it is desired to assign the resonances in the side chains of these residues. The nomenclature of tryptophan side chain atoms is depicted in figure 2.4.

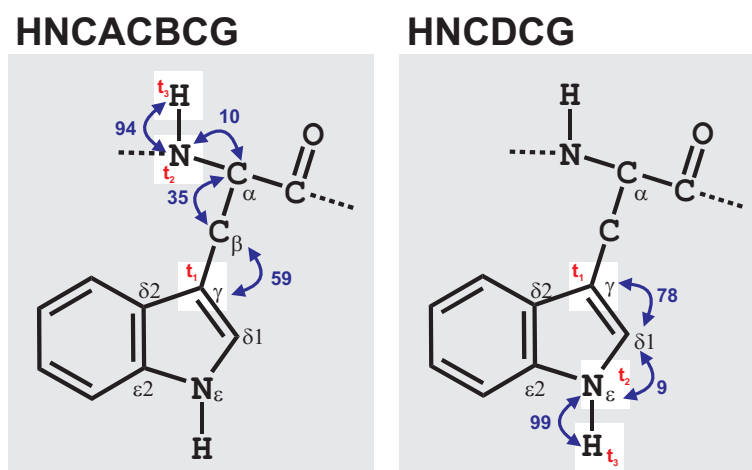


Figure 2.4.: Assignment strategy for the tryptophan side chain resonances. Resonance transfer by scalar coupling is indicated by solid blue arrows. Coupling constants (in Hz) for the transfers are given in blue. Detected nuclei are indicated by white boxes and the symbols for the detection periods are given (t_1 , t_2 and t_3).

A conventional method for the assignment of tryptophan side chains in proteins is based on the combination of side chain specific TOCSY and NOESY spectra (Slupsky *et al.*, 1998). In such an approach, the $^{13}\text{C}_\beta$ resonances are correlated with the aromatic $^1\text{H}_{\delta_1}$ resonances via a selective (HB)CB(CGCD)HD experiment (Yamazaki *et al.*, 1993). The $^1\text{H}_{\delta_1}$ resonances are then correlated to the $^{15}\text{N}_\epsilon$ and $^1\text{H}^{\text{N}}_\epsilon$ resonances by exploiting their scalar and NOE couplings to the $^1\text{H}_{\delta_1}$ nucleus. However, such strategies fail in unfolded proteins due to the massive overlap of resonances in the aromatic side chains.

Another previously published experiment for the sequence-specific assignment of tryptophan aromatic ^1H , ^{13}C and ^{15}N resonances in proteins (Löhr *et al.*, 2002) could also not be applied in the case of non-native lysozyme. The described HN(CDCG)CB experiment correlates the indole protons in tryptophan side chains with $^{13}\text{C}_\beta$ reso-

2.3. Assignment of tryptophan side chains

Table 2.1.: Approximate chemical shifts in tryptophan residues of *all-Ala*-HEWL.

Nucleus	Chemical shift [ppm]
$^1\text{H}^{\text{N}}$	7.2 - 8.0
$^{15}\text{N}^{\text{H}}$	120 - 123
$^{13}\text{C}_{\alpha}$	55
$^{13}\text{C}_{\beta}$	27
$^{13}\text{C}_{\gamma}$	110
$^{13}\text{C}_{\delta 1}$	127
$^{13}\text{C}_{\delta 2}$	129
$^{13}\text{C}_{\epsilon 2}$	138
$^{15}\text{N}_{\epsilon}$	130 - 131
$^1\text{H}_{\epsilon}^{\text{N}}$	9.9 - 10.0

nances in a series of magnetization transfers via one-bond scalar couplings. In the case of non-native proteins like HEWL-S^{Me} or *all-Ala*-HEWL, the $^{13}\text{C}_{\beta}$ resonances are not well enough resolved to allow for the unambiguous assignment of the tryptophan side chain resonances.

Since the existing experiments for the assignment of tryptophan side chain resonances in native proteins are not suitable in the case of some non-native proteins due to the resonance overlap, a set of two new NMR experiments had to be developed in this thesis. In these experiments, the tryptophan side chain indole resonances are assigned by the correlation of both the side chain and the backbone $^{15}\text{N}^{\text{H}}$ resonances with the $^{13}\text{C}_{\gamma}$ resonances of these residues. The $^{13}\text{C}_{\gamma}$ resonances exhibit a reasonable resolution even in non-native HEWL, which is a prerequisite for the unambiguous assignment. In addition, the $^{13}\text{C}_{\gamma}$ resonances are well separated from other aromatic carbon resonances. The two experiments have been entitled HN(CACB)CG and HN(CD)CG and are described in detail in section 2.3.2 on the following page and section 2.3.3 on page 49, respectively. While the HN(CACB)CG experiment correlates the backbone amide $^1\text{H}^{\text{N}}/^{15}\text{N}^{\text{H}}$ resonances with the $^{13}\text{C}_{\gamma}$ resonances in the side chain, the HN(CD)CG experiment correlates the side chain indole $^1\text{H}_{\epsilon}^{\text{N}}/^{15}\text{N}_{\epsilon}$ resonances with the $^{13}\text{C}_{\gamma}$ resonances. This way, the $^{13}\text{C}_{\gamma}$ resonances can be assigned via the HN(CACB)CG experiment assuming the backbone assignment is available. This in turn enables the assignment of the side chain indole resonances using the HN(CD)CG experiment.

2. Materials and Methods

Both experiments are of the *out-and-back* type, are derived from the HNCACB experiment (Muhandiram and Kay, 1994; Wittekind and Mueller, 1993) and use soft-WATERGATE water-suppression pulse schemes (Piotto *et al.*, 1992; Sklenar *et al.*, 1994). The specificity of these experiments for tryptophan side chains highly depends on the selectivity of the applied pulses and transfer delays in the pulse sequences.

Figure 2.4 on page 46 summarizes the courses of both experiments and depicts the exploited scalar couplings along the transfer pathways in Hz, while in table 2.1 on the previous page the typical chemical shifts of the nuclei in tryptophan residues of non-native proteins are listed. ^{13}C , ^{15}N doubly labeled *all-Ala*-HEWL was used to develop this set of two new NMR experiments.

2.3.2. The HN(CACB)CG experiment

Figure 2.5 on page 51 shows the pulse sequence of the HN(CACB)CG experiment. It comprises four successive steps to transfer magnetization via ^1J one-bond scalar couplings from the $^1\text{H}^{\text{N}}$ to the $^{13}\text{C}_{\gamma}$ nuclei. Amplitude modulated shaped pulses are used for the inter-carbon transfers and the decoupling of the $^1\text{J}(\text{C}_{\gamma}, \text{C}_{\delta 1})$ and $^1\text{J}(\text{C}_{\gamma}, \text{C}_{\delta 2})$ couplings during t_1 and to refocus $^{13}\text{C}_{\beta}$ chemical shifts during $2\tau' + t_1$. The delay τ' for the magnetization transfer from the $^{13}\text{C}_{\beta}$ to the $^{13}\text{C}_{\gamma}$ is adjusted to approximately $1/(2 \ ^1\text{J}(\text{C}_{\beta}, \text{C}_{\gamma}))$, conducting a HMQC (Bax *et al.*, 1983) scheme. To gain maximal transfer efficiency, the $^{13}\text{C}_{\alpha}$ - $^{13}\text{C}_{\beta}$ transfer delay is set to $1/(2 \ ^1\text{J}(\text{C}_{\alpha}, \text{C}_{\beta}))$, unlike in a standard HNCACB pulse sequence, where this delay duration is adjusted to $1/(4 \ ^1\text{J}(\text{C}_{\alpha}, \text{C}_{\beta}))$.

Narrow and wide filled bars in figure 2.5 correspond to rectangular 90° and 180° pulses applied with RF field strengths of 23.6 kHz (^1H) and 6.8 kHz (^{15}N), respectively. RF field strengths of 12.9 kHz and 28.7 kHz on ^{13}C resonances are used for the standard Q^3 and Q^5 Gaussian cascades (Emsley and Bodenhausen, 1992). Selective pulses and gradients are indicated by semi-ellipses and the default pulse phase is \times .

Fixed delays are adjusted as follows: $\Delta = 4.6 \text{ ms}$ ($1/(2 \ ^1\text{J}(\text{N}^{\text{H}}, \text{H}^{\text{N}}))$), $\Delta' = 4.9 \text{ ms}$, $\text{T} = 24.8 \text{ ms}$ ($1/(4 \ ^1\text{J}(\text{C}_{\alpha}, \text{N}^{\text{H}}))$), $\tau = 7.2 \text{ ms}$ ($1/(4 \ ^1\text{J}(\text{C}_{\alpha}, \text{C}_{\beta}))$), $\tau' = 8.5 \text{ ms}$ ($1/(2 \ ^1\text{J}(\text{C}_{\beta}, \text{C}_{\gamma}))$). Proton and nitrogen carrier frequencies are centered at the water (4.7 ppm) and the amide ^{15}N region (118 ppm), respectively. The carbon carrier frequency changes during the course of the experiment as indicated by vertical dashed lines and the value of the ^{13}C offset.

The shaped 180° decoupling pulses on carbonyl carbon resonances are Q^3 Gaussian cascades with durations of 2 ms and have offsets of 172 ppm. The shaped 180°

2.3. Assignment of tryptophan side chains

decoupling pulse on the C_β and C_δ resonances during the carbon t_1 time is based on a Q^3 Gaussian cascade and has duration of 1.7 ms.

1 ms water flip-back square pulses are applied after the first INEPT (Morris and Freeman, 1979) step and during the back transfer from the $^{15}\text{N}^{\text{H}}$ to the $^1\text{H}^{\text{N}}$.

Asynchronous GARP decoupling (Shaka *et al.*, 1985) is used to suppress ^{15}N - ^1H heteronuclear scalar coupling during acquisition. Proton decoupling using the DIPSI-2 (Shaka *et al.*, 1988) composite pulse decoupling is applied during most of the pulse sequence.

The pulsed field gradients of 1 ms length are sine-bell shaped, applied along the z-axis and have the following strengths: $G_1 = 27.5 \text{ G cm}^{-1}$, $G_2 = 22 \text{ G cm}^{-1}$, $G_3 = 33 \text{ G cm}^{-1}$, $G_4 = 16.5 \text{ G cm}^{-1}$.

Phase cycling is: $\phi_1 = 4(x), 4(-x)$; $\phi_2 = 16(x), 16(-x)$; $\phi_3 = y, -y$; $\phi_4 = 8(x), 8(-x)$; $\phi_5 = 32(x), 32(-x)$; $\phi_6 = 2(x), 2(-x)$; $\phi_{\text{rec}} = \text{R}, 2(-\text{R}), \text{R}, -\text{R}, 2\text{R}, -\text{R}$, where $\text{R} = 2(x), 4(-x), 2(x)$. In addition, ϕ_4 is incremented in a States-TPPI (Marion *et al.*, 1989) manner to achieve quadrature detection in the ω_1 direction.

The HN(CACB) pulse sequence in the standard Bruker pulse sequence programming language is listed in the appendix (section A.6.3 on page 139).

2.3.3. The HN(CD)CG experiment

Figure 2.6 on page 52 shows the pulse sequence of the HN(CACB)CG experiment. In this experiment, magnetization is transferred from the $^1\text{H}^{\text{N}_e}$ to the $^{13}\text{C}_\gamma$ nuclei in three successive steps: The initial INEPT transfer step from the indole proton to the nitrogen is followed by the transfer to the $^{13}\text{C}_{\delta 1}$ nucleus and the successive transfer to the $^{13}\text{C}_\gamma$. The pulses for the magnetization transfer between $^{13}\text{C}_{\delta 1}$ and the $^{13}\text{C}_\gamma$ are selective as well as the decoupling pulse on the $^{13}\text{C}_\beta$ and the $^{13}\text{C}_{\delta 1/2}$ carbons during t_1 . The offset for this transfer is on the $^{13}\text{C}_\gamma$ and all pulses are off resonance on the $^{13}\text{C}_{\delta 1/2}$ chemical shift region (117 ppm).

Narrow and wide filled bars in figure 2.6 correspond to rectangular 90° and 180° pulses applied with RF field strengths of 23.6 kHz (^1H) and 6.8 kHz (^{15}N), respectively. RF field strengths of 12.9 kHz and 28.7 kHz on ^{13}C resonances are used for the standard Q^3 and Q^5 Gaussian cascades (Emsley and Bodenhausen, 1992). Selective pulses and gradients are indicated by semi-ellipses and the default pulse phase is x.

Fixed delays are adjusted as follows: $\Delta = 4.6 \text{ ms}$ ($1/(2 \text{ } ^1\text{J}(\text{N}^{\text{H}_e}, \text{H}^{\text{N}_e}))$), $\Delta' = 4.9 \text{ ms}$, $T = 27.4 \text{ ms}$ ($1/(4 \text{ } ^1\text{J}(\text{C}_{\delta 1}, \text{N}^{\text{H}_e}))$), $\tau = 6.4 \text{ ms}$ ($1/(4 \text{ } ^1\text{J}(\text{C}_{\delta 1}, \text{C}_\gamma))$). Proton and nitrogen carrier frequencies are centered at the water (4.7 ppm) and the indole ^{15}N region

2. Materials and Methods

(129.2 ppm), respectively. The carbon carrier frequency changes during the course of the experiment as indicated by vertical dashed lines and the value of the ^{13}C offset.

The shaped 180° decoupling pulses on aromatic $^{13}\text{C}_\epsilon$ resonances are Q^3 Gaussian cascades with durations of 2 ms and have offsets of 141 ppm. The shaped 180° decoupling pulse on the $^{13}\text{C}_\beta$ and $^{13}\text{C}_\delta$ carbon resonances during the carbon t_1 time is based on a Q^3 Gaussian cascade and has duration of 1.15 ms.

1 ms water flip-back square pulses are applied after the first INEPT step and during the back transfer from $^{15}\text{N}^{\text{H}}_\epsilon$ to $^1\text{H}^{\text{N}}_\epsilon$.

Asynchronous GARP decoupling (Shaka *et al.*, 1985) is used to suppress ^{15}N - ^1H heteronuclear scalar coupling during acquisition. Proton decoupling using the DIPSI-2 (Shaka *et al.*, 1988) composite pulse decoupling is applied during most of the pulse sequence.

The pulsed field gradients of 1 ms length are sine-bell shaped, applied along the z-axis and have the following strengths: $G_1 = 27.5 \text{ G cm}^{-1}$, $G_2 = 22 \text{ G cm}^{-1}$, $G_3 = 33 \text{ G cm}^{-1}$, $G_4 = 16.5 \text{ G cm}^{-1}$.

Phase cycling is: $\phi_1 = 8(x), 8(-x)$; $\phi_2 = x$; $\phi_3 = y$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = 4(x), 4(-x)$; $\phi_{\text{rec}} = 2\text{R}, 2(-\text{R})$, where $\text{R} = 2(x), 2(-x)$. In addition, ϕ_2 and ϕ_3 are incremented in a States-TPPI (Marion *et al.*, 1989) manner to achieve quadrature detection in the ω_1 direction.

The HN(CD)CG pulse sequence in the standard Bruker pulse sequence programming language is listed in the appendix (section A.6.4 on page 144).

2.3. Assignment of tryptophan side chains

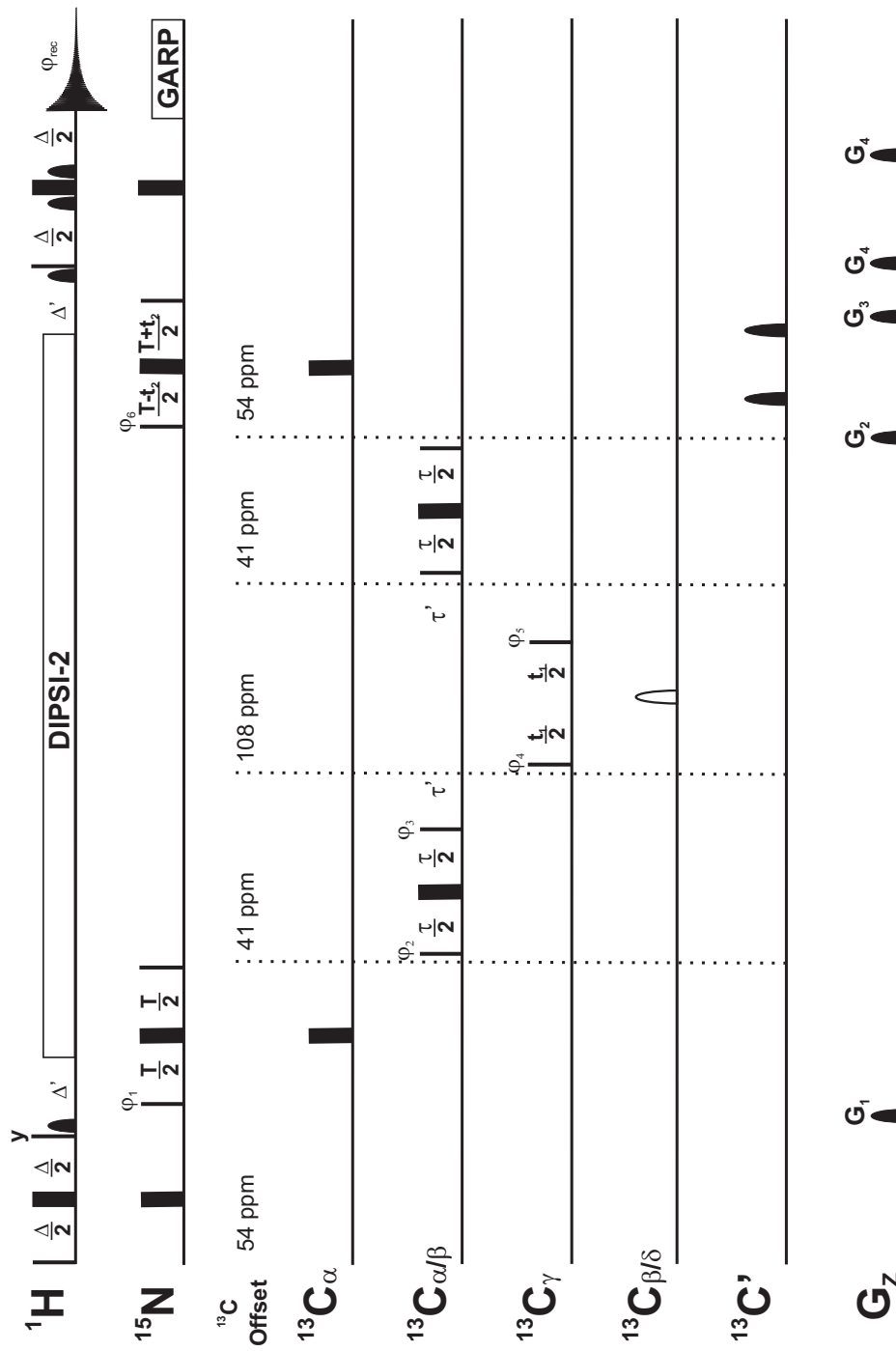


Figure 2.5.: Pulse sequence of the HN(CACB)CG experiment.

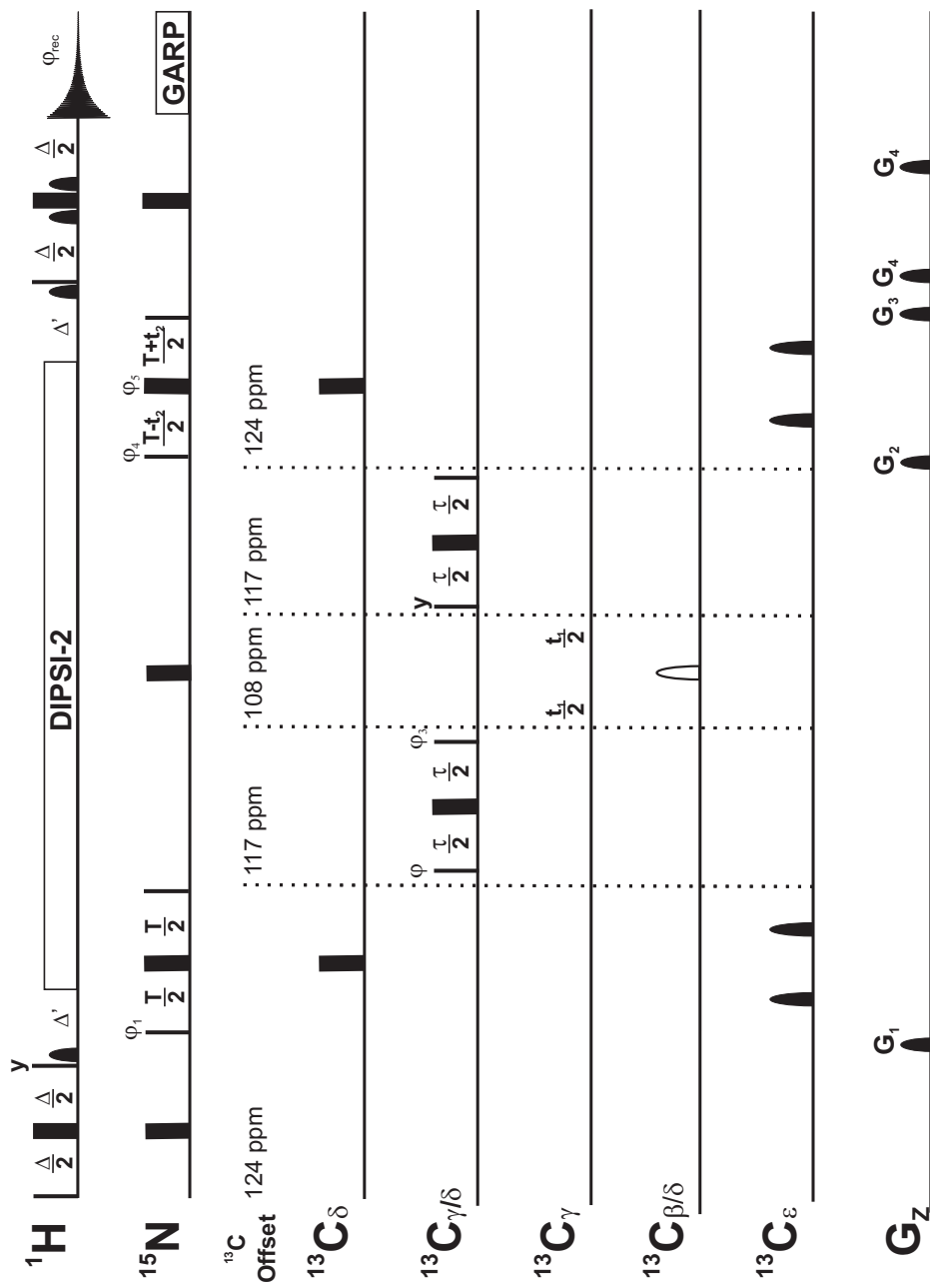


Figure 2.6.: Pulse sequence of the HN(CD)CG experiment.

2.4. Photo-CIDNP experiments

For the characterization of non-native states of proteins and the investigation of protein folding, photo-CIDNP experiments have turned out to be very useful tools (see introductory section 1.2.2 on page 14 and references therein). In the context of the structural analysis of non-native lysozyme, the differential surface exposure of the six tryptophan residues is especially interesting. However, one-dimensional spectra are not well enough resolved in the indole region of the tryptophan side chains of non-native lysozyme to discriminate between these residues in HEWL-S^{Me} or *all-Ala*-HEWL. Therefore, two-dimensional ¹⁵N-¹H photo-CIDNP experiments have been performed following a method that exploits the larger CIDNP enhancement of ¹⁵N compared to ¹H (Lyon *et al.*, 1999). The magnetization of the ¹⁵N nuclei is transferred in a single reverse INEPT step and detected on the ¹H, the water signal is suppressed by a hard-Watergate pulse scheme. The pulse sequence of this experiment is shown in figure 2.7 and listed in the standard Bruker pulse sequence programming language in section A.6.1 on page 133.

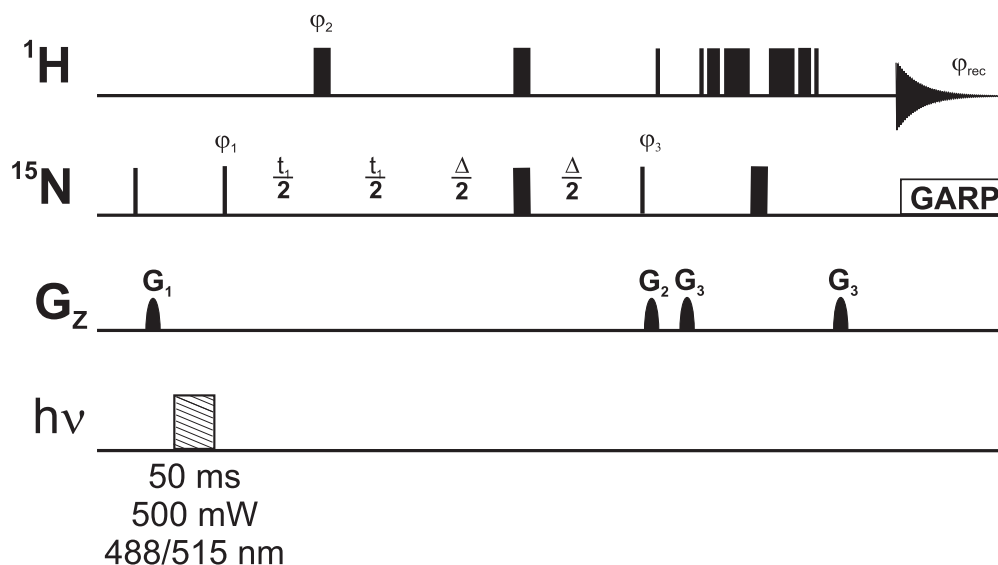


Figure 2.7.: Pulse sequence of the 2D photo-CIDNP experiment. $\Delta = 5.5$ ms, ϕ_1 : x, -x; ϕ_2 : 2(x), 2(-x); ϕ_3 : 4(y), 4(-y); ϕ_{rec} : x, -x, x, -x, -x, x, -x, x.

The scheme of the laser set-up used for the photo-CIDNP experiments is shown in figure 2.8 on the next page. A continuous-wave argon ion laser (Spectra-Physics, Darmstadt, Germany; model 2017) emitting at 488/515 nm has been used as the light source. The laser beam was transferred to the NMR tube via a set of mirrors and

2. Materials and Methods

lenses and an optical fiber (Ceram Optec, Bonn, Germany; core diameter 1 mm) (cp. Wirmer *et al.*, 2001). A spectrometer-controlled shutter was used and the fiber was coupled into standard NMR tubes through a 2 mm stem coaxial insert (Wilma-LabGlass, Buena, NJ, USA), which was inserted into the NMR tube so that the end of the fiber was dipping into the sample not more than 2 mm (cp. Scheffler *et al.*, 1985). The output power of the laser has been adjusted to 500 mW at the end of the fiber.

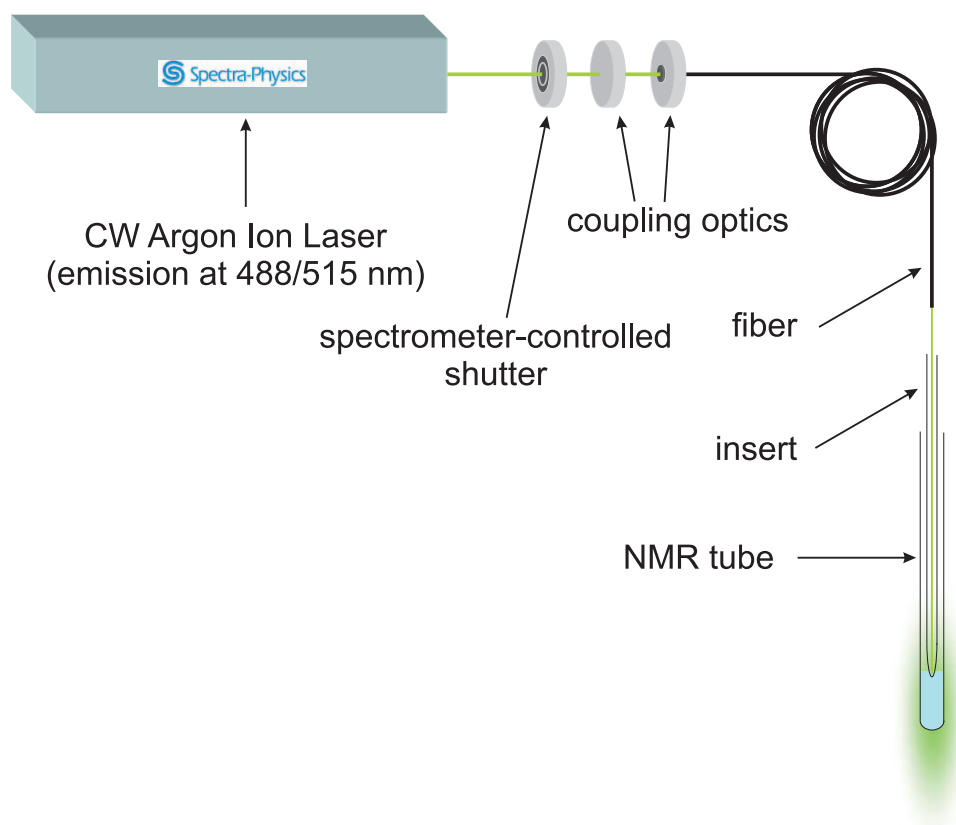


Figure 2.8.: Laser set-up for photo-CIDNP experiments.

The samples contained 200 μM of ^{15}N -labeled *all-Ala*-HEWL or W62G-*all-Ala*-HEWL and the equimolar concentration of the photo-sensitizer *riboflavin 5'-mononucleotide* (FMN) (see figure 2.9 on the facing page). The pH of the samples was adjusted to 2.0 and 10 % D_2O was added.

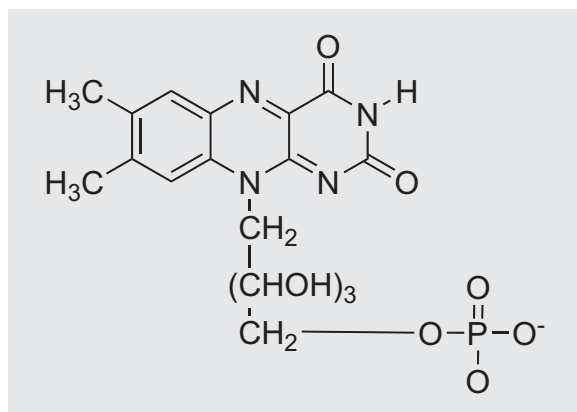


Figure 2.9.: Riboflavin 5'-mononucleotide (FMN).

The ^{15}N - ^1H two-dimensional photo-CIDNP experiments have been carried out at a temperature of 303 K with laser pulses of 50 ms length at a 700 MHz Bruker spectrometer equipped with a 5 mm ^1H , ^{13}C , ^{15}N cryogenic probe. Two scans have been conducted for each increment and 64 increments have been recorded in the indirect dimension. The carrier frequency of the indirect dimension was set to the center of the expected signals and a sweep width of 355 Hz was chosen.

Unlike for conventional one-dimensional photo-CIDNP spectra, where the final spectrum is the difference between the *light* (i.e. with laser irradiation) and the *dark* spectrum (i.e. without laser irradiation), no such correction of the baseline is necessary for the two-dimensional spectra, since the *dark* spectra have very low intensities.

The photo-CIDNP set-up and the pulse sequence for the two-dimensional experiments have been optimized in exploratory experiments on samples containing ^{15}N -labeled tryptophan.

In addition to the ^{15}N - ^1H two-dimensional photo-CIDNP experiment, an analogous ^{13}C - ^1H experiment has been developed and applied to ^{13}C , ^{15}N -labeled *all-Ala*-HEWL. The associated pulse sequence is listed in section A.6.2 on page 136. A spectrum has been recorded on a 600 MHz Bruker spectrometer equipped with a 5 mm ^1H , ^{13}C , ^{15}N cryogenic probe under the same conditions and with the same settings as for the ^{15}N , ^1H case. The sweep width was set to 3771 Hz.

For each two-dimensional photo-CIDNP spectra, a corresponding HSQC spectrum has been recorded with the same carrier frequencies and sweep widths to allow for the comparison of CIDNP vs. non-CIDNP peak intensities.

2.5. Protein Dynamics

2.5.1. Heteronuclear relaxation rates

Heteronuclear relaxation rates are routinely used to characterize the pico- to nano-second dynamics of biomacromolecules. For *all-Ala*-HEWL, several different approaches have been employed in this work. A general discussion of protein dynamics and a detailed description of the NMR experiments used for its elucidation can be found in the introductory section 1.2.2 on page 23, whereas the results are discussed in section 3.6 on page 83.

The pulse sequences used for the determination of the heteronuclear ^{15}N R_1 , $R_{1\rho}$ and R_2 relaxation rates and the heteronuclear ^1H - ^{15}N -NOE of the backbone amides are standard implementations provided by Bruker with the TOPSPIN 1.3 software (Kay *et al.*, 1989; Palmer and Case, 1992). All experiments have been carried out at a temperature of 293 K on a Bruker 600 MHz spectrometer equipped with a 5 mm ^1H , ^{13}C , ^{15}N pulsed field XYZ-triple-gradient probe. A 0.5 mM sample of *all-Ala*-HEWL at pH 2 has been used. The R_1 , $R_{1\rho}$ and R_2 experiments have been conducted in their pseudo-three-dimensional versions, while the heteronuclear NOEs have been determined with the NOE and no-NOE spectra recorded in an interleaved manner. For the R_2 and $R_{1\rho}$ (Dayie and Wagner, 1994; Deverell *et al.*, 1970) experiments, 15 different relaxation delays have been used each, ranging from 17 to 304 ms and from 0.01 to 0.5 s, respectively. For the determination of R_1 relaxation rates, seven different relaxation times have been used, ranging from 0.01 to 1.5 s. In all three cases, three relaxation times have been repeated to account for the error. A total of 128 real points at a spectral width of 2189 Hz in the ^{15}N dimension has been recorded in each case.

The respective relaxation rates have been obtained by fitting the peak intensities measured as a function of the relaxation delay to a two-parameter single-exponential decay with the Sparky 3.112 software (Goddard and Kneller, 2006).

Heteronuclear transverse ^{15}N R_2 relaxation rates have also been determined for the tryptophan side chains of *all-Ala*-HEWL under the same conditions as described for the amide backbone ^{15}N rates, but with 64 increments in the indirect dimension at a sweep width of 355 Hz. In addition, backbone ^{15}N R_2 rates have been determined for the W108G mutant of *all-Ala*-HEWL. Intensities for the tryptophan side chains have been extracted in the Felix 2004 software (Accelrys Inc. San Diego, CA, USA) and fitted to a two-parameter single-exponential decay with SigmaPlot 10.0 (Systat Software, Erkrath, Germany).

2.5.2. Relaxation dispersion

A single-quantum coherence ^{15}N relaxation dispersion experiment (Tollinger *et al.*, 2001) for *all-Ala*-HEWL has been conducted to study micro- to millisecond time scale exchange processes. The pseudo three-dimensional implementation of the pulse sequence provided by Bruker with the TOPSPIN 1.3 software was used. A total of 16 experiments with different CPMG field strengths — ranging from 50 to 1000 Hz — have been performed. Two experiments have been repeated to account for the error. The constant time delay was set to 80 ms. A spectrum with no CPMG field applied has been recorded as the reference. All spectra have been recorded at a temperature of 298 K with 128 real points and a sweep width of 24 ppm in the ^{15}N dimension. For the analysis of the relaxation dispersion data for any given peak, the effective relaxation rates (R_2^{eff}) are plotted against the CPMG field strength. R_2^{eff} is calculated from the peak intensities of the peaks with $I(v_{\text{CPMG}})$ and without (I_0) the CPMG periods and the length of a single CPMG train ($T/2$) as described in equation (2.1) (Mulder *et al.*, 2001).

$$R_2^{\text{eff}}(v_{\text{CPMG}}) = -\frac{1}{T} \ln \frac{I(v_{\text{CPMG}})}{I_0} \quad (2.1)$$

A general discussion of relaxation dispersion in the context of protein dynamics and a description of the related NMR methods can be found in the introductory section 1.2.2 on page 23, whereas the results are discussed in section 3.6 on page 89.

2.6. Residual dipolar couplings

2.6.1. Experimental determination of RDCs

Sample preparation

In order to determine residual dipolar couplings, the molecules in the sample have to be aligned with respect to the static magnetic field. For the theoretical background and a discussion of the different alignment methods, see section 1.2.2 on page 17. For unfolded proteins at low pH, stretched or compressed polyacrylamide gels (Sass *et al.*, 2000; Tycko *et al.*, 2000) have turned out to be the most feasible alignment media.

Here, stretched polyacrylamide gels have been used to align *all-Ala*-HEWL and W62G-*all-Ala*-HEWL. An apparatus (New Era Enterprises Inc., Vineland, NJ, USA; obtained via CortecNet, Paris, France) for casting, stretching and transferring the gels into special NMR tubes has been employed to prepare the samples (Chou *et al.*, 2001).

260 μL of a 7 % (w/v) acrylamide solution was prepared from a stock solution containing 40 % (w/v) acrylamide and 1.07 % (w/v) bisacrylamide. This solution was allowed to polymerize in the gel casting apparatus with an inner diameter of 6 mm upon addition of N,N,N',N'-tetramethylethylenediamine (TEMED) and 0.5 % (w/v) ammonium persulfate (APS). Two hours after initialization of the polymerization reaction, the gel was transferred into a 50 mL washing solution of water pH 2. After 20-24 hours of washing, the gel was dried at 37 °C for 24 hours. The dried gel was allowed to reswell for more than 24 hours in 400 μL of protein solution (500 μM , pH 2) in the casting chamber. Following the swelling, the gel was transferred via the attached funnel into the NMR tube, which has a diameter of only 4.2 mm.

NMR experiments

$^1\text{J}(\text{H}^{\text{N}}, \text{N}^{\text{H}})$ scalar couplings have been determined for the unaligned (i.e. isotropic) samples using the standard Bruker implementation of an in-phase/anti-phase (IPAP) HSQC experiment (Ishii *et al.*, 2001; Ottiger *et al.*, 1998), where in a set of two subspectra, the in-phase and anti-phase signals are added or subtracted, respectively. The coupling is read out from the difference of the respective peak positions in the two-spectra. The sum of the scalar $^1\text{J}(\text{H}^{\text{N}}, \text{N}^{\text{H}})$ and dipolar $^1\text{D}(\text{H}^{\text{N}}, \text{N}^{\text{H}})$ couplings in the aligned (i.e. anisotropic) sample is measured in the same way, the $^1\text{D}(\text{H}^{\text{N}}, \text{N}^{\text{H}})$ couplings therefore result from the subtraction of the values of the isotropic case

from the values in the anisotropic case. All IPAP-HSQC spectra have been recorded with four scans and 512 real points and processed with 4096 points, yielding a resolution of 0.4 Hz in the ^{15}N dimension.

2.6.2. Simulation of RDCs

A series of simulations of residual dipolar couplings of structural ensembles representing peptide chains with the amino acid sequence of *all-Ala*-HEWL in a neutral alignment medium has been conducted in order to compare it with the experimental RDC values. These simulations have been performed in collaboration with Dr. Martin Blackledge (*Institut de biologie structurale (IBS) Jean-Pierre Ebel*, Grenoble, France), who — together with his coworkers — developed and provided the software for these calculations based on the "flexible-meccano" algorithm (Bernadó *et al.*, 2005a,b).

In this approach, statistical ensembles of coil conformations are created and averaged RDCs are calculated over the entire population of these ensembles (Bernadó *et al.*, 2005b; Jha *et al.*, 2005). The "flexible-meccano" algorithm sequentially generates peptide chains with randomly selected ϕ, ψ backbone pairs. These angle pairs are extracted from a database of amino acid specific conformations occurring in loop regions of a set of 500 high-resolution X-ray structures (Lovell *et al.*, 2003). During the generation of the ensembles of conformers, steric overlap is avoided by the use of a simple volume-exclusion model (Levitt, 1976). The alignment tensor of each conformer is predicted employing a model that is based on the hydrodynamic shape of the molecule and that makes use of the similarity between the *alignment* tensor and the *radius of gyration* tensor in the case where the alignment is solely caused by steric exclusion (Almond and Axelsen, 2002; Zweckstetter and Bax, 2000). RDCs are then calculated from these tensors using the relationship given in equation (2.2), where A_a and A_r are the axial and rhombic components of the alignment tensor, and θ and ϕ are the polar angles of the vector with respect to these axes (see the description of equation (1.2) on page 17 for further explanations and the meaning of all symbols).

$$D_{ij}(\theta, \phi) = -\frac{\gamma_i \gamma_j \mu_0 \hbar}{16\pi^3 r_{ij,eff}^3} [A_a(3\cos^2\theta - 1) + \frac{3}{2}A_r \sin^2\theta \cos 2\phi] \quad (2.2)$$

The generated ensembles generally consisted of 50,000 conformers and the predicted individual RDCs have been averaged over the whole ensemble to ensure convergence.

Furthermore, calculations have been performed with additional restraints in independent approaches. Firstly, one or more restrained contacts between residues

2. *Materials and Methods*

have been implemented to take the clustering into account. Secondly, a tryptophan-tryptophan ϕ,ψ sampling has been incorporated for the residue pair W62/W63 to consider the particular ϕ,ψ distribution of this motif. This backbone angle distribution has been extracted from a database of known crystal and solution structures and provided by Jun.-Prof. Dr. Holger Gohlke (*Institute for Cell Biology and Neuroscience, Johann Wolfgang Goethe-University, Frankfurt, Germany*).

2.7. Residual structure in organic solvents

The influence of the organic solvents ethanol and 2,2,2-trifluoroethanol (TFE) on the secondary structure of HEWL-S^{Me} and *all-Ala*-HEWL was studied by circular dichroism (CD) and NMR spectroscopy. HEWL-S^{Me} has been prepared from native HEWL obtained from Fluka by reduction of the disulfide bridges and the successive S-methylation of the cysteine residues as described in section 2.1.3 on page 38. The preparation of *all-Ala*-HEWL followed the protocol discussed in section 2.1.2 on page 36.

The samples of HEWL-S^{Me} and *all-Ala*-HEWL for CD spectroscopy had concentrations in the range from 20 to 45 μM , while the temperature during the experiments was set to 293 K. The CD spectra have been recorded on a Jasco J-810 (Jasco GmbH, Groß-Umstadt, Germany) CD spectrometer equipped with a Jasco PTC-423S Peltier type temperature control system using quartz cuvettes with 0.2 mm path lengths. The ellipticity was calibrated with (+)-10-camphorsulfonic acid and is reported as mean molar residual ellipticity, $[\theta]$ ($\text{deg cm}^2 \text{ dmol}^{-1}$). Data were collected at 0.2 nm increments from 260 to 190 nm.

Samples have been prepared with various concentrations (all percentages are *volume of organic solvent per total volume*) of ethanol (30 %, 60 %, 90 %) and 2,2,2-trifluoroethanol (10 %, 30 %, 50 %) at pH 2. A sample with 90 % ethanol was prepared at pH 10 and a spectrum of native HEWL has been recorded at pH 3.8. Samples with the highest respective concentrations of organic solvents have been monitored for changes over a period of 96 hours.

The mean molar residual ellipticity at 222 nm, θ_{222} , is assumed to be linearly related to the mean helix content, f_H , following the modified Lifson-Roig model for helix-coil transitions (Lifson and Roig, 1961; Rohl *et al.*, 1996). The mean helix content is calculated using equation (2.3) (together with equations (2.4) and (2.5), where θ_C and θ_H are the baseline ellipticities of the random coil and the complete helix, respectively, T is the temperature in $^\circ\text{C}$ and N_r is the chain length in residues) given by Rohl and Baldwin (Rohl and Baldwin, 1997).

$$f_H = \frac{\theta_{222} - \theta_C}{\theta_H - \theta_C} \quad (2.3)$$

$$\theta_C = 2220 - 53T \quad (2.4)$$

2. Materials and Methods

$$\theta_H = (-44000 + 250T)\left(1 - \frac{3}{N_r}\right) \quad (2.5)$$

For the recording of ^1H - ^{15}N -HSQC spectra, samples of 0.5 mM *all-Ala*-HEWL with various concentrations of deuterated ethanol-d6 (30 %, 60 %, 90 %) and 2,2,2-trifluoroethanol-d3 (50 %) have been prepared. All spectra have been recorded at a temperature of 293 K on a Bruker 600 MHz spectrometer equipped with a 5 mm ^1H , ^{13}C , ^{15}N cryogenic probe. All spectra were of the FHSQC-type (Mori *et al.*, 1995) and comprised 128 real points at a sweep width of 28 ppm in the indirect dimension.

3. Results and Discussion

3.1. Sample preparation

The isotope labeling of HEWL is required in order to investigate its non-native states with NMR spectroscopy at atomic resolution. Here, the expression of HEWL in *E. coli* using an isotope enriched medium is reported. The expression in *E. coli* usually is less time-consuming, easier to implement, and isotope enrichment is less cost-intensive due to a less complex medium than in eukaryotic expression systems such as *Pichia pastoris* (Mine *et al.*, 1999) and *Aspergillus niger* (MacKenzie *et al.*, 1996; Spencer *et al.*, 1999) that have been traditionally used for the expression of HEWL.

The expression of HEWL and *all-Ala*-HEWL in *E. coli* typically yielded about 15 mg ^{15}N singly or $^{13}\text{C},^{15}\text{N}$ doubly labeled protein per liter M9 minimal medium. Most of the protein was produced in the first hour after induction. The bulk amount of the overproduced lysozyme was expressed into *inclusion bodies* as shown in the SDS-PAGE analysis of the different steps of the washing and solubilization process (see figure 3.1 on the following page), where the supernatants of the washing steps (S1 to S3) did not contain any significant amounts of lysozyme, while the solubilized *inclusion bodies* (S4) primarily consisted of this protein. After purification with CM sepharose ion exchange chromatography, the protein was already more than 95 % pure as judged from SDS-PAGE analysis (see figure 3.2 on the next page).

HEWL expressed in *E. coli* and purified from *inclusion bodies* can be refolded into the native and fully active form of this enzyme by rapid dilution into a refolding buffer (containing 5 mM reduced glutathione and 0.5 mM oxidized glutathione) (Schlörb *et al.*, 2005). Figure 3.3 on page 65 shows ^1H - ^{15}N -HSQC spectra of successfully refolded HEWL (left) and of *all-Ala*-HEWL (right). The spectrum of refolded HEWL is identical to a spectrum of HEWL expressed in *A. niger* (Buck *et al.*, 1995a).

Following the reduction and S-methylation of the cysteine residues, HEWL-S^{Me} partially aggregated and redissolved in water pH 2 after several days and dialysis steps. The reduction and S-methylation renders the hen egg white lysozyme permanently non-native, as the four disulfide bridges that stabilize the native structure can-

3. Results and Discussion

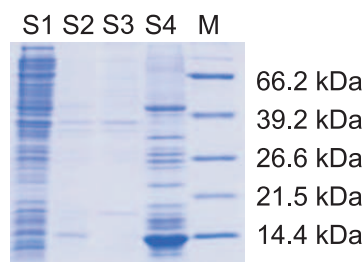


Figure 3.1.: Coomassie-stained 15 % SDS polyacrylamide gel of various samples containing *all-Ala*-HEWL during *inclusion body* washing and solubilization. M: Roche Low Range Marker (Roche Diagnostics, Mannheim, Germany); S1: supernatant 1; S2: supernatant 2; S3: supernatant 3; S4: solubilized inclusion bodies. The molecular weights of the marker proteins are indicated on the right.

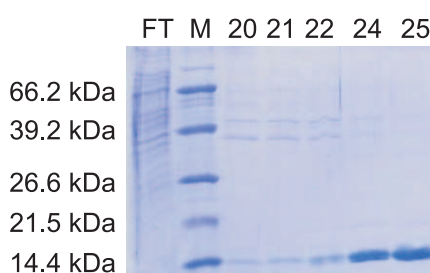


Figure 3.2.: Coomassie-stained 15 % SDS polyacrylamide gel of various samples containing *all-Ala*-HEWL during ion exchange chromatography. FT: flow-through; M: Roche Low Range Marker; 20-25: various fractions after the CM ion exchange column. The molecular weights of the marker proteins are indicated on the left.

not be formed. A mimic of this situation is the *all-Ala* mutant, in which all of the eight cysteines have been replaced by alanines. The low pH value of used throughout the experiments in this work, is a denaturing factor and helps to prevent the proteins from aggregating. It has turned out that the preparation of *all-Ala*-HEWL samples is less time consuming and more reproducible than the generation of HEWL-S^{Me} samples. Therefore, *all-Ala*-HEWL and single point mutants thereof are used for the characterization of the structural and dynamical properties of non-native states of lysozyme.

The NMR samples of HEWL-S^{Me}, *all-Ala*-HEWL and the mutants in water pH 2 turned out to be stable over many months. Only after several months, faint peaks start to appear in the spectra, indicating a slow degradation of the proteins.

The generation of the single-point mutants W62G and W108G of *all-Ala*-HEWL using the Stratagene mutagenesis kit was straightforward. The DNA sequences of the resulting plasmids have been verified (SRD GmbH, Oberursel, Germany), the

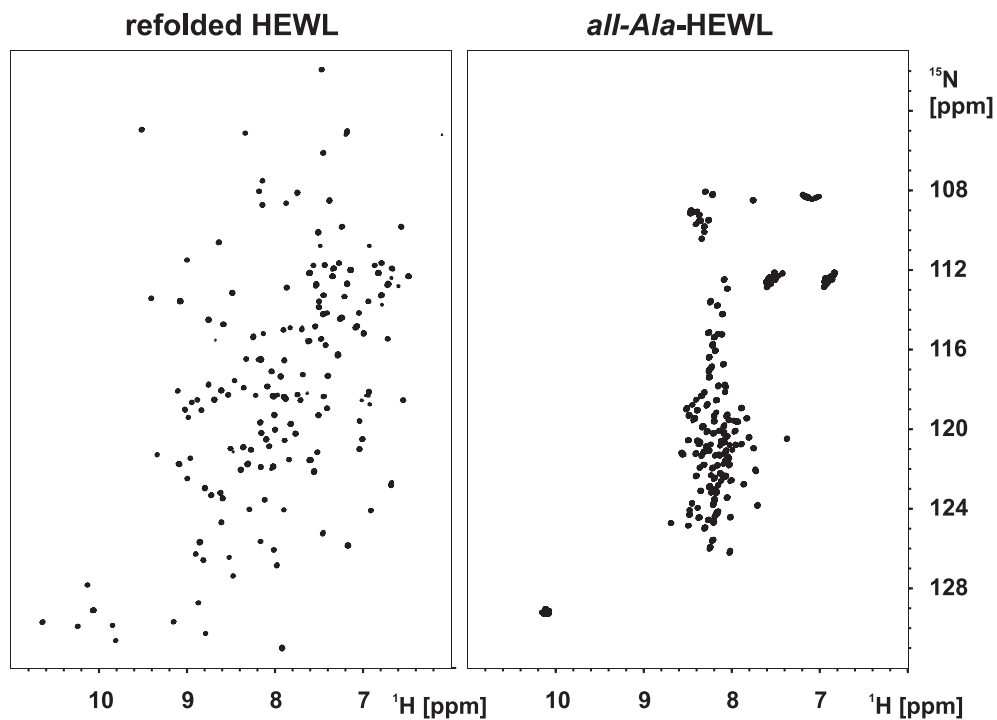


Figure 3.3.: ^1H - ^{15}N -HSQC spectra of refolded HEWL at pH 3.8 (left) and *all-Ala*-HEWL at pH 2.0 (right).

expression yield and the purification procedures have been identical to those of *all-Ala*-HEWL.

3.2. Backbone resonance assignment

The assignment of the backbone resonances of HEWL-S^{Me}, *all-Ala*-HEWL and mutants thereof is the prerequisite to study their properties with NMR spectroscopy at atomic resolution. The right panel in figure 3.3 illustrates the difficulties for the unambiguous assignment of resonances caused by the low dispersion and signal overlap in non-native states as compared to folded proteins. Nevertheless, for 120 of the 130 residues (i.e. 94 % of all residues without the two prolines and the N-terminal methionine) of HEWL-S^{Me} the ¹H^N, ¹⁵N^H, ¹³C_α, ¹³C_β and ¹³C' backbone resonances have been completely assigned in water pH 2 at 293 K. This assignment rate is comparable or even higher, than the rates for S-methylated HEWL in water or reduced and oxidized HEWL in 8 M urea that have been published before (Grimshaw, 1999; Hennig *et al.*, 1999; Schwalbe *et al.*, 1997). For *all-Ala*-HEWL the rate of backbone assignment is even higher, with 125 residues being assigned (98 % of the 127 in principle assignable residues). The assignments for HEWL-S^{Me} and *all-Ala*-HEWL have been deposited in the *Biological Magnetic Resonance Data Bank* (BMRB) under the accession codes 6622 and 15198, respectively.

In general, the HNCACB experiment proved to be the most important spectra for the assignment of the backbone residues of HEWL-S^{Me} and *all-Ala*-HEWL. Combinations of alanine, glycine, serine or glycine residues in the amino acid sequences in most cases have been the most valuable starting points for the sequential correlation of the resonances because of their very characteristic peak patterns especially in the HNCACB spectra. An example for the sequential assignment using strip representations from the HNCACB spectrum of *all-Ala*-HEWL is shown in figure 3.4 on the next page for the residues N39 to N46.

The lower assignment rate for HEWL-S^{Me} compared to *all-Ala*-HEWL mainly results from the additional HNN experiment carried out for the latter, which helped to overcome ambiguities in the other spectra. Especially the high number of alanine residues (20) in *all-Ala*-HEWL made the assignment difficult, yet only two could not be assigned unambiguously. Figure 3.5 on page 68 shows a sequential walk through aligned strips from the HNN spectrum of *all-Ala*-HEWL for residues G22 to V29. An important advantage of the HNN experiment is the presence of cross-peaks for both *i*-1 and *i*+1 residues, which allows for an unambiguous assignment in most cases.

In addition to the ¹H^N, ¹⁵N^H, ¹³C_α and ¹³C_β, which have been assigned from the HNCACB spectra, ¹³C' resonances have been picked and assigned from the HNCO (for HEWL-S^{Me}) or HN(CA)CO (for *all-Ala*-HEWL) experiments. The side chain ¹H_α

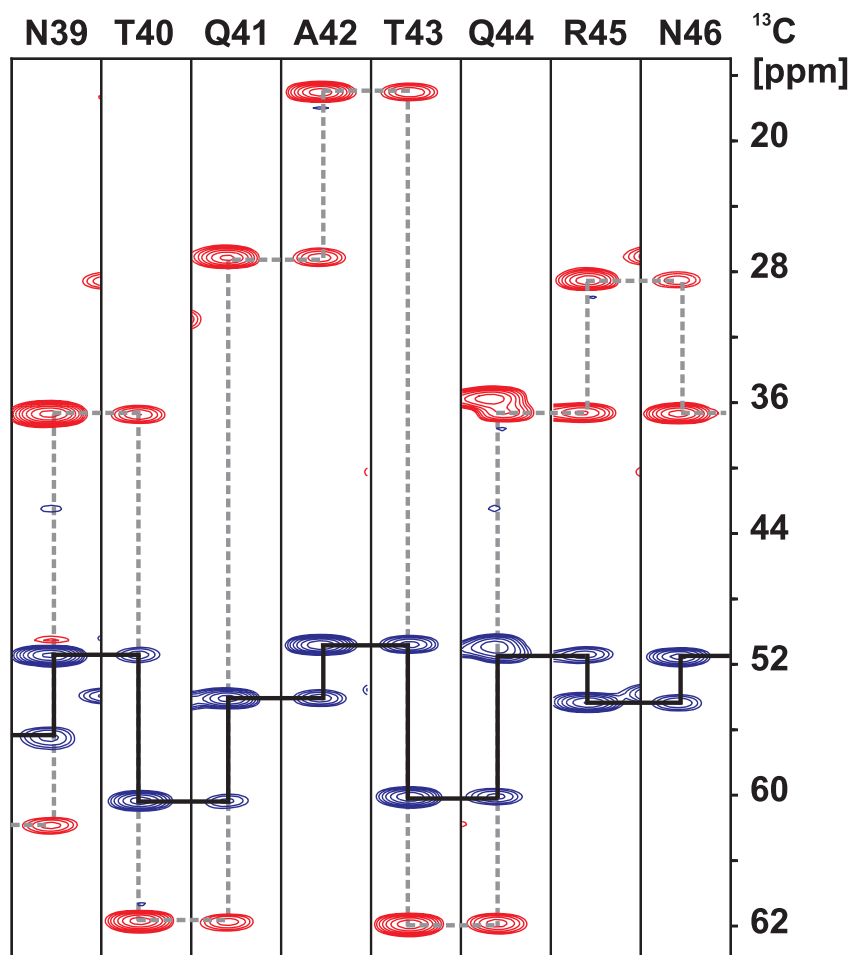


Figure 3.4.: Sequential backbone resonance assignment using the HNCACB experiment (shown exemplarily for residues N39 to N46 of *all-Ala*-HEWL). The walk between adjacent C_α peaks (blue) is indicated by black lines, the walk between adjacent C_β peaks (red) is indicated by dashed grey lines.

3. Results and Discussion

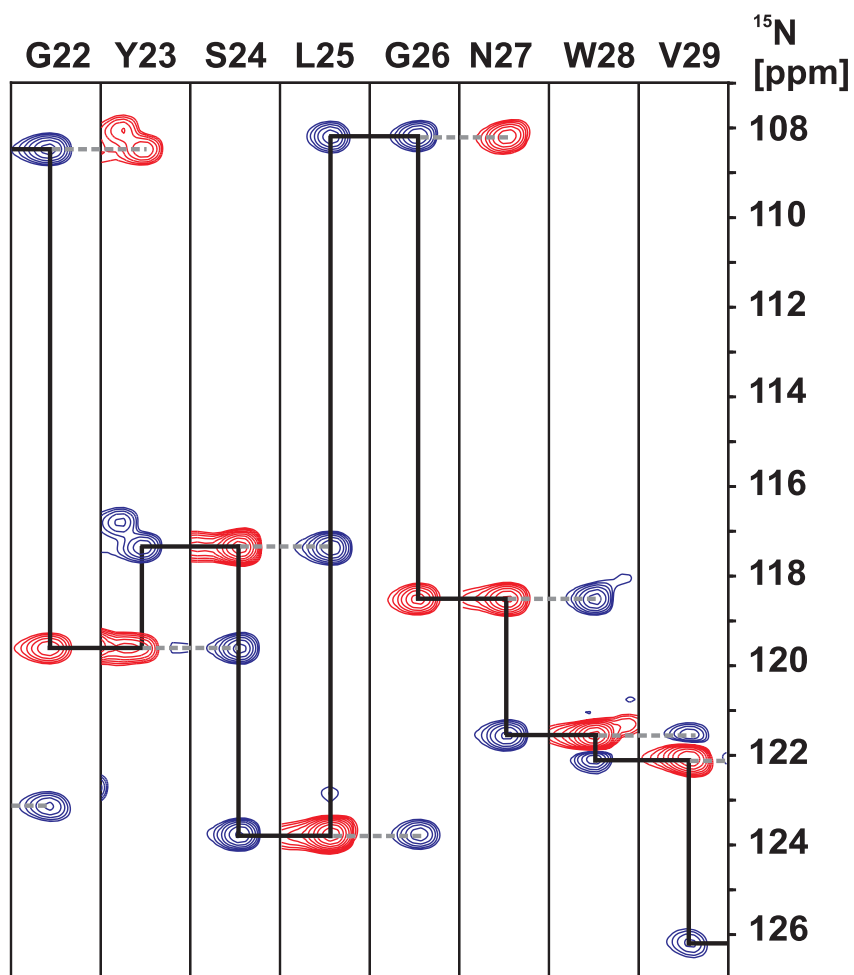


Figure 3.5.: Sequential backbone resonance assignment using the HNN experiment (shown exemplarily for residues G22 to V29 of *all-Ala*-HEWL). The walk between N_{i-1} and N_i peaks is indicated by black lines, while N_i and N_{i+1} peaks are connected by dashed grey lines. Please note, that the $i+1$ cross-peak to a glycine residue has an inverted phase.

3.2. Backbone resonance assignment

and $^1\text{H}_\beta$ resonances have been assigned for the sake of completeness with the help of HBHA(CO)NH experiments for HEWL-S^{Me} and *all-Ala*-HEWL.

Figure 3.6 on the next page shows the result for the assignment of *all-Ala*-HEWL annotated in a ^1H - ^{15}N -HSQC spectrum. All of the major peaks have been assigned. However, a very small set of weak peaks, remains unassigned. These peaks might stem from copurified peptides, degradation products or weakly populated alternative conformations.

The $^1\text{H}^{\text{N}}$ and $^{15}\text{N}^{\text{H}}$ resonances of the single point mutants W62G- and W108G-*all-Ala*-HEWL have been assigned by thorough comparison of their ^1H - ^{15}N -HSQC spectra with the *all-Ala*-HEWL-HSQC spectrum. This way, 117 (W62G) and 122 (W108G) residues have been assigned, respectively.

3. Results and Discussion

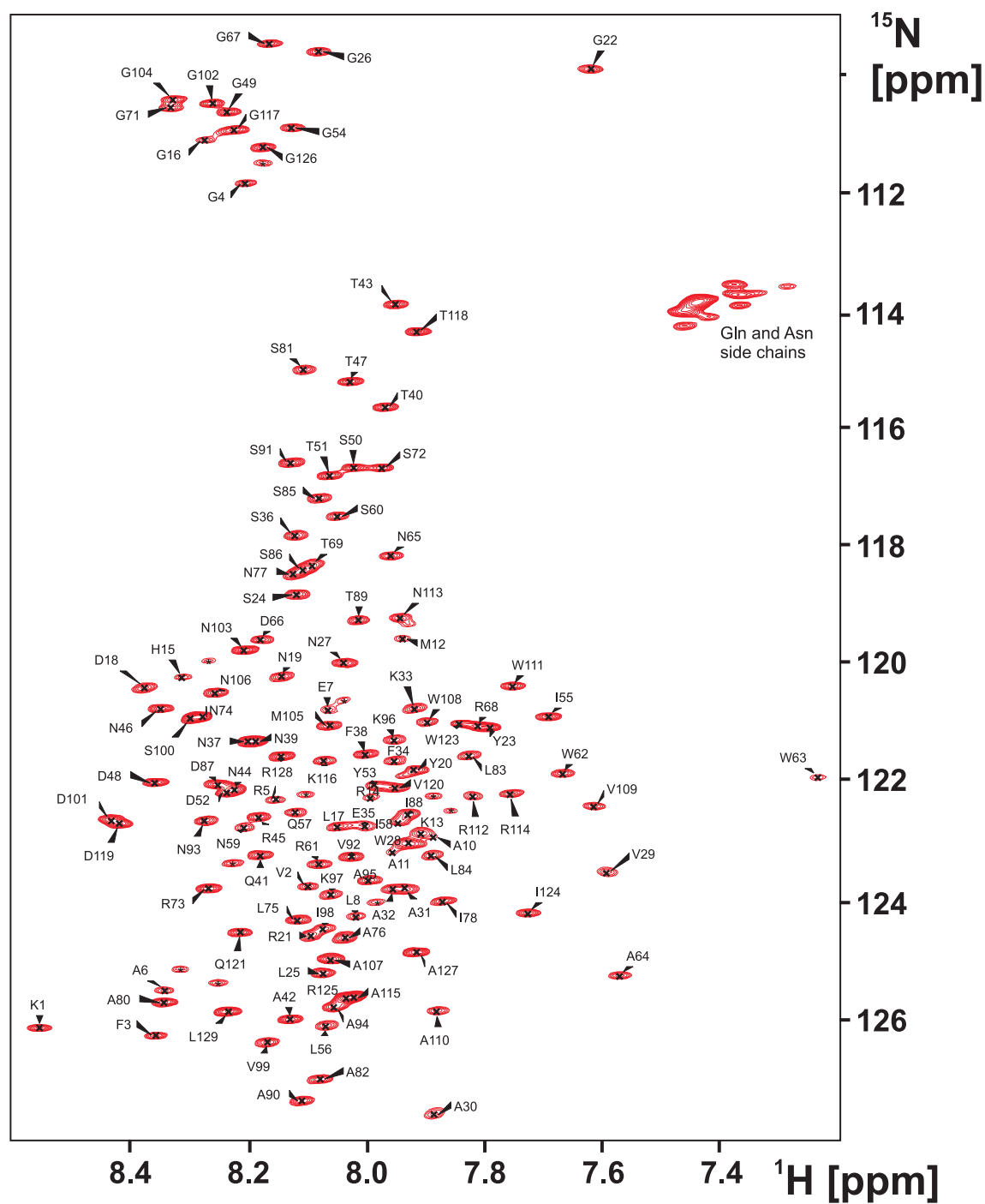


Figure 3.6.: Annotated ^1H - ^{15}N -HSQC spectrum of 0.5 mM *all-Ala*-HEWL in water pH 2.0 at 293 K. Unassigned peaks indicated by stars (*) may stem from copurified peptides, degradation products or weakly populated alternative conformations.

3.3. Chemical shift analysis

As pointed out in section 1.2.2 on page 12, secondary chemical shifts can provide information on the (residual) structure of proteins. Secondary chemical shifts ($\Delta\delta$) are the difference of the measured and correctly referenced chemical shifts (δ_{exp}) and *random coil* chemical shifts (δ_{rc}) from the literature (e.g. from Schwarzsinger *et al.*, 2000). The $^1\text{H}_\alpha$ and $^{13}\text{C}_\alpha$ secondary chemical shifts of *all-Ala*-HEWL are represented in figure 3.7 on the next page against the residue number. The most prominent deviations from *random coil* chemical shifts are situated in three parts of the protein's sequence: residues 5 to 12 (sequence: RAELAAAM), residues 60 to 67 (sequence: SRWWANDG), and residues 107 to 113 (sequence: AWVAVRN). The same pattern of secondary chemical shifts has been observed for HEWL-S^{Me} (Schlörb, 2003; Wirmer *et al.*, 2004). Averaging over all (assigned) residues, the fractional helicity, i.e. the proportion of α -helical-like populations, was found to be 12.3 %, while circular dichroism measurements yielded a helicity of 13.9 % for the whole protein HEWL-S^{Me} (see section 3.8 on page 96). Interestingly, these short segments coincide with the hydrophobic clusters 1, 3 and 5, known already from combined relaxation and mutational studies (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004). The direction of these deviations from *random coil* chemical shifts in these three segments is towards chemical shifts of α_{R} -helical conformations (cf. RefDB, Zhang *et al.*, 2003). In the case of the residues 5 to 12, this is already known from the analysis of scalar couplings of lysozyme based peptides (Graf *et al.*, 2007). In this study, two HEWL-S^{Me} based peptides have been investigated at pH 2 by a joint NMR and molecular dynamics (MD) approach: The first peptide consisted of residues 6 to 14 (9mer with the amino acid sequence: C^{Me}ELAAAMKR) and the second contained residues 1 to 19 (19mer with the amino acid sequence: KVFGRC^{Me}ELAAAMKRHGLDN) of the HEWL sequence, where C^{Me} stands for a S-methylated cysteine. While the conformational distribution in the central three alanine residues of the 9mer was found to be similar to the mainly poly-L-proline II helix-like structures of the short poly-alanine peptides Ala₃ to Ala₇, the 19mer significantly samples α_{R} -helical structures.

In addition, the secondary $^1\text{H}_\alpha$ and $^{13}\text{C}_\alpha$ chemical shifts of the 9mer and the 19mer support these findings and indicate a higher α -helical population for the longer peptide (Graf, 2006). Figure 3.8 on page 73 shows an overlay of zooms into the alanine region of the ^1H - ^{15}N -HSQC spectra of the 9mer (blue peaks) and 19mer (red) peptides and of HEWL-S^{Me} (black). The peaks for alanine 9 to 11 nicely overlay for the 19mer peptide and HEWL-S^{Me}, but not for the 9mer. This also indicates a simi-

3. Results and Discussion

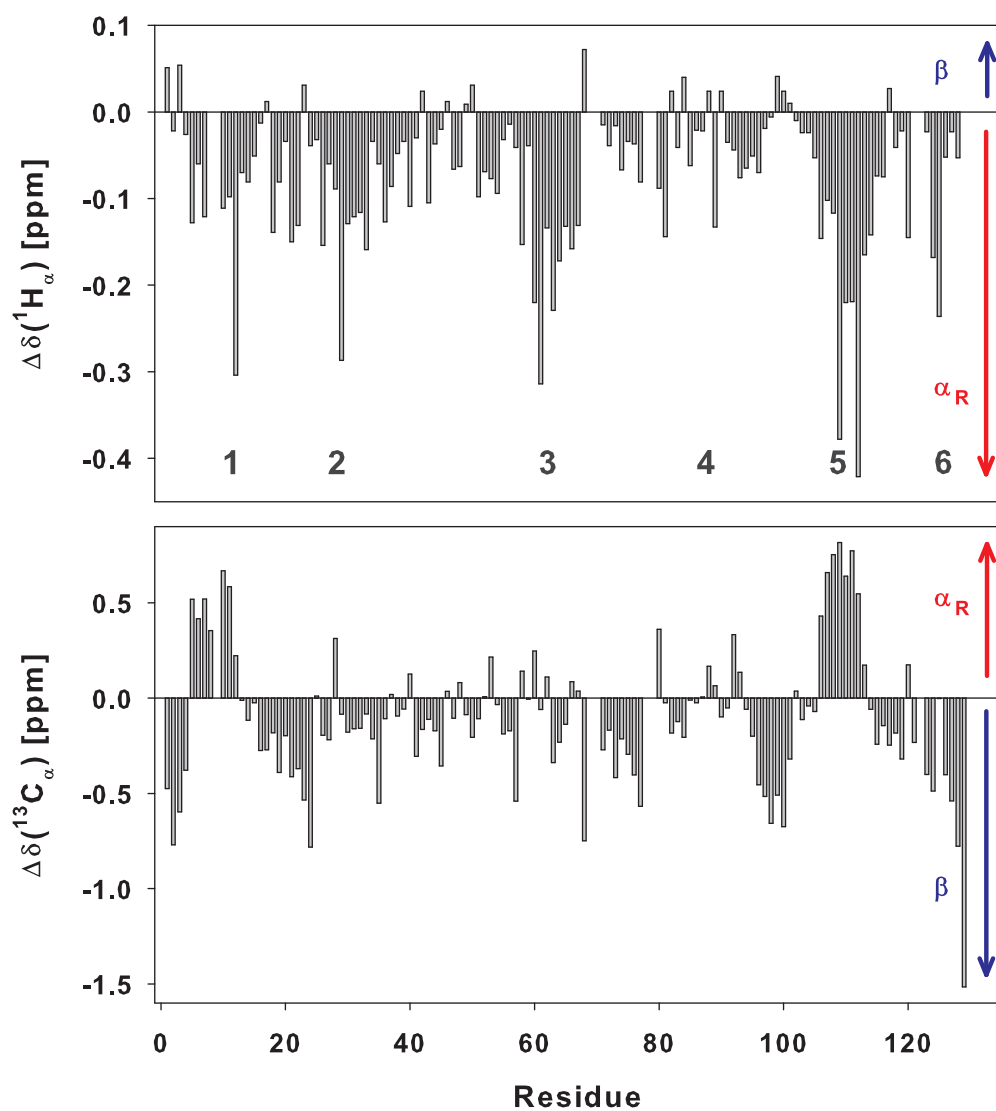


Figure 3.7.: Secondary chemical shift plots for $^1\text{H}_\alpha$ and $^{13}\text{C}_\alpha$ of *all-Ala-HEWL*. The difference of experimental chemical shifts and *random coil* chemical shifts (Schwarzinger *et al.*, 2000) is plotted against the residue numbers. Residues preceding prolines have not been included in the $^{13}\text{C}_\alpha$ plot, due to their unusual chemical shifts. Numbers from 1 to 6 are given for the hydrophobic clusters. α -helical or β -sheet propensities are indicated by red and blue arrows, respectively.

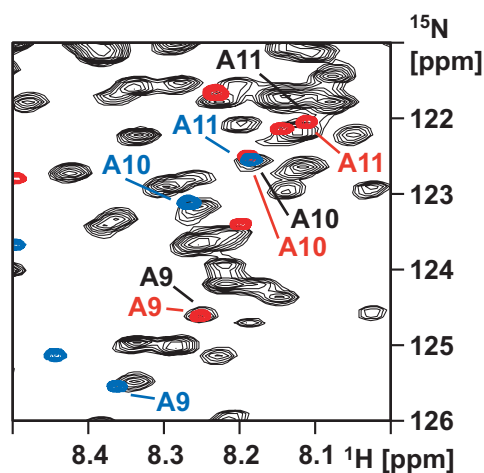


Figure 3.8.: Overlay of zooms into the alanine regions of the ^1H - ^{15}N -HSQC spectra of the 9mer (blue peaks) and 19mer (red) peptides and of HEWL-S^{Me} (black). The spectra for the peptides are taken from Graf (2006).

lar structure for the longer peptide and the corresponding section of the non-native HEWL. Together, these findings indicate, that at least for the sequence around the three alanine residues in cluster 1, the propensity for α -helical conformations might be intrinsic to this specific amino acid sequence. It is not known so far, to which degree the α -helix-like conformations in clusters 3 and 5 are sequence dependent (i.e. a local effect) or whether they are induced by long-range interactions.

3.4. Assignment of tryptophan side chains

The assignment of the resonances in tryptophan side chains in non-native lysozyme enables new possibilities to study structural and dynamical properties of the protein that occur at the level of these side chains, such as hydrophobic clustering and other possible interactions. As mentioned in section 2.3 on page 46, the major obstacle for the unambiguous assignment of the tryptophan side chain nuclei in non-native proteins is the low dispersion of their signals due to the considerably lower degree of structure in non-native proteins compared to native proteins. The dispersion of the side chain indole peaks in *all-Ala*-HEWL (0.07 ppm in the ^1H and 0.25 ppm in the ^{15}N dimension; compare table A.8 on page 115) is about tenfold lower than the dispersion of the backbone amide signals (around 0.7 ppm in the ^1H and 2.7 ppm in ^{15}N dimension, see table 3.1).

Table 3.1.: Chemical shifts of tryptophan side chain nuclei of *all-Ala*-HEWL.

Residue	$\delta(^1\text{H}^{\text{N}}_{\epsilon})$ [ppm]	$\delta(^{15}\text{N}_{\epsilon})$ [ppm]	$\delta(^{13}\text{C}_{\gamma})$ [ppm]
W28	10.109	130.62	110.36
W62	10.182	130.68	110.41
W63	10.108	130.76	110.22
W108	10.131	130.51	110.57
W111	10.141	130.72	110.23
W123	10.109	130.62	110.40

However, the dispersion of the six $^{13}\text{C}_{\gamma}$ peaks (less than 0.4 ppm, see table 3.1) was promising enough to use these nuclei as the connection between the assigned backbone resonances and the $^1\text{H}^{\text{N}}_{\epsilon}$ and $^{15}\text{N}_{\epsilon}$ indole resonances of the side chains. By expanding and adjusting the known HNCACB experiment to a HN(CACB)CG experiment, the $^{13}\text{C}_{\gamma}$ peaks have readily been assigned. In the left half of figure figure 3.9 on the facing page, the correlation of the tryptophan backbone amide peaks in a ^1H - ^{15}N -HSQC spectrum with the peaks in the $^1\text{H}^{\text{N}}$ - $^{13}\text{C}_{\gamma}$ plane of the HN(CACB)CG spectrum is indicated by blue lines.

The HN(CD)CG experiment is also a modification of the basic HNCACB experiment but has been optimized for the situation in tryptophan side chains. Figure 3.9 on the next page illustrates the correlation of the indole $^1\text{H}^{\text{N}}_{\epsilon}$ and $^{15}\text{N}_{\epsilon}$ peaks in a ^1H - ^{15}N -HSQC spectrum of the tryptophan side chain region with the $^1\text{H}^{\text{N}}$ - $^{13}\text{C}_{\gamma}$ plane of the HN(CD)CG spectrum. Their connection is indicated by red lines. Unfortu-

3.4. Assignment of tryptophan side chains

nately, the chemical shifts of W28 and W123 heavily overlap in the $^1\text{H}^{\text{N}}_{\epsilon}$ and the $^{15}\text{N}_{\epsilon}$ dimensions of the spectra and therefore cannot be resolved properly. The correlation of the $^{13}\text{C}_{\gamma}$ in the ^1H - ^{13}C planes of the HN(CACB)CG and the HN(CD)CG spectra is demonstrated by black lines.

The HN(CACB)CG and HN(CD)CG pulse sequence are shown in figures 2.5 and 2.6 on pages 51 and 52 and listed in the standard Bruker pulse sequence programming language in the appendix (section A.6.3 on page 139 and section A.6.4 on page 144).

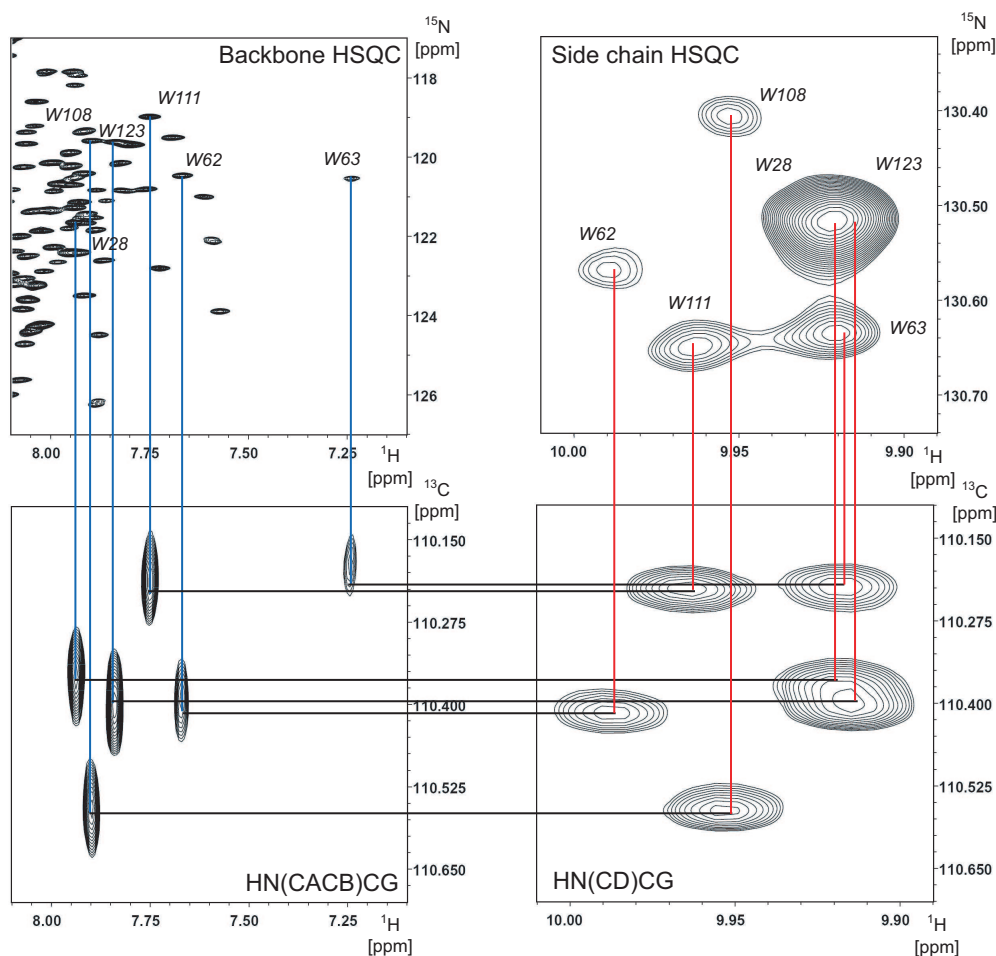


Figure 3.9.: Results of the assignment strategy for the tryptophan side chain indole $^1\text{H}/^{15}\text{N}_{\epsilon}$ resonances by correlation with the aromatic $^{13}\text{C}_{\gamma}$ resonances.

In principle, this set of two experiments is suitable for the assignment of tryptophan side chains in any protein. In most cases, like for *all-Ala*-HEWL, only the two two-dimensional planes mentioned above are needed. Together, the recording

3. Results and Discussion

of these two spectra took less than 11 hours of measurement time in the case of *all-Ala*-HEWL. However, in cases where peaks are overlapping in one dimension, the three-dimensional versions are also available and have been used to double-check and verify the assignments of the tryptophan side chains in *all-Ala*-HEWL.

The four tryptophan residues most involved in the hydrophobic clusters, namely W62, W63, W108 and W111 (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004), are well resolved and unambiguously assignable. Interestingly, the peaks of the tryptophans 62 and 108 in the ^1H - ^{15}N -HSQC spectrum of the side chain region are considerably broadened compared to the other peaks. This coincides with their crucial role in hydrophobic clustering in non-native lysozyme.

3.5. Photo-CIDNP experiments

3.5.1. Two-dimensional ^{15}N - ^1H photo-CIDNP

As pointed out in the introductory section 1.2.2 on page 14, photo-CIDNP experiments can be used to probe solvent accessibilities of aromatic side chains in proteins. Here, the results of a ^{15}N - ^1H photo-CIDNP experiment of the tryptophan side chain region of *all-Ala*-HEWL and its W62G mutant are described. The recorded ^1H - ^{15}N -HSQC spectra of the tryptophan side chain region of *all-Ala*-HEWL (see figure 3.10 on the next page) are very similar to the ^{15}N - ^1H photo-CIDNP spectra in terms of peak distribution and signal-to-noise ratio. The same is true for the comparison of the two spectra of the W62G mutant. The difference between the spectra of the *wild-type*-like *all-Ala*-HEWL and the W62G-*all-Ala*-HEWL are the disappearance of the W62 and W63 peaks. The absence of a W62 peak confirms the introduction of the mutation at this position. A closer look at the peak intensities of the spectra reveals that the W63 peak moved to the position of the W108 peak in the mutant. This is not unexpected, since W63 is right next to the mutated position and it is believed that the W62G mutation breaks the clusters. Therefore, the change of the chemical shift resulting in the overlapping with W108 is not surprising. Unfortunately, the W28 and W123 peaks overlap as well, preventing a separated treatment of these residues.

However, the relative intensity distribution among the individual peaks in the photo-CIDNP spectra is different from the intensity distribution in the HSQC spectra. This is due to the dependency of the CIDNP spectra on the accessibility of the tryptophan side chains to the photochemically activated FMN. In figure 3.11 on page 79 the intensity ratios of CIDNP versus HSQC peak intensities ($I_{\text{CIDNP}}/I_{\text{HSQC}}$) are plotted, where the highest CIDNP/HSQC ratio - resulting from the combined W28/W123 peak - is normalized to 1.0 and the other ratios are scaled accordingly. The peak intensities have been used for resolution purposes instead of the integrals. Where possible, the integration of the resolved peaks yielded comparable results. The errors for the peak intensities have been extracted from the signal-to-noise ratio of the spectra and the error-bars in figure 3.11 on page 79 have been calculated by error propagation from the intensity errors. The analysis of the CIDNP/HSQC ratio in terms of accessibility to the solvent is semi-quantitative as discussed by Morozova et al. (Morozova *et al.*, 2004). Only time-resolved CIDNP data would allow for a quantitative analysis.

The ratio for W111 in *all-Ala*-HEWL is only about half of the value for W28/W123 and the other tryptophan residues also exhibit significantly lower ratios. This agrees

3. Results and Discussion

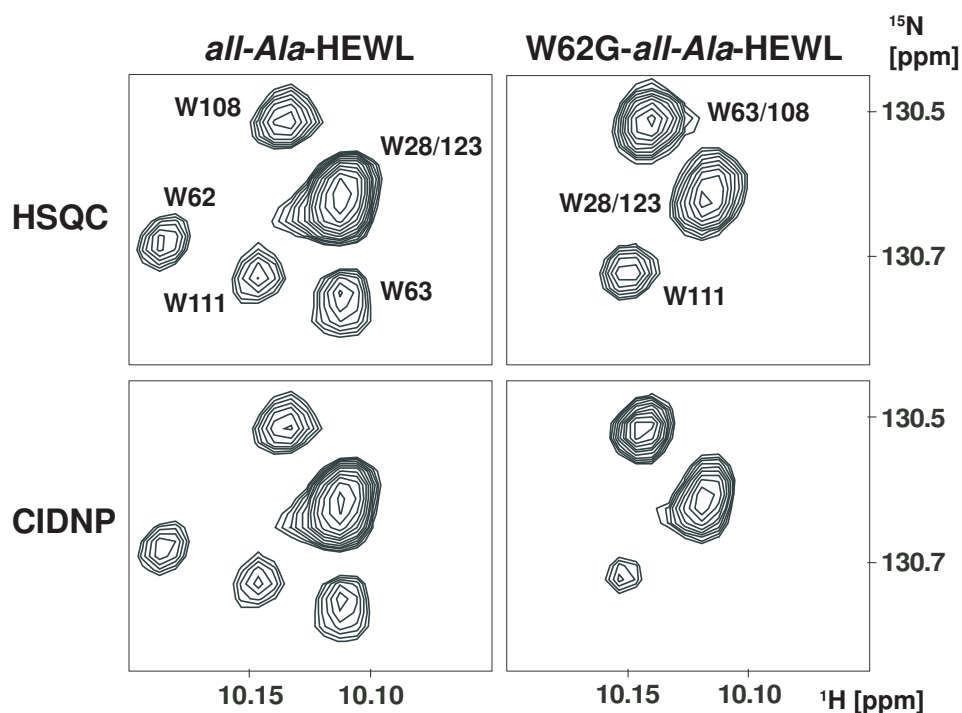


Figure 3.10. ^1H - ^{15}N -HSQC (upper row) and 2D photo-CIDNP (lower row) spectra of *all-Ala*-HEWL (left column) and *W62G-all-Ala*-HEWL (right column).

with the formation of hydrophobic clusters and long-range interactions seen in non-native HEWL (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004). The tryptophans involved in the most pronounced clusters are W62, W63 (designated cluster 3) and W111 (designated cluster 5) as judged from R_2 transverse relaxation rates. The structural findings from solvent accessibility information obtained by the two-dimensional photo-CIDNP studies therefore support the data on the dynamics (see section 1.3 on page 27 and section 3.6 on page 83). The CIDNP/HSQC ratio distribution in the *W62G* mutant of *all-Ala*-HEWL is different from the distribution in *all-Ala*-HEWL: Only W111 shows a similar ratio, while the other ratios are significantly elevated, suggesting a similar solvent accessibility for all tryptophans in the *W62G* mutant but for W111. This is in accordance with mutational studies (Wirmer *et al.*, 2004) where it has been shown that the *W62G* mutation abolishes most of the clustering in non-native lysozyme. In the same publication, it has been reported that the breakage of the long-range interactions by the *W62G* mutation results in an increase of 12 % in the measured *radius of hydration* (R_h), which reflects a shift towards more extended conformations. This, in turn, explains the lower relative accessibilities of the trypto-

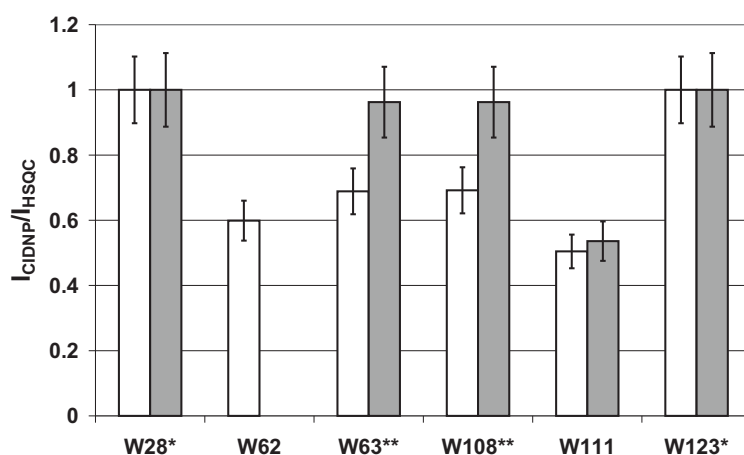


Figure 3.11.: CIDNP/HSQC normalized intensity ratios of *all-Ala-HEWL* (white) and *W62G-all-Ala-HEWL* (grey). *: The W28 and W123 peaks are degenerated in *all-Ala-HEWL* and *W62G-all-Ala-HEWL*. **: The W63 and W108 peaks are degenerated in *W62G-all-Ala-HEWL*.

phans involved in the hydrophobic clustering in the *wild-type-like all-Ala-HEWL*. Interestingly, the W62G mutation neither influences the accessibility of W111 to FMN nor does it change the relaxation properties of the backbone amide and side chain indole ^{15}N nuclei (see results in section 3.6 on page 83). This indicates that the cluster around W111 does not interact with the other clusters, despite its sequential proximity to W108. Thus, W108 and W111 may either be involved in two different clusters or the mutations disrupt interactions of only some of the participating side chains. In general, the applicability of two-dimensional photo-CIDNP experiments for the structural characterization of non-native states of a protein — even at low protein concentrations — has been proven with this work. In addition, the data from such experiments might help to test the results from molecular dynamics simulations of non-native states of proteins.

In the publication first reporting on the application of heteronuclear two-dimensional photo-CIDNP (Lyon *et al.*, 1999), hen egg white lysozyme has been investigated in its native and urea-unfolded states. In native HEWL, only W62 and W123 are accessible to the solvent (see figure 1.12 on page 28) and thus are visible in the photo-CIDNP spectra. The HSQC and the photo-CIDNP spectra of HEWL in 10 M urea look similar to the spectra of *all-Ala-HEWL* in terms of peak distribution and signal-to-noise. However, these spectra seem to be inverted along the ^{15}N axis, most likely due to improper processing of the data and no assignment was available for

3. Results and Discussion

the tryptophan side chains at the time of the publication. The experiments by Lyon et al. have been conducted at a higher temperature (318 vs. 303 K) and a higher pH value (3.6 vs. 2.0) than the ones in this thesis. Indeed, the photo-CIDNP effect was found to be very temperature dependent in the exploratory experiments of this thesis (the signal intensities have found to be more than 25 % higher at 303 K than at 293 K). Therefore, these experiments have been performed at a higher temperature than all other experiments described in this work. In contrast to the findings of Lyon et al. but in line with earlier publications (Tsentalovich *et al.*, 2002), the low pH maintained during the photo-CIDNP experiments did not seem to deteriorate the quality of the tryptophan region of the spectra. Furthermore, in contrast to the earlier studies cited above no addition of hydrogen peroxide to the samples was necessary to prolong the lifetime of the sample by the reoxidizing of reduced FMN. Rather, it was possible to record at least two two-dimensional spectra with 64 increments each, without the observation of any bleaching of the flavin. It should be stressed, that only a very moderate laser illumination (500 mW output power at the tip of the fiber and 50 ms duration of each pulse) had to be used for these experiments. However, no significant CIDNP enhancement has been observed in the 2D experiments in contrast to the one-dimensional CIDNP spectra.

3.5.2. Two-dimensional ^{13}C - ^1H photo-CIDNP

After successfully employing the ^{15}N - ^1H photo-CIDNP experiment on *all-Ala*-HEWL, the question arose, whether this technique was transferable to a ^{13}C - ^1H correlated two-dimensional photo-CIDNP experiment. The pulse scheme for this experiment is analogous to the one listed in figure 2.7 on page 53. In figure 3.12 ^1H - ^{13}C -HSQC and ^{13}C - ^1H 2D photo-CIDNP of the aromatic region are shown for *all-Ala*-HEWL. Neither in the HSQC nor in the photo-CIDNP spectrum, the peaks could be assigned to individual residues. Most likely, the respective tryptophan side chain peaks of the six tryptophans overlap. In addition, the complex multiplet pattern primarily caused by $^1\text{J}(\text{C},\text{C})$ couplings in the aromatic side chain also contributes to the complexity of the HSQC and photo-CIDNP spectra. However, from the characteristic chemical shifts taken from the BMRB (Seavey *et al.*, 1991) distinct types of nuclei have been assigned to groups of peaks where possible. The nomenclature of tryptophan side chains is illustrated in figure 3.13 on the following page.

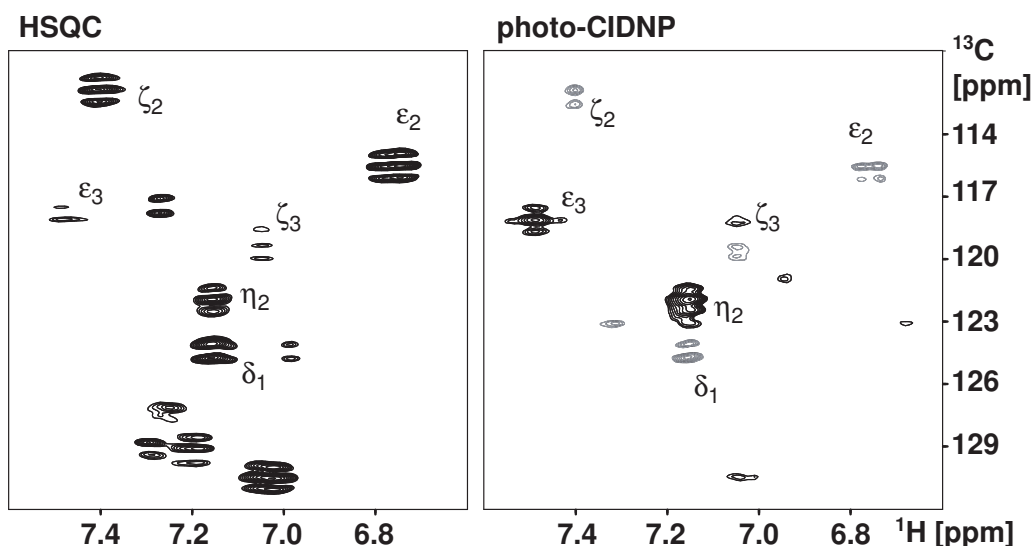


Figure 3.12.: ^1H - ^{13}C -HSQC (left) and ^{13}C - ^1H 2D photo-CIDNP (right) spectra of *all-Ala*-HEWL. Some peaks have been assigned to distinct side chain nuclei of tryptophan residues but not to individual tryptophan residues. Unassigned peaks in the HSQC spectrum may arise from other aromatic residues.

Absorptive (positive) signals are observed for the $^{13}\text{C}_{\epsilon_3}$ and $^{13}\text{C}_{\zeta_2}$ signals, whereas the $^{13}\text{C}_{\delta_1}$, $^{13}\text{C}_{\epsilon_2}$, $^{13}\text{C}_{\eta_2}$ and $^{13}\text{C}_{\eta_3}$ are emissive (negative). The signs of the photo-CIDNP peaks depend on the signs of the hyperfine coupling constant and whether the g-factor of the amino acid radical is larger or smaller than that of the dye radical

3. Results and Discussion

(Kaptein, 1971).

The two-dimensional ^{13}C - ^1H photo-CIDNP spectrum presented here is the first of its kind in solution NMR. The signal-to-noise ratio of this spectrum is very promising given the comparable low protein concentration and the short and weak laser irradiation. Unfortunately, the resolution of the peaks is not sufficient for the analysis of structural properties of the non-native HEWL, though an application to folded proteins, where the signal dispersion is better and only few residues are exposed to the solvent, might turn out to be very useful, e.g. to investigate ligand binding or to characterize surface accessibilities and contact surfaces of binding partners.

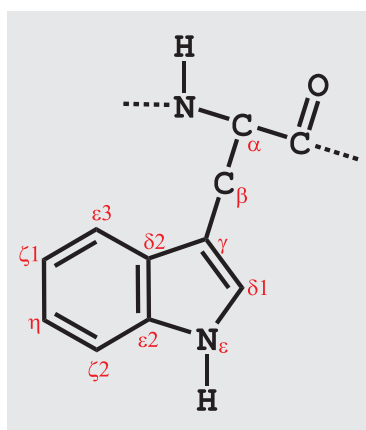


Figure 3.13.: Nomenclature of tryptophan side chains.

3.6. Protein dynamics by heteronuclear relaxation

Heteronuclear relaxation rates can report on the dynamics in proteins. They are routinely used to probe the flexibility of the peptide backbone. Most frequently the amide ^{15}N nuclei are utilized for this purpose. All R_1 , R_2 and $R_{1\rho}$ heteronuclear relaxation rates presented here for the backbone amides of *all-Ala*-HEWL, have been obtained by fitting peak heights to two-parameter single-exponential decay functions. This has been shown to be the most accurate way to determine these rates (Viles *et al.*, 2001). Particularly when overlapping peaks are present, the use of peak heights rather than peak volumes is preferable. The errors from the fits are negligible in general, given the spectra are of good quality. Therefore, the errors have been estimated to be $\pm 2\%$ in the case of the relaxation rates and $\pm 5\%$ for the heteronuclear NOEs. In figure 3.14 on the next page, the results for all relaxation rates mentioned above are combined and plotted over the sequence of *all-Ala*-HEWL. These relaxation profiles closely resemble the profiles for HEWL-S^{Me} in 8 M urea (Schwalbe *et al.*, 1997) and the R_2 profile is very similar to the profile of HEWL-S^{Me} in water (see figure 3.15 on page 85; Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004).

In this work R_2 rates have been determined for more residues than in the previous work for HEWL-S^{Me} and less scattering is observed. The R_2 rates for HEWL-S^{Me} are slightly higher than for *all-Ala*-HEWL, but this effect can most probably be attributed to a different temperature calibration during acquisition of the experiment. If the temperature in the previous experiments was to be lower by only a few degrees, the rates would be expected to be somewhat higher due to a lesser degree of mobility.

Due to significantly higher protein concentrations (3-4 mM) and/or the use of high concentrations of urea (8 M) in some of the earlier works (Buck *et al.*, 1995a; Grimshaw, 1999; Schwalbe *et al.*, 1997), the amplitudes of the R_2 relaxation rates determined therein for various non-native states of HEWL are not directly comparable to the data presented here.

The R_2 rate profile for *all-Ala*-HEWL has a baseline slightly above 2 s^{-1} and six more or less defined regions with elevated rates: Region 1 (around residue A10), region 2 (around W28 and A30), region 3 (around W62/W63), region 4 (around A82), region 5 (around W108/A110/W111) and region 6 (around W123/I124). Of these six regions, regions 3 and 5 are the most pronounced with peaks around 5 s^{-1} and 4 s^{-1} , respectively. The $R_{1\rho}$ rates are almost identical to the R_2 rates, indicating the absence of chemical exchange on the micro- to millisecond time-scale (Dayie and Wagner, 1994; Deverell *et al.*, 1970; Szyperski *et al.*, 1993).

3. Results and Discussion

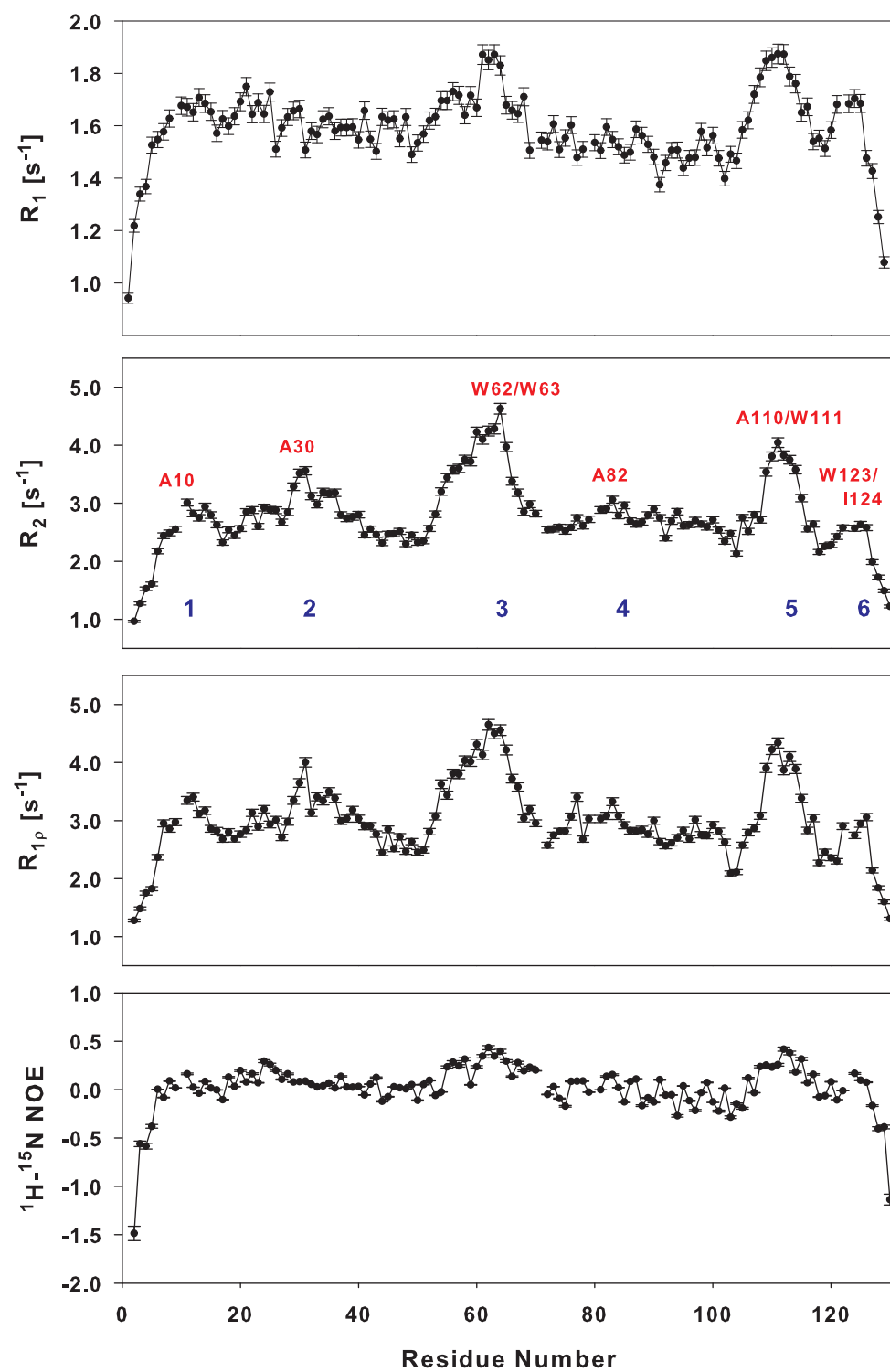


Figure 3.14.: Heteronuclear relaxation rates for *all-Ala*-HEWL

3.6. Protein dynamics by heteronuclear relaxation

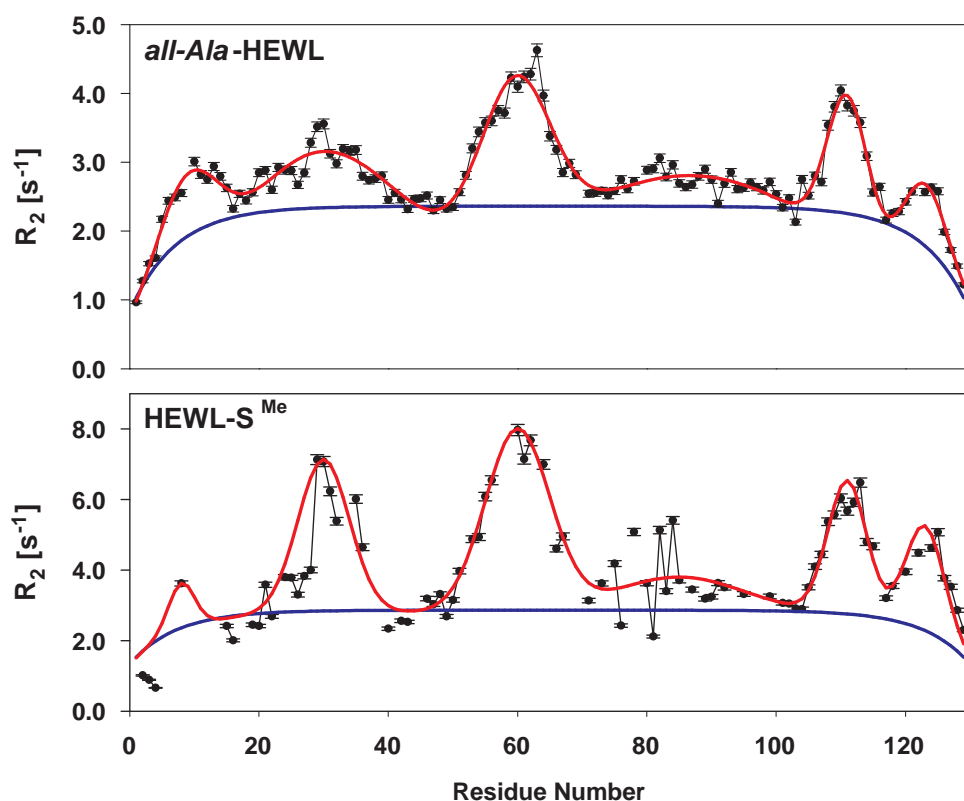


Figure 3.15.: Comparison of R_2 relaxation rates of *all-Ala*-HEWL and HEWL- S^{Me} . The baselines from the segmental motion model (see equation (3.1) on the next page) are given in blue, the Gaussian fits are indicated by red curves (data for HEWL- S^{Me} are taken from Klein-Seetharaman *et al.*, 2002).

The R_1 rates and the steady-state heteronuclear ^1H - ^{15}N -NOE in principal follow the profile of the R_2 rates but with much less pronounced amplitudes for they are less sensible to motional restrictions. However, the hetNOE values are close to zero, which indicates a general flexibility of the chain. For rigid parts in folded proteins, higher values are to be expected.

However, the analysis of the R_1 and R_2 relaxation rates together with the heteronuclear ^1H - ^{15}N -NOEs using the *model-free* approach (Lipari and Szabo, 1982a,b) to calculate the order parameters along the sequence is not possible for *all-Ala*-HEWL, as the assumption of the separability of global and internal motions is not true for non-native proteins (also see page 26 in section 1.2.2). Nevertheless, elevated relaxation rates can be interpreted in terms of decreased flexibility and the regions of raised relaxation rates in the sequence of *all-Ala*-HEWL in figure 3.14 correspond

3. Results and Discussion

to the hydrophobic clusters already identified for HEWL-S^{Me} in previous studies (Klein-Seetharaman *et al.*, 2002; Wirmer *et al.*, 2004). The R_2 rates of *all-Ala*-HEWL and HEWL-S^{Me} in figure 3.15 on the preceding page have been fitted to a Gaussian equation (equation (3.2)), which includes a term for the hydrophobic clusters (red line in figure 3.15).

$$R_2^{rc}(i) = R_{int} \sum_{j=i}^N e^{-\frac{|i-j|}{\lambda_0}} \quad (3.1)$$

$$R_2^{rc}(i) = R_{int} \sum_{j=i}^N e^{-\frac{|i-j|}{\lambda_0}} + \sum_{cluster} R_{cluster} e^{-\frac{1}{2} \left(\frac{i-\chi_{cluster}}{\lambda_{cluster}} \right)^2} \quad (3.2)$$

For a *random coil* peptide chain without any tertiary contacts, the relaxation profile is expected to follow a *segmental motion* model (Schwalbe *et al.*, 1997), which is based on equation (3.1) and gives a baseline (blue line in figure 3.15) for the relaxation rates. The *segmental motion* model has been derived from polymer theory and assumes that the relaxation properties of a given amide (i) are only determined by segmental motions of parts of the polypeptide chain, independent from the overall tumbling of the molecule and the type of the neighboring residues (j). In equation (3.1), R_{int} is the intrinsic relaxation rate, that depends on the temperature and viscosity of the solution, λ_0 is the persistence length of the polypeptide chain and N is the total number of residues. Equation (3.2) takes the presence of hydrophobic clusters into account, with $\chi_{cluster}$ as the position of each cluster in the chain, $\lambda_{cluster}$ as the width of each cluster and $R_{cluster}$ as the distinct relaxation rate for each cluster.

As mentioned above, the $R_{1\rho}$ experiment did not reveal any significant dynamics in *all-Ala*-HEWL on slower (micro- to millisecond) timescales. The same is true for the CPMG relaxation dispersion measurements conducted. Figure 3.17 on page 89 exemplarily shows the dispersion curves (effective R_2 rate as a function of the CPMG frequency) for six residues, including the hydrophobic residues in the most pronounced clusters. The dispersion curves for all other analyzed residues gave the same picture. If conformational exchange is present, the dispersion curves should show significant dependence on the CPMG field.

As discussed in section 1.3 on page 31 and shown in figure 1.14 on page 30, various single-point mutants have been shown to significantly alter the relaxation profile of HEWL-S^{Me} (Wirmer *et al.*, 2004). The most pronounced effect has been observed for the exchange of the tryptophan at position 62 for a glycine (W62G), where in all six clusters the relaxation rates are reduced and the central cluster 3 (around the positions G(W)62 and W63) completely disappears. In the W111G mutant, clusters

3.6. Protein dynamics by heteronuclear relaxation

2 (around W28) and 3 are weekendened, while clusters 5 (W108/W111) and 6 (W123) completely disappear. In the W123G mutant, clusters 2, 3 and 5 exhibit decreased relaxation rates, while the C-terminal cluster six around the mutation vanishes (the numbering of the clusters is also shown in figure 3.16). Taken together, these data suggest a crosstalk between these clusters, since effects in one hydrophobic cluster are translated to the other clusters. Interestingly, the W108G mutant of *all-Ala*-HEWL produced in the course of this thesis, shows a relaxation profile very similar to that of the wild-type HEWL-S^{Me} and *all-Ala*-HEWL, except for cluster 5, which is modulated such that only residues N106 and A107 show increased relaxation rates figure 3.16. This suggests, that the tryptophan 108 might not be as critical as W62 or W111 for the overall hydrophobic clustering in non-native HEWL.

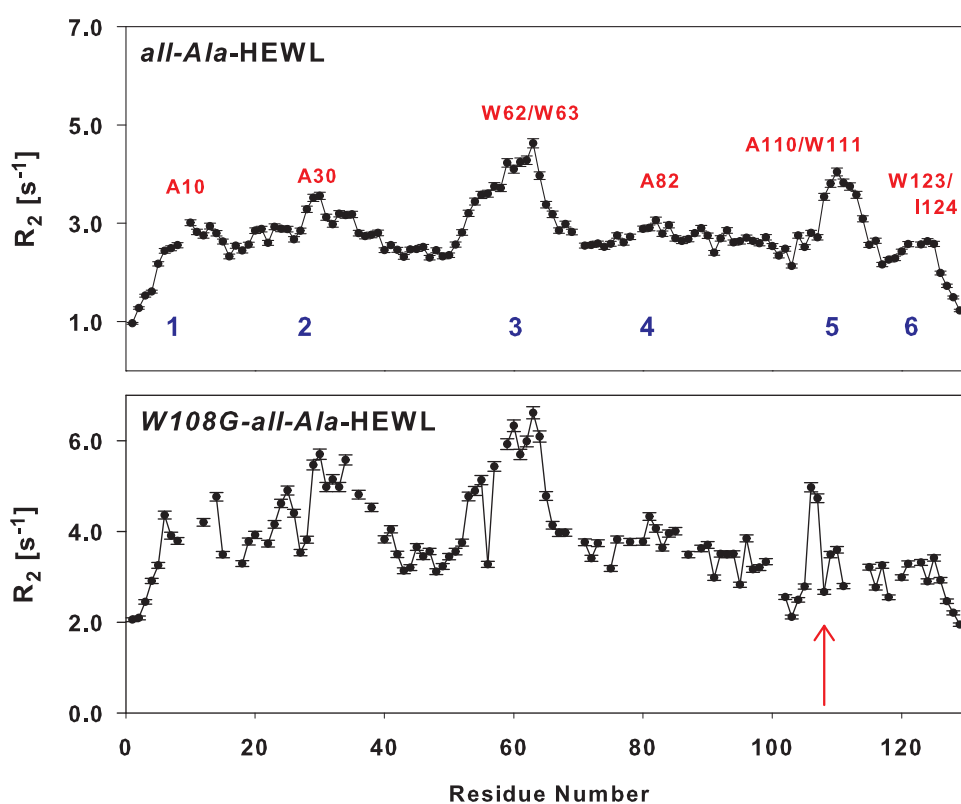


Figure 3.16.: Comparison of R_2 relaxation rates of *all-Ala*-HEWL and W108G-*all-Ala*-HEWL. The location of the single point mutation is indicated by a red arrow.

At the tryptophan side chain level, ^{15}N - R_2 relaxation rates have been determined for *all-Ala*-HEWL and its W62G mutant. Together with the assignment of the tryptophan side chain resonances (see section 2.3 on page 46), this gives a more detailed

3. Results and Discussion

picture. Table 3.2 lists these rates for the six tryptophan residues. Due to the overlap of the tryptophan 28 and 123 peaks, only a combined rate is given for these residues. W62 and W63 show significantly higher rates than the other tryptophans, the rate for W111 is only slightly higher. In the W62G single-point mutant, the rates are more uniform, with the rate of W63 being clearly smaller than in the wild-type like *all-Ala*-HEWL. Backbone and side chain data nicely converge at this point. In general, the rates for the tryptophan side chains are in the order of 2 to 3 s⁻¹, and thus on average around 1 s⁻¹ smaller than the rates of the respective backbone signals. This might be due to a higher flexibility in the side chains and differences in the NH bond lengths. Unfortunately, no reliable data is available to directly compare these bond lengths, although in general XH bond lengths in aromatic systems are smaller than in aliphatic systems (Lide, 1996). However, following equation (1.11) on page 24, a smaller NH bond distance would result in a larger rate.

Table 3.2.: Heteronuclear ¹⁵N transverse relaxation rates in the tryptophan side chains of *all-Ala*-HEWL and W62G-*all-Ala*-HEWL.

Residue	R ₂ in [s ⁻¹]	
	<i>all-Ala</i> -HEWL	W62G- <i>all-Ala</i> -HEWL
W28	2.3 ± 0.1	2.0 ± 0.1
W62	3.0 ± 0.1	—
W63	3.1 ± 0.1	2.2 ± 0.1
W108	2.4 ± 0.1	2.2 ± 0.1
W111	2.6 ± 0.1	2.4 ± 0.1
W123	2.3 ± 0.1	2.0 ± 0.1

3.6. Protein dynamics by heteronuclear relaxation

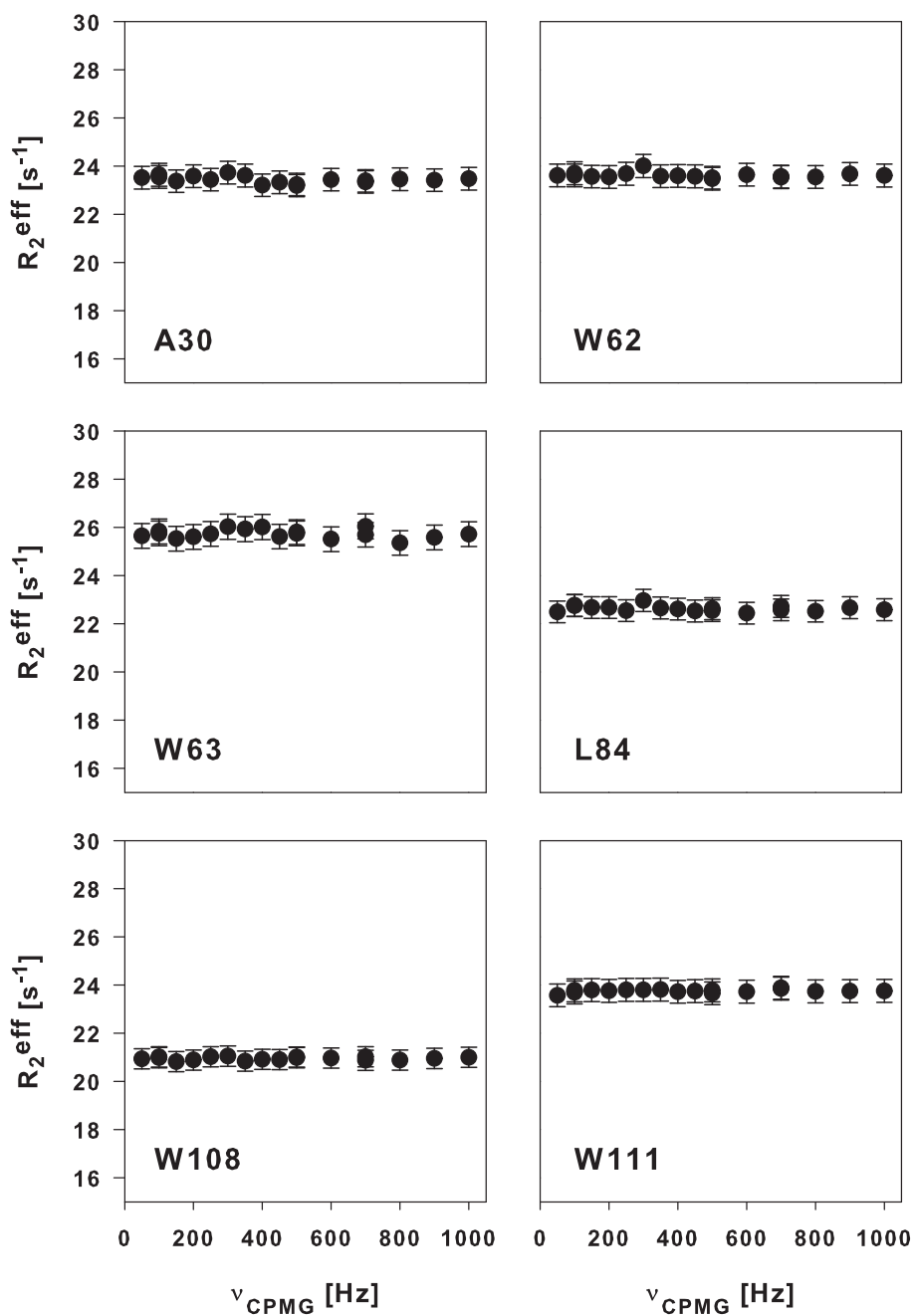


Figure 3.17.: ¹⁵N single-quantum relaxation rates as a function of CPMG frequency for various residues of *all-Ala*-HEWL.

3.7. Residual dipolar couplings

In order to further characterize the structural aspects of the non-native ensemble of *all-Ala*-HEWL in water at pH 2, NH-RDCs have been determined for this protein in stretched 7 % polyacrylamide gels. The preparation of these aligning gels is despite the availability of a convenient casting apparatus (see *Materials and Methods*, page 58), not a routine procedure and multiple trials were needed in order to produce properly aligned samples. The integrity of the gels in the NMR tube and the process of the gel transfer into the tubes, was found to be the limiting step. The proper shimming of these samples in the magnet was challenging, but the suppression of natural-abundance gel signals in the two-dimensional ^1H - ^{15}N -HSQC and IPAP spectra was not critical, for these peaks did not overlap with the signals in the amide region of the protein. The alignment of the samples was checked with one-dimensional ^2H spectra, in which the ^2H quadrupolar splitting in the HDO signals can be detected. Typically, splittings in the range of 5 to 20 Hz have been observed (figure 3.18 gives an example). The NH-RDCs have been calculated from the dif-

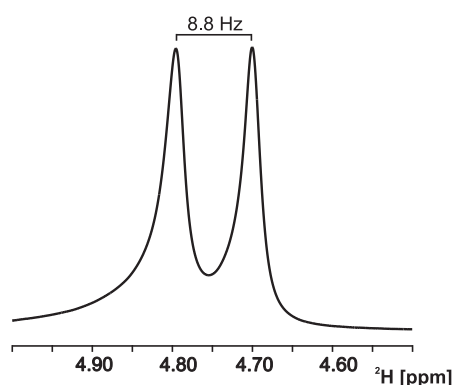


Figure 3.18.: Quadrupolar splitting in HDO in a ^2H 1D spectrum of a stretched 7 % polyacrylamide gel to assess the degree of alignment of the sample.

ference of the couplings in the isotropic medium (water pH 2) and the anisotropic medium (7 % PAG, pH 2), with the latter being the sum of the scalar $^1\text{J}(\text{H}^{\text{N}},\text{N}^{\text{H}})$ and the dipolar $^1\text{D}(\text{H}^{\text{N}},\text{N}^{\text{H}})$ couplings. The couplings themselves have been determined in the ^{15}N dimension by picking the two respective peaks in the ^1H - ^{15}N -IPAP experiments. The final resolution of the processed spectra was 0.4 Hz. For 95 of the 125 assigned residues in *all-Ala*-HEWL unambiguous NH-RDCs have been extracted, figure 3.19 on the facing page shows the occurrence distribution of their values. In addition to the bell-shaped distribution of the bulk RDCs around -2 Hz, a number

of positive RDCs are present. When plotted over the amino acid sequence of HEWL (upper panel in figure 3.20 on the next page), most of these positive RDCs are located in the regions, where hydrophobic clustering had been identified by the relaxation experiments. Most pronounced is the occurrence of these positive RDCs in the clusters 2, 3 and 5, which are centered around W28, W62/W63 and W108/W111, respectively. Large negative RDCs are observed for the residues preceding the two prolines in HEWL. Similarly, NH-RDCs have been determined for the W62G mutant of *all-Ala*-HEWL (lower panel in figure 3.20). The overall RDC profile plotted over the sequence is very similar to the profile found for *all-Ala*-HEWL. However, in W62G positive RDCs are only observed in the regions identified as clusters 5 and 6 at the C-terminus of the protein. The NH-RDC profile for non-native HEWL in no

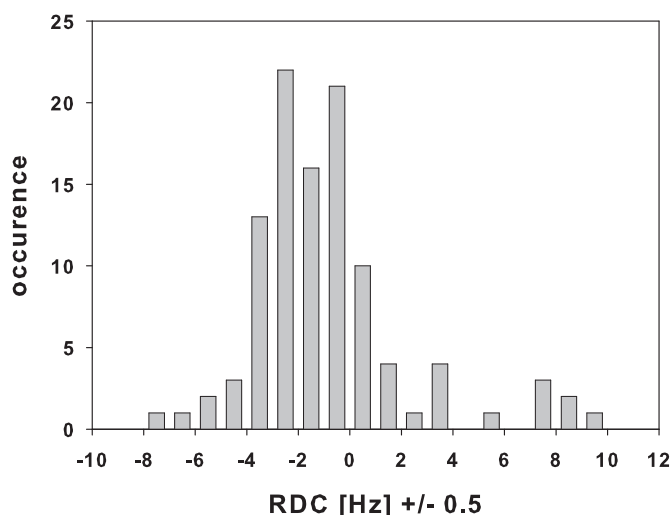


Figure 3.19.: Occurrence of distinct NH-RDC-values in *all-Ala*-HEWL aligned in a 7 % polyacrylamide gel. The values have been rounded up or down to the nearest whole number.

way resembles the profile determined for the native protein in two different bicellar alignment media (figure 3.21 on page 93, data taken from Schwalbe *et al.* (2001)), which is not surprising, since no considerable native-like structural motifs are expected to be present in non-native HEWL and the alignment tensor in the different media should be independent. As for the other NMR parameters, in order to interpret the residual dipolar couplings in unfolded proteins, the measured values may be compared to predicted values based on models. The observed RDCs are averaged over the whole ensemble of interconverting conformers. Therefore, in simula-

3. Results and Discussion

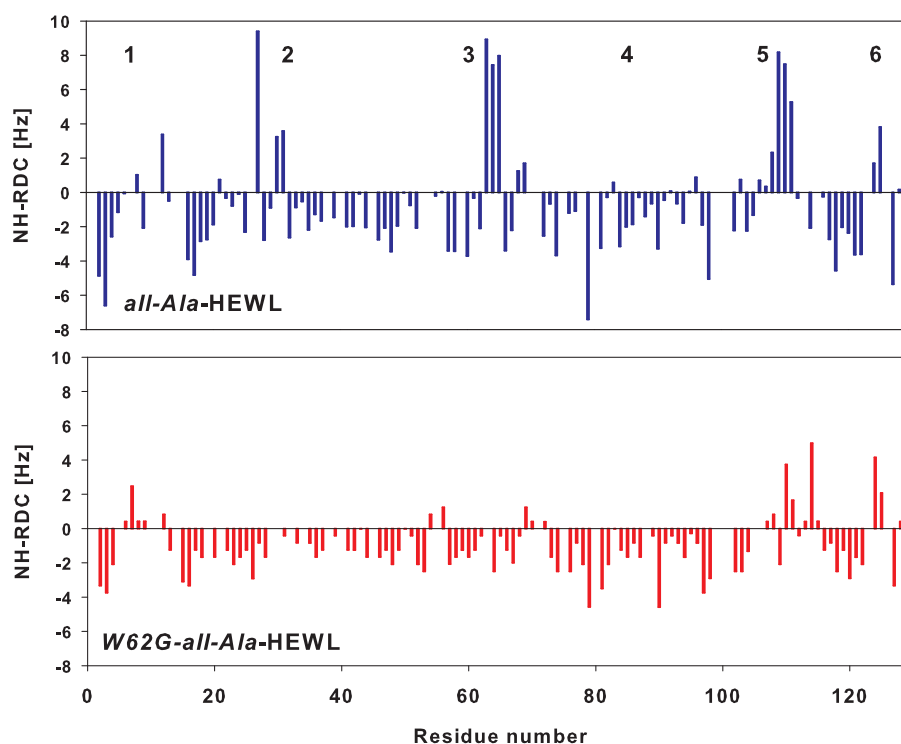


Figure 3.20.: NH-RDCs plotted over the sequence of *all-Ala-HEWL* and the W62G single point mutant aligned in a 7 % polyacrylamide gel. The position of the six previously identified hydrophobic clusters is indicated by their numbers.

tion approaches such ensembles have to be created and averaged residual dipolar couplings have to be calculated. As described in section 2.6.2 on page 59, the *flexible-meccano* algorithm (Bernadó *et al.*, 2005a,b) can be used for this purpose. In contrast to smooth bell-like distributions of RDCs that are expected for ideal *random flight* chains (Louhivuori *et al.*, 2003; Obolensky *et al.*, 2007), the *flexible-meccano* approach includes ϕ, ψ backbone angle propensities extracted from databases of amino-acid specific conformations occurring in loop regions of proteins. When applied for the amino acid sequence of HEWL, the *flexible-meccano* algorithm yields an RDC distribution as shown in figure 3.22A, which, however, does not resemble the experimental data. This is not surprising, since no long-range contact information is included in this calculations. Therefore, distance restraints have been implemented into the *flexible-meccano* algorithm, in order to generate more realistic ensembles of conformations. Figure 3.22B-D shows the results for single distance restraints between two sites at a time: A 15 Å contact restraint between clusters 2 (around W28) and

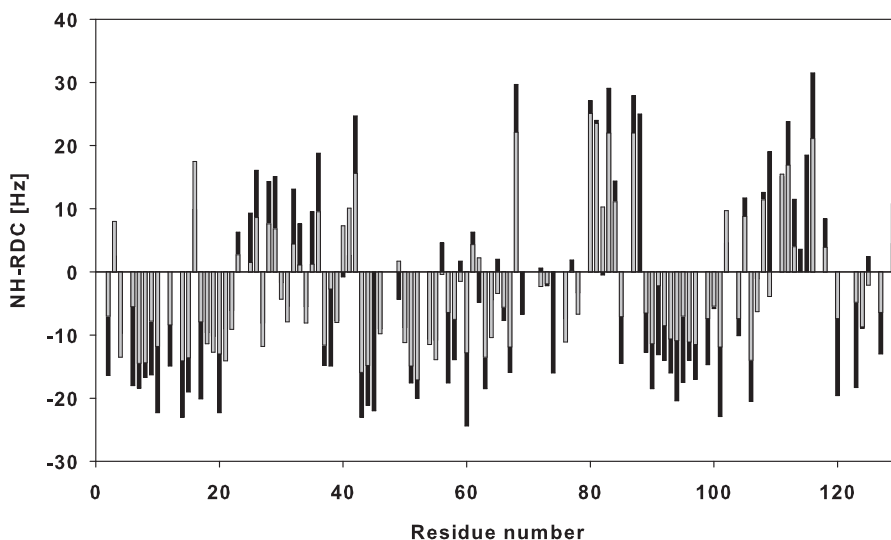


Figure 3.21.: NH-RDCs in native HEWL aligned in 5 % DMPC:DHPC (black bars) and 7.5 % DMPC: DHPC:CTAB (grey bars) (data taken from Schwalbe *et al.*, 2001).

3 (around W62/W63) (B), a 15 Å contact restraint between clusters 2 and 5 (around W108/W111) (C) and an 10 Å contact restraint between clusters 3 and 5 (D). All simulations are averaged over 50,000 conformers to allow for convergence. None of these simulations runs resulted in a relaxation profile similar to the experimental data and no significant positive RDCs have been predicted at all. The introduction of a single contact between two clusters at a time consequently leads to the introduction of a turn and a kink in the RDC profiles halfway between the respective contacting clusters. The effort to include multiple contacts between clusters at a time, was not feasible, since the generation of such ensembles takes too much computer time. An attempt to restrict the ϕ, ψ backbone angles of the tryptophan-tryptophan motif in the central hydrophobic cluster to values extracted from a database in order to generate ensembles which resemble the non-native *all-Ala*-HEWL, did not give any useful results. The residual structure in *all-Ala*-HEWL seems to be less easy to simulate than for other non-native or natively unfolded proteins, for which long-range contacts (Bernadó *et al.*, 2005a) or the presence of β and β -turn structural motifs (Mukrasch *et al.*, 2007a,b) have been picked up by the *flexible-meccano* approach.

3. Results and Discussion

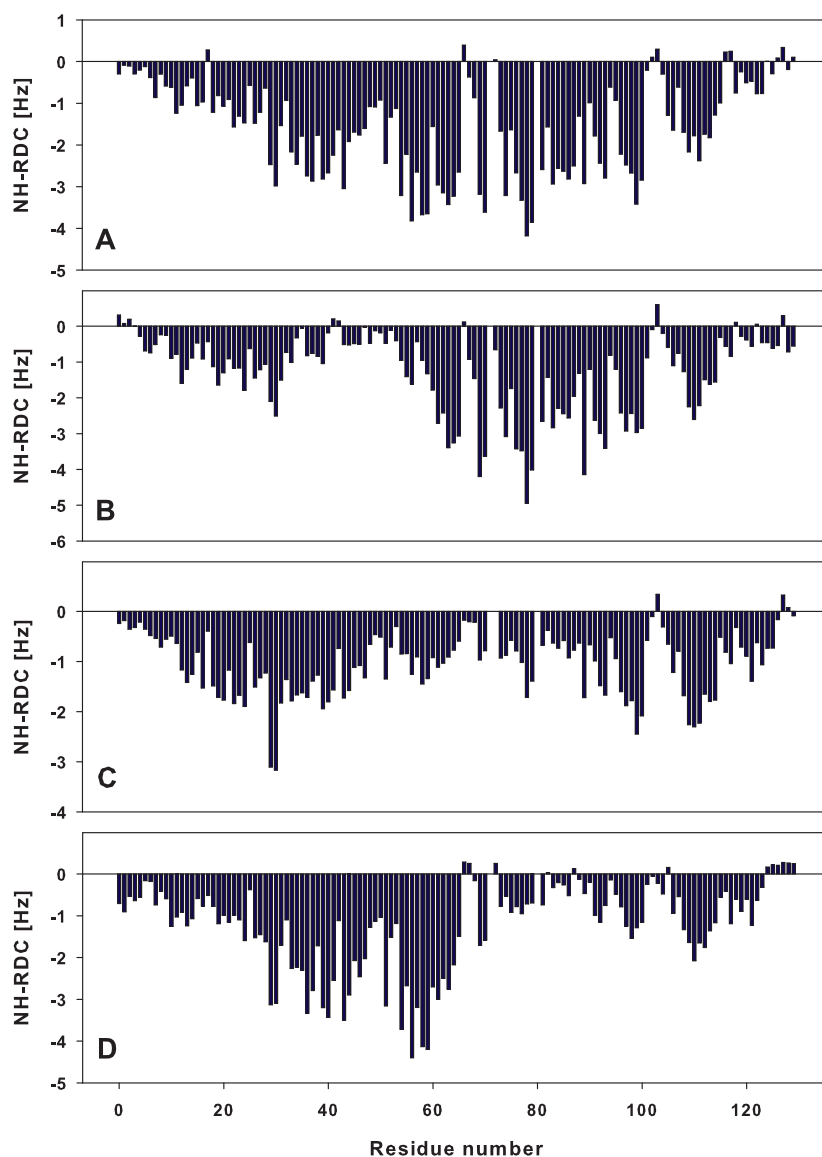


Figure 3.22.: Simulated NH-RDCs from *flexible-meccano*. A: no further restraints, B: with 15 Å contact restraint between clusters 2 and 3, C: with 15 Å contact restraint between clusters 2 and 5, D: with 10 Å contact restraint between clusters 3 and 5. All simulations averaged over 50,000 conformers

Instead, another way of generating ensembles of non-native conformers of HEWL has to be thought of in order to simulate RDCs that match the experimental data not only for the wild-type like *all-Ala*-HEWL but also for its single point mutants. The most promising way to interpret the experimental RDCs for *all-Ala*-HEWL, seems to

3.7. Residual dipolar couplings

be the calculation of RDCs from ensembles of conformers generated using molecular dynamics simulations with restraints from other experimental data, such as relaxation rates in combination with mutational studies, the solvent accessibilities from the photo-CIDNP experiments or the distance information from PREs (Argirevic, 2006).

3.8. Residual structure in organic solvents

Organic solvents such as ethanol and 2,2,2-trifluoroethanol (TFE) are known to induce specific secondary structures in proteins and in some cases induce the formation of amyloid fibrils. The structural changes of HEWL-S^{Me} and *all-Ala*-HEWL induced in the presence of various concentrations of ethanol and TFE in the solvent have been studied here by circular dichroism (CD) and nuclear magnetic resonance (NMR) spectroscopy. In CD spectroscopy, the ellipticity of originally circularly polarized light is measured after it has been partially absorbed by the sample (Kelly and Price, 2000). The peptide bonds are the main chromophores in far-UV CD spectroscopy (240 to 180 nm). Different secondary structure motifs, such as α -helices and β -sheets, absorb the light differently and therefore yield characteristic CD spectra (Johnson, 1990). The spectra of real proteins, native or (partially) unfolded, can be interpreted as combinations of spectra of the different possible secondary structure elements, including *random coil*. In addition, ellipticities at distinct wavelengths can be used to estimate the fraction of a certain secondary structure in a protein. The ellipticity at 222 nm is particularly characteristic for α -helices (Kelly *et al.*, 2005) and has been used in this work to characterize non-native lysozyme. The formulae used to convert the measured ellipticities into α -helical fractions (Rohl and Baldwin, 1997) are discussed in section 2.7 on page 61. However, the values derived from single wavelengths only allow for a rough estimate of secondary structure content.

In figure 3.23 on the next page, the CD spectra of native HEWL and HEWL-S^{Me} are compared. As expected, the spectrum of the non-native protein, compared to the native state, is shifted towards lower wavelengths, resembling a mostly *random coil* spectrum with only about 14 % α -helical content (see table 3.3 on page 98). The native protein in turn is mostly in an α -helical conformation, which is in complete agreement with the available three-dimensional structures of this enzyme (see e.g. figure 1.12 on page 28). With roughly 18 %, the fractional helicity of *all-Ala*-HEWL is very similar to the one in HEWL-S^{Me}, given the limited accuracy of this method.

Figure 3.24 on page 99 shows the far-UV CD spectra of HEWL-S^{Me} in solutions containing 0 %, 10 %, 30 % and 50 % 2,2,2-trifluoroethanol (TFE). An increase in the concentration of the organic solvent leads to a left shift in the minima of the spectra, indicating an increase in α -helicity. The calculated helicity values (table 3.3 on page 98) for 10 - 50 % TFE in the solution are all around 30 %, suggesting, that a small fraction of TFE is sufficient to induce helix formation.

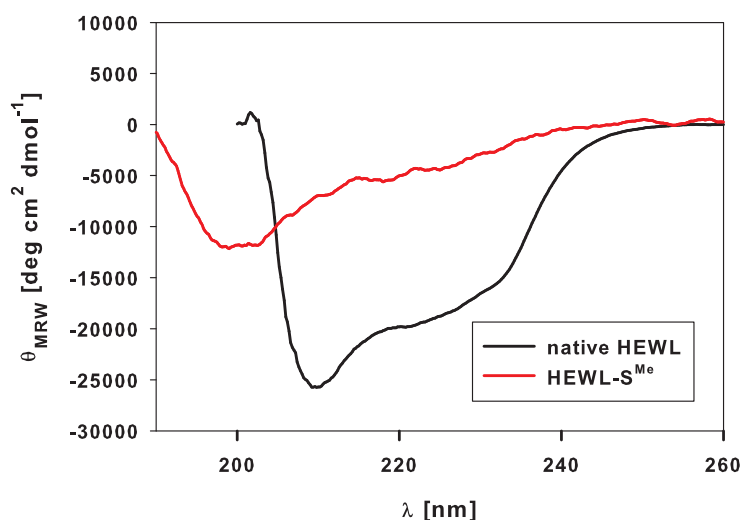


Figure 3.23.: Circular dichroism spectra of HEWL-S^{Me} and native HEWL. θ_{MRW} : mean residue molar ellipticity in [$\text{deg cm}^2 \text{dmol}^{-1}$]; λ : wavelength in [nm].

Figure 3.25 on page 100 shows the far-UV CD spectra of HEWL-S^{Me} in solutions containing 0 %, 30 %, 60 % and 90 % ethanol (EtOH). Increasing concentrations of ethanol lead to increasing helical fractions, ranging from around 30 % helicity at 30 % EtOH to nearly 60 % helical content at 90 % EtOH. Interestingly, the samples containing 50 % TFE also reach an α -helical content of around 60 % after incubation for several hours (see table 3.3 on the next page). The helicity of HEWL-S^{Me} in 90 % EtOH is nearly constant for at least 30 hours. After 4 days, the helicity is decreasing to under 50 %. However, the related spectra (figure 3.26 on page 101) show a shift to higher β -sheet-like fractions after several hours of incubation, indicated by spectral minima at around 215 nm, in contrast to the α -helical double minima at 208 and 222 nm. Most interestingly, a titration of the sample containing 90 % ethanol with NaOH to a pH of 10, lead to spectra (figure 3.27 on page 102) indicating a very pronounced transition to β -sheet-like structures which at the same time is accompanied by a dramatic decrease of the α -helical fraction.

Figure 3.28 on page 102 shows a ^1H - ^{15}N -HSQC spectrum of *all-Ala*-HEWL in water pH 2, while figure 3.29 on page 103 shows spectra of the same protein in the presence of 30 % and 60 % EtOH and figure 3.30 on page 103 in the presence of 90 % EtOH and 50 % TFE, respectively. With increasing concentrations of the organic solvents, the peaks get broader, most dramatically at 90 % EtOH, where some peaks are already broadened beyond detection. More strikingly, the signal dispersion es-

3. Results and Discussion

Table 3.3.: Fractional helicity for various samples of HEWL as calculated from the molecular ellipticity at 222 nm.

Sample	pH	c _{org} ^a	t _{inc} ^b	f _H ^c
native HEWL	2.0	–	–	53.0 %
HEWL-S ^{Me}	2.0	–	–	13.9 %
<i>all-Ala</i> -HEWL	2.0	–	–	17.5 %
HEWL-S ^{Me}	2.0	10 % TFE	–	27.4 %
HEWL-S ^{Me}	2.0	30 % TFE	–	39.3 %
HEWL-S ^{Me}	2.0	50 % TFE	–	32.0 %
HEWL-S ^{Me}	2.0	30 % EtOH	–	25.9 %
HEWL-S ^{Me}	2.0	60 % EtOH	–	45.0 %
HEWL-S ^{Me}	2.0	90 % EtOH	–	59.7 %
HEWL-S ^{Me}	10.0	90 % EtOH	–	27.5 %
HEWL-S ^{Me}	2.0	50 % TFE	23 h	55.7 %
HEWL-S ^{Me}	2.0	50 % TFE	30 h	56.7 %
HEWL-S ^{Me}	2.0	50 % TFE	96 h	56.6 %
HEWL-S ^{Me}	2.0	90 % EtOH	5 h	64.8 %
HEWL-S ^{Me}	2.0	90 % EtOH	23 h	54.8 %
HEWL-S ^{Me}	2.0	90 % EtOH	30 h	54.2 %
HEWL-S ^{Me}	2.0	90 % EtOH	96 h	47.5 %

^aconcentration of organic solvent

^btime of incubation with organic solvent

^cFractional helicity

pecially in the ¹H dimension of the spectra is increasing with the concentration of the organic solvent, indicating the induction of additional structure. Moreover, the peak pattern changes in a way, that an unambiguous identification of specific peaks is almost unfeasible in most cases already at 30 % EtOH. Interestingly, the peak distribution in the spectrum of *all-Ala*-HEWL in TFE clearly looks different from the spectra of this protein in water and in the various percentages of ethanol.

The line broadening in the presence of organic solvents cannot completely be attributed to slower tumbling due to the higher viscosity of these solutions, since the viscosity of azeotropic ethanol-water mixtures has a maximum at around 50 % EtOH (Lide, 1996; Wensink *et al.*, 2003) and therefore line broadening should be maximal in this range and not at 90 % as observed in the experiment. Instead, fast chemi-

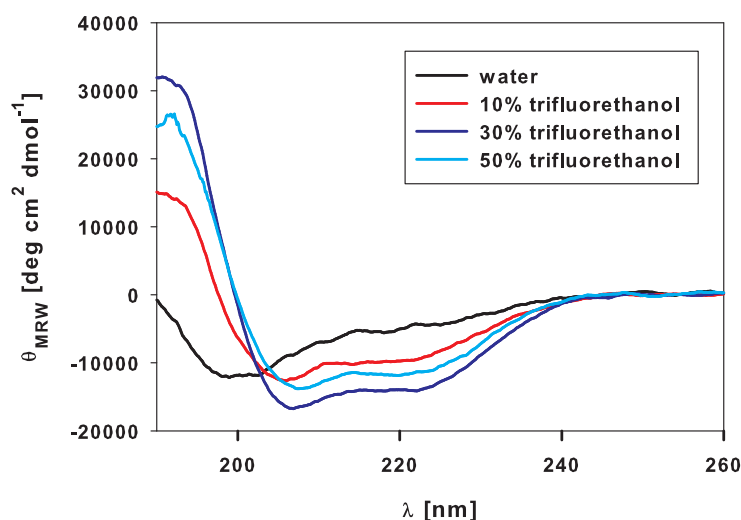


Figure 3.24.: Circular dichroism spectra of HEWL-S^{Me} in the presence of various concentrations of trifluoroethanol. θ_{MRW} : mean residue molar ellipticity in [deg cm² dmol⁻¹]; λ : wavelength in [nm].

cal exchange between different — mostly α -helical — conformers must be the most important contributing factor.

NMR spectroscopy has been used in numerous other studies to elucidate the influence of organic solvents on structural and dynamical properties of HEWL in its oxidized — disulfides intact — form. It has been shown by $^1\text{H}_\alpha$ chemical shift perturbations, NOEs, $^3\text{J}(\text{H}^N, \text{H}_\alpha)$ scalar couplings, hydrogen exchange protection and ^{15}N R_2 relaxation rates that HEWL in 50-70 % TFE is partially folded with more α -helical content than in the native protein, but exhibits few tertiary interactions (Buck *et al.*, 1993, 1995b, 1996; Povey *et al.*, 2007). The refolding kinetics of HEWL from the state in 60 % TFE are similar to the refolding kinetics from 6 M guanidine hydrochloride (GdnHCl), but low concentrations of TFE present during refolding accelerate the process (Lu *et al.*, 1997). In contrast to TFE, unfolding in GdnHCl results in states with very little secondary and tertiary structure present. In 10-50 % dimethylsulfoxide (DMSO), HEWL forms another intermediate state with residual structure present (Bhattacharjya and Balaram, 1997). The effects of different organic solvents to HEWL are compared in a paper by Knubovets *et al.* (1999), suggesting that the native-like structure retention in HEWL is favored in very hydrophilic solvents with strong hydrogen bonding propensities.

The effect of induced α -helical structure in HEWL in the presence of TFE, has also

3. Results and Discussion

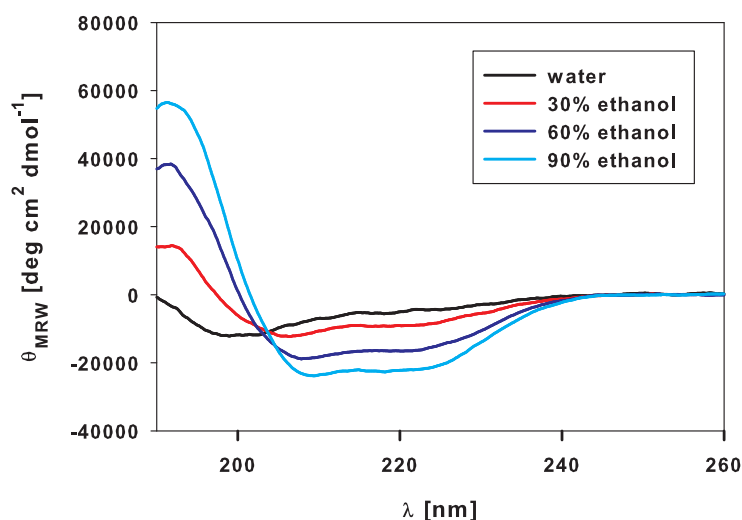


Figure 3.25.: Circular dichroism spectra of HEWL-S^{Me} in the presence of various concentrations of ethanol. θ_{MRW} : mean residue molar ellipticity in [$\text{deg cm}^2 \text{ dmol}^{-1}$]; λ : wavelength in [nm].

been observed by CD spectroscopy and small angle X-ray scattering (Hoshino *et al.*, 1997). Stopped-flow CD and differential scanning calorimetry (DSC) have been used to characterize the folding transitions in the presence of various organic solvents (Lai *et al.*, 2000).

The induction of α -helical structure by TFE in reduced HEWL has also been shown using CD spectroscopy by Prabha and Rao (2004), who in addition found that TFE does not facilitate the oxidative refolding of this protein.

The influence of TFE and other organic solvents on the secondary and tertiary structure has also been shown by NMR for other proteins and peptides (Fan *et al.*, 1993; Jaravine *et al.*, 2001; Live *et al.*, 1984; Thanabal *et al.*, 1994). These findings are all in line with the results in this thesis and the articles cited above.

In addition, it has been shown in the literature, that hen egg white lysozyme can form amyloid fibrils at low pH in the presence of organic solvents when incubated for several hours at high temperatures or several days to weeks at room temperature (Goda *et al.*, 2000; Krebs *et al.*, 2000).

However, the investigation of the formation of amyloid fibrils was not subject of this thesis, since this topic has been investigated in-depth in the references cited above. Similar projects on the fibril formation starting from non-native states of proteins — especially variants of the prion protein and lysozyme — are currently

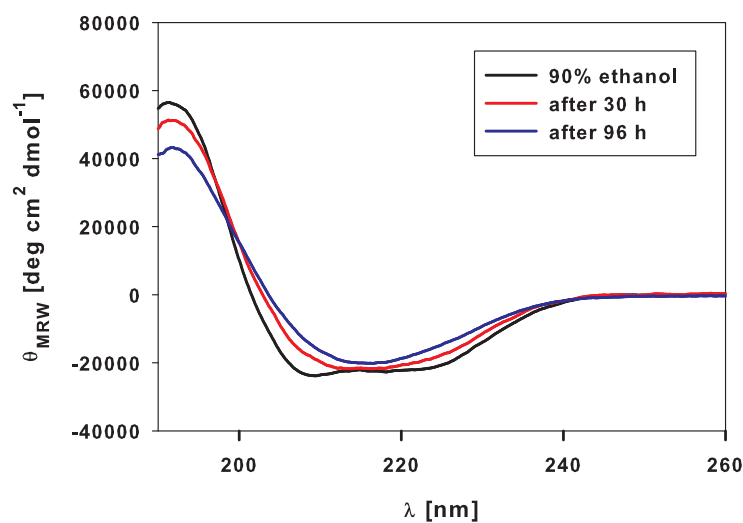


Figure 3.26.: Circular dichroism spectra of HEWL-S^{Me} in the presence of 90 % ethanol at different times of incubation. θ_{MRW} : mean residue molar ellipticity in [deg cm² dmol⁻¹]; λ : wavelength in [nm].

dealt with in our group by other researchers.

In summary, the findings from the CD and NMR spectroscopical investigations of the non-native HEWL-S^{Me} and *all-Ala*-HEWL proteins in organic solvents described in this thesis are very similar to the results found in the literature that started from the native HEWL protein. For the first time, the structural changes in a non-native protein induced by organic solvents have been studied here using two-dimensional heteronuclear NMR. Unfortunately, an unambiguous assignment under this conditions was not possible by simple comparison with the spectrum in pure water. It has been shown with this work that the influence of organic solvents on non-native HEWL can be monitored using 2D NMR spectroscopy. This enables in principle time-resolved studies of the structural changes at atomic resolution, possibly even including fibril formation. However, at least a partial assignment of the resonances under these conditions is needed to gain more valuable information on these processes.

3. Results and Discussion

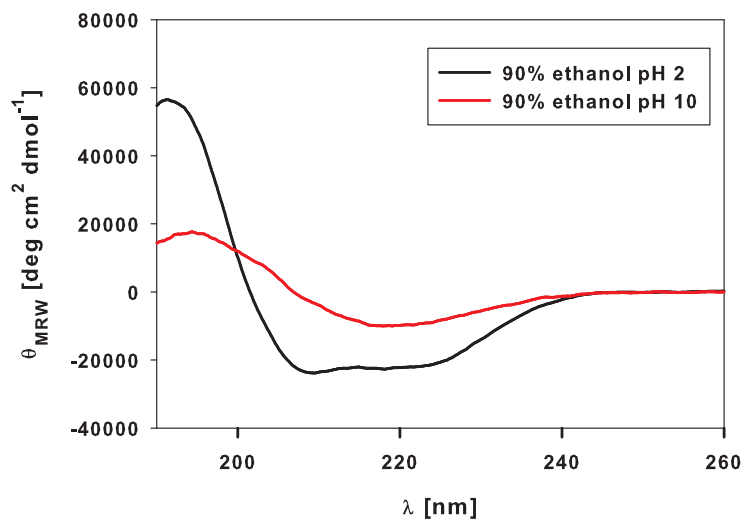


Figure 3.27.: Circular dichroism spectra of HEWL-S^{Me} in the presence of 90 % ethanol at pH 2 and 10. θ_{MRW} : mean residue molar ellipticity in [deg cm² dmol⁻¹]; λ : wavelength in [nm].

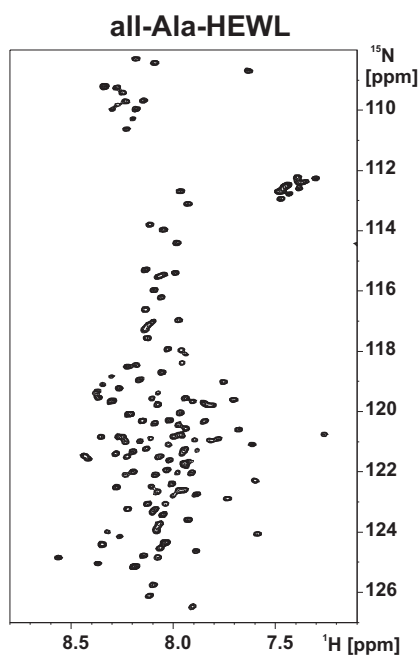


Figure 3.28.: ¹H-¹⁵N-HSQC spectrum of *all-Ala*-HEWL.

3.8. Residual structure in organic solvents

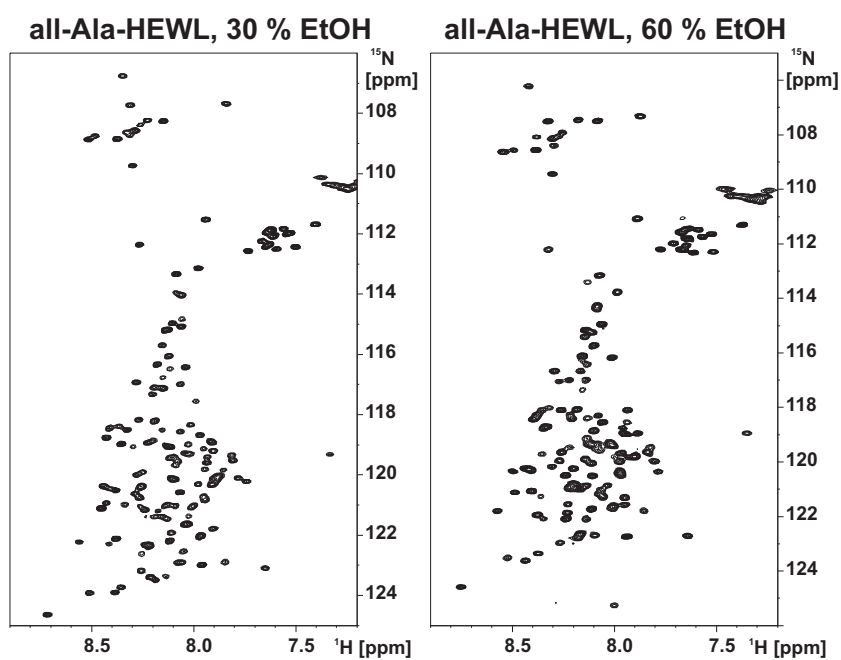


Figure 3.29.: ^1H - ^{15}N -HSQC spectra of *all*-Ala-HEWL in 30 % (left) and 60 % (right) ethanol.

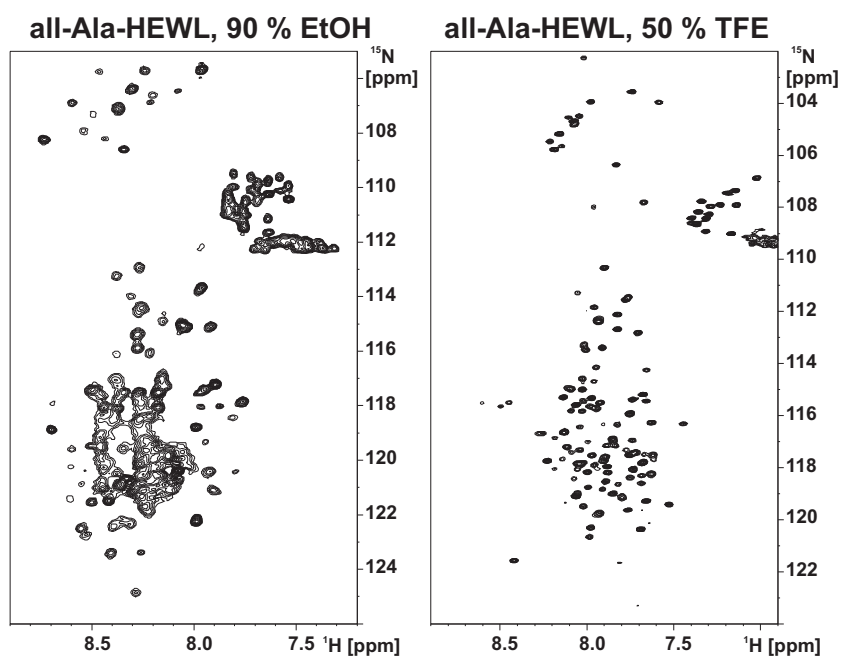


Figure 3.30.: ^1H - ^{15}N -HSQC spectra of *all*-Ala-HEWL in 90 % ethanol (left) and in 50 % TFE (right).

3.9. Conclusions and outlook

Taken together, the results presented in this thesis contribute to the further understanding of the structural and dynamical properties of non-native HEWL, which in this context serves as a model protein. The routine production of non-native HEWL for these investigations was facilitated by the introduction of *all-Ala*-HEWL, the production of which is reliable and reproducible. Single-point mutants can readily be produced by site-directed mutagenesis and be used to modulate the properties of non-native states of this protein.

Unlike for folded native proteins, a direct structure determination for more or less unfolded proteins is not feasible, because these states exist as ensembles of interconverting conformers. The observed properties, e.g. the measured NMR parameters, are therefore averaged over the whole ensemble and over time, depending on the timescale in which the respective method is sensitive. For the analysis, the experimental data can be compared to predictions from model calculations and conclusions can be drawn from deviations of experimental data from the predicted values. The ultimate goal would be to generate ensembles of conformers, for which the predicted properties completely resemble the experiments.

As the prerequisite for the investigation of HEWL-S^{Me} and *all-Ala*-HEWL the NMR resonances of these proteins have been assigned almost to completion despite the difficulties caused by the limited dispersion of the peaks and signal overlap. Suitable combinations of three-dimensional spectra have been used for the purpose of assignment. The use of multiple lines of correlation between the residues in the sequence of the protein, e.g. the combination of HNCACB and HNN experiments, greatly enhanced quality and rate of assignment.

Secondary chemical shifts and CD spectroscopy revealed the presence of residual secondary structure in non-native HEWL at pH 2, particularly the population of α -helix-like conformations in parts of the sequence. Comparison with peptide data (Graf *et al.*, 2007) revealed that for the triple-aline motif in cluster 1, the structural propensities towards α -helical-like conformations are intrinsic to the sequence and not induced by long-range interactions. It is not known so far, to which degree this is also the case for other regions of the sequence that exhibit such propensities, e.g. clusters 3 and 5. An approach to solve this open question would be to synthesize and characterize peptides corresponding to this segments of the lysozyme sequence. Another possibility would be to investigate the secondary chemical shifts for the single point mutants W62G, W108G and W111G of HEWL-S^{Me} or *all-Ala*-HEWL.

Unfortunately, ^{13}C , ^{15}N doubly-labeled samples of these mutants do not exist so far and therefore the most indicative $^1\text{H}_\alpha$ and $^{13}\text{C}_\alpha$ secondary chemical shifts are not available.

The importance of at least some of the six tryptophan side chains for the properties of non-native states of lysozyme has been pointed out repeatedly in this thesis. In an effort to assign the NMR resonances of the tryptophan side chains, a set of two new pulse sequences has been developed, which taken together for the first time allowed for the unambiguous and residue specific assignment of the $^{13}\text{C}_\gamma$, $^{15}\text{N}_\epsilon$ and $^1\text{H}^{\text{N}}_\epsilon$ resonances. This knowledge was eventually exploited in the interpretation of two-dimensional ^{15}N - ^1H photo-CIDNP spectra, which revealed a differential solvent accessibility of the tryptophan residues in *all-Ala*-HEWL but not in W62G-*all-Ala*-HEWL. These findings are in line with the data from the combined relaxation and mutational studies and the determined radii of hydration, which suggested more extended conformations for W62G and other mutants with disrupted hydrophobic clusters (Wirmer *et al.*, 2004). In a more extended conformation caused by the decrease of long-range contacts between hydrophobic residues, the accessibility of single tryptophan side chains for the CIDNP dye will be more uniform than in a situation, where some of the tryptophans are more involved in hydrophobic clustering than others. This is also supported by the comparison of the heteronuclear R_2 relaxation rates of the indole $^{15}\text{N}_\epsilon$ nuclei in *all-Ala*-HEWL with the corresponding rates in it W62G mutant. While in the wild-type like *all-Ala*-HEWL, the rates are different among the six tryptophan residues, in W62G they are more uniform.

A W108G mutant of *all-Ala*-HEWL has been created and R_2 relaxation have been determined. W108, together with W111, forms the core of the second most pronounced hydrophobic cluster in non-native HEWL. Interestingly, this mutation does — unlike W108G or W111G — not significantly alter the relaxation profile of the protein. Apparently, W108 is to a much lesser degree involved in the hydrophobic clustering than W62 or the neighboring W111.

The various relaxation rates and the relaxation dispersion data show no evidence for a slow conformational exchange in *all-Ala*-HEWL, as it is for example observed around the disulfide bridges in fully and partially oxidized HEWL (Collins *et al.*, 2005).

The differences of the W62G and the wild-type like non-native HEWL is also picked up in NH-RDCs of these proteins aligned in polyacrylamide gels. Significant positive RDCs are observed in the regions of the hydrophobic clusters in *all-Ala*-HEWL, but to a much lesser degree in W62G. So far, all attempts to simulate

3. Results and Discussion

RDCs from generated non-native ensembles failed even when including long-range contacts or specific ϕ, ψ backbone angle propensities. Hopefully, approaches will be developed that can more accurately generate ensembles of conformers, that resemble the properties of non-native HEWL. Most promising are molecular dynamics approaches, that use restraints from the relaxation data and the surface accessibilities determined by the photo-CIDNP method. RDCs from such approaches should closely match the experimental data not only for *all-Ala*-HEWL, but also for the single-point mutants.

Another source of information about the structure in non-native proteins are distance restraints derived from paramagnetic resonance enhancements (PREs) caused by paramagnetic spin labels. First data are available for a paramagnetic spin label covalently attached at position 73 of the R73C mutant of *all-Ala*-HEWL (Argirevic, 2006). The results of this experiment showed that residues in cluster 1 and 5 are close to the spin label.

In the meantime, RDC data can be acquired for additional single-point mutants of *all-Ala*-HEWL and paramagnetic spin labels can be attached at various positions in the sequence of the protein in order to gain new distance restraints from PREs.

From the literature it is known, that non-native HEWL can form amyloid fibrils under certain conditions, including the presence of organic solvents (Goda *et al.*, 2000; Krebs *et al.*, 2000) and sodium maleate (Mishima *et al.*, 2006). Most interestingly, the rate of amyloid formation has been found to differ among various single-point mutants, with W62G not forming fibrils under these conditions (Mishima *et al.*, 2006). However, molecular dynamics simulations showed that the W62G mutant is less stable than wild-type HEWL (Zhou *et al.*, 2007). An overall picture that unifies all these findings is still desired, but findings from this thesis may contribute to a better understanding of the properties of non-native HEWL and its mutants.

A. Appendix

A.1. List of buffers, media and primers

A.1.1. M9 minimal media

Table A.1.: M9 minimal media.

¹⁵ N M9 medium	¹⁵ N, ¹³ C M9 medium
42 mM Na ₂ HPO ₄	42 mM Na ₂ HPO ₄
22 mM KH ₂ PO ₄	22 mM KH ₂ PO ₄
8.5 mM NaCl	8.5 mM NaCl
1 g/L NH ₄ Cl (¹⁵ N)	1 g/L NH ₄ Cl (¹⁵ N)
0.1 mM CaCl ₂	0.1 mM CaCl ₂
2 mM MgSO ₄	2 mM MgSO ₄
10 g/L D-glucose (sterile)	2.5 g/L ¹³ C ₆ D-glucose (sterile)
1 mL/L trace elements stock ^a	1 mL/L trace elements stock ^a
30 μM FeCl ₃	30 μM FeCl ₃
5 ng/L thiamine	5 ng/L thiamine
1 mL/L ampicillin stock ^b	1 mL/L ampicillin stock ^b

^atrace elements stock solution contains 1 mM Cu²⁺, Zn²⁺, Mn²⁺, Co²⁺, Ni²⁺, MoO⁴⁻

^bstock solution contains 100 mg/mL ampicillin in water

A. Appendix

A.1.2. Protein purification buffers

Table A.2.: Inclusion bodies solubilization buffers.

Sonication buffer	Washing buffer	Denaturation buffer
50 mM tris/HCl pH 7.5 25 % sucrose (w/v) 1 mM EDTA	20 mM tris/HCl pH 7.5 1 % Triton X-100 (w/v) 1 mM EDTA	20 mM tris/HCl pH 7.5 0.1 M DTT 5 mM EDTA 50 mM NaCl 8 M urea

Table A.3.: Ion exchange buffers.

Low salt buffer (buffer A)	High salt buffer (buffer B)
20 mM tris/HCl pH 7.5 10 mM 2-mercaptoethanol 1 mM EDTA 4 M urea	20 mM tris/HCl pH 7.5 10 mM 2-mercaptoethanol 1 mM EDTA 4 M urea 300 mM NaCl

A.1.3. Primer for site-directed mutagenesis

Table A.4.: Primers for site-directed mutagenesis.

Primer	Sequence
W62G forward	5' -cagatcaacagccgcgggtgggcgaacgacgg-3'
W62G reverse	5' -ccgtcgttcgcccacccgcggctggtgatctg-3'
W108G forward	5' -gcaacggcatgaacgcgggggtggcctggcgtaacc-3'
W108G reverse	5' -ggttacgccaggccacccccgcggtcatgccggttgc-3'

A.2. NMR acquisition and processing parameters

Table A.5.: NMR acquisition and processing parameters for the assignment of HEWL-S^{Me}.

Experiment	T ^a	B ₀ ^b	t ₁ ^c	sw ₁ ^d	t ₂ ^e	sw ₂ ^f	t ₃ ^g	sw ₃ ^h	ns ⁱ	d1 ^j	ω _{15N} ^k	ω _{13C} ^k	SI ₁ ^l	SI ₂ ^l	SI ₃ ^l
¹⁵ N, ¹ H-HSQC	293	21.1	2189.262	128	12626.26	1024	-	-	4	1.0	119.0	-	512	1024	-
HNCACB	293	14.1	10559.66	96	1337.44	32	8389.26	512	16	1.3	117.5	39.0	256	512	2048
CBCA(CO)NH	293	14.1	10559.66	70	1337.44	32	8389.26	512	8	1.8	117.5	39.0	512	256	2048
HNCO	293	16.4	2465.18	45	1561.04	35	9842.52	1024	8	1.4	117.5	172.0	256	256	1024
(H)CC(CO)NH	293	21.1	13888.89	100	2000	28	12626.26	1024	16	1.3	117.5	40.5	512	256	1024
H(CC)(CO)NH	293	21.1	6301.197	100	2000	28	12626.26	1024	8	1.3	117.5	40.5	512	256	1024
HBHA(CO)NH	293	16.4	4201.24	128	1561.04	36	9765.63	512	8	1.4	117.5	37.4	512	256	1024

^atemperature in K^bmagnetic field strength in T^cnumber of complex points in F1 dimension^dsweep width in Hz for F1 dimension^enumber of complex points in F2 dimension^fsweep width in Hz for F2 dimension^gnumber of complex points in F3 dimension^hsweep width in Hz for F3 dimensionⁱnumber of transients per increment^jwaiting delay between recording of two successive transients in s^koffset of carrier frequencies for ¹⁵N and ¹³C in ppm^lnumber of points after Fourier transformation in the respective dimension

Table A.6.: NMR acquisition and processing parameters for the resonance assignment of *all-Ala*-HEWL.

Experiment	T ^a	B ₀ ^b	t ₁ ^c	sw ₁ ^d	t ₂ ^e	sw ₂ ^f	t ₃ ^g	sw ₃ ^h	ns ⁱ	dI ^j	ω _{15N} ^k	ω _{13C} ^k	SI ₁ ^l	SI ₂ ^l	SI ₃ ^l
¹⁵ N, ¹ H-HSQC	293	16.4	128	1702.93	1024	11261.26	-	-	4	1.4	118.5	-	1024	2048	-
HNCACB	293	16.4	92	12678.29	40	1703.00	1024	9765.63	8	1.3	118.5	42.0	512	256	1024
HNCA	293	16.4	80	4225.65	40	1703.00	1024	9842.52	4	1.3	118.5	51.0	512	256	1024
CBCA(CO)NH	293	16.4	68	10565.24	40	1703.00	1024	9842.52	4	1.3	118.5	41.0	512	256	1024
HN(CA)CO	293	18.8	128	4444.44	36	1621.67	1024	11160.71	8	1.5	117.5	174.0	512	256	1024
(H)CC(CO)NH	293	14.1	80	11312.22	20	1459.00	1024	9615.39	16	1.8	118.5	39.0	256	128	1024
HNN	293	18.8	72	1946.09	72	1946.09	1024	11160.71	16	1.1	117.0	54.0	512	256	1024
HBHA(CO)NH	293	16.4	90	5000.00	40	1703.00	1024	11261.26	4	1.3	118.5	40.0	512	256	1024
HN(CACB)CG	293	16.4	128	1760.64	512	9842.52	-	-	64	1.6	118.0	27.0	1024	1024	-
HN(CD)CG	293	16.4	128	528.23	1024	9842.52	-	-	16	1.5	129.2	108.0	1024	1024	-

^atemperature in K^bmagnetic field strength in T^cnumber of complex points in F1 dimension^dsweep width in Hz for F1 dimension^enumber of complex points in F2 dimension^fsweep width in Hz for F2 dimension^gnumber of complex points in F3 dimension^hsweep width in Hz for F3 dimensionⁱnumber of transients per increment^jwaiting delay between recording of two successive transients in s^koffset of carrier frequencies for ¹⁵N and ¹³C in ppm^lnumber of points after Fourier transformation in the respective dimension

A.3. Resonance assignments

A.3.1. HEWL-S^{Me}

Table A.7.: Resonance assignment for HEWL-S^{Me}.

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^{13}\text{C}_\alpha)$ [ppm]	$\delta(^{13}\text{C}_\beta)$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]	$\delta(^1\text{H}_\alpha)$ [ppm]	$\delta(^1\text{H}_\beta)$ [ppm]
M-1 ^{1,2}	—	—	—	—	—	—	—
K1	8.694	126.10	56.15	32.96	171.88	—	—
V2	8.260	124.10	61.73	32.89	176.35	4.070	1.946
F3	8.522	126.32	57.75	39.77	175.57	4.623	3.053
G4	8.368	111.87	44.90	—	176.08	3.913	—
R5	8.279	121.90	56.30	30.52	173.84	—	—
A6 ¹	—	—	—	—	—	4.541	2.907
E7	8.527	124.53	56.07	28.37	175.37	4.347	2.014
L8	8.206	124.48	55.53	42.07	176.08	4.264	1.595
A9	8.247	125.55	53.01	18.66	177.42	4.362	1.385
A10	8.158	123.27	53.09	18.58	176.64	4.205	1.400
A11	8.180	123.54	53.16	18.66	178.37	4.212	1.393
M12	8.116	119.79	56.07	32.66	178.37	—	—
K13 ¹	—	—	—	—	—	—	—
R14 ¹	—	—	—	—	—	—	—
H15 ¹	—	—	—	—	—	—	—
G16	8.444	111.29	45.13	—	174.75	3.943	—
L17	8.223	122.76	55.38	42.30	173.89	4.324	1.572
D18	8.525	120.51	52.78	37.86	177.26	4.623	2.769
N19	8.309	120.22	53.01	38.40	174.69	4.638	2.686
Y20	8.069	121.78	57.91	38.47	174.69	4.489	2.963
R21	8.247	124.60	55.99	30.13	175.75	4.182	1.673
G22	7.752	109.96	44.98	—	176.23	3.808	—
Y23	7.948	120.98	57.60	38.85	173.55	4.556	2.930
S24	8.262	118.76	57.98	63.95	175.80	4.399	3.786
L25	8.218	125.15	55.53	63.87	174.25	4.287	1.595
G26	8.194	109.71	45.21	—	177.67	3.786	—
N27	8.169	119.79	53.01	38.63	173.73	4.668	2.672
W28	8.073	122.76	57.52	29.21	174.89	4.616	3.225
V29	7.844	122.83	62.34	32.66	176.16	3.920	1.909
A30	8.134	123.88	55.30	36.87	175.67	4.362	2.814

¹Backbone unassigned

²The N-terminal methionine has been introduced for expression in *E. coli*

A. Appendix

Table A.7 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^{13}\text{C}_\alpha)$ [ppm]	$\delta(^{13}\text{C}_\beta)$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]	$\delta(^1\text{H}_\alpha)$ [ppm]	$\delta(^1\text{H}_\beta)$ [ppm]
A31	8.385	127.82	52.93	18.81	174.85	4.205	1.370
A32	8.111	123.62	52.78	18.81	177.51	4.182	1.325
K33	8.027	120.48	56.38	32.66	177.94	4.257	1.774
F34	8.219	122.54	55.69	42.30	173.91	4.310	1.561
E35	8.278	122.50	55.61	28.99	177.21	4.317	1.972
S36	8.278	117.81	58.59	63.57	175.71	4.324	3.793
N37	8.347	121.29	53.01	38.85	174.19	4.646	2.686
F38	8.149	121.46	58.06	39.08	174.94	4.549	3.038
N39	8.345	121.25	53.16	38.55	175.51	4.698	2.754
T40	8.113	115.62	65.86		175.46		
Q41	8.329	123.22	56.15	29.06	174.75	4.302	2.040
A42	8.280	125.98	52.78	18.89	175.91	4.317	1.385
T43	8.101	113.87	61.96	69.69	177.99	4.302	
N44	8.382	121.97	53.09	38.17	174.37	4.683	2.757
R45	8.337	122.55	56.22	30.52	175.14	4.317	1.793
N46	8.494	120.73	53.24	38.55	176.05	4.736	2.799
T47	8.178	115.18	62.04	69.54	175.71	4.302	
D48	8.494	122.01	52.93	37.86	174.69	4.758	2.881
G49	8.410	110.49	45.36	—	175.64	3.965	—
S50	8.170	116.65	58.36	63.80	174.28	4.481	3.876
T51	8.211	116.75	61.88	69.54	174.89	4.272	4.085
D52	8.380	122.10	52.78	38.09	174.32	4.683	2.716
Y53	8.131	121.99	58.52	38.32	174.73	4.459	2.990
G54	8.278	110.88	45.28	—	176.37	3.868	—
I55	7.838	120.78	61.43	38.40	174.18	4.145	1.856
L56	8.228	126.11	48.80		176.32		
Q57 ¹	—	—	—	—	—	4.287	1.980
I58	8.109	122.68	61.65	38.47	176.16	4.003	1.804
N59	8.354	122.67	53.16	38.55	176.03	4.683	2.769
S60	8.228	117.29	58.59	63.64	177.69	4.272	3.827
R61	8.194	123.05	56.22	29.67	174.82	4.010	1.430
W62	7.786	121.59	57.60	29.21	176.33	4.541	3.217
W63	7.398	121.60	57.37	28.99	176.12	4.466	2.960
A64	7.728	121.93	55.30	36.87	175.76		
N65 ¹	—	—	—	—	—	4.623	2.761
D66	8.352	119.96	53.24	37.86	175.05	4.631	2.866

¹Backbone unassigned

A.3. Resonance assignments

Table A.7 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^{13}\text{C}_\alpha)$ [ppm]	$\delta(^{13}\text{C}_\beta)$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]	$\delta(^1\text{H}_\alpha)$ [ppm]	$\delta(^1\text{H}_\beta)$ [ppm]
G67	8.331	109.57	45.43	—	175.48	3.838	—
R68	7.964	120.86	55.92	30.75	173.55	4.407	1.804
T69	8.247	118.29	59.67	69.46	174.78		
P70 ¹	—	—	—	—	—	4.706	2.821
G71	8.485	110.55	45.13	—	175.98	3.943	—
S72	8.115	116.61	58.59	63.87	174.28	4.414	3.868
R73	8.420	123.69	56.15	30.52	174.69	4.317	1.763
N74	8.424	120.73	53.16	38.55	175.87	4.661	2.765
L75	8.297	123.87	61.96	41.99	175.09		1.587
A76	8.295	120.91	55.30	37.10	177.12	4.406	2.874
N77	8.507	122.29	52.78	38.47	174.59	4.691	2.739
I78	8.020	123.87	58.59	38.40	174.41		
P79 ¹	—	—	—	—	—	4.414	2.253
A80	8.499	121.60	56.07	36.79	176.76	4.444	2.911
S81	8.515	119.15	58.98	63.41	175.66	4.347	3.876
A82	8.212	126.92	52.78	18.96	174.55	4.302	1.378
L83	7.960	121.35	55.30	41.99	177.78	4.287	1.587
L84	8.045	123.15	55.23	42.07	177.55	4.362	1.617
S85	8.231	117.12	58.59	63.57	177.67	4.407	3.876
S86	8.247	118.29	58.59	63.64	176.26	4.451	3.876
D87	8.406	121.98	53.01	37.94	174.35		
I88 ¹	—	—	—	—	—	4.190	1.901
T89	8.174	119.20	61.88	69.54	176.55		
A90	8.262	127.49	52.55	18.96	174.27	4.332	1.393
S91	8.285	116.47	58.36	63.72	177.76		
V92	8.070	122.37			175.18	4.107	2.096
N93	8.477	122.76	45.78		175.96	4.661	2.709
A94	8.220	121.73			174.07	4.459	2.907
A95	8.382	127.04	52.55	18.81	174.85	4.302	1.385
K96	8.122	123.14	56.38	32.66	176.64	4.242	1.774
K97	8.262	124.31	56.22	32.89	176.33	4.287	1.744
I98	8.269	124.60	60.89	38.32	176.23	4.152	1.826
V99	8.340	126.41	62.27	32.66	176.25	4.152	2.043
S100	8.447	120.91	58.06	63.80	175.96		3.823
D101	8.583	122.63	52.78	38.01	174.34	4.706	2.911
G102	8.393	110.63	45.28	—	175.64	3.928	—

¹Backbone unassigned

A. Appendix

Table A.7 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^{13}\text{C}_\alpha)$ [ppm]	$\delta(^{13}\text{C}_\beta)$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]	$\delta(^1\text{H}_\alpha)$ [ppm]	$\delta(^1\text{H}_\beta)$ [ppm]
N103	8.359	119.71	53.24	38.63	174.05	4.706	2.814
G104	8.474	110.36	45.51	—	176.00	3.928	—
M105	8.215	121.00	55.92	32.20	174.59	4.407	2.021
N106	8.407	120.44	53.77	38.40	176.39	4.586	2.709
A107	8.218	124.94	53.39	18.66	175.55	4.197	1.310
W108	8.049	120.91	58.44	28.99	178.01	4.504	3.292
V109	7.793	122.19	63.57	32.35	176.80	3.711	1.879
A110	8.015	125.60	53.54	18.50	176.41	4.063	1.348
W111	7.899	120.35	58.52	28.91	178.64	4.414	3.292
R112	7.984	121.94	57.22	29.83	177.16	3.905	1.558
N113	8.063	118.93	53.47	38.40	176.57	4.571	2.758
R114	7.899	121.68	56.76	30.21	175.46	4.227	1.767
A115	8.203	121.25	55.69	36.79	176.35	4.429	2.802
K116	8.374	124.39	56.22	32.50	175.14	4.257	1.782
G117	8.360	110.96	45.13	—	176.87	3.980	—
T118	8.069	114.42	61.96	69.69	174.18	4.347	4.227
D119	8.549	122.66	52.70	37.78	174.35	4.728	2.933
V120	8.098	122.07	62.88	32.20	175.12	3.950	2.013
Q121	8.365	124.29	56.30	28.68	176.26	4.160	1.879
A122	8.144	125.33	52.86	18.89	175.91	4.205	1.318
W123	7.984	120.86	57.45	29.14	177.51	4.608	3.217
I124	7.899	123.79	61.43	38.63	176.35	3.988	1.722
R125	8.173	125.17	56.61	30.13	175.94	4.107	1.767
G126	8.302	110.63	45.05	—	176.73	3.928	—
A127	8.129	120.96	55.38	37.25	173.86	4.474	2.814
R128	8.507	124.99	55.76	30.44	174.71	4.279	
L129	8.337	125.97	54.23	41.76	175.78		

A.3.2. *all-Ala*-HEWLTable A.8.: Resonance assignment for *all-Ala*-HEWL.

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}_{\alpha})$ [ppm]	$\delta(^1\text{H}_{\beta})$ [ppm]		$\delta(^{13}\text{C}_{\alpha})$ [ppm]	$\delta(^{13}\text{C}_{\beta})$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]
M-1 ^{1,2}	—	—	—	—	—	—	—	—
K1	8.706	126.11	4.411	1.731		56.23	32.82	173.33
V2	8.253	123.71	4.138	2.011		61.83	32.79	173.41
F3	8.508	126.24	4.704	3.167	3.036	57.50	39.80	173.95
G4	8.357	111.82	3.994			45.02		171.91
R5	8.308	122.25	4.252	1.876		57.02	30.31	174.81
A6	8.492	125.48	4.290	1.442		53.22	18.52	176.44
E7	8.217	120.72	4.299	2.071		56.62	28.01	174.69
L8	8.170	124.22				55.85	41.92	175.66
A9 ¹	—	—	—	—	—	—	—	—
A10	8.039	122.87	4.239	2.754		53.47	18.47	176.48
A11	8.108	123.13	4.252	1.465		53.39	18.41	176.47
M12	8.091	119.51	4.216	1.454		56.02	32.46	176.49
K13	8.060	122.81	4.290	1.862		56.69	32.51	174.74
R14	8.145	122.21	4.299	1.824		56.38	30.38	174.15
H15	8.464	120.17	4.739	3.349	3.235	55.38	28.70	172.65
G16	8.425	111.09	4.007			45.13		171.83
L17	8.201	122.70	4.392	1.624		55.23	42.20	175.13
D18	8.531	120.36	4.681	2.896	2.826	52.82	37.64	172.56
N19	8.297	120.16	4.709	2.738		52.91	38.30	172.57
Y20	8.072	121.73	4.546	3.027		58.10	38.30	173.62
R21	8.246	124.54	4.230	1.799	1.650	56.09	30.11	174.14
G22	7.773	109.88	3.889			45.03		171.46
Y23	7.943	121.02	4.611	3.041	2.938	57.77	38.78	173.69
S24	8.270	118.77	4.471	3.861		57.92	63.90	172.17
L25	8.225	125.17	4.348	1.645		55.51	42.01	175.62
G26	8.233	109.59	3.866			45.21		171.79
N27	8.192	119.93	4.730	2.749		53.08	38.41	172.87
W28	8.088	122.98	4.611	3.297		57.91	29.07	174.10
V29	7.746	123.47	3.873	1.941		62.52	32.58	173.57
A30	8.038	127.55	4.221	1.409		52.62	18.66	175.69
A31	8.098	123.75	4.229	1.409		52.64	18.70	175.93
A32	8.116	123.79	4.234	1.368		52.64	18.69	175.93

¹Unassigned²The N-terminal methionine has been introduced for expression in *E. coli*

A. Appendix

Table A.8 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}_{\alpha})$ [ppm]	$\delta(^1\text{H}_{\beta})$ [ppm]		$\delta(^{13}\text{C}_{\alpha})$ [ppm]	$\delta(^{13}\text{C}_{\beta})$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]
K33	8.071	120.70	4.201	1.712		56.62	32.54	174.35
F34	8.105	121.59	4.616	3.162	3.072	57.89	39.25	173.59
E35	8.155	122.68	4.360	2.099	1.964	55.55	28.53	173.60
S36	8.273	117.77	4.383	3.866		58.59	63.67	172.09
N37	8.351	121.28	4.704	2.738		53.32	38.54	172.82
F38	8.154	121.48	4.602	3.167	3.031	58.01	39.09	173.40
N39	8.346	121.25	4.756	2.882	2.747	53.24	38.49	173.38
T40	8.120	115.60	4.321	4.294		62.13	69.50	172.64
Q41	8.331	123.18	4.350	2.151	2.029	55.90	28.96	173.78
A42	8.282	125.97	4.374	1.442		52.64	18.82	175.89
T43	8.103	113.86	4.325	4.171		61.89	69.70	172.26
N44	8.381	122.09	4.753	2.863		53.13	38.46	173.04
R45	8.334	122.54	4.360	1.908	1.782	56.14	30.34	173.96
N46	8.500	120.70	4.802	2.906	2.840	53.34	38.49	173.57
T47	8.180	115.17	4.364	4.344		61.90	69.47	172.57
D48	8.510	121.95	4.757	2.976		53.08	37.67	173.49
G49	8.388	110.62	4.029			45.31		172.15
S50	8.175	116.63	4.541	3.931		58.50	63.79	172.77
T51	8.214	116.76	4.332	4.173		61.89	69.59	172.00
D52	8.388	122.12	4.751	2.780		53.01	38.33	172.60
Y53	8.139	121.98	4.503	3.112	2.972	58.52	38.11	174.25
G54	8.279	110.88	3.926			45.37		172.09
I55	7.847	120.85	4.178	1.908		61.41	38.30	174.20
L56	8.221	126.06	4.366	1.563		55.33	42.01	175.15
Q57	8.274	122.46	4.339	2.081	1.972	55.66	28.83	174.04
I58	8.099	122.63	4.057	1.858		61.74	38.21	173.95
N59	8.359	122.71	4.751	2.840		53.29	38.41	173.27
S60	8.201	117.45	4.290	3.903	3.833	58.95	63.54	172.83
R61	8.234	123.34	4.066	1.517		56.44	29.50	174.34
W62	7.820	121.80	4.566	3.240		57.71	29.02	174.16
W63	7.396	121.87	4.471	3.139	2.980	57.26	28.96	173.82
A64	7.724	125.22	4.178	1.251		52.57	18.66	175.30
N65	8.113	118.12	4.658	2.887	2.710	53.16	38.41	173.11
D66	8.331	119.53	4.662	2.953	2.844	53.09	37.56	173.35
G67	8.316	109.46	3.889			45.44		171.94
R68	7.964	120.99	4.452	3.942		55.75	30.63	174.13
T69	8.251	118.25				59.82	69.51	170.91

A.3. Resonance assignments

Table A.8 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}_\alpha)$ [ppm]	$\delta(^1\text{H}_\beta)$ [ppm]		$\delta(^{13}\text{C}_\alpha)$ [ppm]	$\delta(^{13}\text{C}_\beta)$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]
P70 ¹	—	—	—	—	—	—	—	—
G71	8.483	110.54	4.005			45.13		172.15
S72	8.126	116.63	4.471	3.926		58.53	63.83	172.58
R73	8.421	123.75	4.364	1.890	1.773	56.08	30.42	173.72
N74	8.440	120.85	4.723	2.892	2.752	53.09	38.39	173.02
L75	8.269	124.28	4.346	1.647		55.21	41.98	174.92
A76	8.190	124.58	4.313	1.409		52.40	18.87	175.17
N77	8.282	118.42	4.709	2.840	2.733	52.73	38.54	172.44
I78	8.023	123.96				58.66	38.43	172.23
P79 ¹	—	—	—	—	—	—	—	—
A80	8.494	125.66	4.262	1.456		53.16	18.53	176.36
S81	8.259	114.97	4.366	3.939		58.68	63.41	172.69
A82	8.232	126.97	4.374	1.433		52.62	18.78	175.75
L83	7.979	121.50	4.339	1.645		55.38	41.97	175.42
L84	8.043	123.18	4.420	1.670		55.30	42.04	175.59
S85	8.232	117.14	4.448	3.945		58.69	63.57	172.70
S86	8.271	118.36	4.489	3.953		58.68	63.59	172.30
D87	8.403	121.99	4.798	2.971		53.01	37.58	173.12
I88	8.082	122.49	4.234	1.969		61.77	38.16	174.51
T89	8.168	119.20	4.297	4.270		62.07	69.54	172.27
A90	8.262	127.34	4.374	1.456		52.70	18.84	175.91
S91	8.280	116.55	4.475	3.930		58.65	63.57	172.98
V92	8.178	123.21	4.116	2.136		62.93	32.24	174.18
N93	8.426	122.59	4.714	2.887	2.780	53.44	38.48	173.23
A94	8.207	125.74	4.285	1.450		52.74	18.74	175.60
A95	8.149	123.60	4.299	1.442		52.60	18.70	175.81
K96	8.105	121.23	4.290	1.843		56.25	32.73	174.38
K97	8.214	123.85	4.341	1.806		56.18	32.69	174.16
I98	8.226	124.42	4.204	1.891		60.94	38.30	174.17
V99	8.323	126.35	4.201	2.096		62.09	32.53	173.90
S100	8.457	120.92	4.534	3.898		58.03	63.75	172.19
D101	8.583	122.59	4.830	2.962		52.68	37.70	173.47
G102	8.412	110.46	4.010			45.44		171.91
N103	8.359	119.71	4.766	2.876		53.19	38.62	173.86
G104	8.479	110.41	3.996			45.36		172.46
M105	8.216	120.99	4.467	2.071		55.73	32.19	174.23

¹Unassigned

A. Appendix

Table A.8 (continued)

Residue	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}_{\alpha})$ [ppm]	$\delta(^1\text{H}_{\beta})$ [ppm]		$\delta(^{13}\text{C}_{\alpha})$ [ppm]	$\delta(^{13}\text{C}_{\beta})$ [ppm]	$\delta(^{13}\text{C}')$ [ppm]
N106	8.408	120.43	4.644	2.766		53.73	38.29	173.35
A107	8.213	124.94	4.248	1.377		53.46	18.51	175.81
W108	8.049	120.93	4.583	3.325		58.35	28.90	174.63
V109	7.768	122.35	3.782	1.927		63.42	32.34	174.21
A110	8.033	125.82	4.130	1.413		53.44	18.33	176.46
W111	7.905	120.32	4.481	3.339		58.37	28.75	175.03
R112	7.972	122.18	3.959	1.684	1.545	57.05	29.92	174.34
N113	8.095	119.17	4.625	2.887	2.710	53.47	38.36	173.26
R114	7.909	122.14	4.238	1.931		56.44	30.17	173.99
A115	8.183	125.59	4.276	1.368		52.56	18.72	175.67
K116	8.224	121.58	4.285	1.858		56.56	32.65	175.07
G117	8.376	110.91	4.047			45.15		172.13
T118	8.067	114.32	4.389	4.266		61.82	69.75	172.23
D119	8.573	122.63	4.798	2.887		52.68	37.71	172.94
V120	8.104	122.04	4.015	2.067		62.77	32.19	174.12
Q121	8.367	124.49				55.97	28.67	173.69
A122 ¹	—	—	—	—	—	—	—	—
W123	7.995	120.96	4.677	3.274		57.20	29.08	174.09
I124	7.879	124.16	4.042	1.772		61.11	38.43	173.73
R125	8.191	125.59	4.144	1.836	1.754	56.50	30.07	174.63
G126	8.327	111.20	3.968			45.00		171.53
A127	8.068	124.83	4.327	1.382		52.26	19.13	175.24
R128	8.296	121.51	4.327	1.866	1.740	55.72	30.36	173.98
L129	8.387	125.83				53.98	41.69	176.97

¹Unassigned

A.3.3. W62G- and W108G-*all-Ala*-HEWLTable A.9.: Resonance assignment for W62G-*all-Ala*-HEWL and W108G-*all-Ala*-HEWL.

Residue	W62G- <i>all-Ala</i> -HEWL		W108G- <i>all-Ala</i> -HEWL	
	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]
M-1 ¹	—	—	—	—
K1	8.690	125.96	8.71	126.22
V2	8.232	123.64	8.24	123.67
F3	8.499	126.20	8.49	126.13
G4	8.362	111.78	8.34	111.87
R5	8.287	122.09	8.31	122.35
A6	8.478	125.59	8.49	125.31
E7	8.246	120.75	8.19	120.61
L8	8.171	124.24	8.15	124.12
A9	—	—	—	—
A10	—	—	8.09	123.00
A11	8.116	123.22	—	—
M12	8.089	119.56	8.09	119.39
K13	8.075	122.89	—	—
R14	8.164	122.30	8.15	122.43
H15	8.485	120.32	8.42	119.93
G16	8.430	111.16	8.39	111.01
L17	8.205	122.65	8.19	122.78
D18	8.499	120.63	8.53	120.27
N19	8.301	120.10	8.28	120.14
Y20	8.075	121.74	8.08	121.76
R21	8.232	124.51	8.24	124.47
G22	7.755	109.87	7.79	109.91
Y23	7.939	120.92	7.94	121.08
S24	8.253	118.73	8.27	118.74
L25	8.205	125.10	8.22	125.19
G26	8.219	109.64	8.23	109.58
N27	8.185	119.83	8.17	119.71
W28	8.068	123.02	8.05	122.93
V29	7.720	123.55	7.76	123.22
A30	8.041	127.77	8.02	127.27
A31	8.082	123.83	—	—
A32	8.116	123.88	8.08	123.57
K33	—	—	8.05	120.51

¹The N-terminal methionine has been introduced for expression in *E. coli*

A. Appendix

Table A.9 (continued)

Residue	W62G- <i>all-Ala</i> -HEWL		W108G- <i>all-Ala</i> -HEWL	
	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]
F34	8.116	121.70	8.08	121.39
E35	8.157	122.89	—	—
S36	8.273	117.82	8.25	117.67
N37	8.348	121.27	8.32	121.25
F38	8.157	121.46	8.14	121.45
N39	—	—	8.34	121.23
T40	8.116	115.59	8.12	115.58
Q41	8.328	123.19	8.33	123.17
A42	8.280	126.00	8.27	125.93
T43	8.103	113.89	8.09	113.82
N44	—	—	8.37	122.05
R45	8.335	122.50	8.32	122.51
N46	8.505	120.74	8.50	120.72
T47	8.185	115.13	8.17	115.19
D48	8.478	122.02	8.52	121.97
G49	8.382	110.58	8.39	110.67
S50	8.185	116.62	8.17	116.71
T51	8.205	116.68	8.21	116.80
D52	8.369	122.19	8.39	122.13
Y53	8.137	122.03	8.13	122.00
G54	8.287	110.93	8.28	110.93
I55	7.836	120.79	7.85	120.88
L56	8.219	126.29	8.24	125.92
Q57	8.294	122.65	8.27	122.41
I58	8.109	122.78	8.09	122.53
N59	8.430	122.86	8.35	122.73
S60	8.219	117.53	8.20	117.48
R61	8.287	123.32	8.24	123.40
W62/G62	8.273	110.28	7.82	121.83
W63	—	—	7.39	121.77
A64	8.123	126.09	7.72	125.24
N65	8.219	118.50	8.12	118.20
D66	8.355	119.64	8.34	119.59
G67	8.362	109.66	8.32	109.55
R68	7.980	120.85	7.96	121.09
T69	8.260	118.30	8.24	118.23
P70	—	—	8.48	110.65

A.3. Resonance assignments

Table A.9 (continued)

Residue	W62G- <i>all-Ala</i> -HEWL		W108G- <i>all-Ala</i> -HEWL	
	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]
G71	8.410	110.41	8.13	116.72
S72	8.137	116.56	8.41	123.79
R73	8.423	123.70	8.44	120.84
N74	—	—	8.18	122.72
L75	8.266	124.25	8.25	124.26
A76	8.191	124.70	8.19	124.54
N77	8.280	118.56	8.27	118.41
I78	8.041	123.98	8.00	123.88
P79	—	—	—	—
A80	—	—	8.49	125.71
S81	8.253	115.11	8.25	114.85
A82	8.253	126.98	8.21	126.94
L83	8.000	121.58	7.96	121.44
L84	8.055	123.31	8.03	123.08
S85	8.232	117.19	8.22	117.11
S86	8.266	118.45	8.25	118.35
D87	8.389	122.02	8.41	121.97
I88	8.075	122.41	8.08	122.48
T89	8.164	119.13	8.16	119.13
A90	8.253	127.30	8.26	127.38
S91	8.273	116.45	8.28	116.58
V92	8.157	123.08	8.17	123.17
N93	8.410	122.58	8.43	122.68
A94	8.191	125.70	8.21	125.87
A95	8.137	123.58	8.16	123.70
K96	8.096	121.23	8.11	121.32
K97	8.212	123.86	8.22	123.95
I98	—	—	8.23	124.45
V99	8.307	126.48	8.32	126.33
S100	8.437	120.85	8.45	120.92
D101	8.560	122.67	8.59	122.61
G102	8.492	110.39	8.42	110.57
N103	8.342	119.88	8.36	119.80
G104	8.464	110.33	8.48	110.47
M105	8.205	120.89	8.21	120.88
N106	8.396	120.38	8.19	120.01
A107	8.191	124.86	8.21	124.72

A. Appendix

Table A.9 (continued)

Residue	W62G- <i>all-Ala</i> -HEWL		W108G- <i>all-Ala</i> -HEWL	
	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]	$\delta(^1\text{H}^{\text{N}})$ [ppm]	$\delta(^{15}\text{N}^{\text{H}})$ [ppm]
W108/G108	8.082	120.80	8.40	109.01
V109	7.734	122.26	7.89	122.57
A110	8.014	125.81	8.36	127.53
W111	7.877	120.14	7.94	120.28
R112	7.932	122.10	8.04	122.54
N113	8.089	119.13	8.03	122.52
R114	7.932	121.92	8.18	125.62
A115	8.171	125.52	8.26	121.83
K116	8.225	121.55	8.40	111.08
G117	8.362	110.87	8.40	111.08
T118	8.055	114.25	8.08	114.45
D119	8.539	122.76	8.59	122.66
V120	8.089	121.95	8.10	122.02
Q121	8.348	124.40	8.37	124.53
A122	—	—	—	—
W123	8.034	120.80	8.00	120.84
I124	7.857	124.06	7.88	124.14
R125	—	—	8.20	125.72
G126	8.314	111.12	8.34	111.29
A127	8.055	124.73	8.08	124.93
R128	8.280	121.46	8.30	121.55
L129	8.314	126.32	8.40	125.69

A.4. Heteronuclear relaxation rates

Table A.10.: Relaxation rates (with errors) for *all-Ala*-HEWL

Residue	R_1 [s^{-1}]		R_2 [s^{-1}]		$R_{1\rho}$ [s^{-1}]		hetNOE	
M-1 ^{1,2}	—		—		—		—	
K1	0.94	± 0.02	0.96	± 0.02	1.28	± 0.03	-1.487	± -0.074
V2	1.22	± 0.02	1.27	± 0.03	1.48	± 0.03	-0.560	± -0.028
F3	1.34	± 0.03	1.53	± 0.03	1.75	± 0.04	-0.584	± -0.029
G4	1.37	± 0.03	1.61	± 0.03	1.83	± 0.04	-0.379	± -0.019
R5	1.53	± 0.03	2.17	± 0.04	2.37	± 0.05	0.003	± 0.000
A6	1.55	± 0.03	2.44	± 0.05	2.95	± 0.06	-0.081	± -0.004
E7	1.58	± 0.03	2.49	± 0.05	2.86	± 0.06	0.090	± 0.004
L8	1.63	± 0.03	2.55	± 0.05	2.97	± 0.06	0.017	± 0.001
A9 ³	—		—		—		—	
A10	1.68	± 0.03	3.01	± 0.06	3.35	± 0.07	0.163	± 0.008
A11	1.67	± 0.03	2.82	± 0.06	3.40	± 0.07	0.025	± 0.001
M12	1.65	± 0.03	2.75	± 0.05	3.11	± 0.06	-0.037	± -0.002
K13	1.71	± 0.03	2.94	± 0.06	3.17	± 0.06	0.084	± 0.004
R14	1.68	± 0.03	2.80	± 0.06	2.86	± 0.06	0.016	± 0.001
H15	1.65	± 0.03	2.63	± 0.05	2.82	± 0.06	-0.003	± 0.000
G16	1.57	± 0.03	2.32	± 0.05	2.68	± 0.05	-0.105	± -0.005
L17	1.63	± 0.03	2.54	± 0.05	2.81	± 0.06	0.129	± 0.006
D18	1.60	± 0.03	2.44	± 0.05	2.69	± 0.05	0.032	± 0.002
N19	1.64	± 0.03	2.56	± 0.05	2.77	± 0.06	0.197	± 0.010
Y20	1.69	± 0.03	2.85	± 0.06	2.84	± 0.06	0.079	± 0.004
R21	1.75	± 0.03	2.88	± 0.06	3.13	± 0.06	0.164	± 0.008
G22	1.64	± 0.03	2.60	± 0.05	2.89	± 0.06	0.071	± 0.004
Y23	1.69	± 0.03	2.93	± 0.06	3.20	± 0.06	0.294	± 0.015
S24	1.64	± 0.03	2.89	± 0.06	2.93	± 0.06	0.260	± 0.013
L25	1.73	± 0.03	2.88	± 0.06	3.01	± 0.06	0.198	± 0.010
G26	1.51	± 0.03	2.67	± 0.05	2.71	± 0.05	0.106	± 0.005
N27	1.59	± 0.03	2.84	± 0.06	2.98	± 0.06	0.163	± 0.008
W28	1.63	± 0.03	3.28	± 0.07	3.35	± 0.07	0.079	± 0.004
V29	1.66	± 0.03	3.52	± 0.07	3.65	± 0.07	0.082	± 0.004
A30	1.66	± 0.03	3.56	± 0.07	4.00	± 0.08	0.088	± 0.004
A31	1.51	± 0.03	3.12	± 0.06	3.13	± 0.06	0.054	± 0.003
A32	1.58	± 0.03	2.98	± 0.06	3.41	± 0.07	0.027	± 0.001

¹Not applicable²The N-terminal methionine has been introduced for expression in *E. coli*³Unassigned

A. Appendix

Table A.10 (continued)

Residue	R_1 [s^{-1}]			R_2 [s^{-1}]			$R_{1\rho}$ [s^{-1}]			hetNOE		
K33	1.57	±	0.03	3.19	±	0.06	3.34	±	0.07	0.037	±	0.002
F34	1.62	±	0.03	3.16	±	0.06	3.50	±	0.07	0.067	±	0.003
E35	1.64	±	0.03	3.18	±	0.06	3.38	±	0.07	0.016	±	0.001
S36	1.58	±	0.03	2.80	±	0.06	2.99	±	0.06	0.137	±	0.007
N37	1.59	±	0.03	2.74	±	0.05	3.03	±	0.06	0.027	±	0.001
F38	1.59	±	0.03	2.76	±	0.06	3.18	±	0.06	0.025	±	0.001
N39	1.59	±	0.03	2.80	±	0.06	3.03	±	0.06	0.031	±	0.002
T40	1.55	±	0.03	2.45	±	0.05	2.90	±	0.06	-0.055	±	-0.003
Q41	1.66	±	0.03	2.55	±	0.05	2.91	±	0.06	0.057	±	0.003
A42	1.55	±	0.03	2.46	±	0.05	2.77	±	0.06	0.125	±	0.006
T43	1.50	±	0.03	2.32	±	0.05	2.45	±	0.05	-0.121	±	-0.006
N44	1.63	±	0.03	2.47	±	0.05	2.85	±	0.06	-0.073	±	-0.004
R45	1.62	±	0.03	2.47	±	0.05	2.52	±	0.05	0.028	±	0.001
N46	1.62	±	0.03	2.52	±	0.05	2.72	±	0.05	0.019	±	0.001
T47	1.55	±	0.03	2.30	±	0.05	2.47	±	0.05	0.008	±	0.000
D48	1.63	±	0.03	2.45	±	0.05	2.64	±	0.05	0.051	±	0.003
G49	1.49	±	0.03	2.32	±	0.05	2.45	±	0.05	-0.111	±	-0.006
S50	1.53	±	0.03	2.35	±	0.05	2.49	±	0.05	0.052	±	0.003
T51	1.57	±	0.03	2.56	±	0.05	2.81	±	0.06	0.095	±	0.005
D52	1.62	±	0.03	2.81	±	0.06	3.07	±	0.06	-0.060	±	-0.003
Y53	1.63	±	0.03	3.20	±	0.06	3.63	±	0.07	-0.025	±	-0.001
G54	1.70	±	0.03	3.44	±	0.07	3.44	±	0.07	0.234	±	0.012
I55	1.70	±	0.03	3.58	±	0.07	3.81	±	0.08	0.283	±	0.014
L56	1.73	±	0.03	3.60	±	0.07	3.80	±	0.08	0.245	±	0.012
Q57	1.72	±	0.03	3.75	±	0.07	4.03	±	0.08	0.314	±	0.016
I58	1.64	±	0.03	3.72	±	0.07	4.01	±	0.08	0.049	±	0.002
N59	1.72	±	0.03	4.23	±	0.08	4.31	±	0.09	0.236	±	0.012
S60	1.67	±	0.03	4.10	±	0.08	4.13	±	0.08	0.345	±	0.017
R61	1.87	±	0.04	4.24	±	0.08	4.65	±	0.09	0.434	±	0.022
W62	1.85	±	0.04	4.28	±	0.09	4.50	±	0.09	0.345	±	0.017
W63	1.87	±	0.04	4.63	±	0.09	4.56	±	0.09	0.396	±	0.020
A64	1.83	±	0.04	3.97	±	0.08	4.22	±	0.08	0.295	±	0.015
N65	1.68	±	0.03	3.38	±	0.07	3.72	±	0.07	0.136	±	0.007
D66	1.66	±	0.03	3.18	±	0.06	3.58	±	0.07	0.278	±	0.014
G67	1.65	±	0.03	2.85	±	0.06	3.04	±	0.06	0.194	±	0.010
R68	1.71	±	0.03	2.98	±	0.06	3.20	±	0.06	0.231	±	0.012
T69	1.51	±	0.03	2.82	±	0.06	2.95	±	0.06	0.202	±	0.010

Table A.10 (continued)

Residue	R_1 [s^{-1}]			R_2 [s^{-1}]			$R_{1\rho}$ [s^{-1}]			hetNOE		
P70 ¹	—			—			—			—		
G71	1.54	±	0.03	2.54	±	0.05	2.58	±	0.05	-0.050	±	-0.002
S72	1.54	±	0.03	2.56	±	0.05	2.75	±	0.05	0.031	±	0.002
R73	1.61	±	0.03	2.59	±	0.05	2.81	±	0.06	-0.092	±	-0.005
N74	1.51	±	0.03	2.52	±	0.05	2.81	±	0.06	-0.171	±	-0.009
L75	1.55	±	0.03	2.58	±	0.05	3.07	±	0.06	0.085	±	0.004
A76	1.60	±	0.03	2.75	±	0.05	3.40	±	0.07	0.087	±	0.004
N77	1.48	±	0.03	2.61	±	0.05	2.68	±	0.05	0.089	±	0.004
I78	1.51	±	0.03	2.72	±	0.05	3.03	±	0.06	-0.027	±	-0.001
P79 ¹	—			—			—			—		
A80	1.54	±	0.03	2.88	±	0.06	3.03	±	0.06	-0.001	±	0.000
S81	1.51	±	0.03	2.91	±	0.06	3.08	±	0.06	0.137	±	0.007
A82	1.60	±	0.03	3.06	±	0.06	3.33	±	0.07	0.154	±	0.008
L83	1.55	±	0.03	2.79	±	0.06	3.08	±	0.06	0.021	±	0.001
L84	1.52	±	0.03	2.96	±	0.06	2.92	±	0.06	-0.127	±	-0.006
S85	1.49	±	0.03	2.69	±	0.05	2.83	±	0.06	0.084	±	0.004
S86	1.50	±	0.03	2.64	±	0.05	2.81	±	0.06	0.109	±	0.005
D87	1.59	±	0.03	2.68	±	0.05	2.84	±	0.06	-0.166	±	-0.008
I88	1.56	±	0.03	2.80	±	0.06	2.77	±	0.06	-0.085	±	-0.004
T89	1.53	±	0.03	2.90	±	0.06	3.00	±	0.06	-0.128	±	-0.006
A90	1.48	±	0.03	2.75	±	0.05	2.64	±	0.05	0.103	±	0.005
S91	1.37	±	0.03	2.40	±	0.05	2.56	±	0.05	-0.058	±	-0.003
V92	1.46	±	0.03	2.69	±	0.05	2.62	±	0.05	-0.055	±	-0.003
N93	1.51	±	0.03	2.85	±	0.06	2.70	±	0.05	-0.271	±	-0.014
A94	1.51	±	0.03	2.61	±	0.05	2.83	±	0.06	0.037	±	0.002
A95	1.44	±	0.03	2.63	±	0.05	2.68	±	0.05	-0.115	±	-0.006
K96	1.48	±	0.03	2.70	±	0.05	3.01	±	0.06	-0.216	±	-0.011
K97	1.48	±	0.03	2.64	±	0.05	2.76	±	0.06	-0.030	±	-0.002
I98	1.58	±	0.03	2.59	±	0.05	2.74	±	0.05	0.073	±	0.004
V99	1.52	±	0.03	2.71	±	0.05	2.92	±	0.06	-0.127	±	-0.006
S100	1.56	±	0.03	2.54	±	0.05	2.81	±	0.06	-0.222	±	-0.011
D101	1.48	±	0.03	2.34	±	0.05	2.63	±	0.05	0.015	±	0.001
G102	1.40	±	0.03	2.48	±	0.05	2.09	±	0.04	-0.284	±	-0.014
N103	1.49	±	0.03	2.13	±	0.04	2.11	±	0.04	-0.145	±	-0.007
G104	1.47	±	0.03	2.75	±	0.06	2.58	±	0.05	-0.192	±	-0.010
M105	1.58	±	0.03	2.52	±	0.05	2.79	±	0.06	0.121	±	0.006
N106	1.62	±	0.03	2.80	±	0.06	2.87	±	0.06	-0.032	±	-0.002

¹Not applicable

A. Appendix

Table A.10 (*continued*)

Residue	R₁ [s⁻¹]			R₂ [s⁻¹]			R_{1ρ} [s⁻¹]			hetNOE		
A107	1.72	±	0.03	2.71	±	0.05	3.08	±	0.06	0.237	±	0.012
W108	1.78	±	0.04	3.54	±	0.07	3.90	±	0.08	0.252	±	0.013
V109	1.85	±	0.04	3.81	±	0.08	4.22	±	0.08	0.231	±	0.012
A110	1.86	±	0.04	4.04	±	0.08	4.34	±	0.09	0.257	±	0.013
W111	1.87	±	0.04	3.82	±	0.08	3.87	±	0.08	0.418	±	0.021
R112	1.87	±	0.04	3.75	±	0.07	4.10	±	0.08	0.379	±	0.019
N113	1.79	±	0.04	3.58	±	0.07	3.89	±	0.08	0.181	±	0.009
R114	1.76	±	0.04	3.09	±	0.06	3.38	±	0.07	0.316	±	0.016
A115	1.65	±	0.03	2.56	±	0.05	2.83	±	0.06	0.070	±	0.004
K116	1.67	±	0.03	2.64	±	0.05	3.04	±	0.06	0.157	±	0.008
G117	1.54	±	0.03	2.16	±	0.04	2.27	±	0.05	-0.076	±	-0.004
T118	1.55	±	0.03	2.26	±	0.05	2.46	±	0.05	-0.063	±	-0.003
D119	1.51	±	0.03	2.28	±	0.05	2.36	±	0.05	0.081	±	0.004
V120	1.58	±	0.03	2.42	±	0.05	2.30	±	0.05	-0.107	±	-0.005
Q121	1.68	±	0.03	2.57	±	0.05	2.91	±	0.06	-0.011	±	-0.001
A122 ¹	—			—			—			—		
W123	1.68	±	0.03	2.57	±	0.05	2.74	±	0.05	0.167	±	0.008
I124	1.70	±	0.03	2.63	±	0.05	2.95	±	0.06	0.097	±	0.005
R125	1.69	±	0.03	2.58	±	0.05	3.06	±	0.06	0.076	±	0.004
G126	1.48	±	0.03	1.99	±	0.04	2.14	±	0.04	-0.165	±	-0.008
A127	1.43	±	0.03	1.73	±	0.03	1.84	±	0.04	-0.404	±	-0.020
R128	1.25	±	0.03	1.49	±	0.03	1.60	±	0.03	-0.387	±	-0.019
L129	1.08	±	0.02	1.22	±	0.02	1.31	±	0.03	-1.137	±	-0.057

¹Unassigned

A.4. Heteronuclear relaxation rates

Table A.11.: ^{15}N transverse relaxation rates (with errors) for W108G-*all-Ala*-HEWL

Residue	R_2 [s^{-1}]			Residue	R_2 [s^{-1}]			Residue	R_2 [s^{-1}]		
M-1	—	—	—	S36	4.81	±	0.10	S72	3.41	±	0.07
K1	2.06	±	0.04	N37	—	—	—	R73	3.74	±	0.07
V2	2.09	±	0.04	F38	4.53	±	0.09	N74	—	—	—
F3	2.45	±	0.05	N39	—	—	—	L75	3.18	±	0.06
G4	2.91	±	0.06	T40	3.82	±	0.08	A76	3.82	±	0.08
R5	3.25	±	0.07	Q41	4.05	±	0.08	N77	—	—	—
A6	4.36	±	0.09	A42	3.50	±	0.07	I78	3.77	±	0.08
E7	3.91	±	0.08	T43	3.13	±	0.06	P79	—	—	—
L8	3.79	±	0.08	N44	3.20	±	0.06	A80	3.77	±	0.08
A9	—	—	—	R45	3.66	±	0.07	S81	4.33	±	0.09
A10	—	—	—	N46	3.45	±	0.07	A82	4.07	±	0.08
A11	—	—	—	T47	3.56	±	0.07	L83	3.64	±	0.07
M12	4.20	±	0.08	D48	3.11	±	0.06	L84	3.95	±	0.08
K13	—	—	—	G49	3.23	±	0.06	S85	4.00	±	0.08
R14	4.76	±	0.10	S50	3.44	±	0.07	S86	—	—	—
H15	3.49	±	0.07	T51	3.55	±	0.07	D87	3.49	±	0.07
G16	—	—	—	D52	3.75	±	0.08	I88	—	—	—
L17	—	—	—	Y53	4.77	±	0.10	T89	3.62	±	0.07
D18	3.29	±	0.07	G54	4.89	±	0.10	A90	3.70	±	0.07
N19	3.78	±	0.08	I55	5.13	±	0.10	S91	2.98	±	0.06
Y20	3.92	±	0.08	L56	3.27	±	0.07	V92	3.50	±	0.07
R21	—	—	—	Q57	5.43	±	0.11	N93	3.49	±	0.07
G22	3.73	±	0.07	I58	—	—	—	A94	3.50	±	0.07
Y23	4.16	±	0.08	N59	5.92	±	0.12	A95	2.83	±	0.06
S24	4.62	±	0.09	S60	6.33	±	0.13	K96	3.85	±	0.08
L25	4.90	±	0.10	R61	5.69	±	0.11	K97	3.16	±	0.06
G26	4.40	±	0.09	W62	5.98	±	0.12	I98	3.20	±	0.06
N27	3.53	±	0.07	W63	6.61	±	0.13	V99	3.33	±	0.07
W28	3.82	±	0.08	A64	6.09	±	0.12	S100	—	—	—
V29	5.46	±	0.11	N65	4.78	±	0.10	D101	—	—	—
A30	5.70	±	0.11	D66	4.14	±	0.08	G102	2.55	±	0.05
A31	4.98	±	0.10	G67	3.97	±	0.08	N103	2.11	±	0.04
A32	5.15	±	0.10	R68	3.97	±	0.08	G104	2.49	±	0.05
K33	4.98	±	0.10	T69	—	—	—	M105	2.78	±	0.06
F34	5.58	±	0.11	P70	—	—	—	N106	4.97	±	0.10
E35	—	—	—	G71	3.76	±	0.08	A107	4.73	±	0.09

A. Appendix

Table A.11 (*continued*)

Residue	R₂ [s⁻¹]			Residue	R₂ [s⁻¹]			Residue	R₂ [s⁻¹]		
W108	2.67	±	0.05	K116	2.77	±	0.06	W123	3.31	±	0.07
V109	3.49	±	0.07	G117	3.25	±	0.07	I124	2.90	±	0.06
A110	3.59	±	0.07	T118	2.55	±	0.05	R125	3.42	±	0.07
W111	2.79	±	0.06	D119	—	—	—	G126	2.92	±	0.06
R112	—	—	—	V120	2.99	±	0.06	A127	2.46	±	0.05
N113	—	—	—	Q121	3.28	±	0.07	R128	2.21	±	0.04
R114	—	—	—	A122	—	—	—	L129	1.95	±	0.04
A115	3.21	±	0.06								

A.5. Residual dipolar couplings

Table A.12.: Residual dipolar couplings for *all-Ala*-HEWL and W62G-*all-Ala*-HEWL in 7% stretched polyacrylamide gels.

Residue	<i>all-Ala</i> -HEWL $^1D(H^N,N)$ [Hz]	W62G- <i>all-Ala</i> -HEWL $^1D(H^N,N)$ [Hz]
M-1	—	—
K1	-4.86	-3.33
V2	-6.60	-3.75
F3	-2.58	-2.08
G4	-1.16	0.00
R5	-0.05	0.41
A6	—	2.48
E7	1.03	0.42
L8	-2.08	0.42
A9	—	—
A10	—	—
A11	3.39	0.83
M12	-0.50	-1.25
K13	—	—
R14	—	-3.09
H15	-3.90	-3.32
G16	-4.82	-1.25
L17	-2.85	-1.67
D18	-2.74	—
N19	-1.88	-1.66
Y20	0.74	0.00
R21	-0.33	-1.25
G22	-0.78	-2.08
Y23	-0.10	-1.66
S24	-2.31	-1.25
L25	—	-2.91
G26	9.40	-0.83
N27	-2.79	-1.66
W28	-0.89	—
V29	3.25	0.00
A30	3.58	-0.41
A31	-2.64	—
A32	-0.88	-0.83
K33	-0.54	—

A. Appendix

Table A.12 (continued)

Residue	<i>all-Ala-HEWL</i> $^1\text{D}(\text{H}^{\text{N}},\text{N})$ [Hz]	W62G- <i>all-Ala-HEWL</i> $^1\text{D}(\text{H}^{\text{N}},\text{N})$ [Hz]
F34	-2.19	-0.83
E35	-1.28	-1.66
S36	-1.66	-1.25
N37	—	—
F38	-1.46	-0.41
N39	—	—
T40	-1.99	-1.25
Q41	-1.97	-1.25
A42	-0.08	-0.01
T43	-2.04	-1.66
N44	—	—
R45	-2.76	-1.66
N46	-2.08	-1.25
T47	-3.45	-2.08
D48	-1.95	-1.25
G49	-0.01	-0.01
S50	-0.75	-0.41
T51	-2.07	-2.09
D52	—	-2.50
Y53	—	0.83
G54	-0.21	0.00
I55	0.04	1.25
L56	-3.41	-2.07
Q57	-3.43	-1.66
I58	—	-1.25
N59	-3.71	-1.66
S60	-0.33	-1.25
R61	-2.09	-0.41
W62	8.93	0.00
W63	7.44	-2.50
A64	7.98	-0.42
N65	-3.40	-1.25
D66	-2.21	-1.99
G67	1.25	-0.41
R68	1.71	1.25
T69	—	0.41
P70	—	—

A.5. Residual dipolar couplings

Table A.12 (continued)

Residue	<i>all-Ala</i> -HEWL $^1D(H^N, N)$ [Hz]	W62G- <i>all-Ala</i> -HEWL $^1D(H^N, N)$ [Hz]
G71	-2.53	0.41
S72	-0.66	-1.66
R73	-3.68	-2.49
N74	—	—
L75	-1.20	-2.49
A76	-1.08	-0.83
N77	—	-2.08
I78	-7.41	-4.57
P79	—	—
A80	-3.25	-3.50
S81	-0.29	-2.08
A82	0.58	-0.01
L83	-3.15	-1.25
L84	-2.01	-1.66
S85	-1.87	-0.83
S86	-0.28	-1.66
D87	-1.40	
I88	-0.65	-0.41
T89	-3.28	-4.58
A90	-0.45	-0.83
S91	0.08	-0.42
V92	-0.66	-0.83
N93	-1.78	-1.67
A94	0.05	-0.29
A95	0.88	-0.83
K96	-1.89	-3.74
K97	-5.06	-2.91
I98	—	—
V99	—	—
S100	—	—
D101	-2.22	-2.49
G102	0.75	-2.49
N103	-2.24	-1.32
G104	-1.33	0.00
M105	0.71	0.00
N106	0.35	0.41
A107	2.33	0.83

A. Appendix

Table A.12 (continued)

Residue	<i>all-Ala</i> -HEWL $^1\text{D}(\text{H}^{\text{N}},\text{N})$ [Hz]	W62G- <i>all-Ala</i> -HEWL $^1\text{D}(\text{H}^{\text{N}},\text{N})$ [Hz]
W108	8.17	-2.08
V109	7.48	3.75
A110	5.28	1.66
W111	-0.33	-0.41
R112	—	0.41
N113	-2.08	4.99
R114	—	0.42
A115	-0.25	-1.25
K116	-2.73	-0.83
G117	-4.57	-2.50
T118	-2.03	-1.25
D119	-2.37	-2.91
V120	-3.63	-1.67
Q121	-3.61	-2.08
A122	—	—
W123	1.71	4.16
I124	3.83	2.08
R125	—	—
G126	-5.36	-3.32
A127	0.16	0.41
R128	-3.77	-2.50
L129	-1.30	-3.75

A.6. Pulse programs

All pulse sequences have been written in the standard Bruker pulse sequence programming language.

A.6.1. The 2D ^{15}N - ^1H photo-CIDNP pulse program

```
; 2dcindpcs2
; 2D photo-CIDNP with 15-N/1-H correlation via inept transfer
; by C. Schlörb, B. Fürtig, C. Richter & H. Schwalbe
; Center for Biomolecular Magnetic Resonance, JWG University of Frankfurt, Germany
; following Lyon et al. (JACS 1999, 121, 6505)

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

"p2=p1*2"
"p22=p21*2"

"d0=3u"
"d26=1s/(cnst4*4)"
"d11=30m"
"d13=4u"

"DELTA1=d26-p2-d0*2"
"DELTA3=d19-p22/2"
"DELTA4=d26-p16-d16-p27*3-d19*5"
"DELTA5=d26-p16-d16-p27*2-p0-d19*5-8u"

1 ze
  d11 p116:f3
2 d1 do:f3 do:f2 p18:f3
  50u UNBLKGRAD
  (p8 ph9):f3
  p17:gp4
  10u setnmr3|14 ;for use on AV700
  d27
  10u setnmr3^14 ;for use on AV700
  d28 p11:f1
3 d11 p13:f3
  (p21 ph3):f3
  d0
  (p2 ph5)
  d0
  DELTA1
  (center (p2 ph1) (p22 ph6):f3 )
  d26
  (p21 ph20):f3
```

A. Appendix

```
4u
p16:gp2
d16
(p1 ph21)
DELTA4
p16:gp3
d16 p118:f1
p27*0.231 ph22
d19*2
p27*0.692 ph22
d19*2
p27*1.462 ph22
DELTA3
(p22 ph6):f3
DELTA3
p27*1.462 ph23
d19*2
p27*0.692 ph23
d19*2
p0*0.231 ph23
4u
p16:gp3
d16
4u BLKGRAD
DELTA5 p116:f3
go=2 ph31 cpd3:f3
dl do:f3 do:f2 mc #0 to 2
    F1PH(ip20 & ip6, id0)
exit

ph1=0
ph3=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph9=0
ph20=1 3
ph21=2
ph22=1
ph23=3
ph31=0 2 0 2 2 0 2 0

;p11 : f1 channel - power level for pulse (default)
;p13 : f3 channel - power level for pulse (default)
;p18 : f3 channel - power level for pulse (default)
;p116: f3 channel - power level for CPD/BB decoupling
;p118: f1 channel - power level for 3-9-19-pulse (watergate)
;p1  : f1 channel - 90 degree high power pulse
;p2  : f1 channel - 180 degree high power pulse
;p8  : f1 channel - 90 degree high power pulse
;p21 : f3 channel - 90 degree high power pulse
```

A.6. Pulse programs

```
;p22 : f3 channel - 180 degree high power pulse
;p16 : homospoil/gradient pulse
;p17 : homospoil/gradient pulse
;p27 : f1 channel - 90 degree pulse at p118
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11 : delay for disk I/O [30 msec]
;d13 : short delay [4 usec]
;d16 : delay for homospoil/gradient recovery
;d19 : delay for binomial water suppression
;d26 : 1/(4J(NH))
;d27 : length of laser pulse
;d28 : delay after laser pulse
;cnst4: J(NH)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio: gp 2 : gp 3 : gp4
; 80 : 30 : 80

;for z-only gradients:
;gpz2: 80%
;gpz3: 30%
;gpz4: 80%

;use gradient files:
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100

; use setnmr0|15 and setnmr0^15 on DRX599
; use setnmr3|14 and setnmr3^14 on AV700

; cs061122
```

A. Appendix

A.6.2. The 2D ^{13}C - ^1H photo-CIDNP pulse program

```
; 2dcindpcs2_13c
; 2D photo-CIDNP with 13-C/H-1 correlation via inept transfer
; by C. Schlörb, B. Fürtig, C. Richter & H. Schwalbe
; Center for Biomolecular Magnetic Resonance, JWG University of Frankfurt, Germany
; following Lyon et al. (JACS 1999, 121, 6505)

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

"p2=p1*2"
"p4=p3*2"

"d0=3u"
"d4=1s/(cnst4*4)"
"d11=30m"
"d13=4u"

"DELTA1=d4-p2-d0*2"
"DELTA3=d19-p4/2"
"DELTA4=d4-p16-d16-p27*3-d19*5"
"DELTA5=d4-p16-d16-p27*2-p0-d19*5-8u"

1 ze
  d11 p112:f2
2 d1 do:f2
  4u p18:f2
  50u UNBLKGRAD
  (p8 ph9):f2
  p17:gp4
  10u setnmr3|14      ;for use on AV700
  d27
  10u setnmr3^14     ;for use on AV700
  d28 p11:f1
3 d11 p12:f2
  (p3 ph3):f2
  d0
  (p2 ph5)
  d0
  DELTA1
  (center (p2 ph1) (p4 ph6):f2 )
  d4
  (p3 ph20):f2
  4u
  p16:gp2
  d16
  (p1 ph21)
  DELTA4
  p16:gp3
```

A.6. Pulse programs

```
d16 p118:f1
p27*0.231 ph22
d19*2
p27*0.692 ph22
d19*2
p27*1.462 ph22
DELTA3
(p4 ph6):f2
DELTA3
p27*1.462 ph23
d19*2
p27*0.692 ph23
d19*2
p0*0.231 ph23
4u
p16:gp3
d16
4u BLKGRAD
DELTA5 p112:f2
go=2 ph31 cpd2:f2
d1 do:f2 mc #0 to 2
    F1PH(ip20 & ip6, id0)
exit
```

```
ph1=0
ph3=0 0 0 0 2 2 2 2
ph5=0 0 2 2
ph6=0
ph9=0
ph20=1 3
ph21=2
ph22=1
ph23=3
ph31=0 2 0 2 2 0 2 0
```

```
;p11 : f1 channel - power level for pulse (default)
;p12 : f2 channel - power level for pulse (default)
;p18 : f2 channel - power level for pulse (default)
;p112: f2 channel - power level for CPD/BB decoupling
;p118: f1 channel - power level for 3-9-19-pulse (watergate)
;p1  : f1 channel - 90 degree high power pulse
;p2  : f1 channel - 180 degree high power pulse
;p3  : f2 channel - 90 degree high power pulse
;p4  : f2 channel - 180 degree high power pulse
;p8  : f1 channel - 90 degree high power pulse
;p16 : homospoil/gradient pulse
;p17 : homospoil/gradient pulse
;p27 : f1 channel - 90 degree pulse at p118
;d0  : incremented delay (2D) [3 usec]
;d1  : relaxation delay; 1-5 * T1
```

A. Appendix

```
;d11 : delay for disk I/O                [30 msec]
;d13 : short delay                       [4 usec]
;d16 : delay for homospoil/gradient recovery
;d19 : delay for binomial water suppression
;d26 : 1/(4J(NH))
;d27 : length of laser pulse
;d28 : delay after laser pulse
;cnst4: J(NH)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;use gradient ratio: gp 2 : gp 3 : gp4
;                      80 : 30 : 80

;for z-only gradients:
;gpz2: 80%
;gpz3: 30%
;gpz4: 80%

;use gradient files:
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100

; use setnmr0|15 and setnmr0^15 on DRX599
; use setnmr3|14 and setnmr3^14 on AV700

; cs070112
```

A.6.3. The HN(CACB)CG pulse program

```

; hncacbcggpwg3d
; 3D sequence for the assignment of tryptophan side-chains
;      F1(H) -> F3(N) -> F2(Ca) -> F2(Cb) -> F2(Cg,t1) ->
; F2(Cb) -> F2(Ca) -> F3(N,t2) -> F1(H,t3)
;
; water suppression using watergate sequence

; by C. Schlörb, C. Richter & H. Schwalbe
; Center for Biomolecular Magnetic Resonance, JWG University of Frankfurt, Germany
; Schlörb et al. (J. Biomol. NMR 2005, 33, 95)

prosol relations=<triple>

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

"p2=p1*2"
"p22=p21*2"

"d0=3u"
"d11=30m"
"d12=20u"
"d13=4u"
"d21=4.9m"
"d23=13.7m"
"d26=2.3m"

"in29=in10"
"in30=in10"

"d10=d23/2-p24/2"
"d29=d23/2-p24/2-p26-d21-4u"
"d30=d23/2-p24/2"

"DELTA1=d23/2-d21-p26-p24/2"
"DELTA2=d26-p16-d16-p11-12u"
"DELTA3=d23/2-p24/2"

"spoff2=0"
"spoff3=0"
"spoff5=0"
"spoff8=0"
"spoff10=bf2*(cnst20/1000000)-bf2*(cnst22/1000000)"

aqseq 321

1 d11 ze
  d11 p116:f3

```

A. Appendix

```
2 d11 do:f3
3 d1 p11:f1
  d12 fq=cnst22(bf ppm):f2
  p1 ph1
  d26 p13:f3
  (center (p2 ph1) (p22 ph1):f3 )
  d26 UNBLKGRAD
  (p1 ph2):f1

4u p10:f1
  (p11:sp1 ph1:r):f1
4u
p16:gp1
d16

(p21 ph3):f3
d21 p119:f1
(p26 ph2):f1
DELTA1 cpds1:f1 ph1
(p24:sp10 ph1):f2
DELTA3
(center (p14:sp3 ph1):f2 (p22 ph1):f3 )
DELTA3
(p24:sp10 ph1):f2
DELTA3
(p21 ph1):f3
d12 fq=cnst23(bf ppm):f2
(p13:sp2 ph9):f2
d22
(p14:sp3 ph1):f2
d22
(p13:sp8 ph4):f2
d19
d12 fq=cnst24(bf ppm):f2
4u
(p23:sp9 ph10):f2
d0
(p25:sp5 ph1):f2
d0
(p23:sp11 ph11):f2
4u
d12 fq=cnst23(bf ppm):f2
d19
(p13:sp2 ph2):f2
d22
(p14:sp3 ph1):f2
d22
(p13:sp8 ph1):f2
4u do:f1
d12 fq=cnst22(bf ppm):f2
(p26 ph7):f1
```


A.6. Pulse programs

```
4u
p16:gp2
d16
(p26 ph2):f1
20u cpds1:f1 ph1

(p21 ph5):f3
d30
(p24:sp10 ph1):f2
d30
(center (p14:sp3 ph1):f2 (p22 ph8):f3 )
d10
(p24:sp10 ph1):f2
d29
4u do:f1
(p26 ph7):f1
d21
(p21 ph1):f3

p16:gp3
d16 p10:f1
(p11:sp1 ph6):f1
4u
4u p11:f1

(p1 ph1)
4u
p16:gp4
d16
DELTA2 p10:f1
(p11:sp1 ph6):f1
4u
4u p11:f1
(center (p2 ph1) (p22 ph1):f3 )
4u p10:f1
(p11:sp1 ph6):f1
4u
DELTA2
p16:gp4
d16 p116:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
d11 do:f3 mc #0 to 2
    F1PH(rd10 & rd29 & rd30 & ip10, id0)
    F2PH(ip5, id10 & id29 & dd30)
exit

ph1=0
ph2=1
ph3=0 0 0 0 2 2 2 2
ph4=1 3
```

A. Appendix

```

ph5=0 0 2 2
ph6=2
ph7=3
ph8=0
ph9=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
      2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph10=0 0 0 0 0 0 0 0 2 2 2 2 2 2 2 2
ph11=0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
      0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
      2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
      2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
ph31=0 0 2 2 2 2 0 0 2 2 0 0 0 0 2 2
      2 2 0 0 0 0 2 2 0 0 2 2 2 2 0 0
      2 2 0 0 0 0 2 2 0 0 2 2 2 2 0 0
      0 0 2 2 2 2 0 0 2 2 0 0 0 0 2 2

;p10 : 120dB
;p11 : f1 channel - power level for pulse (default)
;p13 : f3 channel - power level for pulse (default)
;p116: f3 channel - power level for CPD/BB decoupling
;p119: f1 channel - power level for CPD/BB decoupling
;sp1 : f1 channel - shaped pulse 90 degree (H2O on resonance)
;sp2 : f2 channel - shaped pulse 90 degree (Ca and Cb on resonance)
;sp3 : f2 channel - shaped pulse 180 degree (Ca and Cb on resonance)
;sp5 : f2 channel - shaped pulse 180 degree (Ca and Cb on resonance)
;sp8 : f2 channel - shaped pulse 90 degree (Ca and Cb on resonance)
;sp9 : f2 channel - shaped pulse 90 degree (Cg on resonance)
;sp10: f2 channel - shaped pulse 180 degree (C' on resonance)
;sp11: f2 channel - shaped pulse 90 degree (Cg on resonance)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p11 : f1 channel - 90 degree shaped pulse [1 msec]
;p13 : f2 channel - 90 degree shaped pulse
;p14 : f2 channel - 180 degree shaped pulse
;p16 : homospoil/gradient pulse [1 msec]
;p21 : f3 channel - 90 degree high power pulse
;p22 : f3 channel - 180 degree high power pulse
;p23 : f2 channel - 90 degree shaped pulse on Ca and Cb
;p24 : f2 channel - 180 degree C' [1 msec]
;p25 : f2 channel - 180 degree in t1 time [1 msec]
;p26 : f1 channel - 90 degree pulse at p119
;d0 : incremented delay (F1 in 3D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d10 : incremented delay (F2 in 3D) = d23/2-p14/2
;d11 : delay for disk I/O [30 msec]
;d12 : short delay [20 usec]
;d13 : short delay [4 usec]
;d16 : delay for homospoil/gradient recovery
;d19 : 1/(2J(CbCg)) [12.5 msec]
;d21 : 1/(2J(NH)) [4.9 msec]
;d23 : 1/(4J(NCa)) [13.7 msec]

```

A.6. Pulse programs

```
;d26 : 1/(4J(NH)) [2.3 msec]
;d29 : incremented delay (F2 in 3D) = d23/2-p24/2-p26-d21-4u
;d30 : decremented delay (F2 in 3D) = d23/2-p24/2
;cnst20: C' chemical shift [176 ppm]
;cnst22: Ca chemical shift exact [54 ppm]
;cnst23: Ca/Cb chemical shift exact [41 ppm]
;cnst24: middle of C_delta and C_gamma [108 ppm]
;o2p: Cb chemical shift [27 ppm]
;in0: 1/(2 * SW(Ca)) = DW(Ca)
;nd0: 2
;in10: 1/(4 * SW(N)) = (1/2) DW(N)
;nd10: 4
;in29: = in10
;in30: = in10
;NS: 8 * n
;DS: >= 16
;td1: number of experiments in F1
;td2: number of experiments in F2 td2 max = 2 * d30 / in30
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: States-TPPI (or TPPI) in F2
;cpd31: decoupling according to sequence defined by cpdprg1
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd1: f1 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio: gp 1 : gp 2 : gp 3 : gp 4
; 50 : 40 : 60 : 30

;for z-only gradients:
;gpz1: 50%
;gpz2: 40%
;gpz3: 60%
;gpz4: 30%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100

; cs061122
```

A. Appendix

A.6.4. The HN(CD)CG pulse program

```
; hncdggpwg3d
; 3D sequence for the assignment of tryptophan side-chains
;   F1(H) -> F3(N) -> F2(Cd) -> F2(Cg,t1) -> F2(Cd) -> F3(N,t2) -> F1(H,t3)
;
; water suppression using watergate sequence

; by C. Schlörb, C. Richter & H. Schwalbe
; Center for Biomolecular Magnetic Resonance, JWG University of Frankfurt, Germany
; Schlörb et al. (J. Biomol. NMR 2005, 33, 95)

prosol relations=<triple>

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

"p2=p1*2"
"p22=p21*2"

"d0=3u"
"d11=30m"
"d13=4u"
"d21=4.9m"
"d23=13.7m"
"d26=2.3m"

"in29=in10"
"in30=in10"

"d10=d23/2-p24/2"
"d29=d23/2-p24/2-p26-d21-4u"
"d30=d23/2-p24/2"

"DELTA=d0*2+larger(p25,p22)-p14"
"DELTA1=d23/2-d21-p26-p24/2"
"DELTA2=d26-p16-d16-p11-12u"
"DELTA3=d23/2-p24/2"

"spoff2=bf2*(cnst24/1000000)-bf2*(cnst23/1000000)"
"spoff3=0"
"spoff5=bf2*(cnst24/1000000)-bf2*(cnst23/1000000)"
"spoff8=bf2*(cnst24/1000000)-bf2*(cnst23/1000000)"
"spoff10=bf2*(cnst20/1000000)-bf2*(cnst22/1000000)"
"spoff11=bf2*(cnst21/1000000)-bf2*(cnst23/1000000)"

aqseq 321

1 d11 ze
  d11 p116:f3
```

A.6. Pulse programs

```
2 d11 do:f3
3 d1 p11:f1
  d12 fq=cnst22(bf ppm):f2
  p1 ph1
  d26 p13:f3
  (center (p2 ph1) (p22 ph1):f3 )
  d26 UNBLKGRAD
  (p1 ph2):f1

4u p10:f1
  (p11:sp1 ph1:r):f1
4u
p16:gp1
d16

(p21 ph3):f3
d21 p119:f1
(p26 ph2):f1
DELTA1 cpds1:f1 ph1
(p24:sp10 ph1):f2
DELTA3
(center (p14:sp3 ph1):f2 (p22 ph1):f3 )
DELTA3
(p24:sp10 ph1):f2
DELTA3
(p21 ph1):f3
d12 fq=cnst23(bf ppm):f2
(p13:sp2 ph9):f2
d22
(p14:sp5 ph1):f2
d22
(p13:sp8 ph4):f2
d0
(center (p25:sp11 ph1):f2 (p22 ph8):f3 )
d0
4u
(p14:sp3 ph1):f2
DELTA
(p25:sp11 ph1):f2
4u
(p13:sp2 ph2):f2
d22
(p14:sp5 ph1):f2
d22
(p13:sp8 ph1):f2
4u do:f1
d12 fq=cnst22(bf ppm):f2
(p26 ph7):f1
4u
p16:gp2
d16
```

A. Appendix

```
(p26 ph2):f1
20u cpds1:f1 ph1

(p21 ph5):f3
d30
(p24:sp10 ph1):f2
d30
(center (p14:sp3 ph1):f2 (p22 ph8):f3 )
d10
(p24:sp10 ph1):f2
d29
4u do:f1
(p26 ph7):f1
d21
(p21 ph1):f3

p16:gp3
dl6 p10:f1
(p11:sp1 ph6):f1
4u
4u p11:f1

(p1 ph1)
4u
p16:gp4
dl6
DELTA2 p10:f1
(p11:sp1 ph6):f1
4u
4u p11:f1
(center (p2 ph1) (p22 ph1):f3 )
4u p10:f1
(p11:sp1 ph6):f1
4u
DELTA2
p16:gp4
dl6 p116:f3
4u BLKGRAD
go=2 ph31 cpd3:f3
dll do:f3 mc #0 to 2
    F1PH(rd10 & rd29 & rd30 & ip4 & ip9, id0)
    F2PH(ip5, id10 & id29 & dd30)
exit

ph1=0
ph2=1
ph3=0 0 0 0 0 0 0 2 2 2 2 2 2 2
ph4=1 3
ph5=0 0 2 2
ph6=2
ph7=3
```

A.6. Pulse programs

```
ph8=0 0 0 0 2 2 2 2
ph9=0
ph31=0 0 2 2 0 0 2 2 2 2 0 0 2 2 0 0

;p10 : 120dB
;p11 : f1 channel - power level for pulse (default)
;p13 : f3 channel - power level for pulse (default)
;p116: f3 channel - power level for CPD/BB decoupling
;p119: f1 channel - power level for CPD/BB decoupling
;sp1 : f1 channel - shaped pulse 90 degree (H2O on resonance)
;sp2 : f2 channel - shaped pulse 90 degree (Cg and Cd on resonance)
;sp3 : f2 channel - shaped pulse 180 degree (Cd on resonance)
;sp5 : f2 channel - shaped pulse 180 degree (Cg and Cd on resonance)
;sp8 : f2 channel - shaped pulse 90 degree (Cg and Cd on resonance)
;sp10: f2 channel - shaped pulse 180 degree (Ce on resonance)
;sp11: f2 channel - shaped pulse 180 degree (Cb and Cd on resonance)
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p11 : f1 channel - 90 degree shaped pulse [1 msec]
;p13 : f2 channel - 90 degree shaped pulse
;p14 : f2 channel - 180 degree shaped pulse
;p16 : homospoil/gradient pulse [1 msec]
;p21 : f3 channel - 90 degree high power pulse
;p22 : f3 channel - 180 degree high power pulse
;p24 : f2 channel - 180 degree Ce [1 msec]
;p25 : f2 channel - 180 degree in t1 time [1 msec]
;p26 : f1 channel - 90 degree pulse at p119
;d0 : incremented delay (F1 in 3D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d10 : incremented delay (F2 in 3D) = d23/2-p24/2
;d11 : delay for disk I/O [30 msec]
;d12 : short delay [20 usec]
;d13 : short delay [4 usec]
;d16 : delay for homospoil/gradient recovery
;d21 : 1/(2J(NH)) [4.9 msec]
;d22 : 1/(4J(CdCg)) [3.2 msec]
;d23 : 1/(4J(NCd)) [13.7 msec]
;d26 : 1/(4J(NH)) [2.3 msec]
;d29 : incremented delay (F2 in 3D) = d23/2-p24/2-p26-d21-4u
;d30 : decremented delay (F2 in 3D) = d23/2-p24/2
;cnst20: Ce chemical shift [141 ppm]
;cnst21: Cg chemical shift more [108 ppm]
;cnst22: Cd chemical shift [124 ppm]
;cnst23: Cg chemical shift [108 ppm]
;cnst24: middle of Cg and Cd [117 ppm]
;o2p: Cg chemical shift (cnst21)
;in0: 1/(2 * SW(Cg)) = DW(Cg)
;nd0: 2
;in10: 1/(4 * SW(N)) = (1/2) DW(N)
;nd10: 4
;in29: = in10
```

A. Appendix

```
;in30: = in10
;NS: 8 * n
;DS: >= 16
;td1: number of experiments in F1
;td2: number of experiments in F2      td2 max = 2 * d30 / in30
;FnMODE: States-TPPI (or TPPI) in F1
;FnMODE: States-TPPI (or TPPI) in F2
;cpds1: decoupling according to sequence defined by cpdprg1
;cpd3: decoupling according to sequence defined by cpdprg3
;pcpd1: f1 channel - 90 degree pulse for decoupling sequence
;pcpd3: f3 channel - 90 degree pulse for decoupling sequence

;use gradient ratio:   gp 1 : gp 2 : gp 3 : gp 4
;                       50 :   40 :   60 :   30

;for z-only gradients:
;gpz1: 50%
;gpz2: 40%
;gpz3: 60%
;gpz4: 30%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100
;gpnam4: SINE.100

; cs061122
```


B. References

References are in alphabetical order. References with more than one author are sorted according to the first author.

- D. J. Abraham and A. J. Leo (1987). Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. *Proteins*, 2:130–152.
- K. Ackermann (2003). *NMR-spektroskopische Untersuchungen einer Lysozym-Mutante sowie neu entwickelten Lanthanide-bindenden Peptid-Tags*. Diploma thesis, Johann Wolfgang Goethe-Universität, Frankfurt am Main.
- V. R. Agashe and F. U. Hartl (2000). Roles of molecular chaperones in cytoplasmic protein folding. *Semin. Cell. Dev. Biol.*, 11:15–25.
- H. M. Al-Hashimi, H. Valafar, M. Terrell, E. R. Zartler, M. K. Eidsness, and J. H. Prestegard (2000). Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J. Magn. Reson.*, 143:402–406.
- A. Almond and J. B. Axelsen (2002). Physical interpretation of residual dipolar couplings in neutral aligned media. *J. Am. Chem. Soc.*, 124:9986–9987.
- C. B. Anfinsen (1973). Principles that govern the folding of protein chains. *Science*, 181:223–230.
- V. L. Arcus, S. Vuilleumier, S. M. Freund, M. Bycroft, and A. R. Fersht (1995). A comparison of the pH, urea, and temperature-denatured states of barnase by heteronuclear nmr: implications for the initiation of protein folding. *J. Mol. Biol.*, 254:305–321.
- T. Argirevic (2006). *Charakterisierung des konformationellen Ensembles des entfalteten Lysozyms*. Diploma thesis, Johann Wolfgang Goethe-Universität, Frankfurt am Main.

B. References

- J. Bargon (2006). The discovery of chemically induced dynamic polarization (CIDNP). *Helv. Chim. Acta*, 89:2082–2102.
- J. Bargon, H. Fischer, and U. Johnsen (1967). Kerresonanz-Emissionslinien während rascher Radikalreaktionen. *Z. Naturforschg. A*, 22:1551–1555.
- C. Bartels, T. H. Xia, M. Billeter, P. Güntert, and K. Wüthrich (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR*, 6:1–10.
- A. Bax and D. G. Davis (1985). MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.*, 65:355–360.
- A. Bax, R. H. Griffey, and B. L. Hawkins (1983). Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J. Magn. Reson.*, 55:301–315.
- P. Bernadó, C. W. Bertoncini, C. Griesinger, M. Zweckstetter, and M. Blackledge (2005a). Defining long-range order and local disorder in native α -synuclein using residual dipolar couplings. *J. Am. Chem. Soc.*, 127:17968–17969.
- P. Bernadó, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, and M. Blackledge (2005b). A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA*, 102:17002–17007.
- I. Bertini, I. C. Felli, and C. Luchinat (2000). Lanthanide induced residual dipolar couplings for the conformational investigation of peripheral $^{15}\text{NH}_2$ moieties. *J. Biomol. NMR*, 18:347–355.
- C. W. Bertoncini, Y.-S. Jung, C. O. Fernandez, W. Hoyer, C. Griesinger, T. M. Jovin, and M. Zweckstetter (2005). Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. *Proc. Natl. Acad. Sci. USA*, 102:1430–1435.
- S. Bhattacharjya and P. Balaram (1997). Effects of organic solvents on protein structures: observation of a structured helical core in hen egg-white lysozyme in aqueous dimethylsulfoxide. *Proteins*, 29:492–507.
- N. S. Bhavesh, S. C. Panchal, and R. V. Hosur (2001). An efficient high-throughput resonance assignment procedure for structural genomics and protein folding research by NMR. *Biochemistry*, 40:14727–14735.

- M. Blackledge (2005). Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog. Nucl. Magn. Reson. Spectrosc.*, 46:23–61.
- C. C. Blake, D. F. Koenig, G. A. Mair, A. C. North, D. C. Phillips, and V. R. Sarma (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature*, 206:757–761.
- F. J. Blanco, L. Serrano, and J. D. Forman-Kay (1998). High populations of non-native structures in the denatured state are compatible with the formation of the native folded state. *J. Mol. Biol.*, 284:1153–1164.
- G. Bodenhausen and D. J. Ruben (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.*, 69:185–189.
- D. D. Boehr, H. J. Dyson, and P. E. Wright (2006). An NMR perspective on enzyme dynamics. *Chem. Rev.*, 106:3055–3079.
- R. W. Broadhurst, C. M. Dobson, P. J. Hore, S. E. Radford, and M. L. Rees (1991). A photochemically induced dynamic nuclear polarization study of denatured states of lysozyme. *Biochemistry*, 30:405–412.
- J. Buchner and T. Kiefhaber, eds. (2005). *Handbook of Protein Folding*. Wiley-VCH, Weinheim, Germany.
- M. Buck, J. Boyd, C. Redfield, D. A. MacKenzie, D. J. Jeenes, D. B. Archer, and C. M. Dobson (1995a). Structural determinants of protein dynamics: analysis of ^{15}N NMR relaxation measurements for main-chain and side-chain nuclei of hen egg white lysozyme. *Biochemistry*, 34:4041–4055.
- M. Buck, S. E. Radford, and C. M. Dobson (1993). A partially folded state of hen egg white lysozyme in trifluoroethanol: structural characterization and implications for protein folding. *Biochemistry*, 32:669–678.
- M. Buck, H. Schwalbe, and C. M. Dobson (1995b). Characterization of conformational preferences in a partly folded protein by heteronuclear NMR spectroscopy: assignment and secondary structure analysis of hen egg-white lysozyme in trifluoroethanol. *Biochemistry*, 34:13219–13232.
- M. Buck, H. Schwalbe, and C. M. Dobson (1996). Main-chain dynamics of a partially folded protein: ^{15}N NMR relaxation measurements of hen egg white lysozyme denatured in trifluoroethanol. *J. Mol. Biol.*, 257:669–683.

B. References

- R. Bussell and D. Eliezer (2001). Residual structure and dynamics in Parkinson's disease-associated mutants of α -synuclein. *J. Biol. Chem.*, 276:45996–46003.
- R. E. Canfield (1963). The amino acid sequence of egg white lysozyme. *J. Biol. Chem.*, 238:2698–2707.
- J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton (1996). *Protein NMR Spectroscopy: principles and practice*. Academic Press, San Diego, CA, USA.
- F. Chiti and C. M. Dobson (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.*, 75:333–366.
- M.-K. Cho, H.-Y. Kim, P. Bernado, C. O. Fernandez, M. Blackledge, and M. Zweckstetter (2007). Amino acid bulkiness defines the local conformations and dynamics of natively unfolded α -synuclein and tau. *J. Am. Chem. Soc.*, 129:3032–3033.
- J. J. Chou, S. Gaemers, B. Howder, J. M. Louis, and A. Bax (2001). A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. *J. Biomol. NMR*, 21:377–382.
- G. M. Clore, M. R. Starich, and A. M. Gronenborn (1998). Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am. Chem. Soc.*, 120:10571–10572.
- G. L. Closs (1969). Mechanism explaining nuclear spin polarizations in radical combination reactions. *J. Am. Chem. Soc.*, 91:4552–4554.
- R. T. Clubb, V. Thanabal, and G. Wagner (1992). A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^1\text{H}^{\text{N}}$, ^{15}N , and $^{13}\text{C}'$ chemical shifts in ^{15}N - ^{13}C -labelled proteins. *J. Magn. Reson.*, 97:213–217.
- J. S. Cohen and O. Jardetzky (1968). Nuclear magnetic resonance studies of the structure and binding sites of enzymes. II. Spectral assignments and inhibitor binding in hen egg-white lysozyme. *Proc. Natl. Acad. Sci. USA*, 60:92–99.
- E. S. Collins, J. Wirmer, K. Hirai, H. Tachibana, S. Segawa, C. M. Dobson, and H. Schwalbe (2005). Characterisation of disulfide-bond dynamics in non-native states of lysozyme and its disulfide deletion mutants by NMR. *Chembiochem*, 6:1619–1627.
- T. E. Creighton (1997). How important is the molten globule for correct protein folding? *Trends Biochem. Sci.*, 22:6–10.

- K. T. Dayie and G. Wagner (1994). Relaxation-rate measurements for ^{15}N - ^1H groups with pulsed-field gradients and preservation of coherence pathways. *J. Magn. Reson. A*, 111:121–126.
- C. Deverell, R. E. Morgan, and J. H. Strange (1970). Studies of chemical exchange by nuclear magnetic relaxation in the rotating frame. *Mol. Phys.*, 18:553–559.
- K. Ding, J. M. Louis, and A. M. Gronenborn (2004). Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR. *J. Mol. Biol.*, 335:1299–1307.
- C. M. Dobson (2003). Protein folding and misfolding. *Nature*, 426:884–890.
- C. M. Dobson (2004). Experimental investigation of protein folding and misfolding. *Methods*, 34:4–14.
- C. M. Dobson and P. J. Hore (1998). Kinetic studies of protein folding using NMR spectroscopy. *Nat. Struct. Biol.*, 5 Suppl.:504–507.
- E. Duchardt (2005). *Development of new Parameters for Structure Determination and Dynamic Investigations on Biomacromolecules by NMR*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.
- H. J. Dyson and P. E. Wright (2004). Unfolded proteins and protein folding studied by NMR. *Chem. Rev.*, 104:3607–3622.
- R. J. Ellis (2001a). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.*, 11:114–119.
- R. J. Ellis (2001b). Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.*, 26:597–604.
- R. J. Ellis (2006). Molecular chaperones: assisting assembly in addition to folding. *Trends Biochem. Sci.*, 31:395–401.
- L. Emsley and G. Bodenhausen (1992). Optimization of shaped selective pulses for NMR using a quaternion description of their overall propagators. *J. Magn. Reson.*, 97:135–148.
- P. Fan, C. Bracken, and J. Baum (1993). Structural characterization of monellin in the alcohol-denatured state by NMR: evidence for beta-sheet to alpha-helix conversion. *Biochemistry*, 32:1573–1582.

B. References

- N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*, 33:5984–6003.
- W. Fieber, S. Kristjansdottir, and F. M. Poulsen (2004). Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings. *J. Mol. Biol.*, 339:1191–1199.
- K. M. Fiebig, H. Schwalbe, M. Buck, L. J. Smith, and C. M. Dobson (1996). Toward a description of the conformations of denatured states of proteins. comparison of a random coil model with NMR measurements. *J. Phys. Chem.*, 100:2661–2666.
- B. Fischer (1996). Folding of lysozyme. *EXS*, 75:143–161.
- E. Fischer (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dt. Chem. Ges.*, 27:2985–2993.
- A. Fleming (1922). On a remarkable bacteriolytic element found in tissues and secretions. *Proc. Roy. Soc. Ser. B*, 93:306–317.
- M. K. Frank, G. M. Clore, and A. M. Gronenborn (1995). Structural and dynamic characterization of the urea denatured state of the immunoglobulin binding domain of streptococcal protein G by multidimensional heteronuclear NMR spectroscopy. *Prot. Sci.*, 4:2605–2615.
- R. B. Freedman, P. Klappa, and L. W. Ruddock (2002). Protein disulfide isomerases exploit synergy between catalytic and specific binding domains. *EMBO Rep.*, 3:136–140.
- E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch (2005). *The Proteomics Protocols Handbook*, chap. "Protein Identification and Analysis Tools on the ExPASy Server", 571–607. Humana Press, New Jersey, NJ, USA.
- E. B. Gebel, K. Ruan, J. R. Tolman, and D. Shortle (2006). Multiple alignment tensors from a denatured protein. *J. Am. Chem. Soc.*, 128:9310–9311.
- J. R. Gillespie and D. Shortle (1997a). Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.*, 268:158–169.

- J. R. Gillespie and D. Shortle (1997b). Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.*, 268:170–184.
- S. Goda, K. Takano, Y. Yamagata, R. Nagata, H. Akutsu, S. Maki, K. Namba, and K. Yutani (2000). Amyloid protofilament formation of hen egg lysozyme in highly concentrated ethanol solution. *Protein Sci.*, 9:369–375.
- T. D. Goddard and D. G. Kneller (2006). *SPARKY 3*. University of California, San Francisco, USA.
- J. Graf (2006). *Konformationsanalyse von kurzen alaninbasierten Modellpeptiden*. PhD thesis, Johann Wolfgang Goethe-Universität, Frankfurt am Main.
- J. Graf, P. H. Nguyen, G. Stock, and H. Schwalbe (2007). Structure and Dynamics of the Homologous Series of Alanine Peptides: A Joint Molecular Dynamics/NMR Study. *J. Am. Chem. Soc.*, 129:1179–1189.
- S. B. Grimshaw (1999). *Novel Approaches to Characterising Native and Denatured Proteins by NMR*. PhD thesis, University of Oxford, United Kingdom.
- S. Grzesiek and A. Bax (1992a). Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J. Am. Chem. Soc.*, 114:6291–6293.
- S. Grzesiek and A. Bax (1992b). Improved 3D triple-resonance NMR techniques applied to a 31kDa protein. *J. Magn. Reson.*, 96:432–440.
- S. Grzesiek and A. Bax (1993). Amino acid type determination in the sequential assignment procedure of uniformly $^{13}\text{C}/^{15}\text{N}$ -enriched proteins. *J. Biomol. NMR*, 3:185–204.
- D. Hamada and Y. Goto (2005). *Handbook of Protein Folding* (eds. J. Buchner and T. Kiefhaber), chap. "Alcohol- and Salt-induced Partially Folded Intermediates". Wiley-VCH, Weinheim, Germany.
- M. R. Hansen, L. Mueller, and A. Pardi (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.*, 5:1065–1074.

B. References

- R. L. Henrikson (1971). The selective S-methylation of sulfhydryl groups in proteins and peptides with methyl-p-nitrobenzenesulfonate. *J. Biol. Chem.*, 246:4090–4096.
- J. P. Hendrick and F. U. Hartl (1993). Molecular chaperone functions of heat-shock proteins. *Annu. Rev. Biochem.*, 62:349–384.
- M. Hennig, W. Bermel, A. Spencer, C. M. Dobson, L. J. Smith, and H. Schwalbe (1999). Side-chain conformations in an unfolded protein: χ_1 distributions in denatured hen lysozyme determined by heteronuclear ^{13}C , ^{15}N NMR spectroscopy. *J. Mol. Biol.*, 288:705–723.
- P. J. Hore and R. W. Broadhurst (1993). Photo-CIDNP of biopolymers. *Prog. Nucl. Magn. Reson. Spectrosc.*, 25:345–402.
- M. Hoshino, Y. Hagihara, D. Hamada, M. Kataoka, and Y. Goto (1997). Trifluoroethanol-induced conformational transition of hen egg-white lysozyme studied by small-angle X-ray scattering. *FEBS Lett.*, 416:72–76.
- Y. Ishii, M. A. Markus, and R. Tycko (2001). Controlling residual dipolar couplings in high-resolution NMR of proteins by strain induced alignment in a gel. *J. Biomol. NMR*, 21:141–151.
- V. A. Jaravine, A. T. Alexandrescu, and S. Grzesiek (2001). Observation of the closing of individual hydrogen bonds during tfe-induced helix formation in a peptide. *Protein Sci.*, 10:943–950.
- A. K. Jha, A. Colubri, K. F. Freed, and T. R. Sosnick (2005). Statistical coil model of the unfolded state: resolving the reconciliation problem. *Proc. Natl. Acad. Sci. USA*, 102:13099–13104.
- W. C. Johnson (1990). Protein secondary structure and circular dichroism: a practical guide. *Proteins*, 7:205–214.
- R. Kaptein (1971). Simple rules for chemically induced dynamic nuclear polarization. *Chem. Commun.*, 732–733.
- R. Kaptein, K. Dijkstra, and K. Nicolay (1978). Laser photo-CIDNP as a surface probe for proteins in solution. *Nature*, 274:293–294.
- R. Kaptein and L. J. Oosterhoff (1969). Chemically induced dynamic nuclear polarization III (anomalous multiplets of radical coupling and disproportionation products). *Chem. Phys. Lett.*, 4:214–216.

- M. Karplus (1963). Vicinal proton coupling in nuclear magnetic resonance. *J. Am. Chem. Soc.*, 85:2870–2871.
- L. E. Kay, D. A. Torchia, and A. Bax (1989). Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry*, 28:8972–8979.
- L. E. Kay, G. Y. Xu, and T. Yamazaki (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson. A*, 109:129–133.
- R. Keller (2004). *The Computer Aided Resonance Assignment Tutorial*. CANTINA Verlag, Goldau, Switzerland.
- S. M. Kelly, T. J. Jess, and N. C. Price (2005). How to study proteins by circular dichroism. *Biochim. Biophys. Acta*, 1751:119–139.
- S. M. Kelly and N. C. Price (2000). The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.*, 1:349–384.
- J. Klein-Seetharaman, M. Oikawa, S. B. Grimshaw, J. Wirmer, E. Duchardt, T. Ueda, T. Imoto, L. J. Smith, C. M. Dobson, and H. Schwalbe (2002). Long-range interactions within a nonnative protein. *Science*, 295:1719–1722.
- T. Knubovets, J. J. Osterhout, and A. M. Klibanov (1999). Structure of lysozyme dissolved in neat organic solvents as assessed by NMR and CD spectroscopies. *Biotechnol. Bioeng.*, 63:242–248.
- M. R. Krebs, D. K. Wilkins, E. W. Chung, M. C. Pitkeathly, A. K. Chamberlain, J. Zurdo, C. V. Robinson, and C. M. Dobson (2000). Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the beta-domain. *J. Mol. Biol.*, 300:541–549.
- T. Kühn and H. Schwalbe (2000). Monitoring the kinetics of ion-dependent protein folding by time-resolved NMR spectroscopy at atomic resolution. *J. Am. Chem. Soc.*, 122:6169–6174.
- U. Laemmli (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227:680–685.
- B. Lai, A. Cao, and L. Lai (2000). Organic cosolvents and hen egg white lysozyme folding. *Biochim. Biophys. Acta*, 1543:115–122.

B. References

- R. A. Laskey, B. M. Honda, A. D. Mills, and J. T. Finch (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature*, 275:416–420.
- R. G. Lawler (1967). Chemically induced dynamic nuclear polarization. *J. Am. Chem. Soc.*, 89:5519–5521.
- C. Levinthal (1969). How to fold graciously. In J. T. P. DeBrunner and E. Munck, eds., *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House, Monticello, Illinois*, pp. 22–24. University of Illinois Press, Champaign, IL, USA.
- M. Levitt (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *J. Mol. Biol.*, 104:59–107.
- D. R. Lide, ed. (1996). *CRC Handbook of Chemistry and Physics*. CRC Press, Boca Raton, FL, USA.
- M. A. Lietzow, M. Jamin, H. J. J. Dyson, and P. E. Wright (2002). Mapping long-range contacts in a highly unfolded protein. *J. Mol. Biol.*, 322:655–662.
- S. Lifson and A. Roig (1961). On the helix-coil transition in polypeptides. *J. Chem. Phys.*, 34:1963–1974.
- K. Lindorff-Larsen, S. Kristjansdottir, K. Teilum, W. Fieber, C. M. Dobson, F. M. Poulsen, and M. Vendruscolo (2004). Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc.*, 126:3291–3299.
- G. Lipari and A. Szabo (1982a). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.*, 104:4546–4559.
- G. Lipari and A. Szabo (1982b). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.*, 104:4559–4570.
- R. S. Lipsitz and N. Tjandra (2004). Residual dipolar couplings in NMR structure analysis. *Annu. Rev. Biophys. Biomol. Struct.*, 33:387–413.

- D. H. Live, D. G. Davis, W. C. Agosta, and D. Cowburn (1984). Long range hydrogen bond mediated effects in peptides: ^{15}N NMR study of Gramicidin S in water and organic solvents. *J. Am. Chem. Soc.*, 106:1939–1941.
- F. Löhr, V. Katsemi, M. Betz, J. Hartleib, and H. Rüterjans (2002). Sequence-specific assignment of histidine and tryptophan ring ^1H , ^{13}C and ^{15}N resonances in $^{13}\text{C}/^{15}\text{N}$ - and $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labelled proteins. *J. Biomol. NMR*, 22:153–164.
- J. A. Losonczi, M. Andrec, M. W. Fischer, and J. H. Prestegard (1999). Order matrix analysis of residual dipolar couplings using singular value decomposition. *J. Magn. Reson.*, 138:334–342.
- J. A. Losonczi and J. H. Prestegard (1998). Nuclear magnetic resonance characterization of the myristoylated, N-terminal fragment of ADP-ribosylation factor 1 in a magnetically oriented membrane array. *Biochemistry*, 37:706–716.
- M. Louhivuori, K. Pääkkönen, K. Fredriksson, P. Permi, J. Lounila, and A. Annala (2003). On the origin of residual dipolar couplings from denatured proteins. *J. Am. Chem. Soc.*, 125:15647–15650.
- S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. W. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson (2003). Structure validation by calpha geometry: phi,psi and cbeta deviation. *Proteins*, 50:437–450.
- H. Lu, M. Buck, S. E. Radford, and C. M. Dobson (1997). Acceleration of the folding of hen lysozyme by trifluoroethanol. *J. Mol. Biol.*, 265:112–117.
- P. Lundström and M. Akke (2005). Microsecond protein dynamics measured by $^{13}\text{C}_\alpha$ rotating-frame spin relaxation. *Chembiochem*, 6:1685–1692.
- C. E. Lyon, J. A. Jones, C. Redfield, C. M. Dobson, and P. J. Hore (1999). Two-dimensional ^{15}N - ^1H photo-CIDNP as a surface probe of native and partially structured proteins. *J. Am. Chem. Soc.*, 121:6505–6506.
- C. Ma and S. J. Opella (2000). 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.
- D. A. MacKenzie, J. A. Spencer, M. F. L. Gal-Coëffet, and D. B. Archer (1996). Efficient production from *Aspergillus niger* of a heterologous protein and an individual protein domain, heavy isotope-labelled, for structure-function analysis. *J. Biotechnol.*, 46:85–93.

B. References

- S. C. Makrides (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.*, 60:512–538.
- D. Marion, M. Ikura, R. Tschudin, and A. Bax (1989). Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.*, 85:393–399.
- L. J. Martin, M. J. Hähnke, M. Nitz, J. Wöhnert, N. R. Silvaggi, K. N. Allen, H. Schwalbe, and B. Imperiali (2007). Double-lanthanide-binding tags: Design, photophysical properties, and NMR applications. *J. Am. Chem. Soc.*, 129:7106–7113.
- A. Matagne, E. W. Chung, L. J. Ball, S. E. Radford, C. V. Robinson, and C. M. Dobson (1998). The origin of the alpha-domain intermediate in the folding of hen lysozyme. *J. Mol. Biol.*, 277:997–1005.
- A. Matagne and C. M. Dobson (1998). The folding process of hen lysozyme: a perspective from the 'new view'. *Cell. Mol. Life Sci.*, 54:363–371.
- D. Mathieu (2007). *Entwicklung neuer Methoden zur Messung von Residualen Dipolaren Kopplungen an Naturstoffen*. Diploma thesis, Johann Wolfgang Goethe-Universität, Frankfurt am Main.
- E. R. McCarney, J. E. Kohn, and K. W. Plaxco (2005). Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins. *Crit. Rev. Biochem. Mol. Biol.*, 40:181–189.
- C. C. McDonald and W. D. Phillips (1969). Proton magnetic resonance spectra of proteins in random-coil configurations. *J. Am. Chem. Soc.*, 91:1513–1521.
- D. H. Meadows, J. L. Markley, J. S. Cohen, and O. Jardetzky (1967). Nuclear magnetic resonance studies of the structure and binding sites of enzymes. I. Histidine residues. *Proc. Natl. Acad. Sci. USA*, 58:1307–1313.
- S. Meier, M. Strohmeier, M. Blackledge, and S. Grzesiek (2007). Direct observation of dipolar couplings and hydrogen bonds across a beta-hairpin in 8 M urea. *J. Am. Chem. Soc.*, 129:754–755.
- G. Merlini and V. Bellotti (2005). Lysozyme: a paradigmatic molecule for the investigation of protein structure, function and misfolding. *Clin. Chim. Acta*, 357:168–172.

- S. Mine, T. Ueda, Y. Hashimoto, Y. Tanaka, and T. Imoto (1999). High-level expression of uniformly ^{15}N -labeled hen lysozyme in *Pichia pastoris* and identification of the site in hen lysozyme where phosphate ion binds using NMR measurements. *FEBS Lett.*, 448:33–37.
- T. Mishima, T. Ohkuri, A. Monji, T. Imoto, and T. Ueda (2006). Amyloid formation in denatured single-mutant lysozymes where residual structures are modulated. *Protein Sci.*, 15:2448–2452.
- T. Mittag and J. D. Forman-Kay (2007). Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol.*, 17:3–14.
- R. Mohana-Borges, N. K. Goto, G. J. A. Kroon, H. J. Dyson, and P. E. Wright (2004). Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. *J. Mol. Biol.*, 340:1131–1142.
- K. H. Mok and P. J. Hore (2004). Photo-CIDNP NMR methods for studying protein folding. *Methods*, 34:75–87.
- G. T. Montelione, B. A. Lyons, S. D. Emerson, and M. Tashiro (1992). An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *J. Am. Chem. Soc.*, 114:10974–10975.
- S. Mori, C. Abeygunawardana, M. O. Johnson, and P. C. M. Vanzij (1995). Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson. B*, 108:94–98.
- O. B. Morozova, A. V. Yurkovskaya, R. Z. Sagdeev, K. H. Mok, and P. J. Hore (2004). Time-resolved CIDNP study of native-state bovine and human α -lactalbumins. *J. Phys. Chem. B*, 108:15355–15363.
- G. A. Morris and R. Freeman (1979). Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc.*, 101:760–762.
- D. R. Muhandiram and L. E. Kay (1994). Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson. B*, 103:203–216.

B. References

- M. D. Mukrasch, P. Markwick, J. Biernat, M. von Bergen, P. Bernadó, C. Griesinger, E. Mandelkow, M. Zweckstetter, and M. Blackledge (2007a). Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. *J. Am. Chem. Soc.*, 129:5235–5243.
- M. D. Mukrasch, M. von Bergen, J. Biernat, D. Fischer, C. Griesinger, E. Mandelkow, and M. Zweckstetter (2007b). The "jaws" of the tau-microtubule interaction. *J. Biol. Chem.*, 282:12230–12239.
- F. A. A. Mulder, N. R. Skrynnikov, B. Hon, F. W. Dahlquist, and L. E. Kay (2001). Measurement of slow (μ s-ms) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy: Application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.*, 123:967–975.
- D. Neri, M. Billeter, G. Wider, and K. Wüthrich (1992). NMR determination of residual structure in a urea-denatured protein, the 434-repressor. *Science*, 257:1559–1563.
- P. Neudecker, A. Zarrine-Afsar, W.-Y. Choy, D. R. Muhandiram, A. R. Davidson, and L. E. Kay (2006). Identification of a collapsed intermediate with non-native long-range interactions on the folding pathway of a pair of fyn SH3 domain mutants by NMR relaxation dispersion spectroscopy. *J. Mol. Biol.*, 363:958–976.
- M. Nic, J. Jirat, and B. Kosata (2006). *IUPAC Compendium of Chemical Terminology, Electronic version (<http://goldbook.iupac.org/C01079.html>)*. IUPAC, Research Triangle Park, NC, USA.
- O. Obolensky, K. Schlepckow, H. Schwalbe, and A. Solov'yov (2007). Theoretical framework for NMR residual dipolar couplings in unfolded proteins. *J. Biomol. NMR*, 39:1–16.
- S. Ohnishi, A. L. Lee, M. H. Edgell, and D. Shortle (2004). Direct demonstration of structural similarity between native and denatured eglin C. *Biochemistry*, 43:4064–4070.
- M. Ottiger, F. Delaglio, and A. Bax (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, 131:373–378.
- A. G. Palmer (2004). NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.*, 104:3623–3640.

- A. G. Palmer and D. A. Case (1992). Molecular dynamics analysis of NMR relaxation in a zinc-finger peptide. *J. Am. Chem. Soc.*, 114:9059–9067.
- A. G. Palmer, C. D. Kroenke, and J. P. Loria (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.*, 339:204–238.
- S. C. Panchal, N. S. Bhavesh, and R. V. Hosur (2001). Improved 3D triple resonance experiments, HNN and HN(C)N, for $^1\text{H}^{\text{N}}$ and ^{15}N sequential correlations in (^{13}C , ^{15}N) labeled proteins: application to unfolded proteins. *J. Biomol. NMR*, 20:135–147.
- M. B. Pepys, P. N. Hawkins, D. R. Booth, D. M. Vigushin, G. A. Tennent, A. K. Soutar, N. Totty, O. Nguyen, C. C. Blake, and C. J. Terry (1993). Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*, 362:553–557.
- B. Philipps, J. Hennecke, and R. Glockshuber (2003). FRET-based in vivo screening for protein folding and increased protein stability. *J. Mol. Biol.*, 327:239–249.
- M. Piotto, V. Saudek, and V. Sklenar (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR*, 2:661–665.
- K. W. Plaxco and C. M. Dobson (1996). Time-resolved biophysical methods in the study of protein folding. *Curr. Opin. Struct. Biol.*, 6:630–636.
- K. W. Plaxco, C. J. Morton, S. B. Grimshaw, J. A. Jones, M. Pitkeathly, I. D. Campbell, and C. M. Dobson (1997). The effects of guanidine hydrochloride on the ‘random’ coil conformations and NMR chemical shifts of the peptide series GGXGG. *J. Biomol. NMR*, 10:221–230.
- J. F. Povey, C. M. Smales, S. J. Hassard, and M. J. Howard (2007). Comparison of the effects of 2,2,2-trifluoroethanol on peptide and protein structure and function. *J. Struct. Biol.*, 157:329–338.
- C. R. Prabha and C. M. Rao (2004). Oxidative refolding of lysozyme in trifluoroethanol (TFE) and ethylene glycol: interfering role of preexisting alpha-helical structure and intermolecular hydrophobic interactions. *FEBS Lett.*, 557:69–72.
- J. H. Prestegard, H. M. Al-Hashimi, and J. R. Tolman (2000). NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Quart. Rev. Biophys.*, 33:371–424.

B. References

- J. H. Prestegard, C. M. Bougault, and A. I. Kishore (2004). Residual dipolar couplings in structure determination of biomolecules. *Chem. Rev.*, 104:3519–3540.
- O. B. Ptitsyn (1995). How the molten globule became. *Trends Biochem. Sci.*, 20:376–379.
- C. Redfield and C. M. Dobson (1988). Sequential ^1H NMR assignments and secondary structure of hen egg white lysozyme in solution. *Biochemistry*, 27:122–136.
- J. C. Rochet and P. T. Lansbury (2000). Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.*, 10:60–68.
- C. A. Rohl and R. L. Baldwin (1997). Comparison of NH exchange and circular dichroism as techniques for measuring the parameters of the helix-coil transition in peptides. *Biochemistry*, 36:8435–8442.
- C. A. Rohl, A. Chakrabartty, and R. L. Baldwin (1996). Helix propagation and N-cap propensities of the amino acids measured in alanine-based peptides in 40 volume percent trifluoroethanol. *Protein Sci.*, 5:2623–2637.
- I. J. Ropson and C. Frieden (1992). Dynamic NMR spectral analysis and protein folding: identification of a highly populated folding intermediate of rat intestinal fatty acid-binding protein by ^{19}F NMR. *Proc. Natl. Acad. Sci. USA*, 89:7222–7226.
- G. Saab-Rincón, P. J. Gualfetti, and C. R. Matthews (1996). Mutagenic and thermodynamic analyses of residual structure in the α subunit of tryptophan synthase. *Biochemistry*, 35:1988–1994.
- J. Sambrook and D. W. Russel (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York, NY, USA, 3rd edn.
- N. Sari, P. Alexander, P. N. Bryan, and J. Orban (2000). Structure and dynamics of an acid-denatured protein g mutant. *Biochemistry*, 39:965–977.
- H. J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, and S. Grzesiek (2000). Solution NMR of proteins within polyacrylamide gels: diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. *J. Biomol. NMR*, 18:303–309.
- M. Sattler, J. Schleucher, and C. Griesinger (1999). Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Mag. Res. Sp.*, 34:93–158.

- A. Saupe (1968). Recent results in the field of liquid crystals. *Angew. Chem. Int. Ed. Engl.*, 7:91–112.
- R. M. Scheek, S. Stob, R. Boelens, K. Dijkstra, and R. Kaptein (1984). Applications of two-dimensional ^1H nuclear magnetic resonance methods in photochemically induced dynamic nuclear polarisation spectroscopy. *Faraday Discuss. Chem. Soc.*, 78:245–256.
- R. M. Scheek, S. Stob, R. Boelens, K. Dijkstra, and R. Kaptein (1985). Applications of two-dimensional NMR methods in photochemically induced dynamic nuclear polarization spectroscopy. *J. Am. Chem. Soc.*, 107:705–706.
- J. E. Scheffler, C. E. Cottrell, and L. J. Berliner (1985). An inexpensive, versatile sample illuminator for photo-CIDNP on any NMR spectrometer. *J. Magn. Reson.*, 63:199–201.
- J. Schleucher, M. Sattler, and C. Griesinger (1993). Coherence selection by gradients without signal attenuation: Application to the three-dimensional HNCO experiment. *Angew. Chem. Int. Ed. Engl.*, 32:1489–1491.
- C. Schlörb (2003). *Untersuchungen zur Wechselwirkung von nicht-nativen Zuständen von Proteinen mit molekularen Chaperonen*. Diploma thesis, Johann Wolfgang Goethe-Universität, Frankfurt am Main.
- C. Schlörb, K. Ackermann, C. Richter, J. Wirmer, and H. Schwalbe (2005). Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states. *J. Biomol. NMR*, 33:95–104.
- H. Schwalbe, K. M. Fiebig, M. Buck, J. A. Jones, S. B. Grimshaw, A. Spencer, S. J. Glaser, L. J. Smith, and C. M. Dobson (1997). Structural and dynamical properties of a denatured protein. heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry*, 36:8977–8991.
- H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield, and L. J. Smith (2001). A refined solution structure of hen lysozyme determined using residual dipolar coupling data. *Prot. Sci.*, 10:677–688.
- S. Schwarzingler, G. J. Kroon, T. R. Foss, J. Chung, P. E. Wright, and H. J. Dyson (2001). Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.*, 123:2970–2978.

B. References

- S. Schwarzing, G. J. Kroon, T. R. Foss, P. E. Wright, and H. J. Dyson (2000). Random coil chemical shifts in acidic 8 M urea: implementation of random coil shift data in NMRView. *J. Biomol. NMR*, 18:43–48.
- B. R. Seavey, E. A. Farr, W. M. Westler, and J. L. Markley (1991). A relational database for sequence-specific protein NMR data. *J. Biomol. NMR*, 1:217–236.
- A. J. Shaka, P. B. Barker, and R. Freeman (1985). Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.*, 64:547–552.
- A. J. Shaka, C. J. Lee, and A. Pines (1988). Iterative schemes for bilinear operators; application to spin decoupling. *J. Magn. Reson.*, 77:274–293.
- P. E. Shaw (2002). Peptidyl-prolyl isomerases: a new twist to transcription. *EMBO Rep.*, 3:521–526.
- D. Shortle and M. S. Ackerman (2001). Persistence of native-like topology in a denatured protein in 8 M urea. *Science*, 293:487–489.
- V. Sklenar, M. Piotto, R. Leppik, and V. Saudek (1994). Gradient-tailored water suppression for ^1H - ^{15}N HSQC experiments optimized to retain full sensitivity. *J. Magn. Reson. A*, 102:241–245.
- C. M. Slupsky, L. N. Gentile, and L. P. McIntosh (1998). Assigning the NMR spectra of aromatic amino acids in proteins: analysis of two Ets pointed domains. *Biochem. Cell Biol.*, 76:379–390.
- L. J. Smith, K. A. Bolin, H. Schwalbe, M. W. MacArthur, J. M. Thornton, and C. M. Dobson (1996a). Analysis of main chain torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations. *J. Mol. Biol.*, 255:494–506.
- L. J. Smith, K. M. Fiebig, H. Schwalbe, and C. M. Dobson (1996b). The concept of a random coil. residual structure in peptides and denatured proteins. *Fold. Des.*, 1:R95–106.
- L. J. Smith, M. J. Sutcliffe, C. Redfield, and C. M. Dobson (1991). Analysis of ϕ and χ_1 torsion angles for hen lysozyme in solution from ^1H NMR spin-spin coupling constants. *Biochemistry*, 30:986–996.

- L. J. Smith, M. J. Sutcliffe, C. Redfield, and C. M. Dobson (1993). Structure of hen lysozyme in solution. *J. Mol. Biol.*, 229:930–944.
- A. Spencer, L. A. Morozov-Roche, W. Noppe, D. A. MacKenzie, D. J. Jeenes, M. Joniau, C. M. Dobson, and D. B. Archer (1999). Expression, purification, and characterization of the recombinant calcium-binding equine lysozyme secreted by the filamentous fungus *Aspergillus niger*: comparisons with the production of hen and human lysozymes. *Protein Expr. Purif.*, 16:171–180.
- S. Spera and A. Bax (1991). Empirical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.*, 113:5490–5492.
- H. Sternlicht and D. Wilson (1967). Magnetic resonance studies of macromolecules. I. Aromatic-methyl interactions and helical structure effects in lysozyme. *Biochemistry*, 6:2881–2892.
- L. Stryer (1998). *Biochemistry*. W.H. Freeman and Company, New York, NY, USA, 4th edn.
- T. Szyperski, P. Luginbühl, G. Otting, P. Güntert, and K. Wüthrich (1993). Protein dynamics studied by rotating frame ^{15}N spin relaxation times. *J. Biomol. NMR*, 3:151–164.
- V. Thanabal, D. O. Omecinsky, M. D. Reily, and W. L. Cody (1994). The ^{13}C chemical shifts of amino acids in aqueous solution containing organic solvents: Application to the secondary structure characterization of peptides in aqueous trifluoroethanol solution. *J. Biomol. NMR*, 4:47–59.
- N. Tjandra and A. Bax (1997). Direct measurement of distances and angles in biomolecules by nmr in a dilute liquid crystalline medium. *Science*, 278:1111–1114.
- M. Tollinger, N. R. Skrynnikov, F. A. Mulder, J. D. Forman-Kay, and L. E. Kay (2001). Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.*, 123:11341–11352.
- J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard (1995). Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl. Acad. Sci. USA*, 92:9279–9283.

B. References

- H. T. Tran and R. V. Pappu (2006). Toward an accurate theoretical framework for describing ensembles for proteins under strongly denaturing conditions. *Biophys. J.*, 91:1868–1886.
- Y. P. Tsentalovich, J. J. Lopez, P. J. Hore, and R. Z. Sagdeev (2002). Mechanisms of reactions of flavin mononucleotide triplet with aromatic amino acids. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 58:2043–2050.
- R. Tycko, F. J. Blanco, , and Y. Ishii (2000). Alignment of biopolymers in strained gels: A new way to create detectable dipole-dipole couplings in high-resolution biomolecular NMR. *J. Am. Chem. Soc.*, 122:9340–9341.
- V. N. Uversky (2002). What does it mean to be natively unfolded? *Eur. J. Biochem.*, 269:2–12.
- N. A. J. van Nuland, V. Forge, J. Balbach, and C. M. Dobson (1998). Real-time nmr studies of protein folding. *Acc. Chem. Res.*, 31:773–780.
- J. H. Viles, B. M. Duggan, E. Zaborowski, S. Schwarzingler, J. J. Huntley, G. J. Kroon, H. J. Dyson, and P. E. Wright (2001). Potential bias in NMR relaxation data introduced by peak intensity analysis and curve fitting methods. *J. Biomol. NMR*, 21:1–9.
- S. Vucetic, Z. Obradovic, V. Vacic, P. Radivojac, K. Peng, L. M. Iakoucheva, M. S. Cortese, J. D. Lawson, C. J. Brown, J. G. Sikes, C. D. Newton, and A. K. Dunker (2005). Disprot: a database of protein disorder. *Bioinformatics*, 21:137–140.
- G. Wagner (1993). NMR relaxation and protein mobility. *Cuur. Opin. Struct. Biol.*, 3:748–754.
- H. R. Ward and R. G. Lawler (1967). Nuclear magnetic resonance emission and enhanced absorption in rapid organometallic reactions. *J. Am. Chem. Soc.*, 89:5518–5519.
- E. J. W. Wensink, A. C. Hoffmann, P. J. van Maaren, and D. van der Spoel (2003). Dynamic properties of water/alcohol mixtures studied by computer simulation. *J. Chem. Phys.*, 119:7308–7317.
- J. Wirmer (2005). *Investigations of the kinetics of protein folding and the ensemble of conformations in non-native states of proteins by liquid NMR spectroscopy*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

- J. Wirmer, H. Berk, R. Ugolini, C. Redfield, and H. Schwalbe (2006a). Characterization of the unfolded state of bovine α -lactalbumin and comparison with unfolded states of homologous proteins. *Prot. Sci.*, 15:1397–1407.
- J. Wirmer, T. Kühn, and H. Schwalbe (2001). Millisecond time resolved photo-CIDNP NMR reveals a non-native folding intermediate on the ion-induced refolding pathway of bovine α -lactalbumin. *Angew. Chem.*, 113:4378–4381.
- J. Wirmer, W. Peti, and H. Schwalbe (2006b). Motional properties of unfolded ubiquitin: a model for a random coil protein. *J. Biomol. NMR*, 35:175–186.
- J. Wirmer, C. Schlörb, J. Klein-Seetharaman, R. Hirano, T. Ueda, T. Imoto, and H. Schwalbe (2004). Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew. Chem. Int. Ed. Engl.*, 43:5780–5785.
- J. Wirmer, C. Schlörb, and H. Schwalbe (2005). *Handbook of Protein Folding* (eds. J. Buchner and T. Kiefhaber), chap. "Conformation and Dynamics of Nonnative States of Proteins studied by NMR Spectroscopy". Wiley-VCH, Weinheim, Germany.
- D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, and B. D. Sykes (1995a). ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, 5:67–81.
- D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, and B. D. Sykes (1995b). ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, 6:135–140.
- D. S. Wishart and B. D. Sykes (1994a). The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR*, 4:171–180.
- D. S. Wishart and B. D. Sykes (1994b). Chemical shifts as a tool for structure determination. *Methods Enzymol.*, 239:363–392.
- D. S. Wishart, B. D. Sykes, and F. M. Richards (1991). Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.*, 222:311–333.
- D. S. Wishart, B. D. Sykes, and F. M. Richards (1992). The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry*, 31:1647–1651.

B. References

- M. Wittekind and L. Mueller (1993). HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J. Magn. Reson. B*, 101:201–205.
- J. Wöhnert, K. J. Franz, M. Nitz, B. Imperiali, and H. Schwalbe (2003). Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. *J. Am. Chem. Soc.*, 125:13338–13339.
- P. E. Wright and H. J. Dyson (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, 293:321–331.
- T. Yamazaki, J. D. Forman-Kay, and L. E. Kay (1993). Two-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.
- J. Yao, J. Chung, D. Eliezer, P. E. Wright, and H. J. Dyson (2001). NMR structural and dynamic characterization of the acid-unfolded state of apomyoglobin provides insights into the early events in protein folding. *Biochemistry*, 40:3561–3571.
- Q. Yi, M. L. Scalley-Kim, E. J. Alm, and D. Baker (2000). NMR characterization of residual structure in the denatured state of protein L. *J. Mol. Biol.*, 299:1341–1351.
- J. M. Yon (2002). Protein folding in the post-genomic era. *J. Cell. Mol. Med.*, 6:307–327.
- H. Zhang, S. Neal, and D. S. Wishart (2003). RefDB: a database of uniformly referenced protein chemical shifts. *J. Biomol. NMR*, 25:173–195.
- R. Zhou, M. Eleftheriou, A. K. Royyuru, and B. J. Berne (2007). Destruction of long-range interactions by a single mutation in lysozyme. *Proc. Natl. Acad. Sci. USA*, 104:5824–5829.
- M. Zweckstetter and A. Bax (2000). 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.

Zusammenfassung

Die allermeisten Proteine müssen, um ihre Funktion in ihrer natürlichen Umgebung wahrzunehmen, eine definierte dreidimensionale Struktur einnehmen, sie sind also in ihrem nativen Zustand gefaltet. Daneben gibt es aber sowohl Proteine, die in ihrem nativen und funktionalen Zustand unstrukturiert sind, als auch solche, die missgefaltet sind. Letztere sind häufig an schweren Krankheitsbildern beteiligt. Die Erforschung nicht-nativer Zustände von Proteinen hat bisher viele Erkenntnisse zum besseren Verständnis der Proteinfaltung und -fehlfaltung erbracht. Die Kernmagnetische Resonanzspektroskopie (NMR) ist die einzige bekannte Methode, die zur Aufklärung der strukturellen und dynamischen Eigenschaften solcher Zustände in atomarer Auflösung beitragen kann. Entfaltete und nicht-native Zustände von Proteinen müssen als Ensemble von sich schnell ineinander umwandelten Konformeren beschrieben werden, ihre beobachteten Eigenschaften sind somit populationsgewichtete Mittelungen. In der vorliegenden Arbeit wurde das 129 Aminosäurereste umfassende Lysozym aus dem Hühnereiweiß (HEWL) und Mutanten davon unter Bedingungen untersucht, unter denen diese Proteine permanent in nicht-nativen Zuständen vorliegen. Dies wurde in einem Ansatz durch die Reduktion der vier Disulfidbrücken und der anschließenden Methylierung der acht Cysteinreste („HEWL-S^{Me}“) dieses Proteins und in einem anderen Ansatz durch das Ersetzen der Cysteinreste durch Alanine erreicht („all-Ala-HEWL“). Unter solchen Bedingungen wurde in diesem Protein residuale Sekundär- und Tertiärstruktur beobachtet, es liegt also, anders als einige andere entfaltete Proteine, nicht als Zufallsknäuel (*random coil*) vor. Insbesondere die Existenz von hydrophoben Clustern und Wechselwirkungen zwischen in der Peptidkette weit entfernten Resten ist dabei bemerkenswert und war aus kombinierten Relaxations- und Mutationsexperimenten bekannt. Dabei spielen die Tryptophanreste eine wichtige Rolle. Zu den NMR-spektroskopischen Untersuchungen wurde nicht-natives HEWL in *E. coli*-Zellen in *inclusion bodies* isopenmarkiert exprimiert und daraus aufgereinigt. Nach nahezu vollständiger Zuordnung der $^1\text{H}^{\text{N}}$, $^{15}\text{N}^{\text{N}}$, $^{13}\text{C}_{\alpha}$, $^{13}\text{C}_{\beta}$, $^{13}\text{C}'$, $^1\text{H}_{\alpha}$ und $^1\text{H}_{\beta}$ Resonanzen konnte aus den sekundären chemischen Verschiebungen (d. h. den Differenzen zu chemischen Ver-

Zusammenfassung

schiebungen von *random coil*-Peptiden aus der Literatur) auf Bereiche innerhalb der Sequenz — vor allem rund um die sechs Tryptophanreste — geschlossen werden, die erhöhte α -helix-artige Konformationsanteile enthalten.

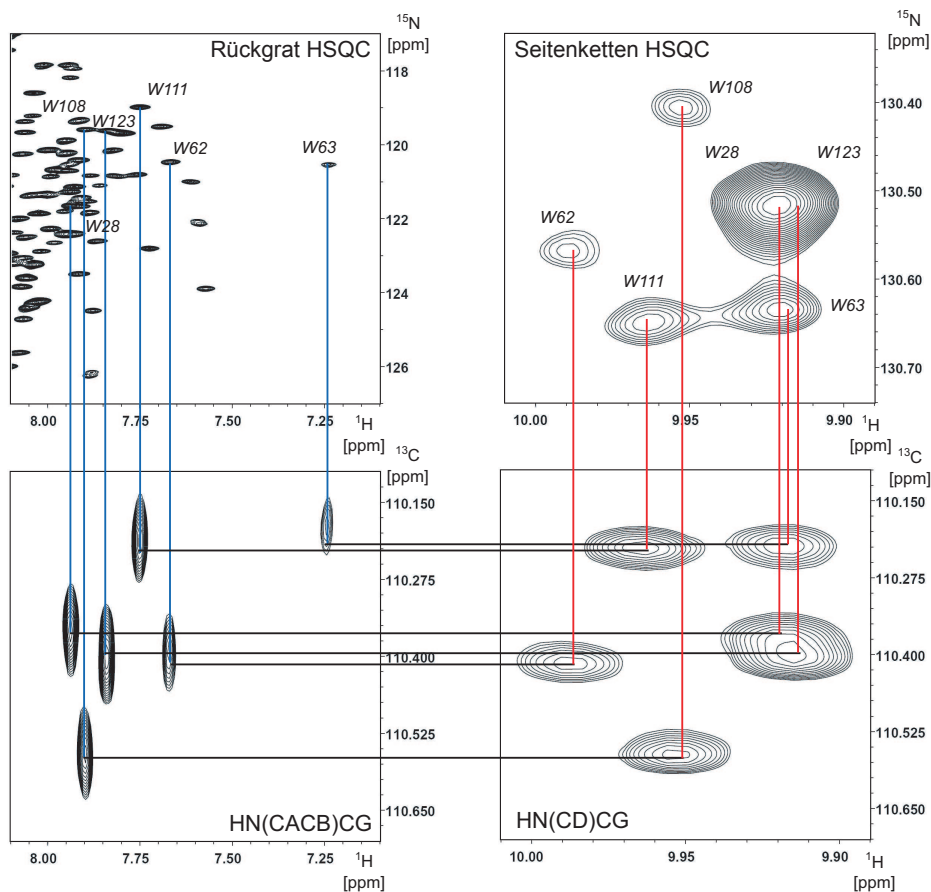


Abbildung B.1.: Ergebnis und Strategie der Zuordnung der Tryptophan-Seitenketten-Resonanzen in *all-Ala*-HEWL.

Um eine genauere Untersuchung der für die residuale Struktur in nicht-nativem Lysozym wichtigen Tryptophanreste mittels NMR zu ermöglichen, wurde eine Kombination aus zwei Pulssequenzen entwickelt, mittels derer sich die $^1\text{H}^{\text{N}}_{\epsilon}$ und $^{15}\text{N}_{\epsilon}$ Resonanzen in der Seitenkette über den $^{13}\text{C}_{\gamma}$ -Kern mit Hilfe der bekannten $^1\text{H}^{\text{N}}$ und $^{15}\text{N}^{\text{H}}$ Rückgratresonanzen zuordnen ließen (siehe Abbildung B.1).

Die Zuordnung der Indolseitenkettenresonanzen ermöglichte die Interpretation von zweidimensionalen $^{15}\text{N}/^1\text{H}$ photo-CIDNP-Experimenten, die über die unterschiedliche Zugänglichkeiten der Tryptophanseitenketten für einen Farbstoff (z.B. FMN) im Lösungsmittel berichteten und somit strukturelle Informationen lieferten.

Dabei wird mittels einem in das NMR-Röhrchen eingekoppelten Laserstrahl das FMN angeregt und schließlich über die Bildung eines Radikalpaares die Tryptophanseitenkette polarisiert. Die unterschiedlichen relativen Intensitäten der photo-CIDNP-Peaks deutet auf eine unterschiedliche Zugänglichkeit der Tryptophanseitenketten in nicht-nativem Lysozym hin. Dagegen sind die Zugänglichkeiten dieser Reste in der W62G-Mutante von HEWL fast alle ähnlich (Abbildung B.2). Dies bestätigte die zentrale Bedeutung dieses Restes für die hydrophoben Cluster, die aus vorherigen Relaxationsstudien bekannt war. Aus Diffusionsmessungen ist bekannt, dass die W62G-Mutante über einen größeren hydrodynamischen Radius verfügt als die nicht-nativen Zustände des Wildtyps, das konformationelle Ensemble zeigt also ausgedehntere Strukturen und somit sind auch die Tryptophanseitenketten exponierter. Eine Ausnahme scheint das Tryptophan 111 zu bilden, welches anders als das benachbarte W108 auch in der W62G-Mutante eine geringere Zugänglichkeit aufweist. Erstmals konnten auch zweidimensionale $^{13}\text{C}/^1\text{H}$ photo-CIDNP-Experimente in Lösung durchgeführt werden.

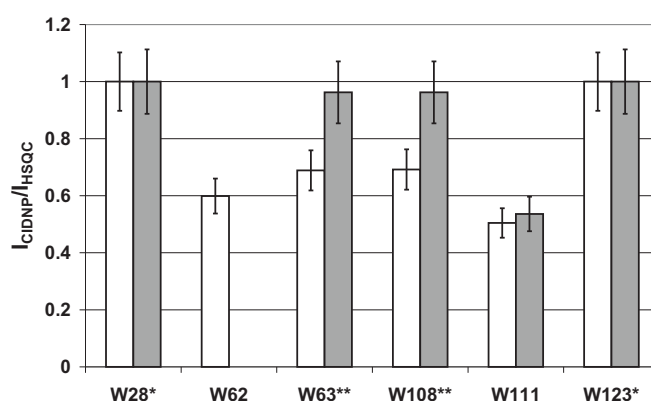


Abbildung B.2.: Normalisierte CIDNP/HSQC Intensitätsverhältnisse von *all-Ala*-HEWL (grau) und W62G-*all-Ala*-HEWL (weiß). *: Die W28 und W123 Signale sind degeneriert in *all-Ala*-HEWL und W62G-*all-Ala*-HEWL. **: Die W63 und W108 Signale sind degeneriert in W62G-*all-Ala*-HEWL.

Für *all-Ala*-HEWL ergab die Bestimmung von heteronuklearen R_1 und R_2 Relaxationsraten und des heteronuklearen $^1\text{H}^{15}\text{N}$ NOE einen Nachweis der verminderten Mobilität in den hydrophoben Clustern. Die Durchführung von Relaxationsdispersionsmessungen sowie die Messung von $R_{1\rho}$ Relaxationsraten erbrachten keinerlei Hinweise zu konformationellen Austausch auf der Mikro- bis Millisekunden-Zeitskala. Im Gegensatz zur W62G-Mutante zeigte eine W108G-Mutante keinen be-

Zusammenfassung

deutenden Einfluss auf das Gesamtbild der hydrophoben Cluster, sondern modulierte nur Cluster 5, in dem es sich selbst befindet (Abbildung B.3).

Die Messung von dipolaren Restkopplungen (RDCs) in partiell in gestreckten Polyacrylamidgelen ausgerichtetem *all-Ala*-HEWL zeigte erhöhte positive RDCs in den Bereichen der hydrophoben Cluster, vorallem in den Bereichen um die Tryptophane 28, 62/62 und 108/111 (Abbildung B.4), dies deutet auf eine Strukturierung in diesen Bereichen hin. In der W62G-Mutante waren die positiven RDCs hingegen größtenteils nicht zu beobachten.

Bisher ließen sich in Simulationen keine Ensembles von Konformeren generieren, deren berechnete RDCs den experimentellen Werten entsprechen. Ein viel versprechender aber zeitaufwendiger Ansatz hierzu ist die Generierung solcher Ensembles mittels Molekulardynamik-Simulationen basierend auf den experimentellen Relaxations-Daten und den Daten zur Lösungsmittelzugänglichkeit der verschiedenen Tryptophanreste. Die gemessenen RDCs können dann zur Validierung der simulierten konformationellen Ensembles genutzt werden.

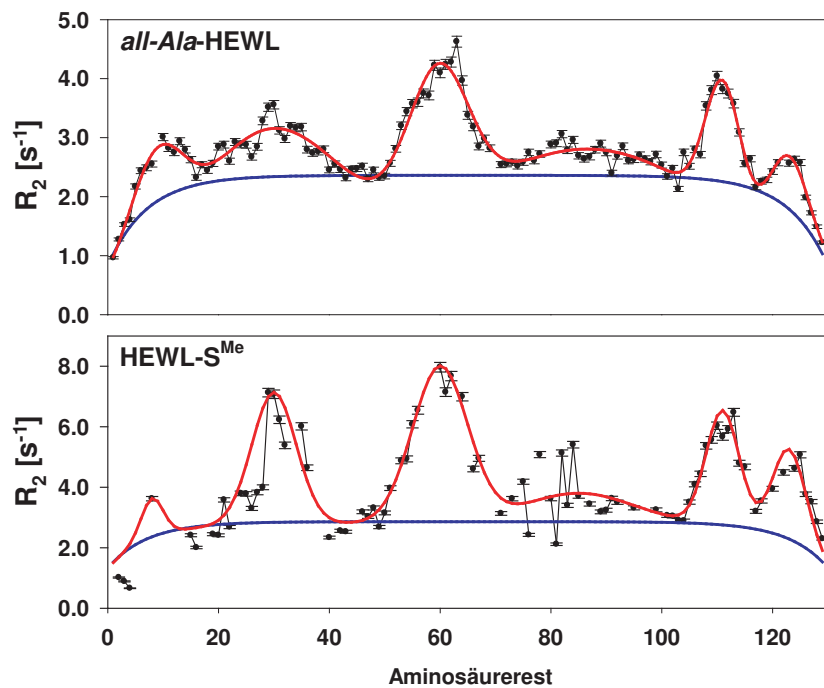


Abbildung B.3.: Vergleich der R_2 Relaxationsraten von *all-Ala*-HEWL und W108G-*all-Ala*-HEWL. Die W108G-Punktmutation ist mit einem roten Pfeil gekennzeichnet. Die rote Kurve zeigt einen Gauss-Fit der hydrophoben Cluster, während die blaue Kurve die nach dem *Segmentellen-Bewegungs-Modell* zu erwartende Kurve für ein Zufallsknäuel darstellt.

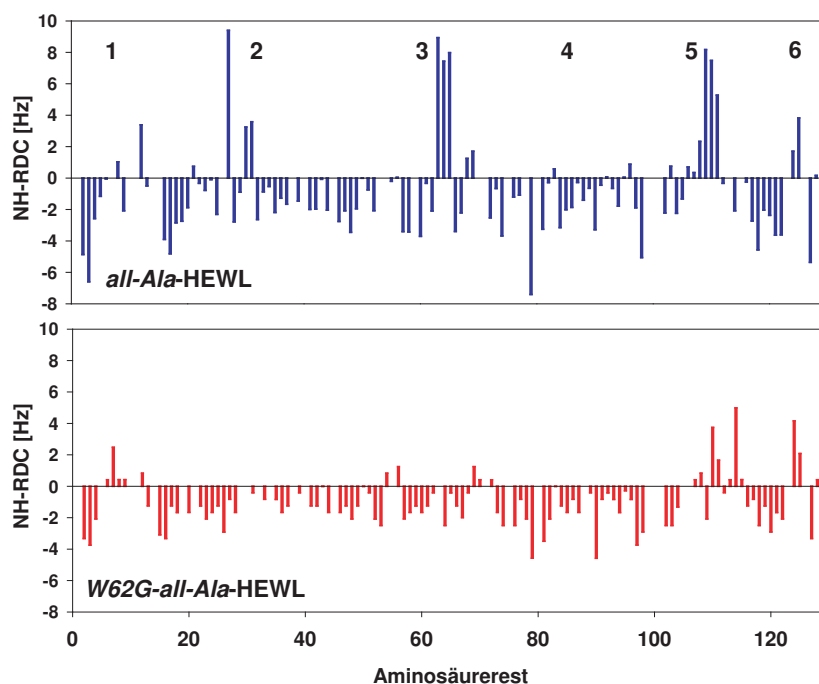


Abbildung B.4.: NH-RDCs gezeigt über die Sequenz von *all-Ala-HEWL* und *W62G-all-Ala-HEWL*. Die Position der sechs hydrophoben Cluster ist durch die jeweiligen Nummern gekennzeichnet.

In der Gegenwart von organischen Lösungsmitteln zeigte Lysozym einen hohen Grad an induzierter α -helicaler Struktur und nach einigen Stunden der Inkubation auch einen großen Anteil β -artiger Struktur, die als Vorstufe zur Fibrillenbildung angesehen werden kann. Mittels zweidimensionaler NMR-Spektroskopie konnte Lysozym in Gegenwart von Ethanol und Trifluorethanol näher charakterisiert werden und damit die prinzipielle Beobachtbarkeit von induzierten Strukturänderungen gezeigt werden.

Zusammenfassend ist festzustellen, dass in dieser Arbeit dynamische und strukturelle Eigenschaften verschiedener nicht-nativer Zustände von HEWL durch die Entwicklung neuer Methoden und die Weiterentwicklung bestehender Experimente mittels NMR-Spektroskopie charakterisiert werden konnten. Dies kommt dem besseren Verständnis dieser Zustände in Proteinen allgemein zu Gute, etwa mit Hilfe von Molekulardynamik-Simulationen.

Danksagung

Zum Abschluss meiner Doktorarbeit möchte ich mich bei einer Reihe von Leuten ganz herzlich für ihre Unterstützung und die abwechslungsreiche und schöne Zeit im AK Schwalbe bedanken:

Meinem Doktorvater Prof. Dr. Harald Schwalbe danke ich für seinen unermüdlichen und erfolgreichen Einsatz, die Bereitstellung einer fantastischen Infrastruktur, für seine Geduld, sein Vertrauen und seinen Enthusiasmus, was alle wissenschaftlichen Fragestellungen angeht.

Christian Richter danke ich für die Hilfe an den Spektrometern und die kurzweilige Zeit, die wir zusammen an diesen verbracht haben.

Julia Wirmer-Bartoschek und Jens Wöhnert danke ich für ihre Unterstützung, sowie die Diskussionen und Anregungen vor allem zu Beginn der Doktorarbeit.

Elke Stirnal und Sarah Mensch möchte ich für die kompetente Durchführung der HPLC sowie für Unterstützung im Labor danken.

Anna Paulus und Elena Hartmann danke ich für die hervorragende Arbeit, die sie bei der Organisation des AK Schwalbe aus dem Sekretariat heraus leisten.

Den wechselnden Mitglieder des legendären *Boys Room* danke ich für ihre Hilfe bei verschiedensten Gelegenheiten, vor allem aber für zahllose lustige und äußerst denkwürdige Momente: Boris Fürtig, Steffen Kaspar Grimm, Jan Ferner, Kai Schlepckow, Jürgen Graf, Serge Ilin, Martin Hähnke und Tomislav Argirevic sowie für jeweils einige Monate auch Stephan Rehm und Anna Fröhlich. Danke für die tolle und unübertreffliche Zeit!

Dank gebührt auch den Mitgliedern des ursprünglichen *Girls Rooms* und einigen ihrer Nachfolgerinnen für das tolle Gruppenklima und viele schöne Momente im AK Schwalbe: Julia Wirmer-Bartoschek, Elke Duchardt, Aphrodite Anastasiadis-Pool, Emily Collins, Karla Werner, Katrin Ackermann und Janina Buck.

Danksagung

Auch den vielen anderen aktuellen und ehemaligen Mitgliedern des AK Schwalbe danke ich für die nette Zeit und ihre Unterstützung. Den Mitgliedern der dritten und vierten Generation von Doktoranden und Diplomanden wünsche ich, dass sie das tolle Gruppenklima und Gemeinschaftsgefühl, das uns in den letzten Jahren geprägt hat, weiterhin aufrechterhalten und genießen können.

Den „üblichen Verdächtigen“ danke ich für die ungezählten Stunden im Kaffeeraum und die spannenden Unterhaltungen, die sich dort ergaben.

Boris Fürtig, Kai Schlepckow und Karla Werner sei auch für die Unterstützung bei den photo-CIDNP-Experimenten gedankt.

Karla Werner und Julia Wirmer-Bartoschek haben dankenswerterweise jeweils große Teile dieser Arbeit korrekturgelesen und mir mit ihren Anmerkungen sehr geholfen.

Johannes Gottfried „Zim“ Zimmermann danke ich für seine Hilfsbereitschaft, die kombiniert mit seinem Improvisationstalent und seinem Fundus an allerlei praktischen Dingen, manchmal sehr nützlich war.

Dr. Martin Blackledge danke ich für die Diskussionen zu den RDC in entfalteten Proteinen und für das zur Verfügung stellen der diversen Versionen des *flexible-meccano* Programms.

Jun.-Prof. Dr. Holger Gohlke danke ich für das Heraussuchen und die Überlassung der Trp-Trp-Torsionswinkel aus seiner Datenbank.

Ganz besonders möchte ich Karla Werner danken, die mir immer eine verständnisvolle ZuhörerIn, unermüdliche MitstreiterIn und vor allem sehr gute Freundin war und ist. Danke für all die schönen Momente und die aufgebrauchte Geduld.

Meinen Eltern danke ich für die vielfältige Unterstützung während meines Studiums und der Zeit als Doktorand, ohne sie wäre dies alles nicht möglich gewesen.

Frankfurt am Main, im Mai 2007

Curriculum Vitae

CHRISTIAN SCHLÖRB

Persönliche Daten:

Geburtsdatum: 26. November 1977
Geburtsort: Schotten
Nationalität: deutsch
Familienstand: ledig

Schulbildung:

1984 – 1988 Friedrich-Ebert-Schule, Schöneck (Grundschule)
1988 – 1997 Otto-Hahn-Schule, Hanau (Gymnasium), Abitur (Note: 1,2)

Zivildienst:

1997 – 1998 Mobiler Sozialer Hilfsdienst beim Arbeiter-Samariter-Bund, Karben

Studium:

1998 – 2003 Studium der Biochemie an der J. W. Goethe-Universität Frankfurt
09/2002 Vordiplom (Note: *sehr gut*)
10/02 – 04/03 Anfertigung der Diplomarbeit unter Anleitung von
Prof. Dr. Harald Schwalbe an der
Johann Wolfgang Goethe-Universität in Frankfurt am Main
Thema: „Untersuchungen zur Wechselwirkung von nicht-nativen
Zuständen von Proteinen mit molekularen Chaperonen“
04/2003 Diplom (Gesamtnote: *sehr gut*)

Promotion:

05/03 Beginn der Doktorarbeit unter Anleitung von
Prof. Dr. Harald Schwalbe an der
Johann Wolfgang Goethe-Universität in Frankfurt am Main

Curriculum Vitae

Meine **akademischen Lehrer** an der Johann Wolfgang Goethe-Universität Frankfurt am Main waren (in alphabetischer Reihenfolge):

Prof. Dr. Hashim Al-Hashimi, Prof. Dr. Ernst Bamberg, Prof. Dr. Jürgen Bereiter-Hahn, Prof. Dr. Karl-Dieter Entian, Prof. Dr. Dr. Hugo Fasold, Prof. Dr. Jürgen Feierabend, Prof. Dr. Werner Gnatzy, Prof. Dr. Michael Göbel, Prof. Dr. Christian Griesinger, Prof. Dr. Stephen K. Hashmi, Prof. Dr. Bernd O. Kolbesen, Prof. Dr. Bernd Ludwig, Prof. Dr. Walter E. Müller, Prof. Dr. Lutz Nover, Prof. Dr. Heinz D. Osiewacz, Prof. Dr. Roland Prinzinger, Prof. Dr. Reiner Protsch von Zieten, Prof. Dr. Hartmut Roskos, Prof. Dr. Heinz Rüterjans, Prof. Dr. Michael Sattler, Prof. Dr. Harald Schwalbe, Prof. Dr. Jörg Soppa, Prof. Dr. Annegret Starzinski-Powitz, Prof. Dr. Robert Tampé, Prof. Dr. Matthias Wagner, Prof. Dr. Dr. h. c. Christian Winter, Prof. Dr. Jens Wöhnert, Prof. Dr. Herbert Zimmermann

List of Publications

- I. Bertini, Y. K. Gupta, C. Luchinat, G. Parigi, C. Schlörb, and H. Schwalbe (2005). NMR spectroscopic detection of protein protons and longitudinal relaxation rates between 0.01 and 50 Mhz. *Angew. Chem. Int. Ed. Engl.*, 44:2223–2225.
<http://dx.doi.org/10.1002/anie.200462344>
- C. Schlörb, K. Ackermann, C. Richter, J. Wirmer, and H. Schwalbe (2005). Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states. *J. Biomol. NMR*, 33:95–104.
<http://dx.doi.org/10.1007/s10858-005-2063-y>
- C. Schlörb, S. Mensch, C. Richter, and H. Schwalbe (2006). Photo-CIDNP reveals differences in compaction of non-native states of lysozyme. *J. Am. Chem. Soc.*, 128:1802–1803.
<http://dx.doi.org/10.1021/ja056757d>
- J. Wirmer, C. Schlörb, J. Klein-Seetharaman, R. Hirano, T. Ueda, T. Imoto, and H. Schwalbe (2004). Modulation of compactness and long-range interactions of unfolded lysozyme by single point mutations. *Angew. Chem. Int. Ed. Engl.*, 43:5780–5785.
<http://dx.doi.org/10.1002/anie.200460907>
- J. Wirmer, C. Schlörb, and H. Schwalbe (2005). *Handbook of Protein Folding* (eds. J. Buchner and T. Kiefhaber), chap. "Conformation and Dynamics of Nonnative States of Proteins studied by NMR Spectroscopy". Wiley-VCH, Weinheim, Germany.